Engaged listeners: shared neural processing of powerful political speeches

Ralf Schmälzle,[1] Frank E. K. Häcker,[1] Christopher J. Honey,[2] and Uri Hasson[3]

[1]Department of Psychology, University of Konstanz, Baden-Württemberg, Germany, [2]Department of Psychology, University of Toronto, Toronto, ON, Canada, and [3]Department of Psychology, Princeton University, Princeton, NJ, USA

Powerful speeches can captivate audiences, whereas weaker speeches fail to engage their listeners. What is happening in the brains of a captivated audience? Here, we assess audience-wide functional brain dynamics during listening to speeches of varying rhetorical quality. The speeches were given by German politicians and evaluated as rhetorically powerful or weak. Listening to each of the speeches induced similar neural response time courses, as measured by inter-subject correlation analysis, in widespread brain regions involved in spoken language processing. Crucially, alignment of the time course across listeners was stronger for rhetorically powerful speeches, especially for bilateral regions of the superior temporal gyri and medial prefrontal cortex. Thus, during powerful speeches, listeners as a group are more coupled to each other, suggesting that powerful speeches are more potent in taking control of the listeners' brain responses. Weaker speeches were processed more heterogeneously, although they still prompted substantially correlated responses. These patterns of coupled neural responses bear resemblance to metaphors of resonance, which are often invoked in discussions of speech impact, and contribute to the literature on auditory attention under natural circumstances. Overall, this approach opens up possibilities for research on the neural mechanisms mediating the reception of entertaining or persuasive messages.

Keywords: rhetoric; speech; listening; fMRI; inter-subject correlation

INTRODUCTION

The rhetorical quality of a speech strongly affects its impact on the audience. During powerful speeches the entire audience listens intently and all attention is focused on the speech. During weak speeches, in contrast, the minds of individual listeners frequently drift away and the audience as a whole attends less to the speech. Alluding to this phenomenon, Plato once portrayed rhetoric as 'the art of ruling the minds of men', and mastering this art has always been considered a prerequisite for aspiring politicians. Indeed, the enormous potential of rhetoric is widely known and almost everyone remembers great speeches by e.g. Martin Luther King, John F. Kennedy or Winston Churchill. However, no previous study examined the neural reception of real-life speeches, and the associated neural processes remain unknown.

Prior work using functional magnetic resonance imaging (fMRI) amassed enormous knowledge about the neural correlates of linguistic processing. Over the last decades, elemental auditory and linguistic functions have been successfully localized to the bilateral auditory cortices, the adjacent superior and middle temporal gyri, angular and supramarginal gyrus, the temporo-parietal junction (TPJ), parietal lobule and the inferior frontal gyrus (IFG; Binder, *et al.*, 1997; Hickok and Poeppel, 2007; Price, 2012). Subsequent studies also began to uncover the neural systems involved in higher-order narrative processing under increasingly realistic conditions. This research revealed additional regions, such as the medial prefrontal cortex (mPFC), the precuneus and the dorsolateral prefrontal cortex, which are likely involved in processing the narrative and social content of a story (Ferstl *et al.*, 2008; Mar, 2011). Within this context, the inter-subject correlation (ISC) analysis technique offers unique opportunities for studying audience-wide brain responses to complex stimuli and

continuous speech in particular (Hasson *et al.*, 2004, 2010; Lerner *et al.*, 2011; Schmälzle *et al.*, 2013). ISC analysis measures the similarity of fMRI time courses across individuals who are exposed to the same stimulus, revealing where in the brain and to what extent neural responses during message reception concur across brains. Several previous studies successfully employed ISC analysis to study language processing (Wilson et al., 2008; Lerner *et al.*, 2011; Brennan *et al.*, 2012; Honey *et al.*, 2012; Boldt *et al.*, 2013; Regev *et al.*, 2013). The results show that listening to extended parts of spoken language evokes neural time courses that are similar across listeners in a widespread network encompassing auditory (e.g. bilateral A1+), linguistic (e.g. angular gyrus, IFG) and extralinguistic brain regions (e.g. mPFC, precuneus). Beyond validating previous results obtained with contrast-based methods and simpler stimuli, these findings add in novel information about temporal processes as well as their inter-subjective similarity: the former is essential for understanding how the brain integrates supra-word-level information essential for comprehension; the latter points to inter-subjective agreement among different persons, which can now be examined at the level of local brain processes. Mapping the ISC of neural time courses can thus provide new insight into the neural basis of higher-order language functions and communication.

This study adopted the ISC approach to examine common brain dynamics evoked by political speeches of varying rhetorical quality. Participants listened to recordings of real-life speeches while we measured continuous whole-brain activity profiles using fMRI. We expected that processing complex, but identical time-varying acoustic stimuli over time would prompt similar, hence correlated neural processes in auditory and linguistic regions. This should be the case for both, rhetorically weak as well as powerful speeches, because the basic processes involved in auditory and linguistic analysis occur in an obligatory manner. However, as argued above, powerful speeches are known to have a substantial impact on their audience by effectively capturing collective attention and commanding deeper processing. We reasoned that this should lead to a more coherent coupling of neural processes across listeners for powerful as compared with weak speeches. Demonstrating such effects would not only reveal the

neural dynamics evoked by an eminent, yet hardly studied stimulus class, but also advance the interdisciplinary study of how rhetoric impacts individuals and groups.

METHODS

Participants

Participants were recruited via posters, email-lists and personal contact. Twelve right-handed participants (mean age = 26.8, s.d. = 4.9, 3 females) took part in the main experiment. One participant did not comply with instructions and was immediately replaced. Six additional right-handed participants (mean age = 26.5, s.d. = 4.4; 3 females) took part in a control experiment, during which they listened to time-reversed versions of the same speeches. All participants had to meet fMRI eligibility criteria and provided written consent to the study protocol, which was approved by the local review board. For compensation, they either received course credits or monetary reimbursement.

Materials and procedure

The experimental stimuli were real-life political speeches that varied in rhetorical quality. To identify candidate speeches, we conducted an extensive search of online and offline archives for original recordings of speeches by German politicians. Based on pilot evaluations, we selected two powerful speeches (high rhetorical quality) that received positive pilot evaluations, prizes or strong reviews. Two corresponding weak speeches (low rhetorical quality) were selected from a database of parliamentary addresses based on pre-testing results. To confirm that participants' perceptions of rhetorical quality matched our preallocation, self-report ratings of rhetorical quality were probed at the end of the experiment. Powerful speeches were evaluated as significantly higher in terms of rhetorical quality ($t_{11} = 3.43$, $P < 0.01$; $M_{powerful} = 5.67$, s.d. = 0.34; $M_{weak} = 3.33$, s.d. = 0.44, Cohen's d = 3.43; question: 'How do you evaluate the speaker's overall rhetoric performance?'—translated from German; 7-point scale with anchors: '1—very bad', '7—very good'). In terms of content, the speeches covered several topics, but the respective main themes were the digitalization revolution, the erosion of freedom rights, pollution control and the escalation of violence in the Balkans. Transcripts of the speeches were also submitted to the 'linguistic analysis of word count' program (Pennebaker *et al.*, 2007; Wolf *et al.*, 2008). Consistent with studies on the rhetorical style of political speeches (e.g. Atkinson, 1984; Whissel and Sigelman, 2001; Pennebaker and Lay, 2002), this revealed that powerful speeches contained, for instance, more first-person personal pronouns, more words related to emotional, social or cognitive states, and more present-related words. Weaker speeches, in contrast, had a larger share of non-fluencies, longer sentences and more long-letter words. While a full rhetoric analysis is beyond the scope of this research, which deals with brain-to-brain coupling rather than stimulus-to-brain mappings, these exploratory analyses confirm and illustrate differences between speeches in terms of rhetoric characteristics.

For functional scanning, participants were instructed to keep their eyes closed and listen to the uninterrupted recordings of the speeches, which were presented using Presentation software (Neurobehavioral Systems Inc.) and high-fidelity MR-compatible earphones (MR Confon GmbH). These headphones function optimally in the bore of the scanner and reduce acoustic scanner noise. Foam padding was used to stabilize participants' heads and to further attenuate external noise. Mean sound volume was balanced across stimuli and the presentation of each speech started with a 15-s blank period (excluded from the analysis). The order of presentation was pseudo-randomized across participants to avoid confounds due to sequence effects. Finally, a post-scan interview revealed that all participants comprehended the speeches.

MRI acquisition and preprocessing

Participants were scanned in a 1.5 T Philips Intera MR System equipped with Power Gradients. Blood oxygenation level dependent (BOLD) contrast was acquired using a T2* weighted Fast Field Echo—Echo Planar Imaging sequence utilizing parallel scanning technique (Pruessmann *et al.*, 1999). In plane resolution of the axially acquired slices was 3×3 mm and the slice thickness was 3.5 mm (32 slices; ascending interleaved acquisition order; no gap; FOV = 240 mm; acquisition matrix: 80×80 voxels; TE = 40 ms; flip angle = $90°$). A TR of 2500 ms was used throughout the experiment. In addition to the functional scans, a T1-weighted high resolution anatomical scan was obtained for each participant (T1TFE; FOV = 256×256 mm; 200 sagittal slices; voxel size = $1 \times 1 \times 1$ mm).

fMRI data were preprocessed using Brainvoyager QX 2.1 (BrainInnovation Inc.) with standard parameters. Functional images were slice time corrected (sinc interpolation), 3D motion corrected (trilinear/sinc interpolation) and spatially smoothed with a Gaussian filter of 6-mm full width at half maximum value. Low frequencies (e.g. slow drift) were filtered out with up to six cycles (using the GLM with Fourier basis set option). All images were coregistered to the coplanar T1 scan and spatially normalized into a shared Talairach coordinate system, and we used the Talairach demon (Lancaster *et al.*, 2000) to determine anatomical locations. Further analyses were performed using the BVQXTools Matlab toolbox and in-house software. To have equal power of detecting differences across speeches, we included 170 continuous functional volumes from each speech, corresponding to 7 min of uninterrupted listening. For all speeches, the first and last 12 recorded time points were discarded to avoid saturation artifacts and signal transients. Functional data from the powerful as well as the weak speeches were concatenated to obtain continuous time courses of 292 functional volumes across which separate ISC analyses were calculated as described below.

ISC analysis

A two-step analysis scheme was adopted to map out regions that showed inter-subjectively coupled responses during listening. In the first step, we produced ISC maps for each speech type (i.e. powerful, weak). For each voxel, ISC is calculated as an average correlation, where the individual r_j are the Pearson correlations between that voxel's BOLD time course in one individual and the average of that voxel's BOLD time course in the remaining individuals. Because of the presence of long-range temporal autocorrelation in the BOLD signal, the statistical likelihood of each observed correlation was assessed using a bootstrapping procedure based on phase-randomization. The null hypothesis was that the BOLD signal in each voxel in each individual was independent of the BOLD signal values in the corresponding voxel in any other individual at any point in time (i.e. that there was no ISC between any time course pair). Phase randomization of each voxel time course was performed by applying a discrete Fourier transform to the signal, then randomizing the phase of each Fourier component and inverting the Fourier transformation. This procedure scrambles the phase of the BOLD time course but leaves its power spectrum intact. A distribution of 1000 bootstrapped average correlations was calculated for each voxel in the same manner as the empirical correlation maps described above, with bootstrapped correlation distributions calculated within subjects and then combined across subjects. The distributions of the bootstrapped correlations for each subject were approximately Gaussian, and thus the mean and standard deviations of the r_j distributions calculated for each subject under the null hypothesis were used to analytically estimate the distribution of the average correlation, R, under the null hypothesis. P values of the empirical average correlations (R values) were then computed by
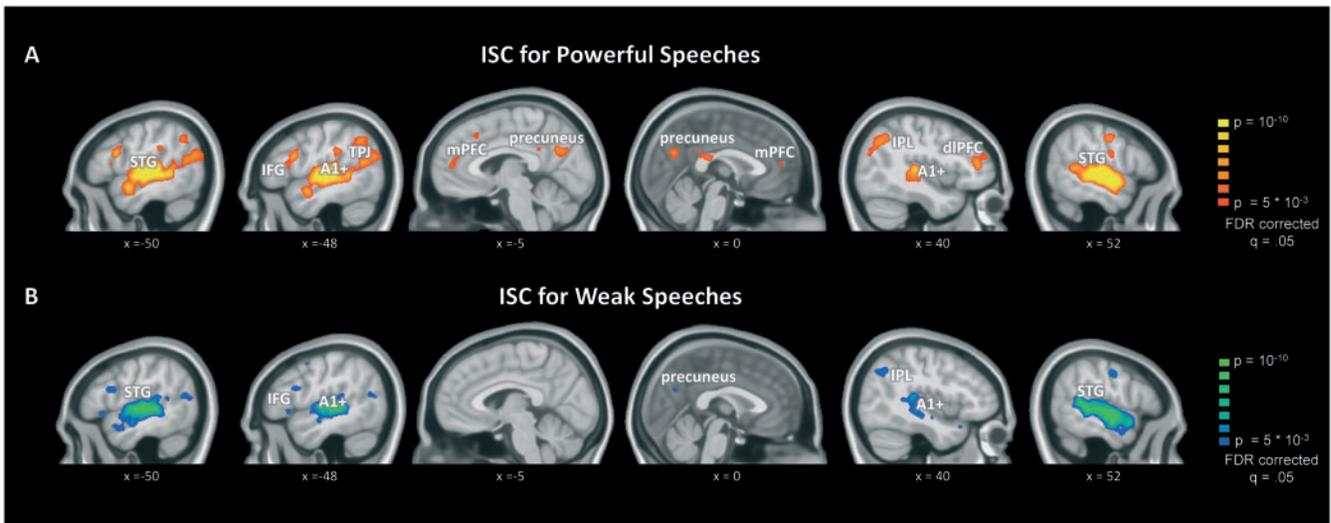
Fig. 1 Inter-subject correlations of brain activity during listening to real-life political speeches. (**A**) ISC result maps for powerful speeches. (**B**) ISC for weak speeches.

comparison with the null distribution of *R* values. Finally, we corrected for multiple statistical comparisons by controlling the false discovery rate (FDR; Benjamini and Hochberg, 1995) of the correlation maps with a q criterion of 0.05.

In the second analysis step, we contrasted the results from the first step to reveal differences in the strength of inter-subject brain coupling, i.e. whether more powerful speeches prompt more similar cross-listener neural responses. To achieve this, we extracted each individual listener's correlation values for the powerful and the weak speeches, transformed them to Fisher-*Z*-values, and submitted them to an un-directed repeated-measures *t*-test (FDR corrected).

Control experiment and analyses

The control experiment used identical procedures, with the only difference that the soundtrack of the speeches was reversed. This eliminates semantic content while retaining acoustic complexity and creates an unintelligible, but still speech-like stimulus. All analyses followed the same steps as the ones described for the main experiment. To ensure that combined analyses of the main (forward) and control (reversed presentation) experiment are robust with respect to differences in sample size, we also computed correlations between all pairs of participants (cf. Hasson *et al.* 2004).

Data were additionally examined within regions of interest (ROI) in order to integrate and compare the main and the control experiment. Independent ROIs were defined based on ISC results from a previous study in which people listened to a different naturalistic story (Honey *et al.*, 2012). Neural time courses were extracted from all four conditions (forward: powerful and weak; reversed: powerful and weak) by averaging across values within a 5-mm radius spherical volume around the following regions: bilateral superior temporal gyrus (STG; Talairach coordinates: mid-STG: ± 59, -16, 7; post-STG: $\pm$ 62, -28, 8), left IFG (-47, 16, 23), TPJ (-48, -59, 21), inferior parietal lobule (IPL: -38, -67, 35), precuneus (-3, 65, 27), and mPFC (2, 55, 9); finally a skull patch (-63, 30, -29) served as control region. To have equal sensitivity for comparing the forward ($N = 12$) and the reversed speech ($N = 6$) condition, and to increase reliability, correlations of ROI time courses were computed by randomly splitting subjects from each experiment into groups of three, averaging the time courses, and computing the average correlation coefficients for each condition.

RESULTS

We presented rhetorically powerful and weaker speeches to individual listeners and compared the strength of the resulting coupling among listeners' neural activity profiles via ISC analysis.

ISCs during listening to political speeches

First, we computed separate ISC analyses for powerful and weak speeches, respectively. Consistent with previous reports (Wilson *et al.*, 2008; Lerner *et al.*, 2011; Honey *et al.*, 2012), these analyses revealed a large network of brain areas exhibiting reliably coupled neural time courses across listeners that is in line with previous studies of narrative processing (Ferstl *et al.*, 2008; Mar, 2011; Specht, 2014). As illustrated in Figure 1, listening to real-life speeches prompted strongly correlated neural responses not only in early auditory areas, but also in linguistic and extralinguistic regions. Early auditory areas include primary auditory cortex. Linguistic areas include the STG, angular gyrus, TPJ and the IFG, and extralinguistic regions include the precuneus, medial and dorsolateral prefrontal cortex. In line with the linguistic nature of the stimuli, the left hemisphere showed overall stronger and more widespread inter-subjective neural coupling.

Contrasting ISC for powerful vs weak speeches

The difference between individual ISC maps (Figure 1) does not constitute a formal statistical test of the differences in ISC strength. To conduct this test, we extracted all individual ISC values for each voxel and statistically compared them between powerful *vs* weak speeches. The results are depicted in Figure 2, demonstrating significantly enhanced ISC in bilateral STG and mPFC.

Analysis of response amplitudes

A crucial difference between ISC-based fMRI and classical fMRI activation analyses is, that the former focuses on the similarity of temporal response profiles whereas the latter detects response amplitude differences between conditions (see Ben-Yakov *et al.*, 2012). To explore the relationship between response similarity (ISC) and response amplitude measures, we also assessed the amplitudes of neural responses toward the speeches. However, in continuous and naturalistic stimulation paradigms one cannot measure amplitude-to-baseline differences, as in studies using discrete stimuli [e.g. contrast: activation(word)−activation(silence)]. As a proxy, we therefore assessed the 'dynamic range'
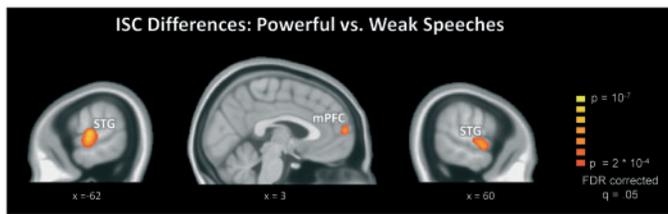
Fig. 2 Differences in ISC between powerful and weak speeches. Significantly enhanced ISC was found in bilateral STG regions and mPFC, indicating that powerful speeches prompt more consistent neural responses across listeners.

Fig. 3 (A) ISC results for the control experiment with unintelligible reversed speeches. (**A**) ISC during listening to reversed powerful speeches. (**B**) ISC during listening to reversed weak speeches.

of the signal, i.e. the standard deviation of the BOLD time series in each voxel. There was no evidence of differences in the 'dynamic range' for powerful *vs* weak speeches at corrected statistical thresholds. However, exploring the ISC within those regions that exhibited significantly different ISC during powerful *vs* weaker speeches (Figure 2: left and right STG; mPFC) revealed that powerful speeches had a higher dynamic range in left and right STG regions (t's = 2.96 and 2.90, P's < 0.05). This suggests that powerful speeches, which were processed more similarly across listeners, also evoked stronger responses in STG. For the mPFC-ROI, in contrast, the dynamic range did not differ between speeches ($t = 0.41$; $P = 0.69$), suggesting that signals in this region have indistinguishable amplitudes, but signals are collectively better aligned during powerful speeches.

Removing peaks and troughs

Inspection of the STG time courses revealed several marked peaks and troughs (cf. dynamic range differences), raising the question whether these prominent samples might influence correlations. To answer this, we removed time points corresponding to peaks or troughs from each region's fMRI signal, and recomputed ISC analyses for the remaining segments. Specifically, we removed all segments that exceeded values of ± 0.75 in the average z-scored time course. Of note, powerful speeches contained more such segments. Furthermore, by tracing back from the signal to the speech, we found that peaks and troughs corresponded well to culmination points of the political speeches. Critically, the results concurred with the earlier findings, showing that enhanced ISC for left and right STG, as well as mPFC (t's = 3.9, 4.5, and 10.7; P's < 0.005) was present even after removing these segments. This suggests that powerful speeches prompt more similar processing across the audience throughout the speech.

ISC during reversed speeches

The data from the control experiment, during which participants listened to time-reversed versions of the same speeches, produced a very different picture. Figure 3 shows that correlated neural time courses are still present in bilateral superior temporal cortex, but the overall ISC is confined to early auditory areas (A1+). Furthermore, directly comparing ISC for the reversed powerful *vs* reversed weak speeches via repeated measures t-tests did not reveal significant differences, suggesting that the enhanced ISC for powerful speeches is not a simple consequence of acoustic feature differences.

Region of interest analysis

The ROI analysis confirmed the previous whole brain analyses (i.e. maps in Figures 1–3). Looking at the solid orange bars in Figure 4 shows that intact (i.e. forward) powerful speeches prompted enhanced ISC compared with all other conditions (reverse powerful, forward and reverse weak speeches). This effect was seen for all regions, except for the control ROI defined in the skull. It was most pronounced for the left STG. These results indicate that when powerful speeches are
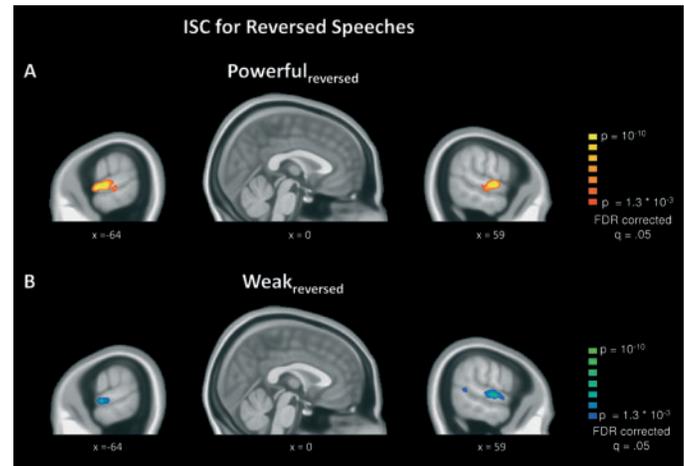
intelligible (i.e. forward listening), they prompt the strongest correlations, even though forward and reversed stimuli are identical in terms of acoustic complexity.

DISCUSSION

Numerous examples document that well-crafted speeches are very engaging for listeners and exert considerable control over the audience as a whole. This study examined this phenomenon from a neural perspective. The results reveal that real-life political speeches prompt correlated neural responses across the brains of listeners, and that this neural coupling is enhanced during more powerful speeches.

All speeches, regardless of rhetorical quality, elicited robust ISC in widespread brain regions involved in auditory, linguistic and higher-order integrative processes (for reviews see Ferstl *et al.*, 2008; Specht, 2014). These results align well with the knowledge gained from thousands of activation fMRI studies on language processing. Although all speeches prompted common neural processing in extended language networks, a direct contrast between ISC during powerful *vs* weaker speeches revealed significantly enhanced ISC in bilateral STG (Figure 2). A well-documented finding in the neuroimaging literature on auditory processing is that attention prompts intensified neural processing of the attended features or stimuli (for reviews see Alho *et al.*, 2014; Lee *et al.*, 2014). However, previous studies in this domain primarily employed instructed attention paradigms that require participants to attend actively and selectively to discrete sounds, syllables or words (e.g. Pugh *et al.*, 1996; Petkov *et al.*, 2004), or to extract relevant information from complex streams (e.g. based on spatial location or speaker identity during dichotic listening or multi-talker situations; Ding and Simon, 2012; Mesgarani and Chang, 2012). The current paradigm differs from these approaches because continuous real-life speeches were presented in a task-free and uninstructed manner, prompting a more natural and spontaneous allocation of attention (e.g. Lang *et al.*, 1997; Schupp *et al.*, 2006; Caporello Bluvas and Gentner, 2013). In particular, listening to the entire speeches may encourage more sustained vigilance rather than temporally focused attention, and it certainly entails concentration on one single stimulus rather than a selection between competing ones. Most importantly, differential effects for powerful *vs* weak speeches appear to result from self-defined relevance rather than task-related top-down or stimulus-related bottom-up factors (cf. Bradley 2009). Comparable relevance-driven attentional modulations of communicative signals have been reported across a very wide variety of tasks and stimuli
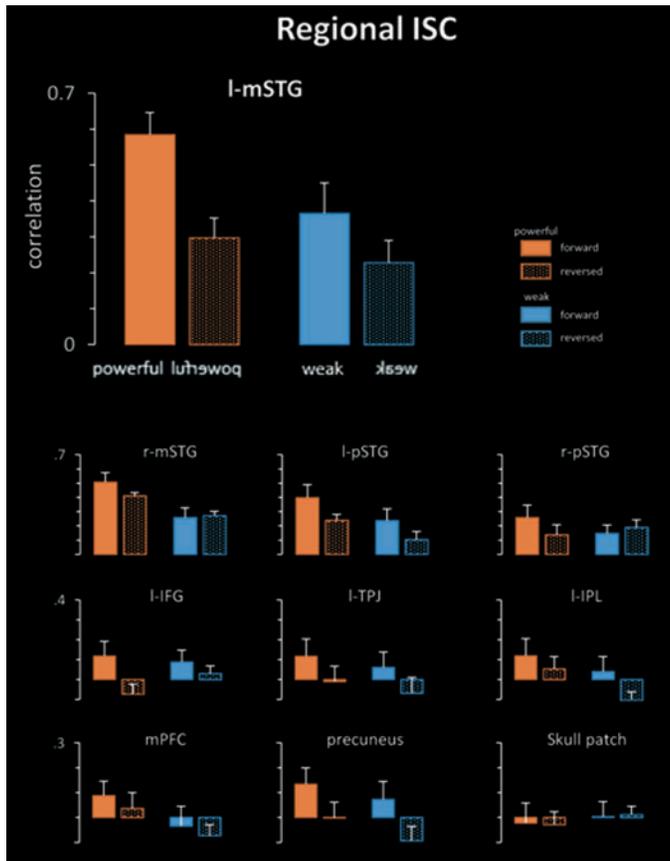
Fig. 4 Inter-subject correlations for powerful and weak speeches that were presented either normal (i.e. forward) or time-reversed, thus rendering the stimulus unintelligible. Bars represent mean ISC and error bars standard deviations. See text for further details.

(e.g. emotional words: Kissler *et al.*, 2007, prosody: Grandjean, *et al.*, 2005; Ethofer *et al.*, 2012; non-verbal vocal expressions: Warren *et al.*, 2006, symbolic gestures: Flaisch *et al.*, 2011, clashing moral statements: Van Berkum *et al.*, 2009). Furthermore, the current findings align well with fMRI data reported by Wallentin *et al.*, (2011), who used continuous self-report ratings of emotional intensity during a narrated story and found enhanced activity in extended regions of superior temporal cortex during emotionally arousing parts of the story. A very recent study extended these results toward inter-subject coupling measures, finding enhanced coupling of neural responses during emotionally arousing scene descriptions (Nummenmaa *et al.*, 2014).

Beyond the STG effects, comparing ISC between powerful and weak speeches also revealed differences in mPFC (Figure 2). Recent reviews and meta-analyses of the growing literature on narrative processing have all suggested a role for the mPFC (Ferstl *et al.*, 2008; Mar, 2011), which is involved in higher-order integrative processes related to information accumulation across time to support comprehension (Lerner *et al.*, 2011), following the social content of the story (Mar, 2011), and mutual understanding in real-life communication paradigms (e.g. Stephens *et al.*, 2010; Honey *et al.*, 2012). Of course, medial prefrontal areas are also discussed within the much broader contexts of self-related and socio-cognitive processing (Amodio and Frith, 2006; van Overwalle, 2009), valuation (Bartra *et al.*, 2013) and autobiographical memory (Spreng *et al.*, 2009), which are all critical for language understanding/communication and likely involved during listening to real-life political speeches. Interestingly, this includes even research on the neural mechanisms of effective health communication,

which linked mPFC activity to persuasion and behavior change (Falk *et al.*, 2010; Chua *et al.*, 2011). Although our study differs in many ways and we did not study possible subsequent effects of the speeches, our results appear compatible with these data, suggesting that differences in rhetorical quality modulate individuals' involvement with the speeches. Unlike STG regions, however, mPFC responses toward the different speeches had indistinguishable amplitudes, but differed in terms of alignment (see Lerner *et al.*, 2011; Pajula *et al.*, 2012). In sum, powerful speeches are characterized by enhanced ISC in mPFC, suggesting that neural processes involved in their receipt are collectively more aligned across the audience and represented more effectively.

The reversed speech control condition evoked only minimal ISC and no differences in ISC between powerful and weak speeches. From a communication perspective, the worst possible speech is one that is completely incomprehensible as this blocks the transmission of ideas. For instance, a recent study showed that when Russian-speaking participants listen to a Russian language recording, they exhibit widespread and strong neural coupling (Honey *et al.*, 2012). However, when English-speaking participants listen to the same stimulus, ISC in higher-order regions is completely abolished and confined to early auditory regions. The reversed speech control condition is similar to the situation of the English listeners for whom the Russian stimulus contains no meaning, and thus fails to engage regions of their brain beyond early auditory processing areas. To illustrate this, Figure 5 presents combined ISC result maps for forward and reversed speeches along with meta-analytic maps corresponding to several auditory and linguistic processes. These maps were derived from the Neurosynth database (Yarkoni *et al.*, 2011), which allows for large-scale automated meta-analysis of neuroimaging studies. Out of Neurosynth's 525 feature maps for a wide array of sensory, cognitive and affective processes, we selected the reverse inference maps for the following auditory and linguistic features: 'phonetic', 'prosodic', 'voice', 'syllables', 'speech', 'listening' and 'sentence' (more maps can be viewed at www.neurosynth.org). Comparing Figures 5A and B demonstrates that brain regions that were reliably engaged by listening to the forward speeches match very well with the combination of the depicted feature maps. This suggests that all speeches effectively recruited widespread language processing networks, with stronger and more widespread effects for powerful speeches. In contrast, the neural coupling evoked by reversed speeches no longer overlaps with higher-order feature maps (Figure 5B and C). Furthermore, although reversal of the soundtrack retains basic acoustic properties, ISC did not differ between reversed powerful *vs* reversed weak speeches. Finally, Figure 4 shows that the ISC advantage of powerful speeches in the l-STG was much larger during forward than during reversed presentation. This speaks against attentional bottom-up effects due to salient stimulus properties, which may result from the inevitably different physical structure of powerful and weaker speeches. Together, these analyses locate our findings, obtained from ISC analysis of real-life political speeches, in the context of the broader language processing literature. In sum, the results show that powerful speeches prompt the strongest and spatially most widespread neural coupling across listeners.

A promising application of the ISC approach may thus lie in assessing population-level effects of individual speeches on neural engagement. For instance, discussions of successful rhetoric often center around metaphors of 'resonance', implying intense and interrelated reactions among listeners. ISC analysis may offer a way to assess the strength of such interrelations across different brains as well as within individual functional brain regions. In previous studies, we found that a better alignment between the speaker's and listener's brain responses can predict the quality of the communication (Stephens *et al.*, 2010). Although we did not look at the

Fig. 5 (**A**) Illustration of ISC results for powerful (orange) and weaker (blue) speeches (overlap in purple). (**B**) Neurosynth reverse inference maps for several linguistic features. For visualization purposes, data were mapped to the PALS-B12 surface atlas (Van Essen, 2005). (**C**) As in (**A**), but for the reversed speech control experiment.

individual elements by manipulating only these while holding constant rhetoric properties of speeches, speakers and presentation modes (cf. Regev *et al.*, 2013). Second, it will be necessary to examine a broader range of speech stimuli, such as famous historic speeches, debates or commencement addresses. The reason for the relatively small sample of stimuli is that real-world speeches last several minutes, and presenting one stimulus consumes about as much time as an entire run in a classical fMRI study. Another interesting possibility would be to select speeches that are received more ambiguously by different listeners. One could then test whether sub-groups of listeners who either are or are not engaged by a given speech exhibit enhanced neural coupling toward the same physical stimulus (cf. Kriegeskorte *et al.*, 2008). In addition to testing more and different stimuli, future studies will have to replicate and extend our findings on a larger sample. Although our results are significant, the low statistical power could conceal true effects in additional brain regions. Finally, ISC represents a very promising for studying engagement into real-world stimuli but it cannot be used blindly to measure audience impact: rather, ISC can only be interpreted relatively compared with appropriate control stimuli since certain stimuli (e.g. extreme and rhythmic volume fluctuations) may trivially elicit high ISC. Overall, more work is needed to disentangle the effects of various form and content parameters and we look forward to future studies examining the neural reception of more rhetorically diverse stimuli.

In summary, we employed the ISC approach to characterize the collective-level neural effects prompted by political speeches. This study reveals enhanced neural coupling across the brains of listeners during powerful speeches. As such, the approach opens up new possibilities for research on the neural mechanisms mediating the reception of entertaining or persuasive messages.

Conflict of Interest

None declared.

REFERENCES

Alho, K., Rinne, T., Herron, T., Woods, D. (2014). Stimulus-dependent activations and attention-related modulations in the auditory cortex: A meta-analysis of fMRI studies. *Hearing Research*, 307, 29–41.

Amodio, D.M., Frith, C.D. (2006). Meeting of minds. The medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, 7, 268–77.

Atkinson, M. (1984). *Our Masters' Voices: The Language and Body Language of Politics*. London: Routledge.

Bartra, O., McGuire, J.T., Kable, J.W. (2013). The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage*, 76, 412–27.

Benjamini, Y., Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. *Series B, Statistical Methodology*, 57, 289–300.

Ben-Yakov, A., Honey, C.J., Lerner, Y., Hasson, U. (2012). Loss of reliable temporal structure in event-related averaging of naturalistic stimuli. *NeuroImage*, 63, 501–6.

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 MRI. *Journal of Neuroscience*, 17, 353–62.

Boldt, R., Malinen, S., Seppä, M., et al. (2013). Listening to an audio drama activates two processing networks, one for all sounds, another exclusively for speech. *PLoS One*, 8, e64489.

Bradley, M.M. (2009). Natural selective attention: orienting and emotion. *Psychophysiology*, 46, 1–11.

Brennan, J., Nir, Y., Hasson, U., Malach, R., Heeger, D.J., Pylkkanen, L. (2012). Syntactic structure building in the anterior temporal lobe during natural story listening. *Brain and Language*, 120, 163–73.

Caporello Bluvas, E., Gentner, T.Q. (2013). Attention to natural auditory signals. *Hearing Research*, 350, 10–8.

Chua, H.F., Ho, S.H., Jasinska, A.J., et al. (2011). Self-related processing of tailored smoking-cessation messages predicts successful quitting. *Nature Neuroscience*, 14, 426–7.

Ding, N., Simon, J.Z. (2012). The emergence of neural encoding of auditory objects while listening to competing speakers, Proceedings of the National Academy of Sciences USA, 109, 11854-59.

responses within the speakers in this study, one can hypothesize that a better speaker is a speaker who has the capacity to take control of his audience, and make them all respond in a similar way during the speech act. These ideas have also been applied to domains beyond rhetoric, such as foreign-language comprehension (Honey *et al.*, 2012) or non-verbal communication (Schippers *et al.*, 2010), all supporting the general notion that inducing shared neural processes is key to the success of communication (Hasson *et al.*, 2012). Moreover, work exists that uses neural measures to study the reception of professionally directed movies (Hasson *et al.*, 2008; Jääskeläinen *et al.*, 2008), or even predict the popularity of different cultural stimuli (Falk *et al.*, 2012a, b; Dmochowski *et al.*, 2014). These encouraging results thus suggest numerous practical implications and may help to advance the emerging field of persuasion neuroscience (Vezich *et al.*, 2015).

To our knowledge, this is the first study on this subject and thus a number of questions remain. One issue is that the current experiment documents effects on neural engagement, but did not aim to identify effective elements of good rhetoric, which are likely to be diverse. For instance, rhetoric schemes distinguish between aspects related to developing and arranging arguments ('inventio' and 'dispositio') from aspects that are related to the style of presentation (e.g. 'actio', 'elocutio' and 'memoria'). In the present experiment, speeches that made a strong impression also contained e.g. more first person pronouns and more emotion words, making it difficult to separate the overall strength from these linguistic features. While one could argue that these dimensions are likely to go hand in hand in most real-world settings, future experiments may specifically attempt to isolate

Dmochowski, J.P., Bezdek, M.A., Abelson, B.P., Johnson, J.S., Schumacher, E.H., Parra, L.C. (2014). Audience preferences are predicted by temporal reliability of neural processing. *Nature Communications*, 5, 4567.

Ethofer, T., Bretscher, J., Gschwind, M., Kreifelts, B., Wildgruber, D., Vuilleumier, P. (2012). Emotional voice areas: anatomic location, functional properties, and structural connections revealed by combined fMRI/DTI. *Cerebral Cortex*, 22, 191–200.

Falk, E.B., Berkman, E., Mann, T., Harrison, B., Lieberman, M.D. (2010). Predicting persuasion-induced behavior change from the brain. *Journal of Neuroscience*, 30, 8421–4.

Falk, E., Berkman, E., Lieberman, M.D. (2012a). From neural responses to population behavior: Neural focus groups predicts population-level media effects. *Psychological Science*, 23(5), 439–45.

Falk., E.B., O'Donnell, M.B., Lieberman, M.D. (2012b). Getting the word out: neural correlates of enthusiastic message propagation. *Frontiers in Human Neuroscience*, 6, 313.

Ferstl, E.C., Neumann, J., Bogler, C., von Cramon, D.Y. (2008). The extended language network: A meta-analysis of neuroimaging studies on text comprehension. *Human Brain Mapping*, 29, 581–93.

Flaisch, T., Häcker, F., Renner, B., Schupp, H.T. (2011). Emotion and the processing of symbolic gestures: an event-related brain potential study. *Social Cognitive and Affective Neuroscience*, 6, 109–18.

Grandjean, D., Sander, D., Pourtois, G., et al. (2005). The voices of wrath: brain responses to angry prosody in meaningless speech. *Nature Neuroscience*, 8, 145–6.

Hasson, U., Ghazanfar, A.A., Galantucci, B., Garrod, S., Keysers, C. (2012). Brain-to-brain coupling: a mechanism for creating and sharing a social world. *Trends in Cognitive Sciences*, 16, 114–21.

Hasson, U., Landsman, O., Knappmeyer, B., Vallines, I., Rubin, N., Heeger, D.J. (2008). Neurocinematics: The neuroscience of film. *Projections*, 2, 1–26.

Hasson, U., Malach, R., Heeger, D.J. (2010). Reliability of cortical activity during natural stimulation. *Trends in Cognitive Sciences*, 14, 40–8.

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science*, 303, 1634–40.

Hickok, G., Poeppel, D. (2007). The cortical organization of speech perception. *Nature Reviews Neuroscience*, 8, 393–402.

Honey, C.J., Thompson, C.R., Lerner, Y., Hasson, U. (2012). Not lost in translation: neural responses shared across languages. *Journal of Neuroscience*, 32, 15277–83.

Jääskeläinen, I.P., Koskentalo, K., Balk, M.H., et al. (2008). Inter-subject synchronization of prefrontal cortex hemodynamic activity during natural viewing. *Open Neuroimaging Journal*, 2, 14–9.

Kissler, J., Herbert, C., Peyk, P., Junghöfer, M. (2007). 'Buzzwords'—early cortical responses to emotional words. *Psychological Science*, 18, 475–80.

Kriegeskorte, N., Mur, M., Bandettini, P.A. (2008). Representational similarity analysis—connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, 2, 4.

Lancaster, J.L., Woldorff, M.G., Parsons, L.M., et al. (2000). Automated Talairach Atlas labels for functional brain mapping. *Human Brain Mapping*, 10, 120–31.

Lang, P.J., Bradley, M.M., Cuthbert, B.N. (1997). Motivated attention: Affect, activation and action. In: Lang, P.J., Simons, R.F., Balaban, M.T., editors. *Attention and Orienting: Sensory and Motivational Processes*. New Jersey: Lawrence Erlbaum Associates, pp. 97–135.

Lee, A.K.C., Larson, E., Maddox, R., Shinn-Cunningham, B. (2014). Using neuroimaging to understand the cortical mechanisms of auditory selective attention. *Hearing Research*, 307, 111–20.

Lerner, Y., Honey, C.J., Silbert, L.J., Hasson, U. (2011). Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *Journal of Neuroscience*, 31, 2906–15.

Mar, R.A. (2011). The neural bases of social cognition and story comprehension. *Annual Review of Psychology*, 62, 103–34.

Mesgarani, N., Chang, E. (2012). Selective cortical representation of attended speaker in multi-talker speech perception, *Nature*, 485, 233–6.

Nummenmaa, L., Saarimäki, H., Glerean, E., et al. (2014). Emotional speech synchronizes brains across listeners and engages large-scale dynamic brain networks. *Neuroimage*, 102, 498–509.

Pajula, J., Kauppi, J.-P., Tohka, J. (2012). Inter-subject correlation in fMRI: method validation against stimulus-model based analysis. *PLoS One*, 8, e41196.

Pennebaker, J.W., Booth, R.E., Francis, M.E. (2007). *Linguistic Inquiry and Word Count: LIWC2007—Operator's Manual*. Austin, TX: LIWC.net.

Pennebaker, J.W., Lay, T.C. (2002). Language use and personality during crises: analyses of mayor Rudolp Giuliani's press conferences. *Journal of Research in Personality*, 36, 271–82.

Petkov, C.I., Kang, X., Alho, K., Bertrand, O., Yund, E.W., Woods, D.L. (2004). Attentional modulation of human auditory cortex. *Nature Neuroscience*, 7, 658–63.

Price, C.J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *NeuroImage*, 62, 816–47.

Pruessmann, K.P., Weiger, M., Scheidegger, M.B., Boesiger, P. (1999). SENSE: sensitivity encoding for fast MRI. *Magnetic Resonance Medicine*, 42, 952–62.

Pugh, K.R., Shaywitz, S.E., Fulbright, R.K., et al. (1996). Auditory selective attention: an fMRI investigation. *Neuroimage*, 4, 159–73.

Regev, M., Honey, C.J., Hasson, U. (2013). Modality-selective and modality-invariant neural responses to spoken and written narratives. *Journal of Neuroscience*, 33, 15978–88.

Schippers, M.B., Roebroeck, A., Renken, R., Nanetti, L., Keysers, C. (2010). Mapping the Information flow from one brain to another during gestural communication. *Proceedings of the National Academy of Sciences USA*, 107, 9388–93.

Schmälzle, R., Häcker, F., Renner, B., Honey, C.J., Schupp, H.T. (2013). Neural correlates of risk perception during real-life risk communication. *Journal of Neuroscience*, 33, 10340–7.

Schupp, H.T., Flaisch, T., Stockburger, J., Junghöfer, M. (2006). Emotion and attention: event-related brain potential studies. *Progress in Brain Research*, 156, 123–43.

Specht, K. (2014). Neuronal basis of speech comprehension. *Hearing Research*, 307, 121–35.

Spreng, R.N., Mar, R.A., Kim, A.S.N. (2009). The common neural basis of autobiographical memory, prospection, navigation, theory of mind and the default mode: a quantitative meta-analysis. *Journal of Cognitive Neuroscience*, 21, 489–510.

Stephens, G.J., Silbert, L.J., Hasson, U. (2010). Speaker-listener neural coupling underlies successful communication. *Proceeding National Academy of Science USA*, 107, 14425–30.

Van Berkum, J.J.A., Holleman, B., Nieuwland, M., Otten, M., Murre, J. (2009). Right or wrong? The brain's fast response to morally objectionable statements. *Psychological Science*, 20, 1092–99.

Van Essen, D.C. (2005). A population-average, landmark- and surface-based (PALS) atlas of human cerebral cortex. *Neuroimage*, 28, 635–62.

Vezich, S., Falk, E.B., Lieberman, M.D. (2015). Persuasion neuroscience: new potential to test dual process theories. In: Harmon-Jones, E., Inzlicht, M., editors. *Social Neuroscience: Biological approaches to social Psychology*. New York: Psychological Press.

Wallentin, M., Nielsen, A.H., Vuust, P., Dohn, A., Roepstorff, A., Lund, T.E. (2011). Amygdala and heart rate variability responses from listening to emotionally intense parts of a story. *Neuroimage*, 58, 963–73.

Warren, J., Sauter, D., Eisner, F., et al. (2006). Positive emotions preferentially engage an auditory–motor "Mirror" system. *Journal of Neuroscience*, 26, 13067–75.

Whissel, C., Sigelman, L. (2001). The times and the man as predictors of emotion and style in the inaugural addresses of US Presidents. *Computers and the Humanities*, 35, 255–72.

Wilson, S.M., Molnar-Szakacs, I., Iacoboni, M. (2008). Beyond superior temporal cortex: intersubject correlations in narrative speech comprehension. *Cerebral Cortex*, 18, 230–42.

Wolf, M., Horn, A., Mehl, M., Haug, S., Pennebaker, J.W., Kordy, H. (2008). Computergestützte quantitative Textanalyse: Äquivalenz und Robustheit der deutschen Version des Linguistic Inquiry and Word Count. *Diagnostica*, 2, 85–98.

Yarkoni, T., Poldrack, R.A., Nichols, T.E., Van Essen, D.C., Wager, T.D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods*, 8, 665–70.