Amplification of local changes along the timescale processing hierarchy

Yaara Yeshurun[a,b,1], Mai Nguyen[a,b], and Uri Hasson[a,b]

[a]Department of Psychology, Princeton University, Princeton, NJ 08540; and [b]Princeton Neuroscience Institute, Princeton University, Princeton, NJ 08540

Small changes in word choice can lead to dramatically different interpretations of narratives. How does the brain accumulate and integrate such local changes to construct unique neural representations for different stories? In this study, we created two distinct narratives by changing only a few words in each sentence (e.g., "he" to "she" or "sobbing" to "laughing") while preserving the grammatical structure across stories. We then measured changes in neural responses between the two stories. We found that differences in neural responses between the two stories gradually increased along the hierarchy of processing timescales. For areas with short integration windows, such as early auditory cortex, the differences in neural responses between the two stories were relatively small. In contrast, in areas with the longest integration windows at the top of the hierarchy, such as the precuneus, temporal parietal junction, and medial frontal cortices, there were large differences in neural responses between stories. Furthermore, this gradual increase in neural differences between the stories was highly correlated with an area's ability to integrate information over time. Amplification of neural differences did not occur when changes in words did not alter the interpretation of the story (e.g., sobbing to "crying"). Our results demonstrate how subtle differences in words are gradually accumulated and amplified along the cortical hierarchy as the brain constructs a narrative over time.

timescale | amplification | fMRI | narrative | hierarchy

Stories unfold over many minutes and are organized into temporarily nested structures: Paragraphs are made of sentences, which are made of words, which are made of phonemes. Understanding a story therefore requires processing the story at multiple timescales, starting with the integration of phonemes to words within a relatively short temporal window, to the integration of words to sentences across a longer temporal window of a few seconds, and up to the integration of sentences and paragraphs into a coherent narrative over many minutes. It was recently suggested that the temporally nested structure of language is processed hierarchically along the cortical surface (1–4). A consequence of the hierarchical structure of language is that small changes in word choice can give rise to large differences in sentence and overall narrative interpretation. Here, we propose that local momentary changes in linguistic input in the context of a narrative (e.g., "he" to "she" or "sobbing" to "laughing") are accumulated and amplified during the processing of linguistic content across the cortex. Specifically, we hypothesize that as areas in the brain increase in their ability to integrate information over time (e.g., processing "word" level content vs. "sentence" or "narrative" level content), the neural response to small word changes will become increasingly divergent.

Previously, we defined a temporal receptive window (TRW) as the length of time during which prior information from an ongoing stimulus can affect the processing of new information. We found that early sensory areas, such as auditory cortex, have short TRWs, accumulating information over a very short period (10–100 ms, equivalent to articulating a phoneme or word), while adjacent areas along the superior temporal sulcus (STS) have intermediate TRWs (a few seconds, sufficient to integrate information at the sentence level). Areas at the top of the processing hierarchy, including the temporal parietal junction (TPJ), angular gyrus, and posterior and frontal medial cortices, have long TRWs (many seconds to minutes) that are sufficient to integrate information at the paragraph and narrative levels (2, 5–7). The ability of an area to integrate information over time may be related to its intrinsic cortical dynamics: Long-TRW areas typically have slower neural dynamic than short-TRW areas (7, 8).

Long-TRW areas, which were identified with coherent narratives spanning many minutes, overlap with the semantic system, which plays an important role in complex information integration (9, 10). It was previously suggested that this network of areas supports multimodal conceptual representation by integrating information from lower level modality-specific areas (11–15). However, the semantic models used in previous work (9–15) do not model the relationship between words over time. In real-life situations, the context of words can substantially influence their interpretation, and thereby their representation in the cortex. For example, we recently demonstrated that a single sentence at the beginning of a story changes the interpretation of the entire story and, consequently, the responses in areas with long TRWs (16). Collectively, this work suggests that long-TRW areas not only integrate information from different modalities as shown in work on the semantic system but also integrate information over time (2).

We propose that the timescale processing hierarchy accumulates and amplifies small changes in word choice, resulting in highly divergent overall narratives (Fig. 1). This process is analogous to the integration of visual features across visual areas with increasing spatial receptive fields (SRFs) in visual cortex (17–19). In the visual domain, cortical areas with larger SRFs integrate and summate information from downstream areas that have smaller SRFs (20,

Significance

Changing a few words in a story can drastically alter the interpretation of the narrative. How does the brain amplify sparse changes in word selection to create a unique neural representation for each narrative? In our study, participants listened to one of two stories while being scanned using fMRI. The stories had an identical grammatical structure but varied in a small number of words, resulting in two distinct narratives. We found that differences in neural responses between the stories were amplified as story information was transferred from low-level regions (that are sensitive to the words' acoustic features) to high-level regions (that accumulate and integrate information across sentences). Our results demonstrate how subtle differences can be accumulated and amplified along the cortical hierarchy.

NEUROSCIENCE

21). As a result, small changes in the visual field, which are detected by low-level areas with small SRFs, can be amplified as they are integrated up the visual hierarchy to areas with successively larger SRFs (22). Analogously, temporally local, momentary changes in the content of linguistic input in the context of a narrative (e.g., he vs. she) will introduce short, transient changes in the responses of areas with short TRWs. However, such temporally transient changes can affect the interpretation of a sentence (e.g., "he built a wall" vs. "she built a wall"), which unfolds over a few seconds, as well as the interpretation of the overall narrative ("they didn't pay for it, a big surprise" vs. "they didn't vote for her, a big surprise"), which unfolds over many minutes. As a result, we predict that sparse and local changes in the content of words, which drastically alter the narrative, will be gradually amplified along the timescale hierarchy from areas with short TRWs to areas with medium to long TRWs.

To test these predictions, we scanned subjects using fMRI while they listened to one of two stories. The two stories had the same grammatical structure but differed on only a few words per sentence, resulting in two distinct, yet fully coherent narratives (Fig. 2A). To test for differences in neural responses to these two stories, we measured the Euclidean distance between time courses for the two stories in each voxel. In line with our hypothesis, we found a gradual divergence of the neural responses between the two stories along the timescale hierarchy. The greatest divergence occurred in areas with long TRWs and concomitant slow cortical dynamics. Neural differences were significantly correlated with TRW length. We did not observe this pattern of divergence when word changes did not result in different narratives or when narrative formation was interrupted. Our results suggest that small neural differences in low-level areas, which arise from local differences in the speech sounds, are gradually accumulated and amplified as information is transmitted from one level of the processing hierarchy to the next, ultimately resulting in distinctive neural representations for each narrative at the top of the hierarchy.

Results

Subjects were scanned using fMRI while listening to either Story1 or Story2. These two stories had identical grammatical structure but differed in one to three words per sentence, giving rise to very distinct narratives. To test that neural differences between stories arose from changes in narrative interpretation, an additional group of subjects listened to Story1_synonyms, which told the same narrative as Story1 but with some words replaced with synonyms, and Story2_scrambled, which disrupted narrative
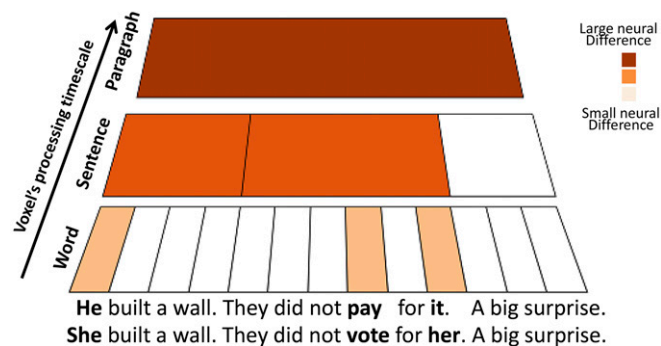

Fig. 1. Hypothesis. There are two paragraphs with three sentences each. Only three words (out of 13) differ between paragraphs, but these small local changes result in large changes in the overall narrative. We hypothesize that voxels with short TRWs will have relatively small neural differences between narratives (bright orange), whereas voxels with long TRWs will have relatively large neural differences (brown).
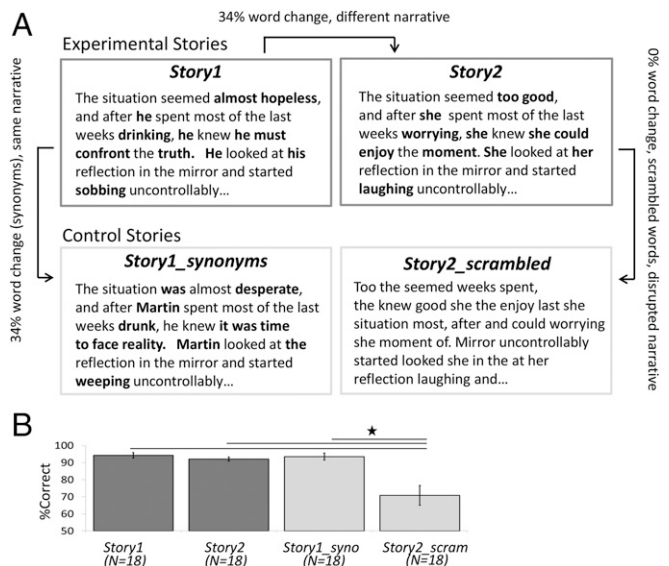

Fig. 2. Stimuli. We scanned subjects while they listened to one of four stories. (A) Two experimental stories had the same grammatical structure but differed in 34% of the words, resulting in two distinct narratives excerpted from Story1 (*Left Upper*) and from Story2 (*Right Upper*). (*Lower Left*) First control story (Story1_synonyms) had the same grammatical structure as Story1 but 34% of the words were changed to their synonyms, resulting in the same narrative as Story1. (*Lower Right*) Second control story had the same words as Story2 but the words were scrambled within a sentence, resulting in a disrupted narrative. (B) Participants' performance on the behavioral questionnaire revealed that the comprehension level for each of the experimental stories was high, with no difference between the two groups. The comprehension for the control Story1_synonyms (Story1_syno) was high and similar to the two experimental stories. The comprehension for the control Story2_scrambled (Story2_scram) was relatively low and impaired compared with the two experimental stories. *$P < 10^{-5}$.

interpretation by scrambling the words within each sentence in Story2.

Behavioral Results: Similar Comprehension of Intact, but Not Scrambled, Stories. Story comprehension was assessed with a 28-question quiz following the scan. Comprehension for both Story1 and Story2 was high (Story1: 94.2 ± 1.6%; Story2: 92.1 ± 1.2%), with no difference between the two story groups [$t(34) = 1.1$, $P = 0.28$], indicating that both stories were equally comprehensible (Fig. 2B).

Furthermore, the comprehension for the control Story1_synonyms was high (93.5 ± 2%) and similar to the two experimental stories [Story1_synonyms vs. Story1: $t(34) = 0.38$, $P = 0.7$; Story1_synonyms vs. Story2: $t(34) = 0.78$, $P = 0.44$]. As predicted, scrambling the order of the words within a sentence significantly reduced the comprehension level for the control Story2_scrambled (70.8 ± 5.74%) compared with the two experimental stories [Story2_scrambled vs. Story1: $t(34) = 7.47$, $P < 10^{-6}$; Story2_scrambled vs. Story2: $t(34) = 7.21$, $P < 10^{-6}$].

Neural Results: Differences and Similarities Among Stories Along Timescale Hierarchy. Differences in neural response across the timescale processing hierarchy among the four stories were then analyzed using both Euclidean distance and intersubject correlation (ISC; details are provided in *Methods*). In the main analysis, we first compared the neural responses between two distinct narratives that only differ in a few words per sentence (Story1 and Story2) across the timescale processing hierarchy. In two control analyses, we then compared processing of Story1 with processing of Story1_synonyms and Story2_scrambled.

Increased Neural Difference Between the Stories from Short- to Long-Timescale Regions. We hypothesized that neural responses to Story1 and Story2, which vary only in a few words per sentence yet convey completely different narratives, will increasingly diverge along the timescale processing hierarchy. To test this hypothesis, we calculated the Euclidean distance between response time courses of the two main story groups within all voxels that responded reliably to both stories (7,591 voxels; *Defining Voxels That Responded Reliably to the Story*). We ranked the 7,591 reliable voxels based on their neural Euclidean distance, and then divided them into five equal-sized bins. These bins are presented in Fig. 3*A*, showing bins with relatively small differences between Story1 and Story2 (−0.05 ± 0.05, light orange color) up to bins with relatively large differences between the stories (0.44 ± 0.11, brown color). Consistent with our hypothesis, we find that voxels with the least difference between stories (Fig. 3*A*, light orange) are predominantly located in areas of the brain previously shown to have short TRWs, including auditory cortex, medial STS, and ventral posterior STS (marked in light orange). The distance in neural responses across the two groups gradually increased along the timescale hierarchy, with the largest Euclidean distances (Fig. 3*A*, dark orange and brown) between stories located in medium- to long-TRW areas, such as the precuneus, bilateral angular gyrus, bilateral temporal poles, and medial and lateral prefrontal cortex (marked in dark orange and brown; details are provided in Table S1).
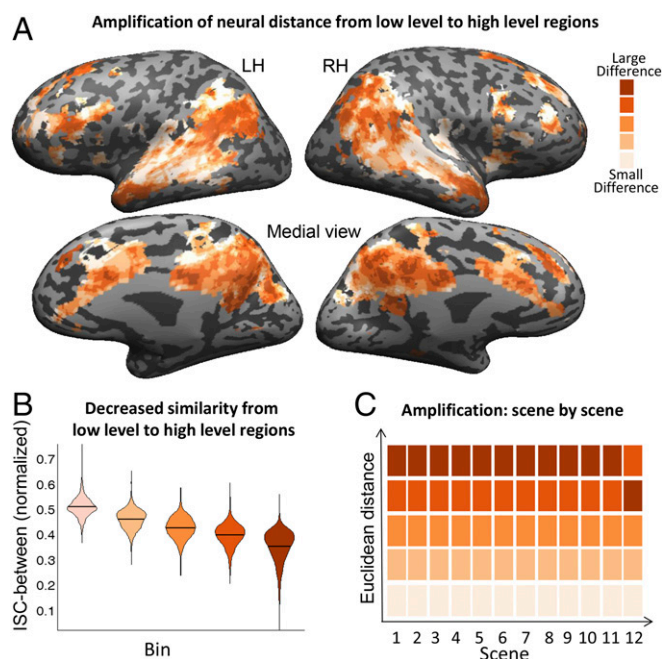
To provide an intuitive sense of the magnitude of the increasing neural divergence, we also measured the similarity of the neural response to Story1 and Story2 using a between-groups ISC analysis (details are provided in *Methods*). We found that the bin with the smallest Euclidean distance had the highest mean ISC similarity between the stories (mean = 0.51 ± 0.04). The range of ISC values among the voxels in this bin was skewed toward greater between-story similarity (range: 0.39–0.73; Fig. 3*B*). In contrast, the bin with the largest Euclidean distance showed very little cross-story ISC similarity (mean = 0.35 ± 0.06), with skewed probability toward areas with close to zero similarity in response patterns across the two stories (range: 0.066–0.52; Fig. 3*B*). These differences between the bins were highly significant as revealed by one-way ANOVA on the normalized ISC [$F(1, 7,585) = 2,622.58$, $P < 10^{-10}$; Scheffé's post hoc comparisons revealed that all of the bins' normalized ISCs differ significantly from each other].

We next tested whether the amplification of neural differences from short- to long-TRW areas occurred in every scene of Story1 and Story2. To that end, we calculated the Euclidean distance between the stories' neural response in each of the 12 scenes, averaged them across voxel bins (defined over the entire story), and then rank-ordered these average distance values. In 11 of the 12 scenes, the ordering of bins was the same as for the overall story, with only a small change in the Euclidean distance between the two bins with larger differences in the last segment (Fig. 3*C*). Thus, the consistency of ordering across the 12 independent segments suggests that the amplification of neural distance from low-level (short-TRW) regions to high-level (long-TRW) regions is robust and stable.

Significant Correlation Between Neural Distance and Capacity to Accumulate Information Over Time. Next, we asked whether the difference in neural response to Story1 and Story2 was related to a voxel's capacity to accumulate information over time. To characterize this capacity, we calculate a TRW index using an independent dataset obtained while subjects listened to intact and word-scrambled versions of another story (details are provided in *Methods*). Comparison of the TRW index brain map (Fig. 4*A*, *Upper*) and the two stories' Euclidean distance map (Fig. 4*A*, *Lower*) revealed high similarity between the maps. Indeed, we found, at a voxel-by-voxel basis, that the larger the difference in the neural responses between the stories, the larger was the voxel's TRW index ($r = 0.426$, $P < 0.001$) (Fig. 4*B*). As an area's TRW size was found to be related to intrinsically slower cortical dynamics in these areas (7, 8), we also calculated the proportion of low-frequency power during a resting state scan. We found that voxels with a larger neural difference between stories also had a greater proportion of low-frequency power ($r = 0.246$, $P < 0.001$) (Fig. 4*C*).

Amplification pattern is dependent on differences in interpretations between stories. In the present work, we observed increasingly divergent neural responses across the timescale processing hierarchy during the processing of two distinct narratives that vary only in a small number of words. However, it is possible that the differences in neural responses at the top of the hierarchy are induced by increased sensitivity to local changes in the semantic content of each word (e.g., laughing vs. sobbing), irrespective of the aggregate effect such local changes have on the overall narrative. To test whether the amplification effect arises from differences in the way the narrative is constructed over time, we added a new condition (Story1_synonyms). Story1 and Story1_synonyms had the same grammatical structure and differed in the exact same number of words as the original Story1 and Story2 did. However, for Story1_synonyms, we replaced words with their synonyms (e.g. sobbing with "weeping") and a few pronouns with their proper nouns (e.g., he with "Martin"), preserving semantic and narrative content in the two stories. This manipulation resulted in similar and high-comprehension performance to both stories (Fig. 1*B*).

Fig. 3. Neural results demonstrating amplification of the neural distance between narratives. (*A*) We calculated the Euclidean distance between the response time courses to Story1 and Story 2, ranked the voxels based on their neural Euclidean distance, and then divided them into five equal-sized bins. Small neural differences are primarily observed in and around primary auditory cortex, while increasingly large neural differences are observed extending toward TPJ, precuneus, and frontal areas. LH, left hemisphere; RH, right hemisphere. (*B*) Normalized between group ISC (detailed explanation is provided in *Methods*) in each of the five bins. The bin with the smallest Euclidean distance (marked in light orange) showed the highest ISC, whereas the bin with the largest Euclidean distance showed the lowest ISC between the stories (brown color). These differences between the bins were highly significant. (*C*) We calculated the Euclidean distance between the stories' neural response in each of the 12 scenes, and then rank-ordered these average distance values. In 11 of the 12 scenes, the ordering of bins was the same as for the overall story, with only a small change in the Euclidean distance between the areas with large differences in the last segment.
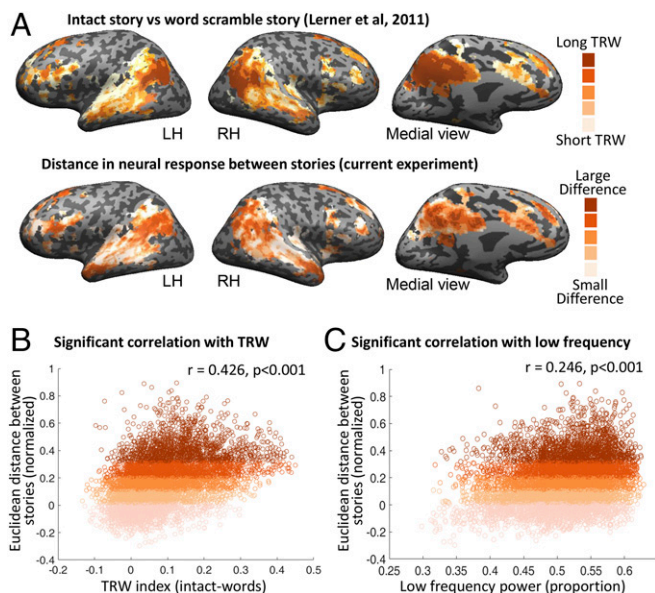
Fig. 4. Correlation with processing timescales. (*A*) TRW index brain map (*Upper*) and Story1 vs. Story2 Euclidean distance map (*Lower*). LH, left hemisphere; RH, right hemisphere. (*B*) Scatter plot of the voxel's TRW index and the Euclidean distance between Story1 and Story2. The larger the voxel's TRW index, the larger was the difference in the neural responses between the stories. (*C*) Scatter plot of the voxel's proportion of low-frequency power during a resting state scan and the Euclidean distance between the stories. Voxels with a greater low-frequency power proportion were correlated with a larger neural difference between stories.

Unlike the neural difference between Story1 and Story2, we found that the neural distance between Story1 and Story1_synonyms did not increase from short-TRW to long-TRW regions (Fig. 5 *B* and *C*). Indeed, we found larger differences between Story1 and Story1_synonyms in short-TRW areas compared with long-TRW areas. These results are consistent with our processing timescale hierarchy model. Some of the acoustic differences between Story1 and Story1_synonyms (e.g., the pronoun he and the proper noun Martin) are larger than the acoustic differences between Story1 and Story2 (e.g. the pronoun he and the pronoun she). However, the semantic differences between Story1 and Story1_synonyms are small (e.g., weeping vs. "crying") relative to the semantic differences between Story1 and Story2 (e.g., weeping vs. laughing). We thus did observe relatively larger differences in neural responses between Story1 and Story1_synonyms in early auditory areas, but not in high-order semantic areas that process long timescales (Fig. 5*B*, *Center*), resulting in a slightly negative correlation ($r = -0.087$, $P < 0.001$) between the neural distance between Story1 and Story1_synonyms and the TRW (Fig. 5*D*). In contrast, while we did observe some small differences in the neural response in auditory areas between Story1 and Story2, we additionally observed progressively larger differences in higher level areas of the processing hierarchy, resulting in a positive correlation between neural difference and TRW size (Fig. 4*B*).

Amplification pattern is dependent on construction of coherent interpretation over time. In a second control experiment, we preserved the semantic identity of all words but scrambled the order of words within each sentence. This manipulation preserved the semantic content of all words in each sentence but impaired the listeners' ability to integrate the words into a coherent story, resulting in significantly lower comprehension performance (Fig. 2 *A* and *B*). In agreement with prior results (2), scrambling the order of words decreased the reliability of responses in areas with middle to long processing timescales (which integrate, respectively, words to sentences and sentences to ideas; Fig. 5*A*, *Right*). The lack of reliable responses to the scrambled story induced a constant neural

distance between Story1 and scrambled Story2 (Fig. 5*B*, *Right*). As a result, we observed no change in the neural distance between Story1 and Story2_scrambled across the timescale hierarchy. Taken together, these two control experiments demonstrate that the amplification of neural distance along the processing hierarchy relies on the ability of the brain to synthesize over time subtle differences in word meaning into a large and coherent difference in narrative.

Discussion

Small, local changes of word choice in a story can completely alter the interpretation of a narrative. In this study, we predicted that the accumulation and amplification of sparse, temporally local word changes that generate a big difference in the narrative will be reflected in increasingly divergent neural response along the timescale processing hierarchy. In line with our predictions, we found that short timescale areas, including primary auditory cortex, showed only small neural differences in response to local word changes. This finding is consistent with observations that these early auditory areas process transient and rapidly changing sensory input (23, 24), such that brief local alternations in the sound structure will only induce brief and local alternations in the neural responses across the two stories. However, neural differences in response across the two groups gradually increased along the processing timescale hierarchy, with the largest neural distance in areas with a long TRW, including TPJ, angular gyrus, posterior cingulate cortex, and dorsomedial prefrontal cortex. This large neural difference in long-TRW areas is consistent with our previous work (25) that suggests long-TRW areas gradually construct and retain a situation model (26, 27) of a narrative as it unfolds
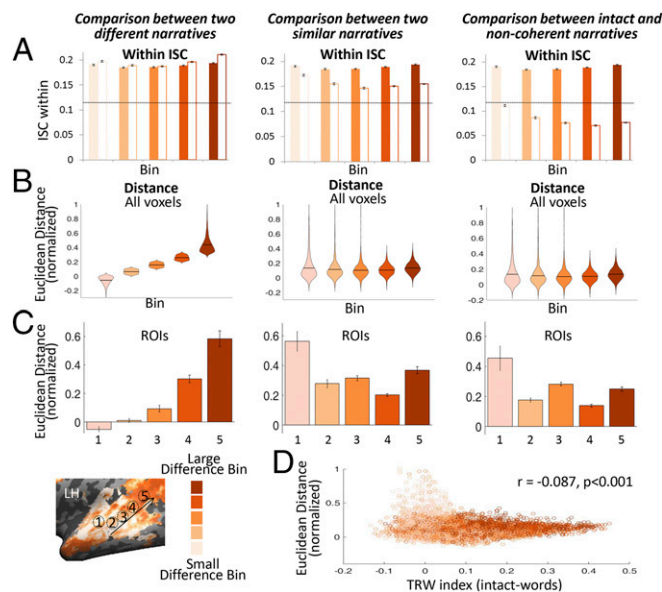

Fig. 5. Amplification pattern is dependent on forming distinct narrative interpretations. (*A*) Within-subject ISC across bins for Story1 (*Left*, full bars) and Story2 (*Left*, white bars), for Story1_synonyms (*Center*, white bars), and for Story2_scrambled (*Right*, white bars). A significant ISC threshold is marked as a dotted black line. The reliability of responses within each of the intact stories was significantly above threshold, whereas reliability for Story2_scrambled was not. (*B*) Euclidean distance in neural response across bins between Story1 and Story2 (*Left*), between Story1 and Story1_synonyms (*Center*), and between Story1 and Story2_scrambled (*Right*). (*C*) Euclidean distance in neural response between stories in five specific regions of interest (ROIs; one from each bin) (*Upper*) illustrated on the brain map (*Lower*). (*D*) Scatter plot of the voxel's TRW index and the Euclidean distance between Story1 and Story1_synonyms. The smaller the voxel's TRW index, the larger was the difference in the neural responses between the stories.

over time (28, 29). Moreover, the difference in neural response between stories was significantly correlated with both TRW length and slower cortical dynamics, suggesting that the increasing divergence in neural response along the hierarchy may arise from an accumulation and amplification of small differences over time. Notably, these effects were only evident when the words changed created two different intact narratives (Fig. 5).

The topographical hierarchy of processing timescales along the cortical surface (2) has been supported and replicated across methodologies, including single-unit analysis (4), electrocorticography analysis (7), fMRI analysis (5, 6, 25), magnetoencephalography analysis (1), computational models (3), and resting state functional connectivity (30, 31). This hierarchy, as well as the divergence of neural responses observed here, is additionally consistent with previously proposed linguistic hierarchies: low-level regions (A1+) represent phonemes (32, 33), syllables (34), and pseudowords (35), while medium-level regions (areas along A1+ to STS) represent sentences (36, 37). At the top of the hierarchy, high-level regions (bilateral TPJ, precuneus, and medial prefrontal cortex) can integrate words and sentences into a meaningful, coherent whole narrative (5, 15, 38–40). We found that the regions at the top of the hierarchy had the most divergent response to the two stories. These regions overlap with the semantic system (9, 10), which is thought to play a key role in combining information from different modalities to create a rich representation (11–14). Our results suggest that the semantic system is also involved in integration of information over time. We found that regions within the semantic network amplify the difference created by local word changes only if (*i*) the words are presented in an order that allow for meaningful integration of information over time and (*ii*) the integration of the words creates a semantic difference in the narratives (Fig. 5).

Does the timescale processing hierarchy capture changes over time (milliseconds to seconds to minutes), changes in information type (phonemes to words to sentences to paragraphs), or both? A previous study that varied the rate of information transmission suggests that TRWs are not dependent on absolute time per se, but on the amount of information conveyed in that time (41), suggesting a tight link between ongoing, temporally extended neural processes, memory, and linguistic content (2). By keeping the speech rate constant, the present study observed that the longer the processing integration window, the greater was the divergence of neural responses to the two stories. It is likely that these differences may additionally arise from differences in the type of information processed across the hierarchy. Further research is necessary to characterize other factors that influence the amplification of neural distance along the processing hierarchy.

Understanding a narrative requires more than just understanding the individual words in the narrative. For example, children with hydrocephaly, a neurodevelopmental disorder that is associated with brain anomalies in regions that include the posterior cortex, have well-developed word decoding but concomitant poor understanding of narrative constructed from the same words (42, 43). In Parkinson's disease (PD) and early Alzheimer's disease, researchers have found that individual word comprehension is relatively intact, yet the ability to infer the meaning of the text is impaired (44). In PD, this deficit in the organization and interpretation of narrative discourse was associated with reduced cortical volume in several brain regions, including the superior part of the left STS and the anterior cingulate (45). Our finding that these regions demonstrated large differences between the stories (Fig. 3) suggests that a reduced volume of cortical circuits with long processing timescales may result in a reduced capability to understand temporally extended narratives.

The phenomenon we described here, in which local changes in the input generate a large change in the meaning, is ubiquitous: A small change in the size of the pupil can differentiate surprise from anger, a small change in intonation can differentiate comradery from mockery, and a small change in hand pressure can differentiate

comfort from threat. What is the neural mechanism underlying this phenomenon? In vision, researchers have shown that low-level visual areas with small SRFs are sensitive to spatially confined changes in the visual field, but that these small changes are accumulated and integrated along the visual processing stream such that high-level areas, with large SRFs, amplify the responses to very small changes in visual stimuli. Analogously, we propose that our results demonstrate how local changes in word choice can be gradually amplified and generate large changes in the interpretation and responses in high-order areas that integrate information over longer time periods.

Methods

Subjects. Fifty-four right-handed subjects (aged 21 ± 3.4 y) participated in the study. Two separate groups of 18 subjects (nine females per group) listened to Story1 and Story2. A third group of 18 subjects (nine females) listened to two control stories: Story1_synonyms and Story2_scrambled. Three subjects were discarded from the analysis due to head motion (>2 mm). Experimental procedures were approved by the Princeton University Committee on Activities Involving Human Subjects. All subjects provided written informed consent.

Stimuli and Experimental Design. In the MRI scanner, participants listened to Story1 or Story2 (experimental conditions) or to Story1_synonyms and Story2_scrambled (control conditions). Story1 and Story2 had the exact same grammatical structure but differed in a few words per sentence, resulting in two distinct narratives (mean word change in a sentence = 2.479 ± 1.7, 34.24 ± 20.8% of the words in the sentence). Story_synonyms had the same narrative and grammatical structure as Story1, differing in only a few words per sentence (mean word change in a sentence = 2.472 ± 1.7, 34.21 ± 20.9% of the words in the sentence), but words were replaced with their synonyms, resulting in the same narrative as Story1 (Fig. 2A). Story2_scrambled had the exact same words as Story2, but the words in each sentence were randomly scrambled (Fig. 2A). This manipulation created a narrative that was very hard to follow. More details are provided in *Stimuli and Experimental Design*.

In addition, 26 of the subjects (13 from the Story1 group and 13 from the Story2 group) underwent a 10-min resting state scan. Subjects were instructed to stay awake, look at a gray screen, and "think on whatever they like" during the scan.

Behavioral Assessment. Immediately following scanning, each participant's comprehension of the story was assessed using a questionnaire presented on a computer. Twenty-eight two-forced choice questions were presented. Two-tailed Student *t* tests (α = 0.05) on the forced choice answers were conducted between the four story groups to evaluate the difference in participants' comprehension.

MRI Acquisition. Details are provided in *MRI Acquisition*.

Data Analysis.
Preprocessing. The fMRI data were reconstructed and analyzed with the BrainVoyager QX software package (Brain Innovation) and with in-house software written in MATLAB (MathWorks). Preprocessing of functional scans included intrasession 3D motion correction, slice-time correction, linear trend removal, and high-pass filtering (two cycles per condition). Spatial smoothing was applied using a Gaussian filter of 6-mm full-width at half-maximum value. The complete functional dataset was transformed to 3D Talairach space (46).
Euclidean distance measure. We were interested in the differences in the neuronal response of subjects listening to Story1 compared with those listening to Story2. To test for such differences, we used a Euclidean distance metric in all of the voxels that reliably responded to both (details are provided in *Euclidean Distance Measure*). Reliable voxels were then ranked based on their normalized Euclidean distance value (D_{norm}) and divided into five equal-sized bins. These bin categories (small to large neural difference) were then projected onto the cortical surface for visualization. Finally, we tested whether the order of these neural difference bins was consistent across sections of the story. Thus, we calculated D_{norm} for each of 12 scenes of the story across all of the included voxels. These voxels were then sorted based on their bin from calculating D_{norm} across the entire story. The D_{norm} values for each bin were averaged across voxels for each scene and then rank-ordered.
ISC between the stories. The normalized Euclidean distance measure of similarity has an arbitrary scale, and thus does not provide an intuitive sense of the magnitude of effects. As a complementary measure, we additionally calculated similarity using ISC (details are provided in *ISC Between the Stories*).

Correlation between neural distance and capacity to accumulate information over time. Next, we calculated an index of each voxel's TRW (i.e., capacity to accumulate information over time) and calculated the correlation between the TRW index and neural response difference (details are provided in *Correlation Between Neural Distance and Capacity to Accumulate Information Over Time*).
Relation between neural distance and timescale of blood oxygen level-dependent signal dynamics. We evaluated the neural dynamics of our reliable voxels, and calculated the Pearson correlation between the proportion of low-frequency power and the neural response difference voxel by voxel (details are provided in *Relation Between Neural Distance and Timescale of BOLD Signal Dynamics*).
Control for word change-induced differences not related to the narrative. Story1 and Story2 differed in 34% of their words, which resulted in very different narratives. To dissociate neural differences that arose from changing word form (e.g., the acoustic differences between words) and neural differences that arose from changing word meaning (e.g., the semantic differences between words), we also included a control group that listened to Story1_synonyms. Story1 and Story1_synonyms differed in 34% of their words.

However, as the words in Story1 were replaced by synonyms, the narratives of the two stories were the same. We calculated the normalized Euclidean distance between Story1 and Story1_synonyms using the same procedure described above in each of the reliable voxels.

We also conducted an additional control, the Story2_scrambled story. This story had the exact same words as Story2, but the words within each sentence were scrambled, preventing the formation of a coherent narrative interpretation. We calculated the normalized Euclidean distance between Story1 and Story2_scrambled using the same procedure described above in each of the reliable voxels.

1. Ding N, Melloni L, Zhang H, Tian X, Poeppel D (2016) Cortical tracking of hierarchical linguistic structures in connected speech. *Nat Neurosci* 19:158–164.
2. Hasson U, Chen J, Honey CJ (2015) Hierarchical process memory: Memory as an integral component of information processing. *Trends Cogn Sci* 19:304–313.
3. Kiebel SJ, Daunizeau J, Friston KJ (2008) A hierarchy of time-scales and the brain. *PLoS Comput Biol* 4:e1000209.
4. Murray JD, et al. (2014) A hierarchy of intrinsic timescales across primate cortex. *Nat Neurosci* 17:1661–1663.
5. 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.
6. Hasson U, Yang E, Vallines I, Heeger DJ, Rubin N (2008) A hierarchy of temporal receptive windows in human cortex. *J Neurosci* 28:2539–2550.
7. Honey CJ, et al. (2012) Slow cortical dynamics and the accumulation of information over long timescales. *Neuron* 76:423–434.
8. Stephens GJ, Honey CJ, Hasson U (2013) A place for time: The spatiotemporal structure of neural dynamics during natural audition. *J Neurophysiol* 110:2019–2026.
9. Binder JR, Desai RH, Graves WW, Conant LL (2009) Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb Cortex* 19:2767–2796.
10. Huth AG, de Heer WA, Griffiths TL, Theunissen FE, Gallant JL (2016) Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature* 532:453–458.
11. Fernandino L, et al. (2016) Concept representation reflects multimodal abstraction: A framework for embodied semantics. *Cereb Cortex* 26:2018–2034.
12. Fernandino L, Humphries CJ, Conant LL, Seidenberg MS, Binder JR (2016) Heteromodal cortical areas encode sensory-motor features of word meaning. *J Neurosci* 36:9763–9769.
13. Binder JR, Westbury CF, McKiernan KA, Possing ET, Medler DA (2005) Distinct brain systems for processing concrete and abstract concepts. *J Cogn Neurosci* 17:905–917.
14. Graves WW, Binder JR, Desai RH, Conant LL, Seidenberg MS (2010) Neural correlates of implicit and explicit combinatorial semantic processing. *Neuroimage* 53:638–646.
15. Price AR, Bonner MF, Peelle JE, Grossman M (2015) Converging evidence for the neuroanatomic basis of combinatorial semantics in the angular gyrus. *J Neurosci* 35:3276–3284.
16. Yeshurun Y, et al. (2017) Same story, different story. *Psychol Sci* 28:307–319.
17. Dumoulin SO, Wandell BA (2008) Population receptive field estimates in human visual cortex. *Neuroimage* 39:647–660.
18. Grill-Spector K, Malach R (2004) The human visual cortex. *Annu Rev Neurosci* 27:649–677.
19. Smith AT, Singh KD, Williams AL, Greenlee MW (2001) Estimating receptive field size from fMRI data in human striate and extrastriate visual cortex. *Cereb Cortex* 11:1182–1190.
20. Kay KN, Winawer J, Mezer A, Wandell BA (2013) Compressive spatial summation in human visual cortex. *J Neurophysiol* 110:481–494.
21. Press WA, Brewer AA, Dougherty RF, Wade AR, Wandell BA (2001) Visual areas and spatial summation in human visual cortex. *Vision Res* 41:1321–1332.
22. Hasson U, Hendler T, Ben Bashat D, Malach R (2001) Vase or face? A neural correlate of shape-selective grouping processes in the human brain. *J Cogn Neurosci* 13:744–753.
23. Okada K, et al. (2010) Hierarchical organization of human auditory cortex: Evidence from acoustic invariance in the response to intelligible speech. *Cereb Cortex* 20:2486–2495.
24. Poeppel D (2003) The analysis of speech in different temporal integration windows: Cerebral lateralization as 'asymmetric sampling in time'. *Speech Commun* 41:245–255.
25. Baldassano C, et al. (2016) Discovering event structure in continuous narrative perception and memory. *bioRxiv*: 081018.
26. van Dijk TA, Kintsch W (1983) *Strategies of Discourse Comprehension* (Academic, New York).
27. Zwaan RA, Radvansky GA (1998) Situation models in language comprehension and memory. *Psychol Bull* 123:162–185.
28. Speer NK, Zacks JM, Reynolds JR (2007) Human brain activity time-locked to narrative event boundaries. *Psychol Sci* 18:449–455.
29. Zacks JM, et al. (2001) Human brain activity time-locked to perceptual event boundaries. *Nat Neurosci* 4:651–655.
30. Sepulcre J, Sabuncu MR, Yeo TB, Liu H, Johnson KA (2012) Stepwise connectivity of the modal cortex reveals the multimodal organization of the human brain. *J Neurosci* 32:10649–10661.
31. Margulies DS, et al. (2016) Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proc Natl Acad Sci USA* 113:12574–12579.
32. Arsenault JS, Buchsbaum BR (2015) Distributed neural representations of phonological features during speech perception. *J Neurosci* 35:634–642.
33. Humphries C, Sabri M, Lewis K, Liebenthal E (2014) Hierarchical organization of speech perception in human auditory cortex. *Front Neurosci* 8:406.
34. Evans S, Davis MH (2015) Hierarchical organization of auditory and motor representations in speech perception: Evidence from searchlight similarity analysis. *Cereb Cortex* 25:4772–4788.
35. Binder JR, et al. (2000) Human temporal lobe activation by speech and nonspeech sounds. *Cereb Cortex* 10:512–528.
36. Pallier C, Devauchelle AD, Dehaene S (2011) Cortical representation of the constituent structure of sentences. *Proc Natl Acad Sci USA* 108:2522–2527.
37. Fedorenko E, et al. (2016) Neural correlate of the construction of sentence meaning. *Proc Natl Acad Sci USA* 113:E6256–E6262.
38. Hickok G, Poeppel D (2007) The cortical organization of speech processing. *Nat Rev Neurosci* 8:393–402.
39. Fletcher PC, et al. (1995) Other minds in the brain: A functional imaging study of "theory of mind" in story comprehension. *Cognition* 57:109–128.
40. 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.
41. Lerner Y, Honey CJ, Katkov M, Hasson U (2014) Temporal scaling of neural responses to compressed and dilated natural speech. *J Neurophysiol* 111:2433–2444.
42. Barnes MA, Dennis M (1992) Reading in children and adolescents after early onset hydrocephalus and in normally developing age peers: Phonological analysis, word recognition, word comprehension, and passage comprehension skill. *J Pediatr Psychol* 17:445–465.
43. Barnes MA, Faulkner HJ, Dennis M (2001) Poor reading comprehension despite fast word decoding in children with hydrocephalus. *Brain Lang* 76:35–44.
44. Chapman SB, Anand R, Sparks G, Cullum CM (2006) Gist distinctions in healthy cognitive aging versus mild Alzheimer's disease. *Brain Impair* 7:223–233.
45. Ash S, et al. (2011) The organization of narrative discourse in Lewy body spectrum disorder. *Brain Lang* 119:30–41.
46. Talairach J, Tournoux P (1988) *Co-Planar Stereotaxic Atlas of the Human Brain* (Thieme, New York).
47. Pelli DG (1997) The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spat Vis* 10:437–442.