Behavioral/Systems/Cognitive

Not Lost in Translation: Neural Responses Shared Across Languages

Christopher J. Honey,[1,2]* **Christopher R. Thompson,**[1]* **Yulia Lerner,**[1,2] and **Uri Hasson**[1,2]
[1]Department of Psychology and [2]Princeton Neuroscience Institute, Princeton University, Princeton, New Jersey 08540

How similar are the brains of listeners who hear the same content expressed in different languages? We directly compared the fMRI response time courses of English speakers and Russian speakers who listened to a real-life Russian narrative and its English translation. In the translation, we tried to preserve the content of the narrative while reducing the structural similarities across languages. The story evoked similar brain responses, invariant to the structural changes across languages, beginning just outside early auditory areas and extending through temporal, parietal, and frontal cerebral cortices. The similarity of responses across languages was nearly equal to the similarity of responses within each language group. The present results demonstrate that the human brain processes real-life information in a manner that is largely insensitive to the language in which that information is conveyed. The methods introduced here can potentially be used to quantify the transmission of meaning across cultural and linguistic boundaries.

Introduction

When we communicate using language, content (meaning) is always encoded within a system of grammar that is in turn encoded in a specific form (physical structure). The diversity of human languages attests that the same content can often be conveyed in more than one grammar and in radically different physical forms. This linguistic diversity also provides an opportunity to identify neural responses that are shared across individuals who hear the same content expressed in very different forms.

A network of cortical areas spanning temporal, parietal, and frontal regions is responsible for linguistic processing (Binder et al., 1997; Hickok and Poeppel, 2007). Within the language processing systems, early sensory regions are specialized for representing the immediate physical properties of the environment (form), while higher order regions extract increasingly abstract content (Davis and Johnsrude, 2003; Quiroga et al., 2005; Okada et al., 2010). Prior studies have shown substantial overlap in the sets of brain regions involved in processing different spoken languages (Perani and Abutalebi, 2005; Klein et al., 2006a) and sign languages (Neville et al., 1998). However, the overlapping responses to language tasks do not imply that specific content is processed in the same way in different languages.

To identify cross-linguistic content representations, studies in bilingual subjects have mapped neural response adaptation to the same word-meanings presented in different languages (Chee et al., 2003; Crinion et al., 2006; Klein et al., 2006b). However, the adaptation studies have produced variable results (Chee, 2009) and the single-word stimuli they used are unlikely to be processed in the same manner as the contextually meaningful expressions of real-life language (Lerner et al., 2011; Ben-Yakov et al., 2012). As a result, it is still unknown to what extent real-life content elicits similar neural responses across speakers of different languages. Determining the similarity of content representations across languages is of great interest, as it provides a clue to how what we say (content) depends on how we say it (form).

Here, we empirically address this issue by directly comparing the time courses of neural responses to a real-life spoken narrative whose form and content are independently manipulated via translation. A group of nine native English speakers (Anglophones, who spoke no Russian) and a group of nine native Russian speakers (Russophones, who were bilingual) listened to an 11 min auditory story told in Russian and to an English translation of the same story while undergoing functional magnetic resonance imaging (fMRI). By comparing blood-oxygenation level-dependent (BOLD) time courses within and across languages, we identified three sets of brain areas. First, we identified the areas that responded similarly across subjects within a language group when they heard the story in a language that they understood. Second, by comparing the brain responses across comprehending Russophones and noncomprehending Anglophones who listened to the same Russian story, we delineated low-level sensory areas that responded to the physical properties (form) of the stimulus independent of the content it conveyed. Finally, we identified brain regions that responded with similar time courses across Russophone and Anglophone listeners, regardless of the language in which the content of the story was presented.

Materials and Methods

Subjects. Twenty-one right-handed subjects (ages 19–33 years) participated in the fMRI study. One Russophone subject was excluded for head motion (>2 mm), another was excluded because of the presence of transients (>5× standard deviation of the time course) in the BOLD signal, and the final Anglophone scanned was excluded to equalize group size (inclusion of this subject did not change the results). Thus, nine Anglophones (4 males and 5 females) and nine Russophones (5 males and 4 females) were included in the analysis.

All nine Russophone subjects were born to native-speaking Russian parents. They all learned Russian as their first language at home, and learned English as their second language between the ages of 4–15 years (mean 8.3 ± 3.5 years). Eight of the Russians spoke Russian daily, and one spoke it weekly. Similarly, all nine Anglophone subjects were born to native-speaking English parents, learned English as their first and home language, and had 13 or more years of general education in English. None of the Anglophone subjects had previously learned any Russian.

Experimental procedures were approved by the Princeton University Committee on Activities Involving Human Subjects. All subjects provided written informed consent.

MRI acquisition. Subjects were scanned using a 3T head-only MRI scanner (Allegra; Siemens). A custom radio-frequency head-coil was used to achieve high-resolution structural scans (NM-011 transmit head coil; Nova Medical). In the English and Russian story scans, 458 volumes were acquired using a T2*-weighted EPI pulse sequence (TR, 1500 ms; TE, 30 ms; flip angle, 76°; FOV, 192 × 192 mm^2; right-to-left phase encoding). Each volume included 25 interleaved slices of 3 mm thickness with an interslice gap of 1 mm (in-plane resolution, 3 × 3 mm^2) for near whole-brain coverage. The final anatomical scan was acquired using a T1-weighted high-resolution (1 × 1 × 1 mm^3) MPRAGE pulse sequence. Stimuli were presented using MATLAB (MathWorks) and Psychophysics Toolbox and were delivered via high-fidelity MR-compatible headphones (MR Confon). These headphones function optimally in the bore of the scanner and reduce acoustic scanner noise.

Stimuli and experimental design. The Russian storyteller in the study was a female native speaker of Russian. To emulate naturalistic dialogue and ensure the intention to communicate, we instructed her to relate an unrehearsed personal anecdote in her native language. For the purpose of a separate experiment, the narration was recorded in the scanner. In-scanner recording was performed with a customized MR-compatible dual channel optical microphone (FOMRI II; Optoacoustics), and scanner audio artifacts were removed as described previously (Stephens et al., 2010).

The Russian speaker produced a time-stamped transcript of the story in English, which preserved the content of the narrative as closely as possible. A male native speaker of American English produced the English recording using Logic Pro digital audio software (Apple) and equipment provided by Princeton Broadcasting Service (Princeton, NJ). In translating the story to English, we tried to preserve its content while altering its physical properties. Thus, although sentence onsets were aligned to maintain the timing at which information arrived in the English and Russian audio, the Russian and English stories differed in their utterances (only 6.6% of the words in the English translation were homophonic with their Russian equivalents), in the voices of the speakers (female Russian speaker, male English speaker), and in the quality of the recording (the Russian story contains additional background noise). Increasing the low-level differences across the English and Russian versions assisted us in identifying brain areas that respond similarly to narrative content, despite marked differences in form across languages.

English sentences ended, on average, 0.48 ± 0.8 s before the corresponding Russian sentence, as many of the English words had shorter vocalizations. Thus, while sentence onsets were precisely aligned across languages, the English version included slightly longer pauses at the end of sentences. The Russian recording and its English translation were each 11 min, 21 s in duration. All subjects heard the Russian version of the narrative followed by the English version.

Behavioral assessment. Attentive listening to the story was confirmed via subjects' written recollections of the story, which were graded against

a rubric by independent raters. Scores were evaluated using 11 questions for personal facts (recollection about characters in the story) and 10 questions for anecdotal facts (recollection of particular events in the story). Overall, subjects recalled 7.5 ± 1.4 personal facts and 7.6 ± 1.6 anecdotal facts. No group difference was observed for the recollection of personal facts (Russophones: mean = 7.4, SD = 1.8; Anglophones mean = 7.5, SD = 1.0; $t_{(8)} = 0.16$, $p = 0.87$) or anecdotal facts (Russophones: mean = 7.8, SD = 1.2; Anglophones: mean = 7.9, SD = 1.9; $t_{(8)} = 0.11$, $p = 0.91$).

When the Anglophones listened to the unintelligible Russian, they were instructed to listen to the story just as if they could understand it. Conversely, the Russophones subjects could understand both versions of the story, and this additional exposure to the content could have affected their performances on the behavioral test. Crucially, however, all Russophone subjects listened to the Russian before the English translation, and none of the neural analyses reported in this study involve the data from Russophones listening to the English story. Thus, the comparison of Russophone and Anglophone neural responses is always based on data collected from their first comprehensible experience of the narrative.

Finally, an independent behavioral assessment of the Russian proficiency of the Russophones was obtained by testing their comprehension of a separate narrative spoken in Russian. Subjects listened to a 1.5 min recording of a classic Russian story (narrated by the same Russian speaker who told the Russian story in the fMRI study) and then provided a written summary of the plot. Story comprehension was scored by three native Russian speaking raters on a six-point scale (0, no comprehension; 3, reasonable comprehension; 5, perfect comprehension). There was high agreement across pairs of raters (mean interrater absolute difference: 0.26) and all Russophone subjects performed extremely well (mean score across raters and subjects: 4.66 ± 0.37).

fMRI preprocessing. fMRI data were preprocessed and aligned using Version 1.8.6 of the BrainVoyager QX software package (Brain Innovation). All analyses were performed using in-house scripts written in MATLAB. Preprocessing of functional scans included 3D motion correction, linear trend removal, slice scan-time correction, and high-pass filtering (3 cycles per scan, i.e., 4.4×10^{-3} Hz). A Gaussian spatial filter of 6 mm width at half-maximum value was applied to correct for any misalignment or structural differences between brains. Anatomical MPRAGE images of subjects' brains, along with functional scans, were aligned to standard Talairach coordinates. To remove transients and nonspecific signal elevation effects, we excluded the first 15 and the last five volumes of each scan from further analysis. Voxels with low mean BOLD values (>4 standard deviations below the gray and white matter mean) were excluded from all analyses.

Intersubject correlation analysis. The central results of this study are derived from the intersubject correlation (inter-SC) analysis. This analysis provides a measure of the reliability of the responses to natural stimuli by comparing the BOLD response time courses within each brain area across different subjects. In a standard GLM analysis, experimenters usually assume a prototypical response profile for each stimulus. The inter-SC analysis method circumvents the need to specify a prior model for the neuronal response by using one subject's brain responses for a given stimulus as a model for the responses to the same content in other subjects.

Correlation maps were produced for both intragroup (e.g., English speakers hearing English) and intergroup (e.g., English speakers hearing English vs Russian speakers hearing Russian) comparisons. The correlation maps provide a measure of the reliability of brain responses to the audio recordings by quantifying the variability of the time course of BOLD activity among subjects listening to the same stimuli (Hasson et al., 2004, 2010; Lerner et al., 2011).

For each voxel, intragroup correlation is calculated as an average correlation $R = \frac{1}{N} \sum_{j=1}^{N} r_j$, where the individual r_j is the Pearson correlation between that voxel's BOLD time course in one individual and the average of that voxel's BOLD time courses in the other individuals. Intergroup correlation is calculated as an average $R = \frac{1}{N} \sum_{j=1}^{N} \tilde{r}_j$ of the correlations, $\tilde{r}_j$,

between the BOLD time courses of the j-th individual from one group and the average BOLD time courses of all the individuals in the other group.

For the intergroup analyses shown in Figure 3, the response time course in each voxel was averaged across all Russophone listeners, and this average was then correlated with the responses in that voxel for each individual Anglophone speaker. The same analysis can also be performed in reverse, with individual Russophone time courses compared against the average Anglophone time course, resulting in nearly identical correlation maps.

Bootstrapping by phase-randomization. Because of the presence of long-range temporal autocorrelation in the BOLD signal (Zarahn et al., 1997), 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 intersubject correlation between any pair of subjects).

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 bootstrap 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. Lastly, p values of the empirical average correlations (R values) were computed by comparison with the null distribution of R values.

False discovery rate correction and cluster size threshold. 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 0.01.

Audio correlations. Although we tried to increase the physical differences across the audio waveforms, some low-level similarities were preserved. In particular, because the time of sentence onsets was preserved, there was low-frequency (<0.3 Hz) correlation between the sound envelopes of the English and Russian stimuli (Fig. 4). To further quantify the effect of low-level acoustic properties on the responses in early auditory areas, we correlated the brain responses of each language group against the audio envelopes of the English and Russian stories.

To compute the correlation between BOLD signals and the audio envelopes of the English and Russian stimuli, we bandpassed the audio signal between 4 and 4000 Hz, extracted the envelope of the signal using a Hilbert transform, and then downsampled the envelope to the sampling rate of the BOLD signal using an anti-aliasing low-pass finite impulse response filter. For the Russian-Russophone and English-Anglophone experimental conditions, we then computed an average BOLD signal for each voxel by z-transforming the BOLD time course within subjects and averaging time courses across all nine subjects. Finally, Pearson correlations between the average BOLD signal and the audio envelopes (shifted in time to account for hemodynamic delay) were computed at every voxel. P values, which were generated via transformation to Student's t values, were subsequently entered into an FDR correction, as described above, to produce the thresholded correlation maps.

Projection of audio amplitude. To diminish the impact of shared acoustic properties across the Russian and English stimuli on the calculation of the shared-content map (Fig. 3B), we projected the envelopes of the audio signals from the BOLD signal in each voxel in each subject. Russian and English audio envelopes (see above), convolved with a hemodynamic response function (Glover, 1999), were entered into a bivariate linear regression to predict the BOLD signal in each voxel. The BOLD signals were then replaced with the residuals resulting from this regres-

sion, and were treated exactly as all other BOLD signals for the remainder of the correlation analysis.

Results

The language network

We began by identifying the complete set of areas that respond reliably to a complex narrative within each of our subject populations. This was done by mapping the intragroup correlation within each language group (Anglophones listening to English, Fig. 1A; Russophones listening to Russian, Fig. 1B). Consistent with previous reports (Lerner et al., 2011), we found that, among both Anglophones and Russophones, responses to the story were reliable not only in early auditory areas but also in linguistic and extralinguistic regions. Early auditory areas include primary auditory cortex and nearby tonotopic areas (A1+) (Romanski and Averbeck, 2009). Linguistic areas include the superior temporal gyrus (STG), angular gyrus, supramarginal gyrus, temporoparietal junction (TPJ), parietal lobule, and the inferior frontal gyrus (IFG), each of which has been linked with one or more aspects of linguistic processing (Hickok and Poeppel, 2007; Sahin et al., 2009). Finally, extralinguistic regions, which are likely involved in processing the narrative and the social content of the story (Fletcher et al., 1995; Xu et al., 2005; Ferstl et al., 2008; Lerner et al., 2011), include the precuneus, inferior occipital gyrus, dorsolateral prefrontal cortex, orbitofrontal cortex, and medial prefrontal cortex.

Although the response reliability of the Russian listeners was slightly weaker overall, very similar networks of regions responded reliably within both Russophone and Anglophone listener groups as they listened to the story in their native languages. This spatial correspondence suggests that similar language processing networks are recruited by English and Russian listeners. However, these maps (Fig. 1) reveal the combination of regions with low-level, language-specific, and language-general response properties. To tease these systems apart, we next identified areas that respond specifically to the form of the story regardless of its content, followed by those regions that respond to the content of the story regardless of its form.

Regions sensitive to form

We identified areas that respond reliably to the low-level properties (form) of the sound file, regardless of its content, by mapping the intersubject correlation within the Anglophone subjects listening to the unintelligible Russian recording. The Russian speech evoked reliable responses across Anglophones in early auditory cortical areas (Fig. 2), indicating that the Anglophones indeed heard the spoken Russian, but there were no reliable responses outside of these sensory areas, as expected for unintelligible speech. This immediately indicates that higher-order cortical areas will respond reliably only in cases where subjects comprehend the narrative (Fig. 1).

Next, we asked whether the responses in early auditory areas were similar across Russophones and Anglophones, even when the stimulus was only comprehensible to one group. To that end, we compared the brain responses of the Russophone and Anglophone listeners as they listened to the Russian story (Fig. 3A). In this situation, both groups received identical physical input, but while one group (Russophones) was able to extract its content, the other group (Anglophones) could not. Thus, any shared brain responses across the groups arises from low-level properties of the audio. Inter-SC among the Russophones and Anglophones listening to the Russian recording was restricted to early auditory regions (Fig. 3A). This result suggests that this small set of regions
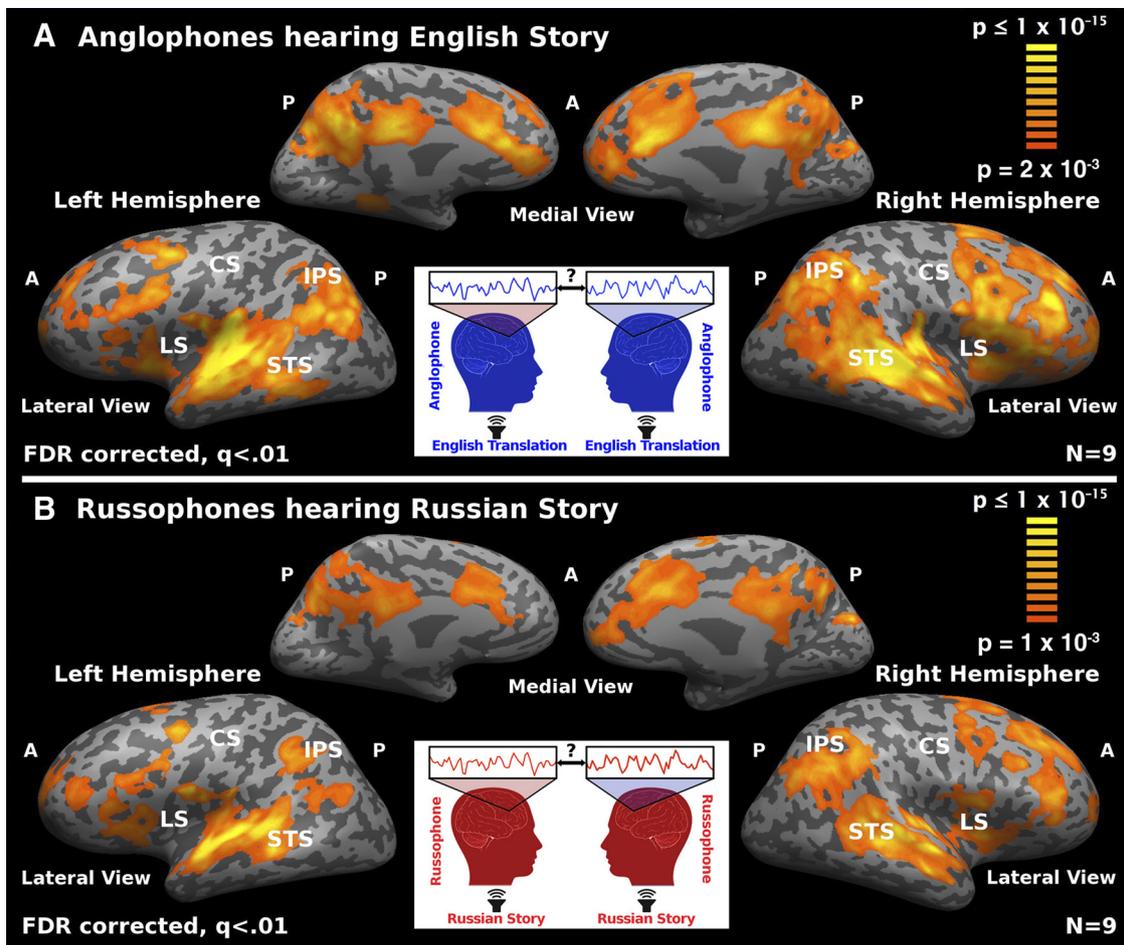
Figure 1. Reliability of brain responses within each language group. The fMRI BOLD time course in each area is correlated across the listeners within each group to produce a map of inter-SC within each language. The surface maps show the areas exhibiting reliable responses for Anglophones listening to the English (**A**) and Russophones listening to the Russian story (**B**). P, Posterior; A, anterior; CS, central sulcus; LS, lateral sulcus; STS, superior temporal sulcus; IPS, intraparietal sulcus.

is primarily sensitive to low-level physical properties of the sound wave, consistent with prior neurophysiology (Nourski et al., 2009) and fMRI (Warren et al., 2005) data. Furthermore, some of the correlation in early auditory areas may be due to nonverbal information, such as the emotional valence of the speaker, which the Anglophones may have extracted from the intonation of the Russian speaker.

Regions sensitive to shared form across languages

Although we tried to amplify the physical differences across the audio waveforms, some low-level similarities were preserved. In particular, because the time of sentence onsets was preserved during the translation, there was some low-frequency (less than ~0.3 Hz) correlation between the sound envelopes of the English and Russian stimuli (see Materials and Methods, above; Fig. 4C). To further quantify the effect of low-level acoustic properties on the responses in early auditory areas, we correlated the brain responses of each language group against the audio power modulations of the English and Russian stories (see Materials
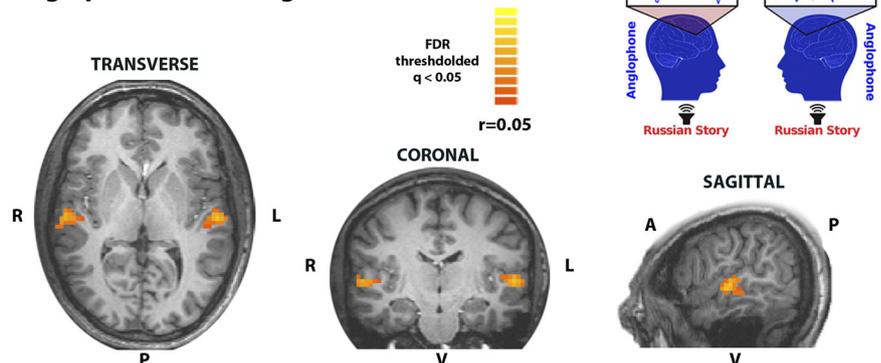
Figure 2. Anglophones listening to Russian share early auditory responses. The component of the response time courses that is shared among Anglophones listening to Russian is shown in transverse, coronal, and axial slices. The shared response is restricted to the earliest stages of cortical auditory processing. R, Right; L, left; P, posterior; V, ventral. Talairach $(x, y, z) = -56, -17, 6$.

and Methods, above). The residual similarities in the audio power modulations induced weak correlations between the audio envelope of the recording in one language and the brain responses of listeners who listened to the other language (Fig. 4D). Importantly, the audio envelopes did not predict BOLD responses outside of early auditory areas in any comparison.
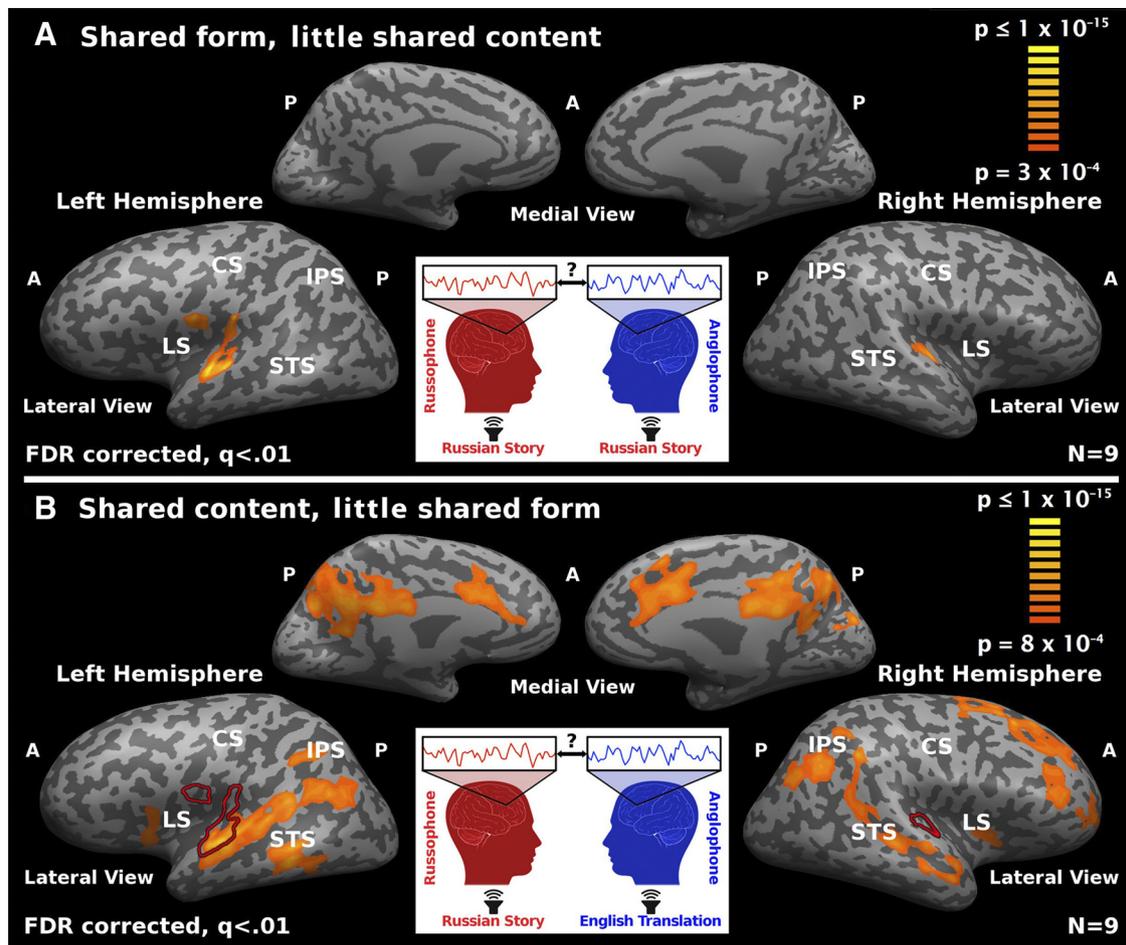
Figure 3. Separating form and content using a real-life spoken narrative. *A*, To isolate regions primarily responsive to the form of the auditory stimulus, rather than its content, we compared cortical responses of Russophones and Anglophones listening to the Russian recording. The map shows the responses shared across Anglophones and Russophones listening to the Russian recording. *B*, To isolate the regions responsive to the content of the narrative, we compared Russophones listening to Russian against Anglophones listening to English, after projecting out the audio amplitude envelopes. In this comparison, the two groups extracted similar content from distinct audio waveforms. The map shows the shared responses across Anglophones listening to English and Russophones listening to Russian. Red border denotes the areas that shared form-related responses, from *A*. P, Posterior; A, anterior; CS, central sulcus; LS, lateral sulcus; STS, superior temporal sulcus; IPS, intraparietal sulcus.

Regions sensitive to shared content across languages

To identify areas that respond similarly to the content that is shared across the translation, regardless of form, we compared the brain responses of Russophones and Anglophones listening to the same story told in their own native language (Fig. 3*B*). In this situation, the two groups perceived inputs that shared the same content but differed radically in their physical forms (i.e., different voice, gender of speaker, number of words, lexical items). To remove low-level acoustic similarities across the stories (Fig. 4*C*), we regressed out the English and Russian audio amplitude time courses from all BOLD signals before comparing the responses across languages (see Materials and Methods, above). After this regression, shared brain responses across listeners should arise primarily from content that is preserved across languages.

Whereas presentation of the identical form induced little shared responses across languages in the absence of shared content (Fig. 3*A*), presentation of the same content evoked widespread shared responses across the two groups, despite the stark differences in form across languages (Fig. 3*B*). The areas that show reliable responses across languages include language-related areas such as the superior temporal sulcus, angular gyrus, supramarginal gyrus, and the IFG, as well as extralinguistic areas

including the precuneus, middle frontal gyrus, and orbitofrontal cortex. The similarity of responses across languages is nearly as strong as the similarity of responses within language groups.

Discussion

The present results highlight the robust capacity of the human brain to extract behaviorally relevant content from the environment while generalizing across low-level physical parameters. We found that the content and the coarse temporal structure of the Russian and English stories was sufficient to evoke shared responses in many higher-order areas across languages (Fig. 3*B*). In contrast, preserving the exact timing and all low-level physical properties of the audio soundtrack while removing the content of the story evoked no reliable responses in these high-order areas (Fig. 3*A*). These results expose a set of neural processes for representing content that unites individuals across languages. Such general processes may be related to the syntactic structures shared across languages, as well as to shared affective responses and shared extralinguistic knowledge about the structure of the world.

We emphasize that the present results should not be taken to indicate that low-level stimulus parameters are unimportant to the brain. The human nervous system is exquisitely sensitive
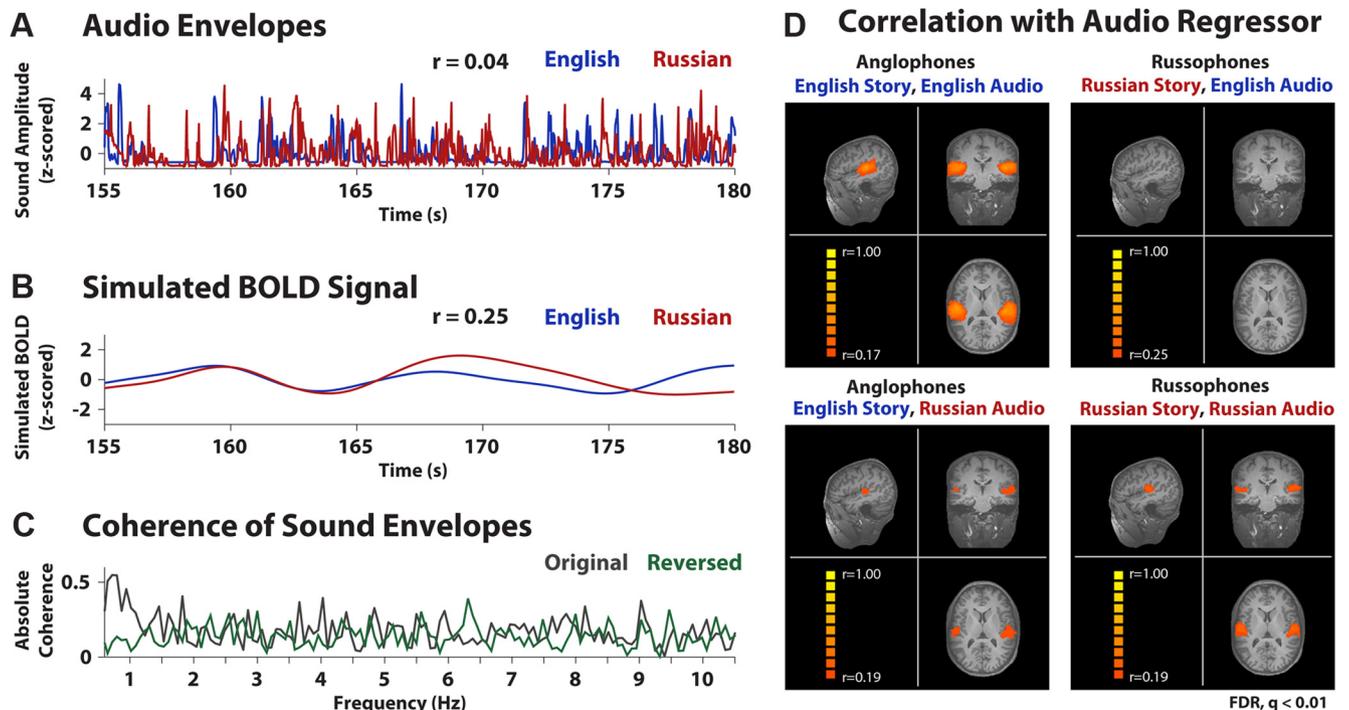
Figure 4. Observed similarities in audio envelope and BOLD signal across stories. ***A***, Plots of the English (blue) and Russian (red) story audio envelopes. The sound envelopes correlated with $r = 0.04$. ***B***, Plots of the English (blue) and Russian (red) sound envelopes convolved with a hemodynamic response function (Glover, 1999) to simulate the BOLD time course. These convolved time courses are highly significantly correlated ($r = 0.25$, $p < 0.01$). ***C***, Plots of the magnitude of the coherence of the sound envelopes in ***A***, both for the intact envelope (blue line) and for a control condition in which the envelope of one story is time-reversed (green line). ***D***, Maps show correlation between BOLD signal and the envelope of the audio stimuli. English and Russian audio envelopes compared against BOLD signal recorded from Anglophones listening to the English story and Russophones listening to the Russian story.

to some parameters of the physical world (Hecht et al., 1942; LeMasurier and Gillespie, 2005), and indeed, our present results indicate that even slight similarities or differences in the form across conditions can influence the BOLD responses of early sensory areas (Fig. 4). However, neural sensitivity to the form of a stimulus affects only the earliest sensory areas unless that stimulus has some ecological or behavioral relevance. Thus, in most of the cerebral cortex, neural responses are insensitive to the form of behaviorally irrelevant stimuli and highly sensitive to the content of meaningful stimuli, regardless of the form in which that content is conveyed.

We note that, although the brain responses in early language areas seem to be mostly driven by the sound envelope (Fig. 4*D*) regardless of the speech content (Fig. 3*A*), some of the responses may be related to nonverbal information that is conveyed by speech intonation (Wildgruber et al., 2004) and is processed by the non-Russian listeners. In addition, although only 6.6% of the corresponding words across languages were homophonic, there may have been individual phonemes that engaged similar processing across languages. However, these effects, if present, were either very weak or did not engage higher-order areas, because when the Anglophone subjects listened to the Russian story, there were not any reliable responses outside of auditory cortex (Fig. 2).

Our experimental design did not allow for the definitive identification of language-processing characteristics that are unique to English or Russian. We performed an analysis to identify brain areas that respond reliably within each language group, but significantly less reliably across the two language groups, and found such effects in the STG, right TPJ, and right IFG (data not shown). However, because we designed the Russian and English recordings to differ across multiple dimensions—including low-level acoustic properties, amplitude of background noise, gender of speakers, and language-specific grammatical features—it was not possible to unambiguously identify the origin of these language-group specific effects.

The strength of this experimental design is in identifying response patterns that are shared across groups despite dramatic differences in linguistic form, and indeed we find widespread shared responses, time-locked across languages to the content of a complex real-life narrative. This is particularly interesting given that Russian and English are from distinct language families within the Indo-European languages (Slavic and Germanic), which are mutually unintelligible. The two languages differ in myriad ways: phonologically, lexically, and syntactically. For example, Russian is a much more synthetic language than English in that words tend to include a greater number of meaningful morphemes. Also, Russian word order is more flexible and dependent on properties of the discourse than is English. Future studies using a wider array of languages will be required to understand which shared language features are necessary for brain responses to be shared extensively across speakers of different languages.

To the extent that temporal profiles of neural activity are indicative of cognitive processing, the finding of shared responses throughout most of the human language system presents a serious challenge to the strong Whorfian view that linguistic peculiarities substantially determine what individuals perceive and think (Whorf, 1956). Furthermore, this study introduces a new methodology that can be used in future studies that manipulate aspects of a narrative (e.g., the grammatical or semantic content) to further isolate the elements that drive shared responses across languages. Finally, we note that regions such as the medial prefrontal cortex and precuneus are linked to narrative understanding (Maguire et al., 1999; Xu et al., 2005; Ferstl et al., 2008; Lerner

et al., 2011), and that the alignment of response time courses in these regions across speakers of different languages may be a marker of successful communication via translation. Future work must test the present findings using other languages, other sensory modalities (auditory, visual, tactile), and other modes of communication (linguistic and nonlinguistic). We hypothesize that the extent of shared brain responses may serve as a metric of the quality of communication across individuals (Pickering and Garrod, 2004; Stephens et al., 2010) who have different levels of language competence or different social and cultural backgrounds.

References

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B 57:289–300.

Ben-Yakov A, Honey CJ, Lerner Y, Hasson U (2012) Loss of reliable temporal structure in event-related averaging of naturalistic stimuli. Neuroimage 63:501–506.

Binder JR, Frost JA, Hammeke TA, Cox RW, Rao SM, Prieto T (1997) Human brain language areas identified by functional magnetic resonance imaging. J Neurosci 17:353–362.

Chee MW (2009) fMR-adaptation and the bilingual brain. Brain Lang 109:75–79.

Chee MW, Soon CS, Lee HL (2003) Common and segregated neuronal networks for different languages revealed using functional magnetic resonance adaptation. J Cogn Neurosci 15:85–97.

Crinion J, Turner R, Grogan A, Hanakawa T, Noppeney U, Devlin JT, Aso T, Urayama S, Fukuyama H, Stockton K, Usui K, Green DW, Price CJ (2006) Language control in the bilingual brain. Science 312:1537–1540.

Davis MH, Johnsrude IS (2003) Hierarchical processing in spoken language comprehension. J Neurosci 23:3423–3431.

Ferstl EC, Neumann J, Bogler C, von Cramon DY (2008) The extended language network: a meta-analysis of neuroimaging studies on text comprehension. Hum Brain Mapp 29:581–593.

Fletcher PC, Happé F, Frith U, Baker SC, Dolan RJ, Frackowiak RS, Frith CD (1995) Other minds in the brain: a functional imaging study of "theory of mind" in story comprehension. Cognition 57:109–128.

Glover GH (1999) Deconvolution of impulse response in event-related BOLD fMRI. Neuroimage 9:416–429.

Hasson U, Nir Y, Levy I, Fuhrmann G, Malach R (2004) Intersubject synchronization of cortical activity during natural vision. Science 303:1634–1640.

Hasson U, Malach R, Heeger DJ (2010) Reliability of cortical activity during natural stimulation. Trends Cogn Sci 14:40–48.

Hecht S, Shlaer S, Pirenne MH (1942) Energy, quanta, and vision. J Gen Physiol 25:819–840.

Hickok G, Poeppel D (2007) The cortical organization of speech processing. Nat Rev Neurosci 8:393–402.

Klein D, Watkins KE, Zatorre RJ, Milner B (2006a) Word and nonword repetition in bilingual subjects: a PET study. Hum Brain Mapp 27:153–161.

Klein D, Zatorre RJ, Chen JK, Milner B, Crane J, Belin P, Bouffard M (2006b) Bilingual brain organization: a functional magnetic resonance adaptation study. Neuroimage 31:366–375.

LeMasurier M, Gillespie PG (2005) Hair-cell mechanotransduction and cochlear amplification. Neuron 48:403–415.

Lerner Y, Honey CJ, Silbert LJ, Hasson U (2011) Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. J Neurosci 31:2906–2915.

Maguire EA, Frith CD, Morris RG (1999) The functional neuroanatomy of comprehension and memory: the importance of prior knowledge. Brain 122:1839–1850.

Neville HJ, Bavelier D, Corina D, Rauschecker J, Karni A, Lalwani A, Braun A, Clark V, Jezzard P, Turner R (1998) Cerebral organization for language in deaf and hearing subjects: biological constraints and effects of experience. Proc Natl Acad Sci U S A 95:922–929.

Nourski KV, Reale RA, Oya H, Kawasaki H, Kovach CK, Chen H, Howard MA 3rd, Brugge JF (2009) Temporal envelope of time-compressed speech represented in the human auditory cortex. J Neurosci 29:15564–15574.

Okada K, Rong F, Venezia J, Matchin W, Hsieh IH, Saberi K, Serences JT, Hickok G (2010) Hierarchical organization of human auditory cortex: evidence from acoustic invariance in the response to intelligible speech. Cereb Cortex 20:2486–2495.

Perani D, Abutalebi J (2005) The neural basis of first and second language processing. Curr Opin Neurobiol 15:202–206.

Pickering MJ, Garrod S (2004) Toward a mechanistic psychology of dialogue. Behav Brain Sci 27:169–190; discussion 190–226.

Quiroga RQ, Reddy L, Kreiman G, Koch C, Fried I (2005) Invariant visual representation by single neurons in the human brain. Nature 435:1102–1107.

Romanski LM, Averbeck BB (2009) The primate cortical auditory system and neural representation of conspecific vocalizations. Annu Rev Neurosci 32:315–346.

Sahin NT, Pinker S, Cash SS, Schomer D, Halgren E (2009) Sequential processing of lexical, grammatical, and phonological information within Broca's area. Science 326:445–449.

Stephens GJ, Silbert LJ, Hasson U (2010) Speaker-listener neural coupling underlies successful communication. Proc Natl Acad Sci U S A 107:14425–14430.

Warren JD, Jennings AR, Griffiths TD (2005) Analysis of the spectral envelope of sounds by the human brain. Neuroimage 24:1052–1057.

Whorf BL (1956) Language, thought, and reality: selected writings. Cambridge: Technology Press of Massachusetts Institute of Technology.

Wildgruber D, Hertrich I, Riecker A, Erb M, Anders S, Grodd W, Ackermann H (2004) Distinct frontal regions subserve evaluation of linguistic and emotional aspects of speech intonation. Cereb Cortex 14:1384–1389.

Xu J, Kemeny S, Park G, Frattali C, Braun A (2005) Language in context: emergent features of word, sentence, and narrative comprehension. Neuroimage 25:1002–1015.

Zarahn E, Aguirre GK, D'Esposito M (1997) Empirical analyses of BOLD fMRI statistics. I. Spatially unsmoothed data collected under null-hypothesis conditions. Neuroimage 5:179–197.