Cortical representation of the constituent structure of sentences

Christophe Pallier[a,b,c,d,1,2], Anne-Dominique Devauchelle[a,c,d,1], and Stanislas Dehaene[a,c,d,e,2]

[a]Institut National de la Santé et de la Recherche Médicale, Cognitive Neuroimaging Unit, U992, [b]Centre National de la Recherche Scientifique, F-75794 Paris, France; [c]NeuroSpin Center, Commissariat à l'Energie Atomique, Département des Sciences de la Vie/Institut d'Imagerie Biomédicale, and [d]Université Paris-Sud, F-91191 Gif/Yvette, France; and [e]Collège de France, F-75005 Paris, France

Contributed by Stanislas Dehaene, December 15, 2010 (sent for review September 23, 2010)

Linguistic analyses suggest that sentences are not mere strings of words but possess a hierarchical structure with constituents nested inside each other. We used functional magnetic resonance imaging (fMRI) to search for the cerebral mechanisms of this theoretical construct. We hypothesized that the neural assembly that encodes a constituent grows with its size, which can be approximately indexed by the number of words it encompasses. We therefore searched for brain regions where activation increased parametrically with the size of linguistic constituents, in response to a visual stream always comprising 12 written words or pseudowords. The results isolated a network of left-hemispheric regions that could be dissociated into two major subsets. Inferior frontal and posterior temporal regions showed constituent size effects regardless of whether actual content words were present or were replaced by pseudowords (jabberwocky stimuli). This observation suggests that these areas operate autonomously of other language areas and can extract abstract syntactic frames based on function words and morphological information alone. On the other hand, regions in the temporal pole, anterior superior temporal sulcus and temporo-parietal junction showed constituent size effect only in the presence of lexico-semantic information, suggesting that they may encode semantic constituents. In several inferior frontal and superior temporal regions, activation was delayed in response to the largest constituent structures, suggesting that nested linguistic structures take increasingly longer time to be computed and that these delays can be measured with fMRI.

Most theories of syntax described sentences as tree-like hierarchical structures of nested phrases. These phrases, or constituents, constitute syntactic units that can be moved or replaced as a whole (for example, a noun phrase can be replaced by a pronoun). Support for the psychological reality of syntactic structures comes from studies showing that speakers tend to reuse the syntactic structure of recently heard sentences, a phenomenon known as syntactic priming (1, 2). Yet, although a variety of approaches using brain imaging methods have sought to characterize the regions implicated in syntactic processes (see, e.g., refs. 3–7), how the human brain computes and encodes syntactic structures remain largely open questions (8–11).

In this article, we propose an experimental paradigm to address this issue. We start with the intuitive hypothesis that a more complex cell assembly should be needed to encode a constituent of three elements, such as "Mary's father's car," than a constituent of two elements such as "Mary's father." Thus, neural activity might increase by a fixed amount each time a new node must be added to the constituent structure constructed on the basis of the preceding words. Arguments in favor of such an "accumulation" model comes from neurophysiological recordings in awake macaque monkeys during sequence-learning tasks, where specific neurons enter into a sustained mode of activity at different points of the sequence, thus creating a cumulative code for the sequences content (12–14). In the human brain, such a pattern of neural discharges has not yet been observed during language processing, but a detailed neural network proposal for structure encoding that could produce such cumulative activity is the tensor product theory put forward by Paul Smolensky (15). This theory proposes that a tree structure can be represented neurally by the superposition of several neural assemblies, each consisting in the tensor product of two neural vectors, a "role" and a "filler." The role vector codes for the syntactic role of a word (i.e., its position in the constituent structure), whereas the filler vector codes for the content being encoded at this position (i.e., the specific lexical item). Assuming that at least one of these vectors is encoded by a sparse neural code (a small proportion of active neurons, the majority being inactive), the superposition principle implies that mean activation increase by an approximately constant amount each time a new tensor product is added to the overall structure. Thus, this model leads to the expectation of a linear increase in activation with the number of nodes of the constituent structure.

In our experiment, we presented human participants with sequences of written words that afforded the construction of constituent structures of variable size (Table 1). All sequences had a fixed length of 12 items. Those belonging to the experimental condition "c01" were simply lists of unrelated words, controlled such that neighboring words could not be combined into larger constituents. In condition c02, sequences could be parsed as a series of constituents of size 2; in c03, the successive constituents comprised 3 words, and so on, all of the way to a full sentence of 12 words (c12).

Fig. 1 exemplifies how the accumulation model might apply to our experiment. For simplicity, we assume that each incoming word elicits a corresponding increase in activation when it can be integrated into a constituent structure, and that activation returns to baseline whenever word integration fails. Note that this model serves only as a first-order predictive tool. In high-level linguistic structures, the number of nodes is likely to differ from the number of words, and future research should take into account the precise syntactic structure afforded by the stimuli. At this stage, however, the number of words should serve as a useful approximation. When constituent size increases, the model shown in Fig. 1 predicts that not only the amplitude of the response, but also its phase, should increase. This prediction is a consequence of the assumption that activation accumulates and is thus stronger toward the end of a constituent than at the beginning. Here, following earlier work (16–19), we show that fMRI can be sensitive to such temporal patterns and that several left-hemispheric regions follow the predicted increase in amplitude and phase.

As an additional experimental manipulation, we introduced a "jabberwocky" condition where content words were systematically replaced with pseudowords while maintaining all of the morphological markers and closed-class words needed to permit parsing (Table 1). This manipulation drastically reduces lexico-semantic content while keeping syntactic constituent structure intact, allowing to disentangle syntactic constituency effects from semantic co-

Author contributions: C.P., A.-D.D., and S.D. designed research; C.P. and A.-D.D. performed research; C.P., A.-D.D., and S.D. analyzed data; and C.P. and S.D. wrote the paper.

The authors declare no conflict of interest.

See Commentary on page 2177.

[1]C.P. and A.-D.D. contributed equally to this work.

[2]To whom correspondence may be addressed. E-mail: christophe.pallier@gmail.com or stanislas.dehaene@cea.fr.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1018711108/-/DCSupplemental.

Table 1. The stimuli were 12 items long sequences obtained by concatenating constituents of fixed sizes extracted from natural or jabberwocky right-branching sentences

Condition	Constituent size	Examples
c12	12 words	I believe that you should accept the proposal of your new associate *I tosieve that you should begept the tropufal of your tew viroate*
c06	6 words	the mouse that eats our cheese two clients examine this nice couch *the couse that rits our treeve fow plients afomine this kice bloch*
c04	4 words	mayor of the city he hates this color they read their names *tuyor of the roty he futes this dator they gead their wames*
c03	3 words	solving a problem repair the ceiling he keeps reading will buy some *relging a grathem regair the fraping he meeps bouding will doy some*
c02	2 words	looking ahead important task who dies his dog few holes they write *troking ahead omirpant fran who mies his gog few biles they grite*
c01	1 word	thing very tree where of watching copy tensed they states heart plus *thang very gree where of wurthing napy gunsed they otes blart trus*

In jabberwocky, all content words were replaced with pseudowords (*italics*). Examples are only illustrative, because the original stimuli were in French.

herence effects (7, 20–22). It also introduces an important control for transition probabilities. In the materials with actual words, constituent size is confounded with the mean transition probability between words, which is lower at constituent boundaries than within constituents. In the jabberwocky condition, however, transition probability is close to zero, so a systematic variation in activation from c01 to c12 conditions would necessarily indicate a sensitivity to constituent structure. A first group of 20 subjects (the "normal-prose group") read stimuli containing actual words, in which syntactic and semantic constituent structures were confounded, whereas a second group of 20 subjects received jabberwocky-type stimuli.

Results

Behavioral Results. During fMRI, participants were required to attentively read the sequences. To ensure attention, they had to detect rare probe sentences that requested them to press a button and were warned that a word memory test would be presented at the end of each run. The participants correctly detected 96.8% of the probes. On the memory task, the percentages of correct responses were 60% (SEM = 1.5%; comparison with chance level of 50%; $P < 0.01$) in the normal-prose group and 59% (SEM = 1.5%; $P < 0.01$). This relatively low level of performance is likely due to the large number of words (420) presented in each run over a period of 8 min. Overall, these behavioral results suggest that participants were actively attending to all stimuli.

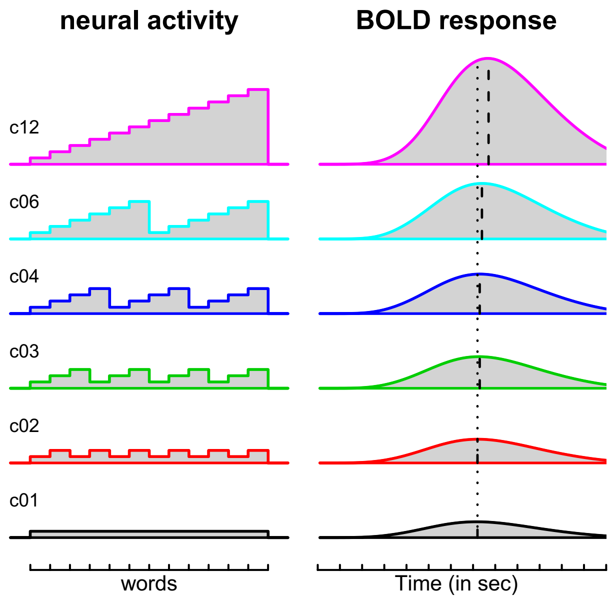


Fig. 1. Simulations of a simple model in which neural activity is assumed to increase by one unit each time a new word is incorporated into a constituent and to return to baseline as soon as a novel word cannot be incorporated into the preceding constituent. This model predicts an increase in both amplitude and phase of activation as a function of constituent size from condition c01 to c12.

Increasing Activation with Constituent Size. We first searched for brain regions showing increasing activation with constituent size. With normal prose, six brain regions showed a significant increase (voxel-based $P < 0.05$, corrected for multiple comparison, familywise error) (Fig. 2 and *SI Appendix*, Table S1). Four were located along the superior temporal sulcus (STS), and two in the left inferior gyrus pars triangularis (IFGtri) and pars orbitalis (IFGorb; small clusters were also seen at other sites such as left putamen). At the same stringent statistical threshold, the jabberwocky group showed increased activations only in IFGorb, which thus appears as a major region encoding constituent structure. With a lower statistical threshold ($P < 0.001$, voxel-based uncorrected), an effect of constituent size was also detected in the left posterior STS (pSTS) and IFGtri.

A direct comparison between the two groups, using the group × constituent size interaction, revealed that the normal-prose group showed significantly larger increases in activation in the temporal pole (TP), the anterior STS (aSTS), and temporo-parietal junction (TPJ; Fig. 2 *B* and *C* and *SI Appendix*, Table S1). These regions were thus sensitive to constituent structure only when content words were present. Conversely, the other regions (IFGorb, IFGtri, pSTS) showed unchanged increases with constituent size even when the constituents were solely denoted by grammatical words and morphology, in the absence of content words. The computation of constituent structure in these regions therefore unfolds independently of the presence of a meaningful content and even when activation is absent in other anterior temporal lobe and TPJ regions.

Other Profiles of Activation. We also searched the whole brain for patterns of activation other than an increase with constituent size. For instance, a region associated with memory and effort might show a greater activation to random lists of words than to sentences, thus showing a decrease with constituent size. A large network of regions showed such a behavior, including the precuneus, posterior cingulate, ventral medial prefrontal cortex, and bilateral inferior parietal lobule (*SI Appendix*, Fig. S3 and Table S2). None of these regions belonged to the classical perisylvian language areas. Instead, they are part of the classical "default mode" network (23) and, indeed, they were globally deactivated

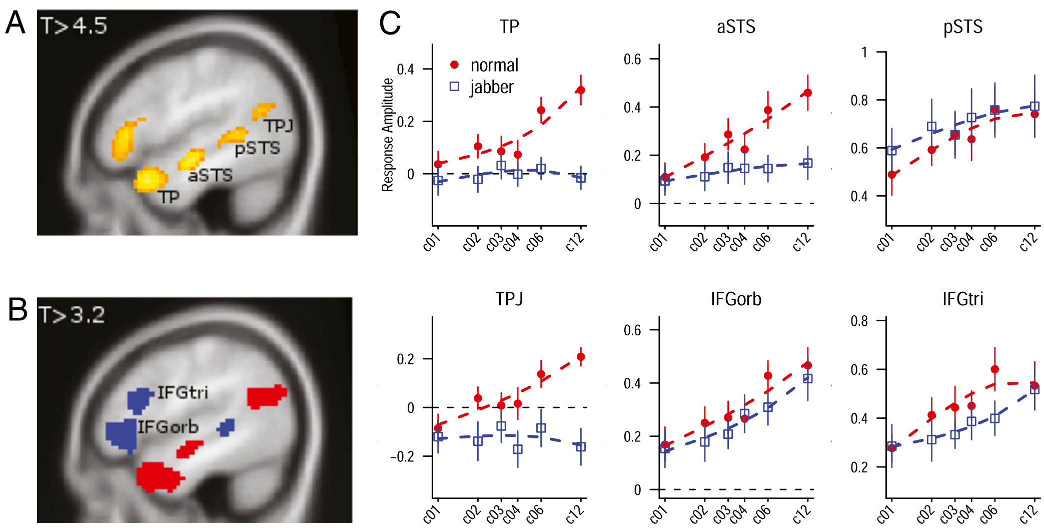


Fig. 2. Brain regions showing a significant increase in activation with constituent size. (*A*) fMRI results from the normal-prose group who read sequences with actual French words (group analysis thresholded at $T > 4.5$, $P < 0.05$ FWE, spatial extent > 10). (*B*) Areas in blue show a significant constituent size effect in the jabberwocky group listening to delexicalized stimuli, whereas regions in red show a significant group by constituent size interaction (reflecting a stronger effect of constituent size in normal prose than in jabberwocky) (maps thresholded at $T > 3.2$, $P < 0.001$ uncorrected, spatial extent > 50). (*C*) Amplitude of activations across conditions in the six regions of interest (error bars represent ± 1 SEM). Conditions c01 to c12 are organized according to a logarithmic scale of constituent size, thus a line on this graph indicates a logarithmic increase of activation. The fitting lines are from a regression analysis including linear and logarithmic predictors.

relative to rest and progressively deactivated more as constituent size increased.

Finally, we looked for regions exhibiting a quadratic, U-shape activation pattern as a function of constituent size. Such a pattern might arise, for instance, in putative regions responsive to syntactic violations. Arguably, a minimal amount of prior syntactic context is required before a syntactic violation can be detected or even defined. Therefore, one might postulate that violation detection is absent from both the lists of words (c01) and the normal sentences (c12). However, no such nonmonotonic response pattern was observed in a whole-brain search.

Nonlinearity of the Response with Constituent Size. The simplified model presented in Fig. 1 predicts that the amplitude of the fMRI activation should increase linearly with the number of words integrated into a constituent. The parametric manipulation of constituent size in our experiment yielded quantitative data bearing on this point (Fig. 2*C*). Regression analyses in fact showed that, in all brain regions showing an increase, except TP, the increase in activation was logarithmic rather than linear as a function of constituent size. First, for each region of interest (ROI) and each group, we performed simple regressions on the group-averaged activation in the six conditions c01 to c12, either with constituent size itself or with log (constituent size). *SI Appendix*, Table S3 reports the coefficients of determination R^2. The fit was higher for the log than for the linear model in all regions except the TP for the normal group and the IFGorb and IFGtri for the jabberwocky group.

As a better test of these effects, taking into interindividual variability, for each ROI and each group, we fitted two hierarchical models with individual-subject activation as the dependent variable, either constituent size (1–4, 6, 12) or its log as a regressor, and subject as a random factor. *SI Appendix*, Table S3 displays the ratio of the likelihoods of the log and the linear models. Values >1 favor the log model by the corresponding ratio. By this test, the log model was generally mildly or strongly favored over the linear model, except for the TP region in the normal group.

Overall, and given the absence of any interaction of group × constituent size in regions IFGorb and IFGtri, we conclude that the logarithmic model provides a better fit in all regions except the TP.

Analyses of Response Phases. The second prediction of the accumulation model was that activation should be increasingly delayed for larger constituent sizes (Fig. 1). Fourier analysis of fMRI data (18, 19) allowed us to compute reliable estimates of activation phase in most regions (Fig. 3). For TP and TPJ, activations were significant only in conditions c06 and/or c12, thus preventing further quantitative analysis of how phase varied with conditions (*SI Appendix, Methods*).

In aSTS, the phase increased significantly from conditions c01 to c12, by a delay of approximately half of the sentence duration [increase = 1.4 ± 0.9 s (95% confidence interval)]. A slower fMRI response, modestly but still significantly delayed by constituent size, was seen in IFGtri and pSTS (increases respectively of 1.0 ± 0.7 s and 0.6 ± 0.3 s). Finally, the slowest BOLD response was in IFGorb, and in this region the phase increase did not reach significance (0.6 ± 0.7 s) in spite of a large change in response amplitude.

Additional Control Condition with Nonconstituents. As a further control, our experiment also comprised two additional conditions called "non-constituents" (nc3 and nc4) where we presented series of 3- and 4-word-long excerpts that were contiguous within sentences, as in the corresponding c03 and c04 conditions, but that straddled across constituent boundaries (e.g., "new car are very," extracted from "the conditions for buying a new car are very attractive"). We reasoned that in a brain region sensitive only to possible sequences of words, these conditions should induce activation similar to the corresponding 3- and 4-word constituent conditions. In a region coding for constituents, however, these nonconstituents conditions should behave like the 1- and 2-word constituents conditions (c01, c02). In each ROI, we therefore compared the condition "nc3" to conditions c03 and c01, as well as condition nc4 to conditions c04 and c02, using Student's paired *t* tests. The results indicated that whenever the test could be meaningfully performed, activation in the nonconstituent condition was lower than in the corresponding constituent condition in pSTS, aSTS, IFGorb, and marginally in IFGtri (*SI Appendix*, Table S4). These results, together with the responsivity of pSTS, IFGorb, and IFGtri to jabberwocky constituents where transition probabilities are very low, suggest that these regions encode the constituent rather than sequential structure of word streams.

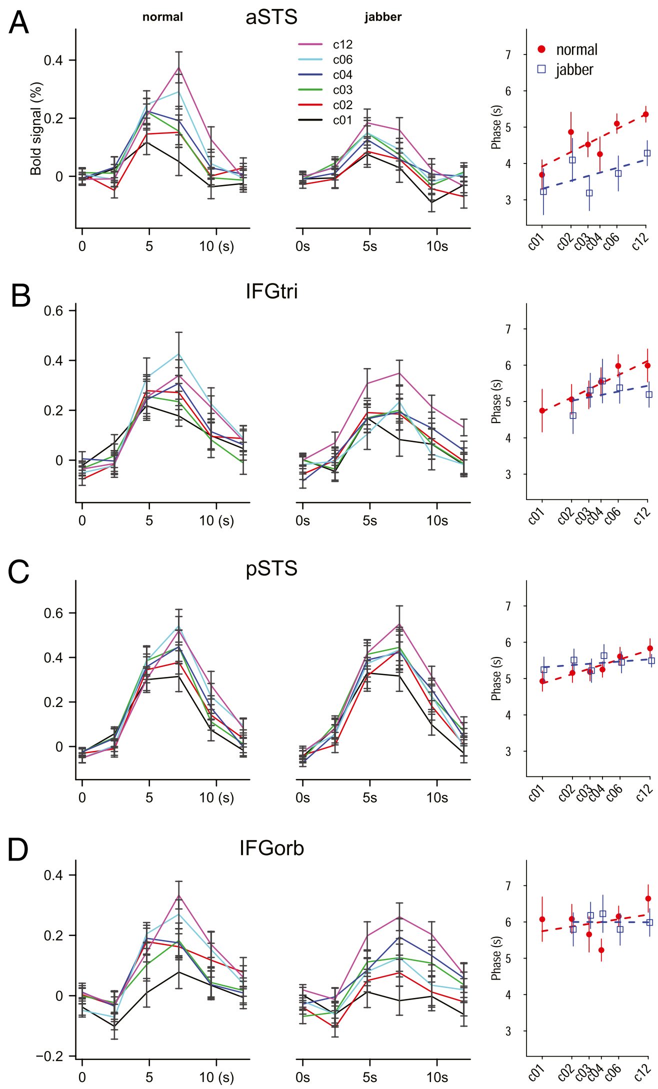


Fig. 3. Temporal profiles of activation (*Left*) and estimates of phase (*Right*) in four ROIs with distinct temporal patterns. (*A*) aSTS: response peaking early, increasing in amplitude and phase, with reduced response to jabberwocky. (*B*) IFGtri: slower response (larger phase), increasing in amplitude and phase for normal prose and jabberwocky. (*C*) pSTS: response similar to that of IFGtri, with a more modest increase in phase with constituent size. (*D*) IFGorb: slowest response (larger phase), increasing in amplitude but not phase, similarly for normal prose and jabberwocky. More detailed information and profiles for other ROIs are provided in section 4 of *SI Appendix*.

Discussion

In the visual system, parametric manipulations of the stimuli using both spatial scrambling (24) and temporal scrambling (25) have played an instrumental role in understanding how pictures are encoded. Here, we introduced a similar parametric approach to linguistic coding (see also refs. 26 and 27). Our research builds on previous studies that contrasted only two extreme conditions, sentences versus word lists, and already showed the involvement of the left anterior temporal regions during speech comprehension (e.g., refs. 7, 20, and 28). Our stimuli comprised a hierarchy of six successive conditions with increasing constituent sizes, always using simple right-branching constituents to minimize potentially confounding factors that are also known to contribute to parsing difficulty, such as movement (29). The results revealed a network of regions where activation increases parametrically in both amplitude and phase with the integration of words into constituents during on-line language comprehension. Moreover, the jabberwocky manipulation dissociated these language areas into two major subsets. The inferior frontal and posterior STS regions showed constituency effects regardless of whether actual content words were used. However, the TP, aSTS, and TPJ regions exhibited robust effects of constituent size only when actual words were presented.

The finding that the IFG is sensitive to syntactic structure in the absence of lexico-semantic information is consistent with studies of natural and artificial grammar learning that reported increased activation in Broca's area for phrase structure violations (30–33). We observed a similar behavior in the pSTS, a region directly connected to the IFG (34). Both regions frequently coactivate in studies designed to isolate syntactic computations (e.g., refs. 9, 10, 29, 35, and 36). Indeed, the pSTS region uncovered in our study (with a peak at −51,−39, 3) is located very near a region labeled as left middle temporal gyrus (−52, −44, 4) and more activated, together with left IFG, when participants read a verb phrase or noun phrase (e.g., a smell, I smell) than the corresponding content word alone (10) (see also ref. 37). Altogether, the evidence is consistent with the notion that IFG and pSTS cooperate in assembling words into linguistic constituents.

The inferior frontal activations that we observed are centered on the pars triangularis and pars orbitalis (According to the atlas of (38), the IFGtri peak has 80% probabilty to be in BA45 and 20% in BA44; The IFGorb is located in BA 45/47). It has been argued (11, 39) that there is a posterior-to-anterior gradient in Broca's area, with syntactic processes involving the posterior part (BA 44; ref. 26) and the more anterior aspects dealing with semantic processes (40). Nevertheless several studies, like the present one, have reported effects of syntactic manipulations in BA 45 or 47 (e.g., refs. 22, 31, 41–44). Most noteworthy is a recent study by Tyler et al. (45) where the activations to grammatically coherent sentences without meaning were remarkably similar to those that we report here. Moreover, contrasting hierarchically structured meaningless mathematical formulas to flat ones also reveals activations in BA 45 and 47 (46), suggesting that mathematics and language processing may share partially similar structure-building operations.

In a parametric design, Friederici et al. (26) manipulated three levels of complexity in the verb arguments of a dative structure and found activations in BA 44 that increased with ungrammaticality ratings, i.e., the opposite of the present result. They suggested that "BA 44 may be functionally related to language-internal processes involved in the reconstruction of a nondirect mapping between linear order and interpretation." As noted by Tyler et al. (45), many fMRI studies "have used stimuli and tasks which make it difficult to separate the effects of the online construction of a syntactic representation from the contribution of variables that may not reflect the normal processes of comprehending language." To mitigate this problem, we placed our subjects in a simpler situation of mere comprehension where the stimuli themselves limited the size of the linguistic representation that could be constructed and did not permit the deployment of complex processes of sentence repair (41, 47). Accordingly, in classical perisylvian areas, we observed only increasing activation amplitudes and delays with constituent size, presumably reflecting online construction with no or minimal involvement of additional repair processes.

In this study, regions TP and TPJ hardly responded to single words or even two-word constituents. Indeed, TPJ was deactivated by jabberwocky stimuli and activated above the resting-state level only for conditions c06 and c12 (Fig. 2*C*), i.e., complete or nearly complete sentences. This finding concurs with previous observations (48) suggesting that TPJ may be primarily engaged in the integration of words into a coherent discourse. Region aSTS, on the other hand, showed a more progressive increase in activation with constituent size, suggesting an involvement in processing smaller-sized constituents. In both regions, the observed nonresponsiveness to jabberwocky might indicate that these temporal regions rely on lexical information to build constituents, in agreement with "lexicalist" parsing models (49). For instance, the unification model postulates that a region in the posterior middle temporal gyrus, located just below our pSTS and TPJ regions, provides lexically stored syntactic frames that are later bound together into larger structures by inferior frontal cortex (50). This proposal, although plausible for normal prose, cannot readily explain the constituent size effect that we observed for jabberwocky in IFG and pSTS, an aspect of our results that underlines the relative independence of syntax from lexico-semantic features (51). The jabberwocky data show that the pSTS and inferior frontal cortex can compute syntactic frames solely on the basis of function words and syntactic category information provided by morphological features. Anterior temporal and TPJ regions could then bind these syntactic roles to lexico-semantic representations provided by other regions of the temporal lobe to form high-level representations of semantic constituent structure.

In most regions except the temporal pole, the increases in activation with constituent size were logarithmic rather than linear with the number of words inside the constituent. This finding is subject to several nonnecessarily incompatible interpretations. Because fMRI does not directly measure neural activation, it is possible that nonlinearities arise from the mapping between neural activation and the fMRI BOLD response. At short stimulus durations (<2 s), nonlinearities in BOLD signals have been reported (52–54): The total activation to a temporally extended stimulus is smaller than the sum of the activations to its shorter-duration components. However, it remains debated whether such nonlinearities reflect neural or vascular effects, because MEG investigations have shown that they are largely imputable to neural adaptation (55). Furthermore, their relevance to the present situation is unclear because all of our experimental conditions involve the same total duration and same number of words. We therefore consider plausible the alternative hypothesis of a genuinely logarithmic variation in average neural activity with constituent size. In *SI Appendix*, we show that the model presented in Fig. 1 can be modified to account for this logarithmic profile (*SI Appendix*, Fig. S1). The sole assumption required is that the additional activation evoked by the incorporation of a word in a constituent is not a constant, but decreases as the constituent increases. Each word successively integrated into a constituent thus would have a progressively smaller impact on brain activation. Such an effect could arise from stimulus predictability: As a constituent builds up, the increasingly richer context of the preceding words makes the syntactic features of the incoming words increasingly predictable, a factor that might reduce their impact and therefore the activation that they induce (56). This proposal would mesh well with the general assumption, made outside of the linguistic domain, that the cortex acts as a predictive system that constantly attempts to improve the predictability of its inputs and that ascending brain activation reflects the residual prediction error or "surprisal" (57).

Note however that prediction error alone cannot explain the increasing activation patterns without the additional assumption that language areas accumulate and hold online the evidence from words forming constituents. If the activation merely reflected prediction error, then the less-probable sequences, with the shortest constituents, should have elicited stronger activations than the longest, most probable ones. Several prefrontal, parietal, and midline areas showed such a decrease in activation with constituent size, but none of them were located in the perisylvian language areas, and all were "deactivated," showing lower activations for sentences than for rest. In classical language areas, increasingly complex constituent structures are primarily encoded by increasing levels of activation. As noted in Introduction, such a finding is compatible with a summation principle according to which, during the encoding of constituent structure, sparse neural codes for the successive nodes are superimposed within the same cortical area (15).

Our study also shows that fMRI presents enough sensitivity to detect not only amplitude changes, but also temporal delays associated with the construction of linguistic structures. The ability to estimate the precise phase of fMRI response, with a sensitivity of 200 ms or less, has been validated by both theory and past experimentation (16–19). In the language domain, a temporal hierarchy is detectable among the perisylvian areas, with faster responses to spoken sentences in the superior temporal cortex near Heschl's gyrus and increasingly slower activations as one moves away anteriorily or posteriorily along the STS and toward inferior frontal cortex (17). Furthermore, fMRI activation in these slower regions accelerates when the sentence is repeated (17). The present results add that, in aSTS, pSTS, and IFGtri, but not IFGorb, fMRI responses get increasingly slower for larger constituent sizes. The observed delays are generally quantitatively compatible with the prediction of the accumulation model (the linear model predicts a 0.56-s increase from conditions c01 to c12, and the revised logarithmic model a 0.40-s increase). Only in aSTS was the observed delay (1.4 ± 0.9 s) marginally larger than predicted This observation might indicate an additional contribution of slow operations that would be primarily deployed toward the end of a sentence (48). Another possibility is that processing in aSTS suddenly stops at the end of the first legal constituent, as soon as the incoming words cannot be integrated with the past context. This "sustained-activity model," detailed in *SI Appendix*, predicts that fMRI phase should increase with a slope of half the total sentence duration, which is close to what we observed.

In conclusion, although much remains to be discovered as to how the structure of sentences is processed in the brain, these results provide quantitative parametric evidence on how linguistic constituents are encoded. Two constraints on any future model of the neural code for constituent structure are (*i*) a logarithmic increase in fMRI activation amplitude with constituent size; and (*ii*) a corresponding increase in activation delay. Future work should search for more precise determinants of syntactic complexity, using response amplitude and delay as direct markers of syntactic coding. Although number of words was used here as a proxy for constituent complexity, the ultimate predictors, which may vary across brain regions, are likely to be more abstract and language-invariant. Higher-temporal-resolution fMRI, combined with slower stimulus presentation, should probe why the temporal patterns of response differ across regions. Finally, equivalent studies using finer temporal methods such as magnetoencephalography (58) should provide detailed information as to the time course of constituent formation, ultimately paving the way to a detailed understanding of the neural code for syntax.

Materials

Participants. Forty native French speakers (23 men and 17 women; age range = 18–37, mean = 24, SD = 6) took part in the experiment, which was approved by the regional ethical committee All gave written informed consent and received 80 euros for their participation.

Stimuli. ***Normal prose.*** The stimuli were sequences of 12 words generated by concatenating sentence fragments. Two hundred right-branching 12-word sentences (listed in *SI Appendix*) were first generated to create the condition c12 (full sentence). Thousands of fragments forming syntactic constituents were then extracted from this set. Fragments of size 2, 3, 4, and 6 were then randomly picked up from distinct sentences and concatenated to construct sequences of 12 words corresponding respectively to the experimental conditions c02, c03, c04, and c06 (see examples in Table 1). Single words from the original sentences were randomly concatenated to create sequences for the c01 condition. Different sequences were generated from the same pool of fragments for each participant. Individual participants were never exposed more than once to a given fragment, thus preventing the use of memory to mentally reconstruct the original sentences. The stimuli were manually verified and, if necessary, reshuffled to ensure that they did not contain, by chance, constituents of greater size than intended. For instance, in the c01 condition, two consecutive words never formed a plausible constituent. As a control over serial word order versus constituent structure, two additional conditions called non-constituents (nc3 and nc4) were included, consisting of sequences of 3- and 4-word-long excerpts that were contiguous within center-embedded sentences (listed in *SI Appendix*), but straddled across constituent boundaries (e.g., "new car are very", extracted from "the conditions for buying a new car are very attractive").

Jabberwocky. For the jabberwocky condition, the experimental lists were modified by replacing all content words with pseudowords of the same length and morphological endings. As an example, the original sentence "Le passant examine le luxueux canap abandonn sur le bord du trottoir" became "Le chevant poisine le permeux relip pingann sur le gord du monchoir" (see Table 1 for examples in English). Function words were not modified (The average proportions of function words in the sequences were the following: c01, 27%; c02, 42%; c03, 36%; c04, 41%; c06, 43%, c12, 47%; nc3, 32%; nc4, 39%.).

Procedure. The 12-word sequences were presented by using rapid serial visual presentation (300 ms per word). After a short familiarization block comprising 9 sequences, each subject took 5 fMRI runs containing 35 sequences each (interstimulus interval = 12 s). In addition to four stimuli belonging to each of eight experimental conditions c01, c02, c03, c04, c06, c12, nc3, and nc4, each run contained three probe sentences were inserted at random positions. These probe sentences explicitly requested the participants to press a button and served to ensure that they were paying attention to the stimuli.

Imaging and Data Analysis. The imaging methods and data analysis procedures are described in *SI Appendix*.

ACKNOWLEDGMENTS. We thank L. Rizzi, A. Bachrach, Biyu J. He, A. Moro., Y. L. Boureau, G. Dehaene-Lambertz, and A. Jobert for their support and helpful suggestions. This research was supported by the Pathfinder program of the European Union (Contract 12738) and was part of a general research program on functional neuroimaging of the human brain sponsored by the Atomic Energy Commission (Denis Le Bihan).

1. Branigan HP, Pickering MJ, Cleland AA (2000) Syntactic co-ordination in dialogue. *Cognition* 75:B13–B25.
2. Bock K, Dell GS, Chang F, Onishi KH (2007) Persistent structural priming from language comprehension to language production. *Cognition* 104:437–458.
3. Just MA, Carpenter PA, Keller TA, Eddy WF, Thulborn KR (1996) Brain activation modulated by sentence comprehension. *Science* 274:114–116.
4. Dapretto M, Bookheimer SY (1999) Form and content: Dissociating syntax and semantics in sentence comprehension. *Neuron* 24:427–432.
5. Noppeney U, Price CJ (2004) An FMRI study of syntactic adaptation. *J Cogn Neurosci* 16:702–713.
6. Friederici AD, Bahlmann J, Heim S, Schubotz RI, Anwander A (2006) The brain differentiates human and non-human grammars: Functional localization and structural connectivity. *Proc Natl Acad Sci USA* 103:2458–2463.
7. Humphries C, Binder JR, Medler DA, Liebenthal E (2006) Syntactic and semantic modulation of neural activity during auditory sentence comprehension. *J Cogn Neurosci* 18:665–679.
8. Grodzinsky Y, Friederici AD (2006) Neuroimaging of syntax and syntactic processing. *Curr Opin Neurobiol* 16:240–246.
9. Kaan E, Swaab TY (2002) The brain circuitry of syntactic comprehension. *Trends Cogn Sci* 6:350–356.
10. Tyler LK, Marslen-Wilson W (2008) Fronto-temporal brain systems supporting spoken language comprehension. *Philos Trans R Soc Lond B Biol Sci* 363:1037–1054.
11. Hagoort P (2005) On Broca, brain, and binding: A new framework. *Trends Cogn Sci* 9: 416–423.
12. Barone P, Joseph JP (1989) Prefrontal cortex and spatial sequencing in macaque monkey. *Exp Brain Res* 78:447–464.
13. Ninokura Y, Mushiake H, Tanji J (2003) Representation of the temporal order of visual objects in the primate lateral prefrontal cortex. *J Neurophysiol* 89:2868–2873.
14. Ninokura Y, Mushiake H, Tanji J (2004) Integration of temporal order and object information in the monkey lateral prefrontal cortex. *J Neurophysiol* 91:555–560.
15. Smolensky P, Legendre G (2006) *The Harmonic Mind* (MIT Press, Cambridge, MA) Vol 1.
16. Menon RS, Luknowsky DC, Gati JS (1998) Mental chronometry using latency-resolved functional MRI. *Proc Natl Acad Sci USA* 95:10902–10907.
17. Dehaene-Lambertz G, et al. (2006) Functional organization of perisylvian activation during presentation of sentences in preverbal infants. *Proc Natl Acad Sci USA* 103: 14240–14245.
18. Sigman M, Jobert A, Lebihan D, Dehaene S (2007) Parsing a sequence of brain activations at psychological times using fMRI. *Neuroimage* 35:655–668.
19. Sigman M, Dehaene S (2008) Brain mechanisms of serial and parallel processing during dual-task performance. *J Neurosci* 28:7585–7598.
20. Mazoyer BM, et al. (1993) The cortical representation of speech. *J Cogn Neurosci* 5: 467–479.
21. Friederici AD, Meyer M, von Cramon DY (2000) Auditory language comprehension: An event-related fMRI study on the processing of syntactic and lexical information. *Brain Lang* 75:289–300.
22. Röder B, Stock O, Neville H, Bien S, Rösler F (2002) Brain activation modulated by the comprehension of normal and pseudo-word sentences of different processing demands: A functional magnetic resonance imaging study. *Neuroimage* 15:1003–1014.
23. Raichle ME, et al. (2001) A default mode of brain function. *Proc Natl Acad Sci USA* 98: 676–682.
24. Grill-Spector K, et al. (1998) A sequence of object-processing stages revealed by fMRI in the human occipital lobe. *Hum Brain Mapp* 6:316–328.
25. 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.
26. Friederici AD, Fiebach CJ, Schlesewsky M, Bornkessel ID, von Cramon DY (2006) Processing linguistic complexity and grammaticality in the left frontal cortex. *Cereb Cortex* 16:1709–1717.
27. Santi A, Grodzinsky Y (2007) Working memory and syntax interact in Broca's area. *Neuroimage* 37:8–17.
28. Vandenberghe R, Nobre AC, Price CJ (2002) The response of left temporal cortex to sentences. *J Cogn Neurosci* 14:550–560.
29. Ben-Shachar M, Palti D, Grodzinsky Y (2004) Neural correlates of syntactic movement: Converging evidence from two fMRI experiments. *Neuroimage* 21:1320–1336.
30. Tettamanti M, et al. (2002) Neural correlates for the acquisition of natural language syntax. *Neuroimage* 17:700–709.
31. Musso M, et al. (2003) Broca's area and the language instinct. *Nat Neurosci* 6:774–781.
32. Bahlmann J, Schubotz RI, Friederici AD (2008) Hierarchical artificial grammar processing engages Broca's area. *Neuroimage* 42:525–534.
33. Bahlmann J, Gunter TC, Friederici AD (2006) Hierarchical and linear sequence processing: An electrophysiological exploration of two different grammar types. *J Cogn Neurosci* 18:1829–1842.
34. Saur D, et al. (2010) Combining functional and anatomical connectivity reveals brain networks for auditory language comprehension. *Neuroimage* 49:3187–3197.
35. Ben-Shachar M, Hendler T, Kahn I, Ben-Bashat D, Grodzinsky Y (2003) The neural reality of syntactic transformations: Evidence from functional magnetic resonance imaging. *Psychol Sci* 14:433–440.
36. Rodd JM, Longe OA, Randall B, Tyler LK (2010) The functional organisation of the fronto-temporal language system: Evidence from syntactic and semantic ambiguity. *Neuropsychologia* 48:1324–1335.
37. Longe O, Randall B, Stamatakis EA, Tyler LK (2007) Grammatical categories in the brain: The role of morphological structure. *Cereb Cortex* 17:1812–1820.
38. Amunts K, et al. (1999) Broca's region revisited: Cytoarchitecture and intersubject variability. *J Comp Neurol* 412:319–341.
39. Bookheimer S (2002) Functional MRI of language: New approaches to understanding the cortical organization of semantic processing. *Annu Rev Neurosci* 25:151–188.
40. Hagoort P, Hald L, Bastiaansen M, Petersson KM (2004) Integration of word meaning and world knowledge in language comprehension. *Science* 304:438–441.
41. Ni W, et al. (2000) An event-related neuroimaging study distinguishing form and content in sentence processing. *J Cogn Neurosci* 12:120–133.
42. Moro A, et al. (2001) Syntax and the brain: Disentangling grammar by selective anomalies. *Neuroimage* 13:110–118.
43. Dronkers NF, Wilkins DP, Van Valin RD, Jr, Redfern BB, Jaeger JJ (2004) Lesion analysis of the brain areas involved in language comprehension. *Cognition* 92:145–177.
44. Santi A, Grodzinsky Y (2010) fMRI adaptation dissociates syntactic complexity dimensions. *Neuroimage* 51:1285–1293.
45. Tyler LK, et al. (2010) Preserving syntactic processing across the adult life span: The modulation of the frontotemporal language system in the context of age-related atrophy. *Cereb Cortex* 20:352–364.
46. Friedrich R, Friederici AD (2009) Mathematical logic in the human brain: Syntax. *PLoS ONE* 4:e5599.
47. Meyer M, Friederici AD, von Cramon DY (2000) Neurocognition of auditory sentence comprehension: Event related fMRI reveals sensitivity to syntactic violations and task demands. *Brain Res Cogn Brain Res* 9:19–33.
48. Humphries C, Binder JR, Medler DA, Liebenthal E (2007) Time course of semantic processes during sentence comprehension: An fMRI study. *Neuroimage* 36:924–932.
49. Hagoort P (2003) How the brain solves the binding problem for language: A neurocomputational model of syntactic processing. *Neuroimage* 20(Suppl 1):S18–S29.
50. Snijders TM, et al. (2009) Retrieval and unification of syntactic structure in sentence comprehension: An FMRI study using word-category ambiguity. *Cereb Cortex* 19: 1493–1503.
51. Chomsky N (1957) *Syntactic Structures* (Mouton, The Hague).
52. Boynton GM, Engel SA, Glover GH, Heeger DJ (1996) Linear systems analysis of functional magnetic resonance imaging in human V1. *J Neurosci* 16:4207–4221.
53. Boynton GM, Finney EM (2003) Orientation-specific adaptation in human visual cortex. *J Neurosci* 23:8781–8787.
54. Birn RM, Saad ZS, Bandettini PA (2001) Spatial heterogeneity of the nonlinear dynamics in the FMRI BOLD response. *Neuroimage* 14:817–826.
55. Tuan AS, Birn RM, Bandettini PA, Boynton GM (2008) Differential transient meg and fMRI responses to visual stimulation onset rate. *Int J Imaging Syst Technol* 18:17–28.
56. Levy R (2008) Expectation-based syntactic comprehension. *Cognition* 106:1126–1177.
57. Kiebel SJ, Daunizeau J, Friston KJ (2008) A hierarchy of time-scales and the brain. *PLOS Comput Biol* 4:e1000209.
58. Bastiaansen M, Magyari L, Hagoort P (2010) Syntactic unification operations are reflected in oscillatory dynamics during on-line sentence comprehension. *J Cogn Neurosci* 22:1333–1347.

Supporting Information for article:

The Cortical Representation of the Constituent Structure of Sentences

Christophe Pallier and Anne-Dominique Devauchelle and Stanislas Dehaene

INSERM-CEA Cognitive Neuroimaging Unit, CEA/SAC/DSV/DRM/Neurospin Center
Bât 145, Point Courier 156, F-91191 Gif-sur-Yvette Cedex, France

November 22, 2010

Contents

List of Tables

List of Figures

1 Imaging and Data analysis

Magnetic Resonance Imaging

The acquisition was performed with a 3 Tesla Siemens Tim Trio system equipped with a twelve channels coil. For each participant, an anatomical image was first taken, using a 3D Gradient-echo sequence and voxel size of 1x1x1.1 mm; Then, a total of 1160 functional scans were acquired during five sessions of 232 scans each, using an Echo-Planar sequence sensitized to the BOLD effect (TR = 2.4 secs, TE = 30 msecs, Matrix = 64x64 ; Voxel size = 3x3x3 mm; 40 slices in ascending order).

Data preprocessing and Analysis

Data processing was performed with SPM5 (Wellcome Department of Cognitive Neurology, software available at http://www.fil.ion.ucl.ac.uk/spm). The anatomical scan was spatially normalized to the avg152 T1-weighted brain template defined by the Montreal Neurological Institute using the default parameters (nonlinear transformation). Functional volumes were corrected for slice timing differences, realigned to correct for motion correction, spatially normalized using the parameters obtained from the normalization of the anatomy, and smoothed with an isotropic Gaussian kernel (FWHM=5mm).

In a first SPM model, experimental effects at each voxel were estimated using a multi-session design matrix modeling the 9 conditions (c01, c02, c03, c04, c06, c12, nc3, nc4, probes) and the 6 movement parameters computed at the realignment stage. Each stimulus was modeled as an epoch lasting 3.6 seconds, corresponding to the duration of stimulus display. The regressors were created by convolving these epochs by the standard SPM hemodynamic response function and its derivative. Constrasts averaging the regression weights associated with each condition were computed and smoothed with a 8x8x8mm Gaussian kernel. These estimates of the individual effect sizes were entered in a second-level analysis of variance model with one regressor per group and experimental condition, as well as one regressor per subject to remove main effects of subjects. Unless otherwise mentioned, the results are reported using a stringent correction for multiple comparisons across the whole brain volume (voxel-based $p<0.05$ Family-Wise-Error threshold).

To search for regions showing an effect of constituent size, that is, where activation increases from conditions c01, c02, c03, c04, c06 to c12, we used a linear contrast testing for increasing activation across the six conditions (with weights=[-5 -3 -1 1 3 5]). We also evaluated a strictly linear constrast respecting the numerical scale 1, 2, 3, 4, 6, 12 (weights=[-11 -8 -5 -2 4 22]) and a logarithmic contrast proportional to log(constituent size). These two contrasts yielded essentially similar results to the simple linear contrast above, although the logarithmic contrast was slightly more sensitive. We also searched the whole brain for linearly decreasing and for quadratic effects of constituent size.

To compute the responses phase, a second SPM model was designed, using constant, sine and cosine basis functions with a period of 12s instead of the standard hemodynamic response function and its derivative. Phases were computed for each subject from the obtained parameters using the arctangent (atan2) function, and further converted in seconds using the corresponding fraction of the period. Thus, a phase of 5 seconds indicates that, when fitted by a sinusoid, the BOLD response peaked approximately 5 seconds after the onset of the stimulus sequence. The group results were computed from individual phases using the circular mean and standard error.

The phase of the BOLD response is only defined for conditions in which a significant activation is observed, and can be meaningfully interpreted as a delayed activation for regions in which a positive BOLD effect is seen. We therefore only report and analyze the conditions in which the mean phases passed a statistical test for significance across subjects (Rayleigh's test, $p<0.05$), and further fell in a reasonable range for a BOLD response (2.5 to 8 seconds). The effect of constituent size on phase was defined as the phase difference between conditions c12 and c01, estimated from the slope of the regression of the group mean phases in the six main conditions of interest (c01, c02, c03, c04, c06, c12) with $\log(n)$ (where n is constituent size). Because phases lie on a circular rather than scalar continuum, significance could not be assessed with classical regression tests. We computed confidence intervals for the slope using a bootstrap approach from the distribution of the circular means in randomly permuted data. To improve the stability of the phase analysis, we pooled together the data from the normal and jabberwocky groups for those regions showing activation in both conditions (IFGorb, IFGtri, pSTS). For the remaining regions (aSTS, TP, TPJ), only the normal-prose group was used. For TP and TPJ, the Rayleigh test was not passed for most conditions, thus preventing further quantitative analysis of phases.

2 Revision of the model based on the observed logarithmic increase

Our empirical observations show that, for a fixed number of words, activation increases as a logarithmic function of their internal constituent size. This observation is incompatible with the simple accumulation model which predict a linear increase. In this section, we examine how the empirical observation leads to a revision of this model.

Definitions and assumptions:

- Let w be the number of words in a constituent.
- Let n be the number of times that a constituent of size w is presented.
- The total number of words therefore is $n \times w$.

  In our experiments, for instance, the condition c02 presents 6 constituents of size 2, hence $w = 2$ and $n = 6$. Generally, our design is such that in condition cw, $n = 12/w$, since a total of 12 words is always presented.
- Let $F(w)$ be the size of the neural activation evoked to encode a single constituent of size w.
- We assume that this activation is itself the sum of the activations $G(i)$ evoked by each of the successive words ($i = 1$ to w). That is, each incoming word gets integrated into the previously established structure by adding to it an additional amount of activation $G(i)$.

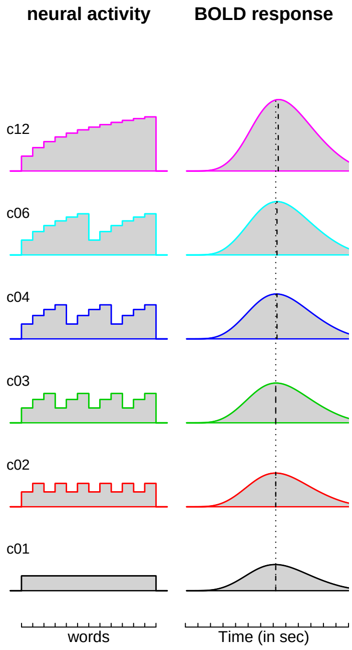


Fig. S1: Simulations from a revised model in which neural activity increases less than linearly whenever a word is added to a constituent. Such a model, although tentative, is needed to account for the observed logarithmic increase of the BOLD response (right column).

An assumption underlying the accumulation model presented in the introduction is that each of the n successive constituents elicits an identical activation $F(w)$. After convolution with the impulse heamodynamic function (HRF), we therefore expect the total activation evoked by condition cw to be approximately proportional to $n \times F(w)$. We empirically observe this activation to be proportional to $\log(w)$. After substituting n with $12/w$, and solving for $F(w)$, we obtain

$$F(w) = kw \log(w)$$

where k is a scaling constant. Because $F(w)$ is itself the sum of the $G(i)$ for $i = 1 \cdots w$, $G(i)$ itself must vary as $k(i \log(i)\check{}(i-1)\log(i-1))$, or approximately as $k \log(i)$.

Thus, we obtain the following conclusion:

If the observed activation is the sum of activations evoked by each of the constituents, then the activation evoked by the ith word of each constituent must be proportional to $\log(i)$.

3 The sustained-activity model

The sustained-activity model relies on two hypotheses. First, a brain region that integrates the successive words forming a linguistic constituent should exhibit sustained activity throughout the presentation of the corresponding words, holding the merged internal representation of their constituent structure in working memory. Second, this activation might collapse as soon as a novel incoming word ceases to be incorporable into the overall structure. Thus, sustained activation would start with the first word and stop when a novel unrelated constituent starts. In other words, only the first constituent contributes to the observed activation.

Such a profile of activation might occur in a putative area coding for the overall integrated structure of the stimuli: the words forming the first constituent can be integrated together into a well-formed structure, but when a second unrelated constituent is presented, they cannot be merged together into a single well-formed assemblage, and activation therefore drops to zero.

Predictions The sustained-activity model predicts a linear increase in fMRI activation amplitude as a function of constituent size, and a large increase in phase, in the order of 1.8s from c01 to c12 (see Fig.S2). This is because the phase approximately reflects the location of the barycenter of activation. Thus, it should vary linearly with constituent size, with a slope equal to half of the total stimulation duration, hence 1.8 second for our 3.6 second-long stimuli. Simulations verified these inferences, showing a linear increase in phase with an increase of 1.65s from c01 to c12.

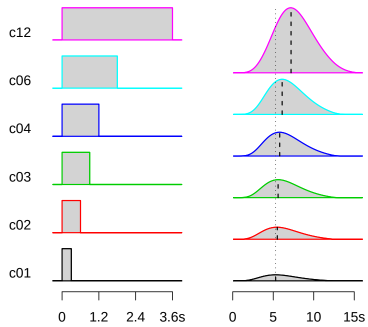


Fig. S2: Response profile of a region where activity is sustained throughout the first constituent, then collapses

Revising the sustained-activity model based on the observed logarithmic increase

The key assumption underlying the sustained activity model is that only the first constituent contributes to the total activation. Thus, the total activation evoked by condition cw is simply proportional to $F(w)$. We empirically observe this activation to be proportional to $\log(w)$: $F(w) = k\log(w)$. As a result, $G(i)$ itself must vary as $G(i) = k/i$

Thus, we obtain the following conclusion:

> **If the observed activation is evoked only by the first constituent, then the activation evoked by the ith word of the first constituent must be proportional to $1/i$.**

Simulations show that such a model predicts a reduced increase in phase from c01 to c12 (0.83s), about half that of the original linear sustained-activity model.

Table S1: Regions showing increasing activations with constituent size

	Cluster size	T (df=266)	x	y	z
Analysis at $p < .05$ family-wise-error corrected					
normal-prose group					
Temp. Mid. L *(aSTS)*	333	9.37	-54	-12	-12
Temp. Pole L. *(TP)*		7.41	-48	15	-27
Temp. Mid. L *(pSTS)*	190	6.93	-39	-57	18
Temp. Mid. L		5.58	-48	-45	3
Frontal Inf. Tri. L	204	6.43	-51	30	6
Frontal Inf. Tri. L		5.76	-54	21	15
Putamen L	9	5.03	-18	6	12
Frontal Sup. Medial L	7	4.72	-6	54	36
Temp. Pole L. (internal)	2	4.61	-27	-3	-39
Temporal Mid. R	1	4.55	51	0	-21
Jabberwocky group					
Frontal Inf Orb L	42	5.53	-45	33	-6
Normal>Jabberwocky					
Temporo-Parietal Junction	53	5.25	-45	-66	24
		5.18	-39	-57	21
Temp. Pole L	28	5.08	-48	15	-30
Temp. Mid. L.	13	4.80	-54	-12	-12
		4.58	-40	3	15
Jabberwocky>Normal					
No suprathreshold activations					
Additional analysis of the Jabberwocky group at $p < .001$ uncorrected, cluster size> 40					
Frontal Inf Orb L *(IFGOrb)*	332	5.53	-45	33	-6
Frontal Inf Tri L *(IFGTri)*		4.41	-51	21	21
Temp. Mid. L *(pSTS)*	75	4.35	-51	-39	3
		3.71	-57	-48	6

Anatomical labels are obtained with the Anatomical Automatic Labeling toolbox (http://www.cyceron.fr/web/aal_anatomical_automatic_labeling.html)

Table S2: Regions where activation decreased with constituent size

	Cluster size	T (df=266)	x	y	z
Analysis at $p < .05$ family-wise-error corrected					
normal-prose					
Precuneus L	159	6.26	-6	-69	39
Precuneus R		4.76	9	-69	39
Parietal Inf. L	51	5.68	-47	-53	47
Frontal Mid. R	34	5.23	36	30	45
Occipital Mid. L	27	5.22	-15	-90	-9
Occipital Inf L		4.81	-27	-81	-9
Cingulum Mid (L & R)	50	5.32	3	-27	39
Cingulum Ant. R	34	4.91	6	45	0
Frontal Mid. Orb. L	6	4.95	-27	45	-12
Frontal Mid. R	6	4.76	42	39	33
Frontal Mid. L	5	4.69	-36	30	39
Fusiform R	4	4.65	27	-72	-9
Frontal Sup. Medial L	1	4.54	-12	51	0
Jabberwocky					
Occipital Mid. L	178	7.75	-18	-87	-6
Occipital Mid. R	28	5.42	21	-84	3
Fusiform R	22	4.98	33	-60	-3
Frontal Sup. R	15	4.73	24	-3	60
Supp. Motor Area R	1	4.71	12	9	54
Lingual R	3	4.60	6	-57	0

Table S3: Coefficients of determination of models predicting response amplitude either from constituent size (R^2_{lin}) or from its log R^2_{log} , and Likelihood ratios comparing the two models

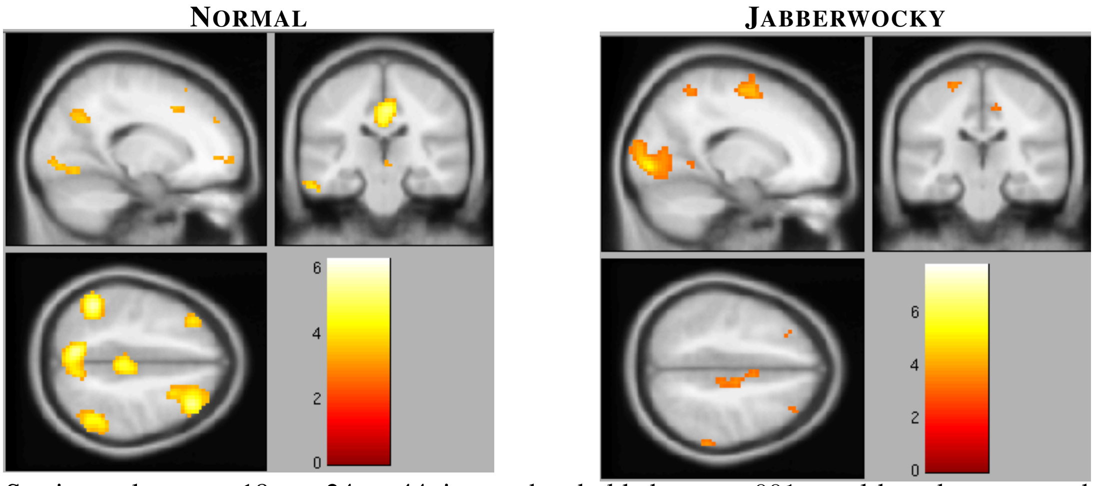
	normal-prose			Jabberwocky		
	R^2_{lin}	R^2_{log}	Likelihood ratio	R^2_{lin}	R^2_{log}	Likelihood ratio
IFGorb	0.83	0.89	10	0.94	0.93	1.6
IFGtri	0.50	0.78	493	0.97	0.91	2
pSTS	0.61	0.87	148	0.65	0.86	32
aSTS	0.82	0.91	45	0.64	0.87	5
TP	0.86	0.78	0.45	0.0	0.09	12
TPJ	0.84	0.88	3	0.1	0.04	3

Table S4: p-values from Student's T tests comparing activations in the non-constituent vs. constituent conditions in the 6 main Regions of Interest.

Region	c03>c01	nc3>c01	c03>nc3	c04>c02	nc4>c02	c04>nc4
TP	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
aSTS	0.008	n.s.	0.02	n.s.	n.s.	n.s.
pSTS	0.02	n.s.	0.01	n.s.	n.s.	n.s.
TPJ	0.07	n.s.	n.s.	n.s.	n.s.	n.s.
IFGorb	n.s.	n.s.	0.02	n.s.	n.s.	n.s.
IFGtri	0.03	n.s.	0.06	n.s.	n.s.	n.s.

n.s. = non significant (p>0.1)

Fig. S3: Regions where activation *decreased* with constituent size



Sections taken at x=18; y=-24, z=44; image thresholded at $p < .001$ voxel-based uncorrected

4 Regions of Interests and their response profiles

Increasing activation with constituent size				
ROI name	x	y	z	size (mm3)
TP	-48	15	-27	3768
aSTS	-54	-12	-12	3512
pSTS	-51	-39	3	2696
TPJ	-39	-57	18	3328
IFGorb	-45	33	-6	3968
IFGtri	-51	21	21	1984
Putamen	-18	16	12	241

Decreasing activation with constituent size				
ROI name	x	y	z	size (mm3)
PrecuneusL	-6	-69	39	4152
InfParL	-47	-53	47	3512
CingAnt	6	45	0	3664
CingMid	3	-27	39	4040
FrontMidR	36	30	45	3384
Occip	-15	-90	-9	2512

Table S5: Regions of Interest

Regions of interest (ROI) were defined as the intersections of spheres of 10mm radius with the clusters identified by the linear contrasts for constituent size thresholded at voxel-based p<.001 (see Tables S1 and S2). The centers of the regions were defined from the normal prose group results for TP, aSTS and TPJ and from the Jabberwocky group resultes for IFGorb, IFGtri, pSTS. The centers of ROIs showing decreasing activations originate from the negative linear contrast in the normal prose group.

For each region of interest, four graphics are provided:

Time course of the event-related responses. Since a slow-event related design was used, with one word sequence every 12s, the event-related response was computed by selective averaging of the BOLD signal in the 12 seconds time-window following each stimulus, separately for each condition. The figure shows the BOLD signal averaged across all voxels in the region of interest. Variables of non-interest accounting for movement artifacts were regressed out. The error bars indicate between-subject standard errors.

BOLD response amplitudes were computed from the first SPM model, which used a standard hemodynamic response function (hrf). The parameters estimates from the hrf model for each condition and each subject were averaged across all voxels in the region of interest. The graph shows the means and standard errors across subjects. Note that on the x axis, conditions c01 to c12 are organized according to a logarithmic scale of constituent size. The fitting lines are from a regression analysis with linear and logarithmic predictors.

Phases were computed within subjects and for each condition, from the parameters estimates of the sin/cos model (see Methods). The graph shows the circular group means and circular standard error, converted into seconds using the corresponding fraction of the period ($\phi_{\text{seconds}} = \phi_{\text{rad}} \times \frac{12}{2\pi}$). Only conditions in which the phase coherence was significant across subjects by a Rayleigh test ($p < 0.05$) appear on this graph, as the mean phase could not be reliably estimated otherwise. Only values in the range 2.5-8s are shown, corresponding to plausible positive activations given the standard hemodynamic response function.

Circular plots of phases in each condition. Each dot corresponds to the phase of one participant's response in a given condition. Numerical values of the circular group means and circular standard error are shown above each graph (the latter is not defined when the dispersion is excessive). For each condition and group, a Rayleigh test for phase coherence was performed, the p-value of which is indicated above each graph. On each circular plot, a vector corresponding to the mean resultant length is shown. The direction of this vector indicates the mean phase, while its length indicates the coherence across subjects. In most regions, a phase increase can be seen as a progressive rotation of the vector across the conditions c01 to c12.

Fig. S4: Response in Left Anterior Superior Temporal Sulcus (aSTS; -54 -12 -12)

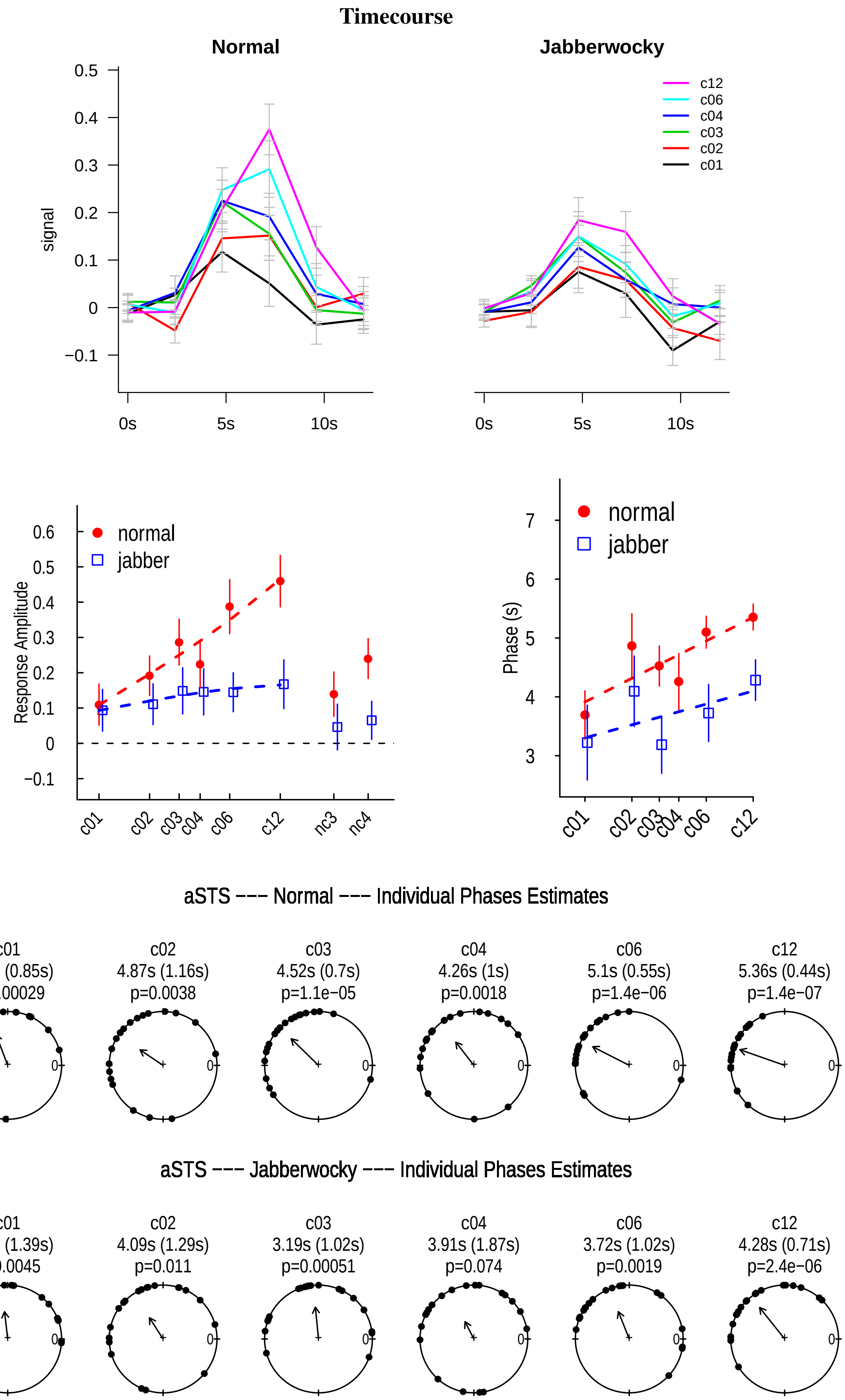


Fig. S5: Response in Left Inferior Frontal Gyrus - Pars Orbitalis (IFGorb; -45 33 -6)]

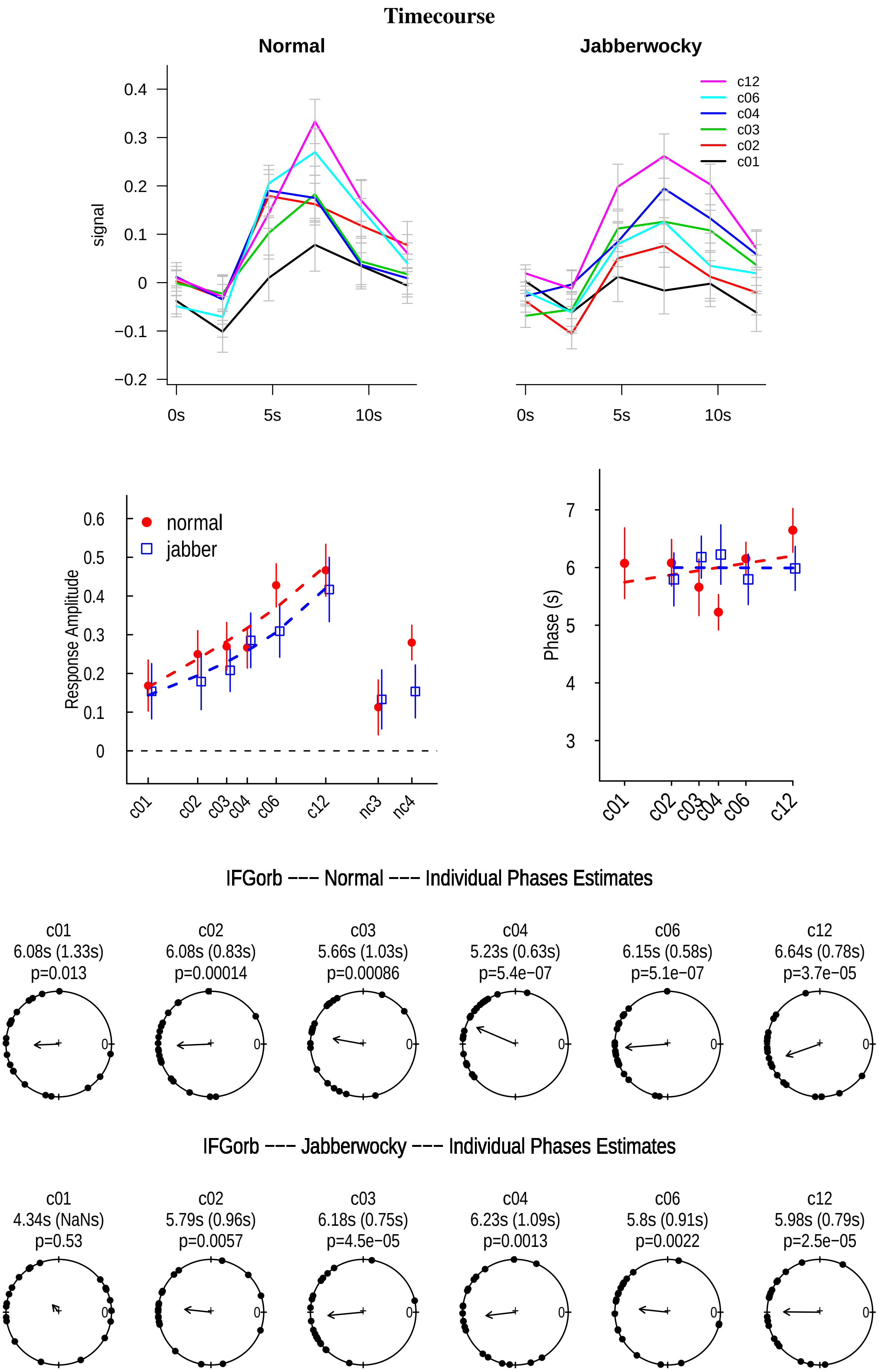


Fig. S6: Response in Left Inferior Frontal Gyrus - Pars Triangularis (IFGtri; -51 21 21)

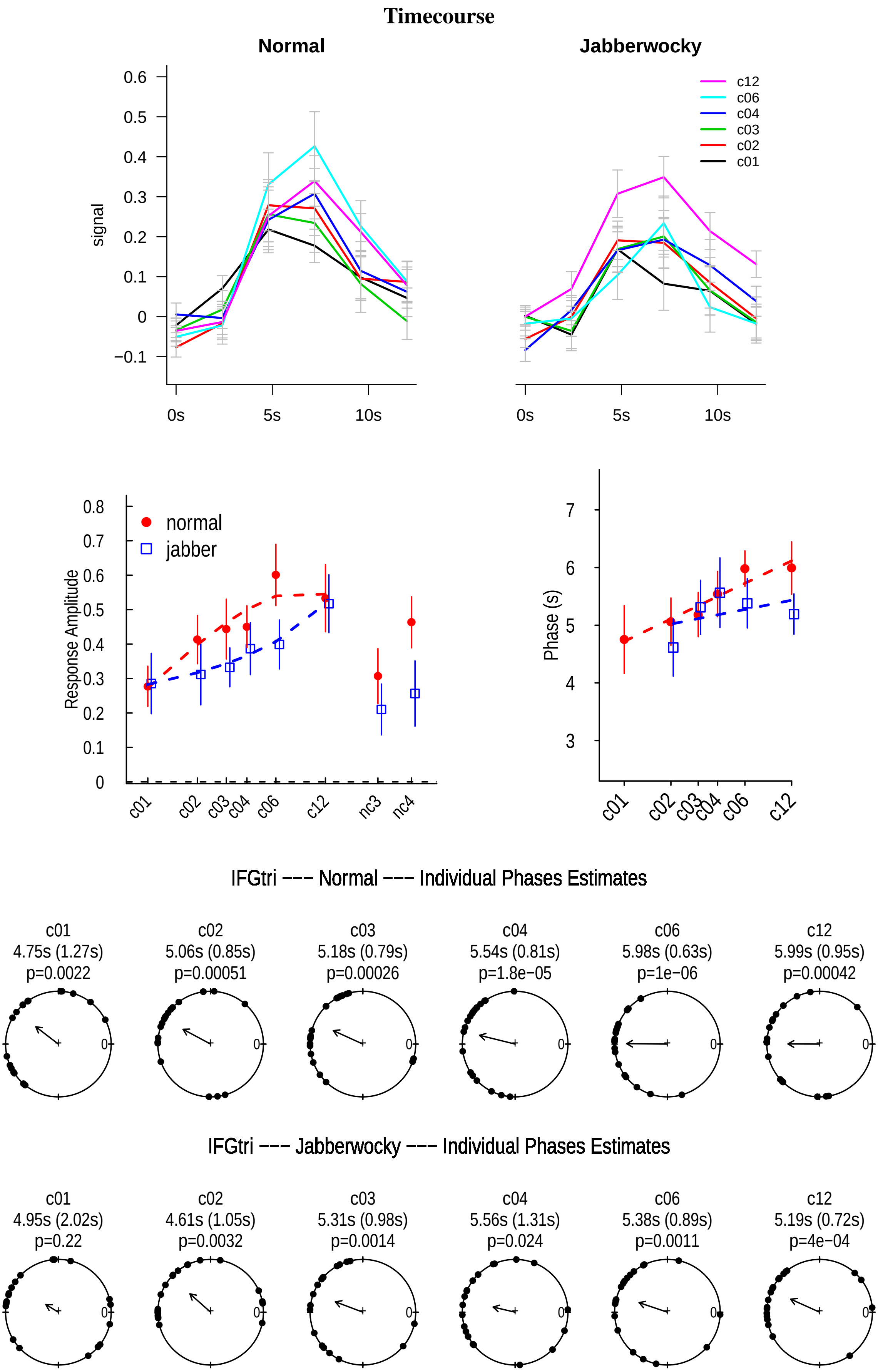


Fig. S7: Response in Left Posterior Superior Temporal Sulcus (pSTS; -51 -39 3)

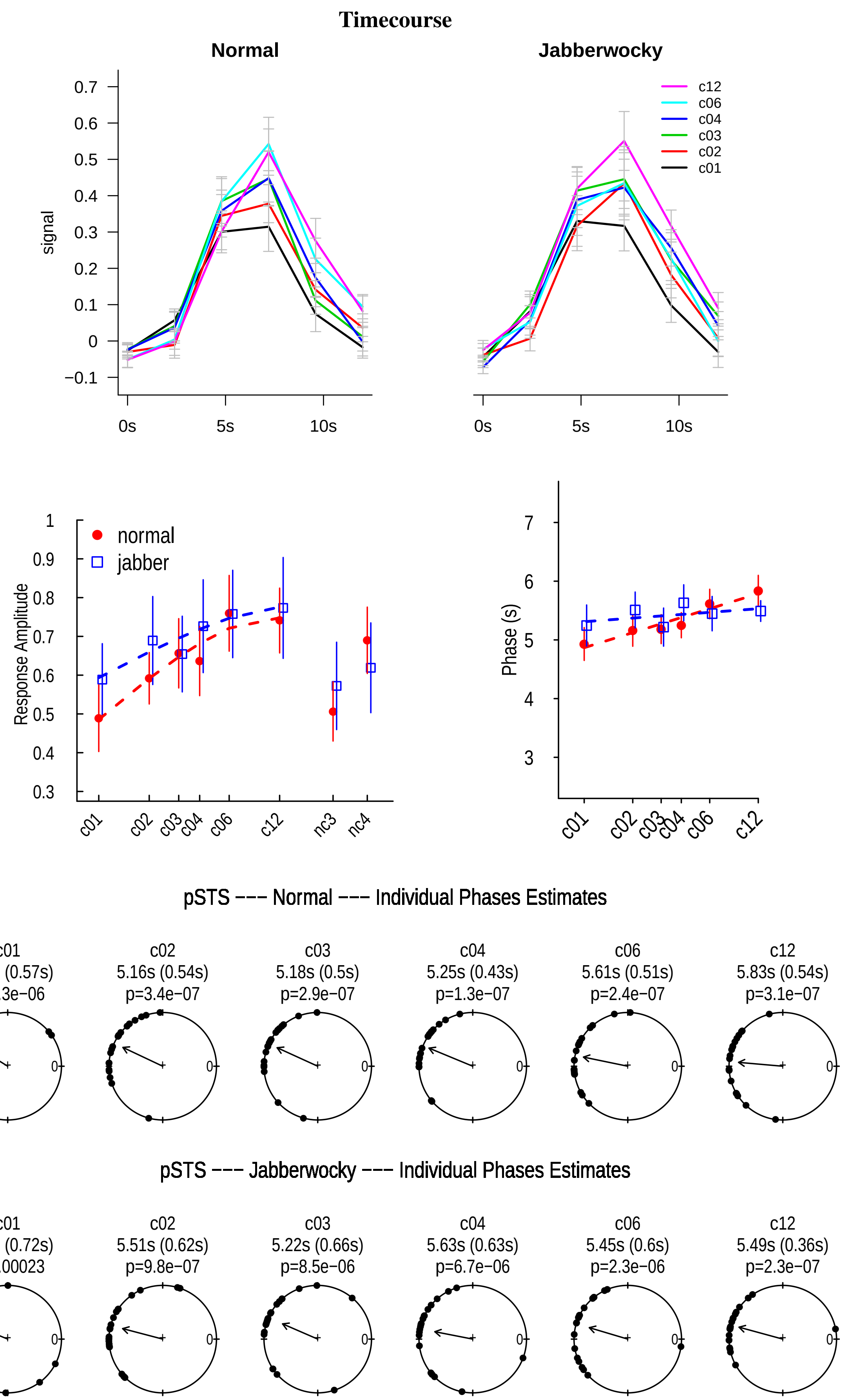


Fig. S8: Response in Left Temporal Pole (TP; -48 15 -27)

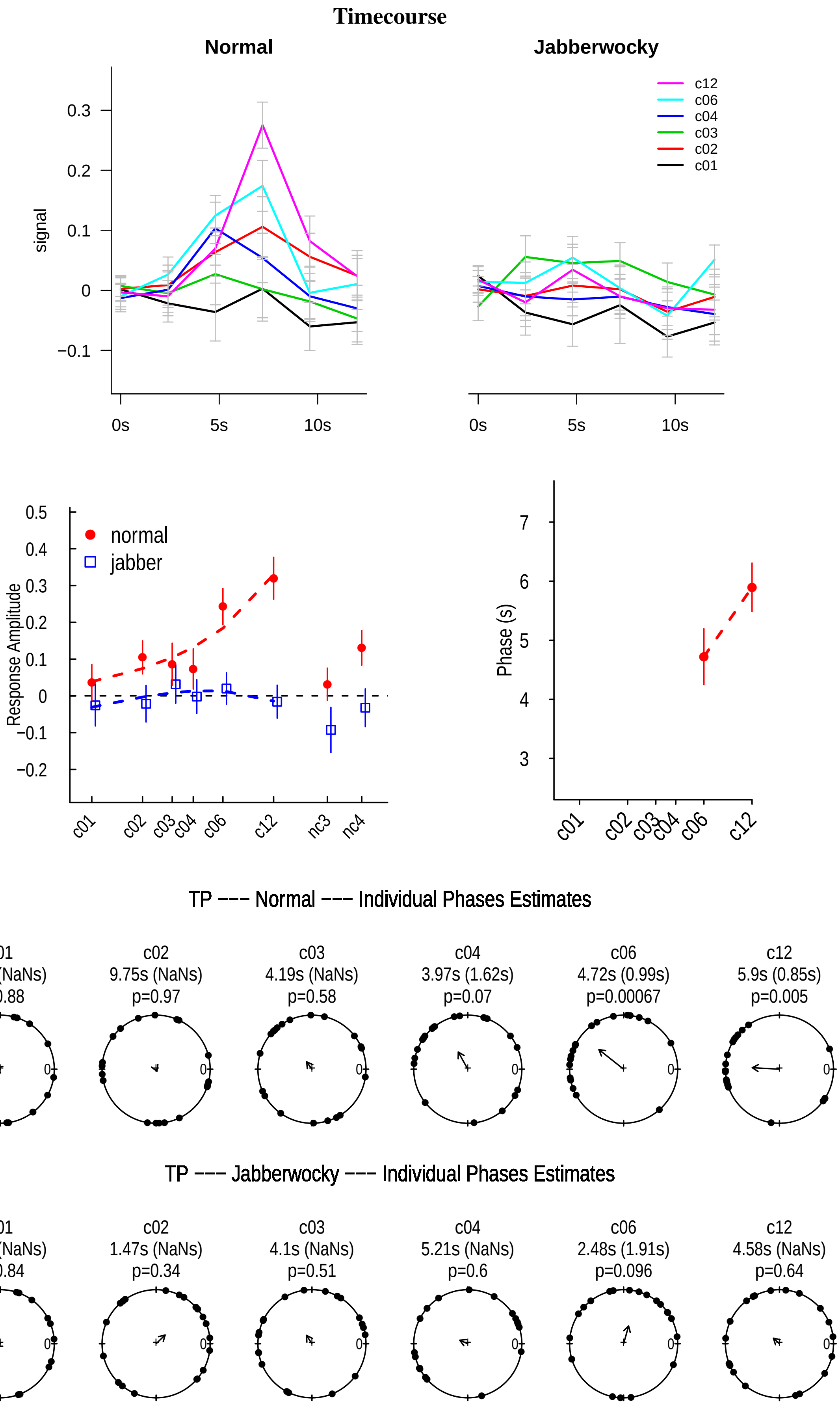


Fig. S9: Response in Left Temporo-Parietal Junction (TPJ; -39 -57 18)

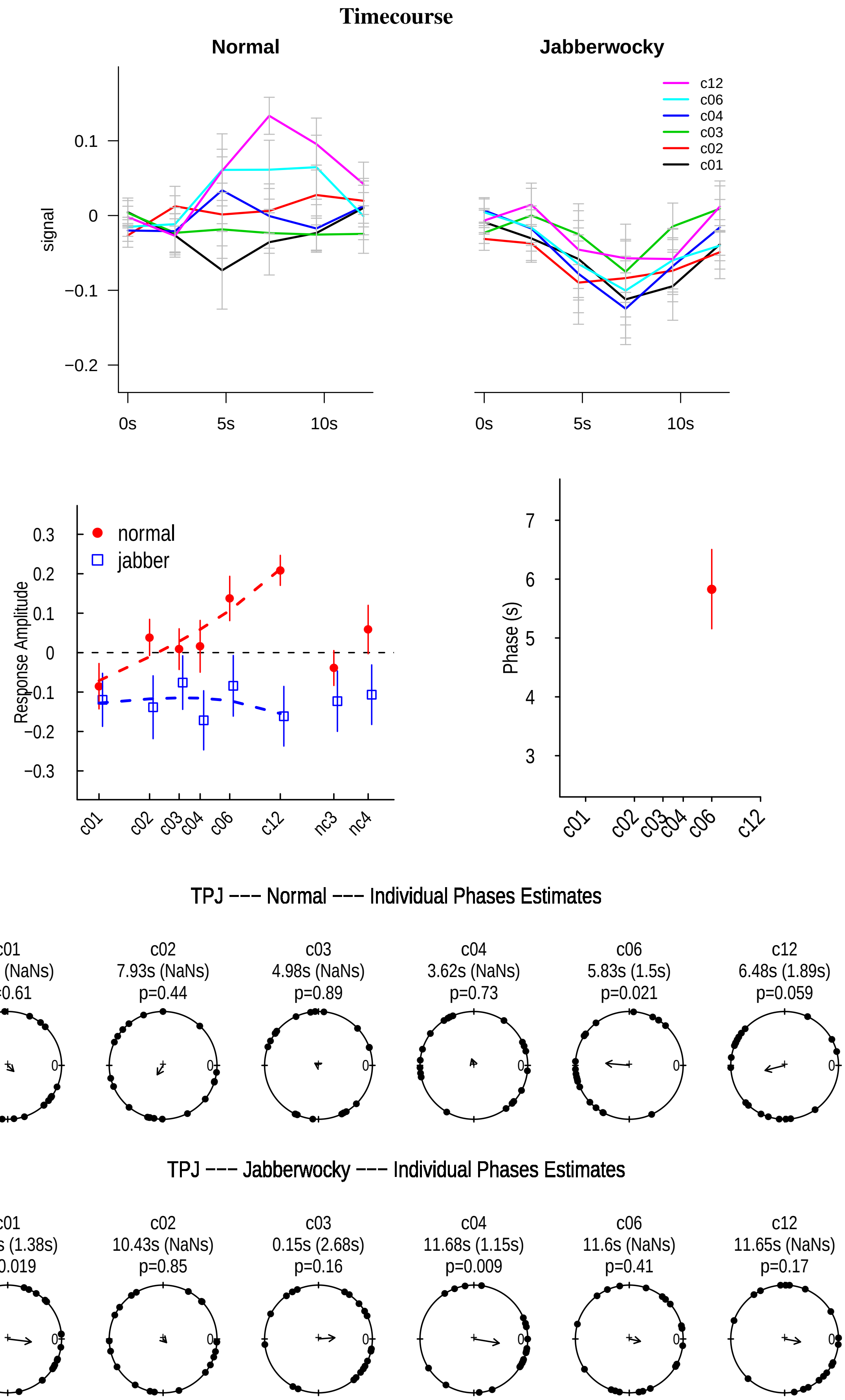


Fig. S10: Response in Left Putamen (-18 16 12)

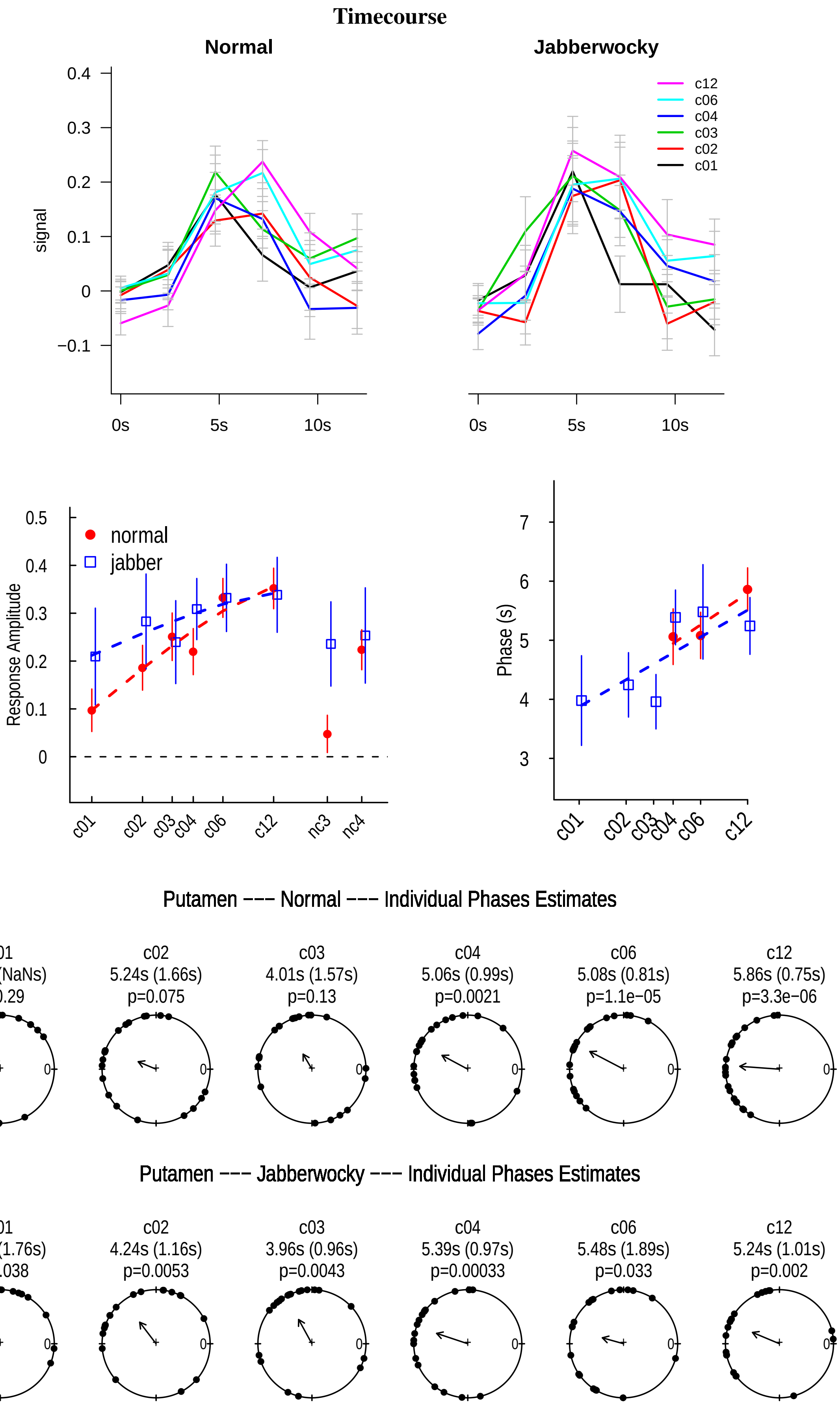


Fig. S11: Response in Precuneus (-6 -59 39)

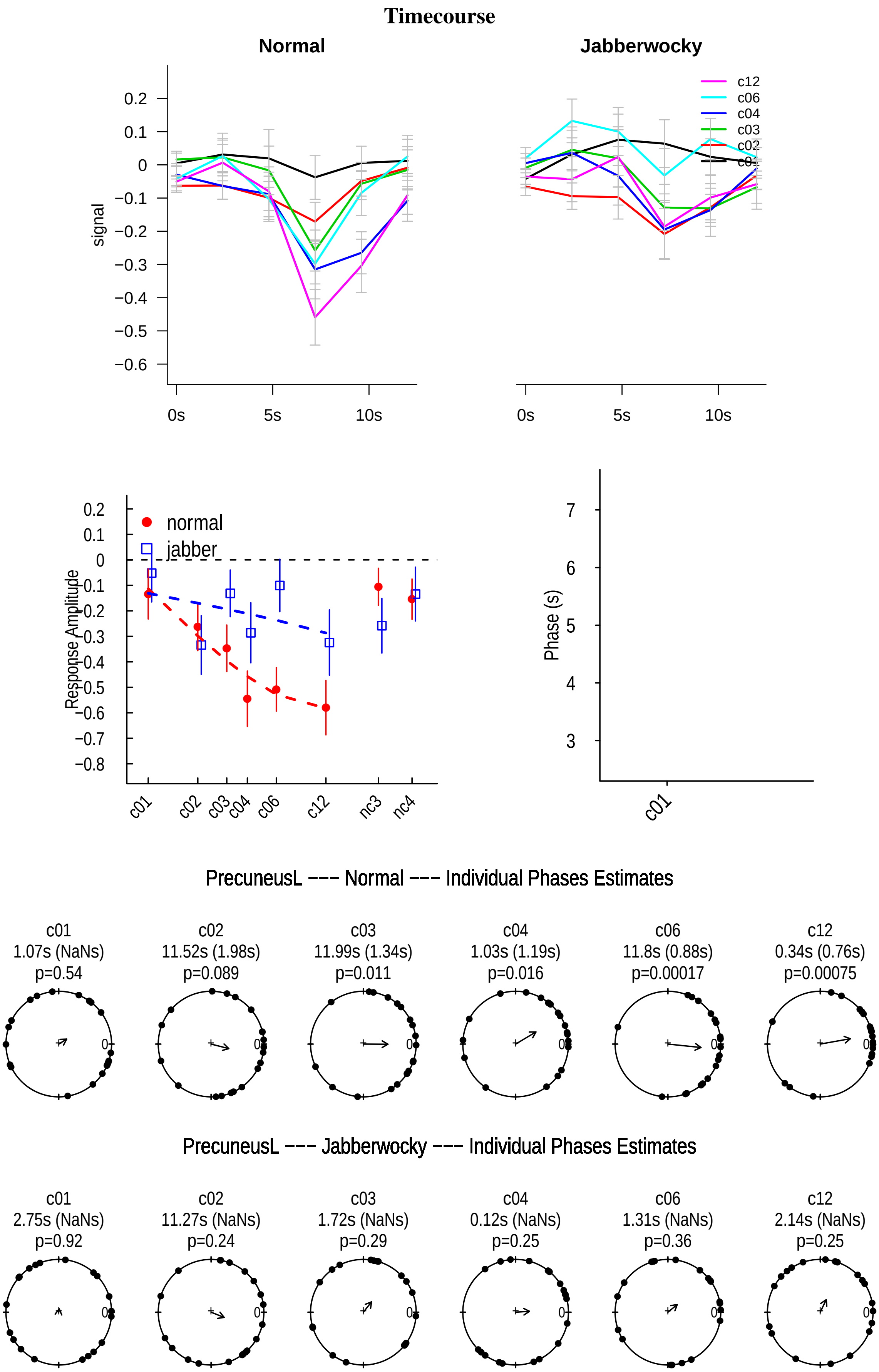


Fig. S12: Response in Left Inferior Parietal Lobule (-47 -53 47)

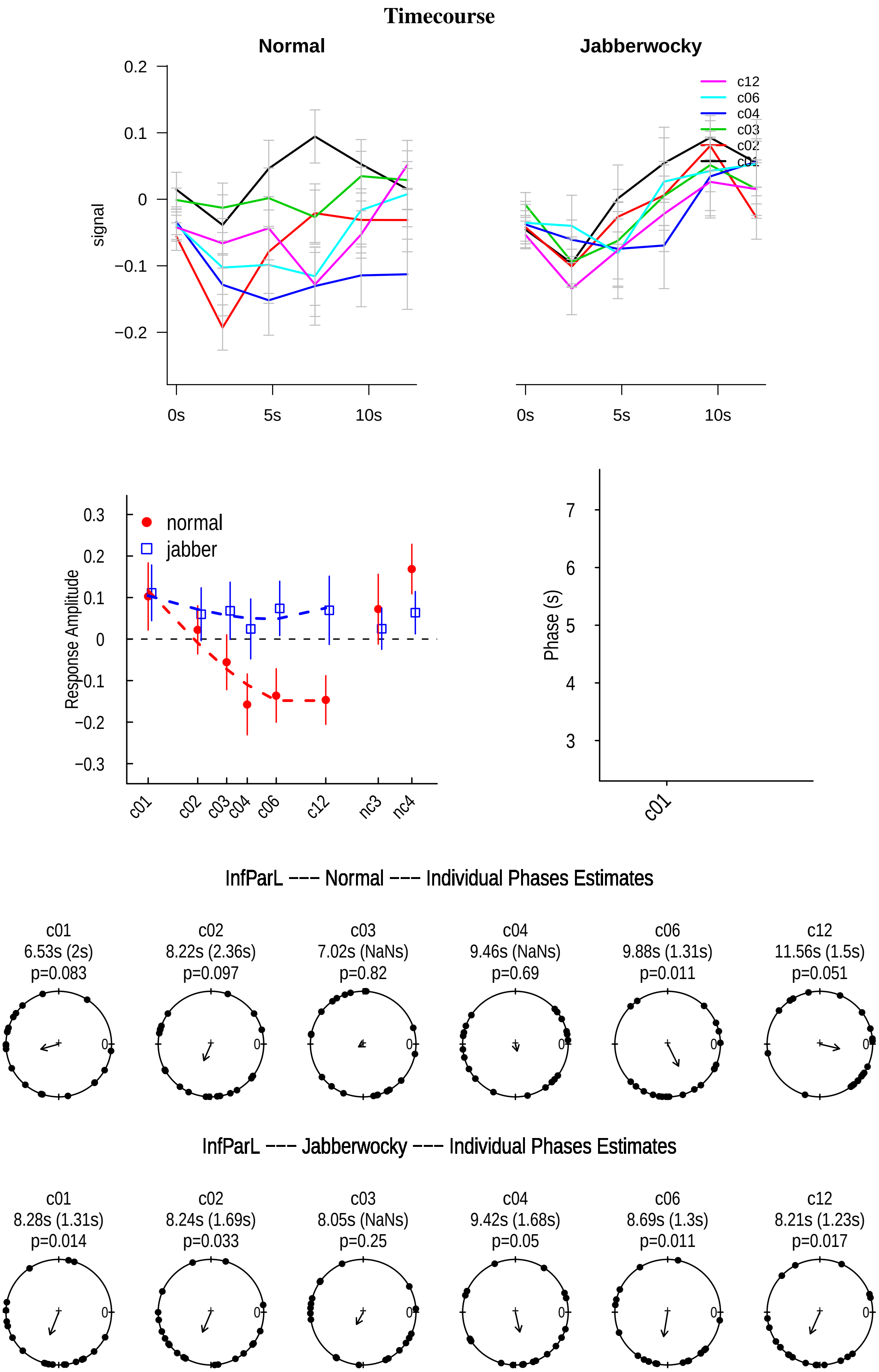


Fig. S13: Response in Anterior Cingulate (6 45 0)

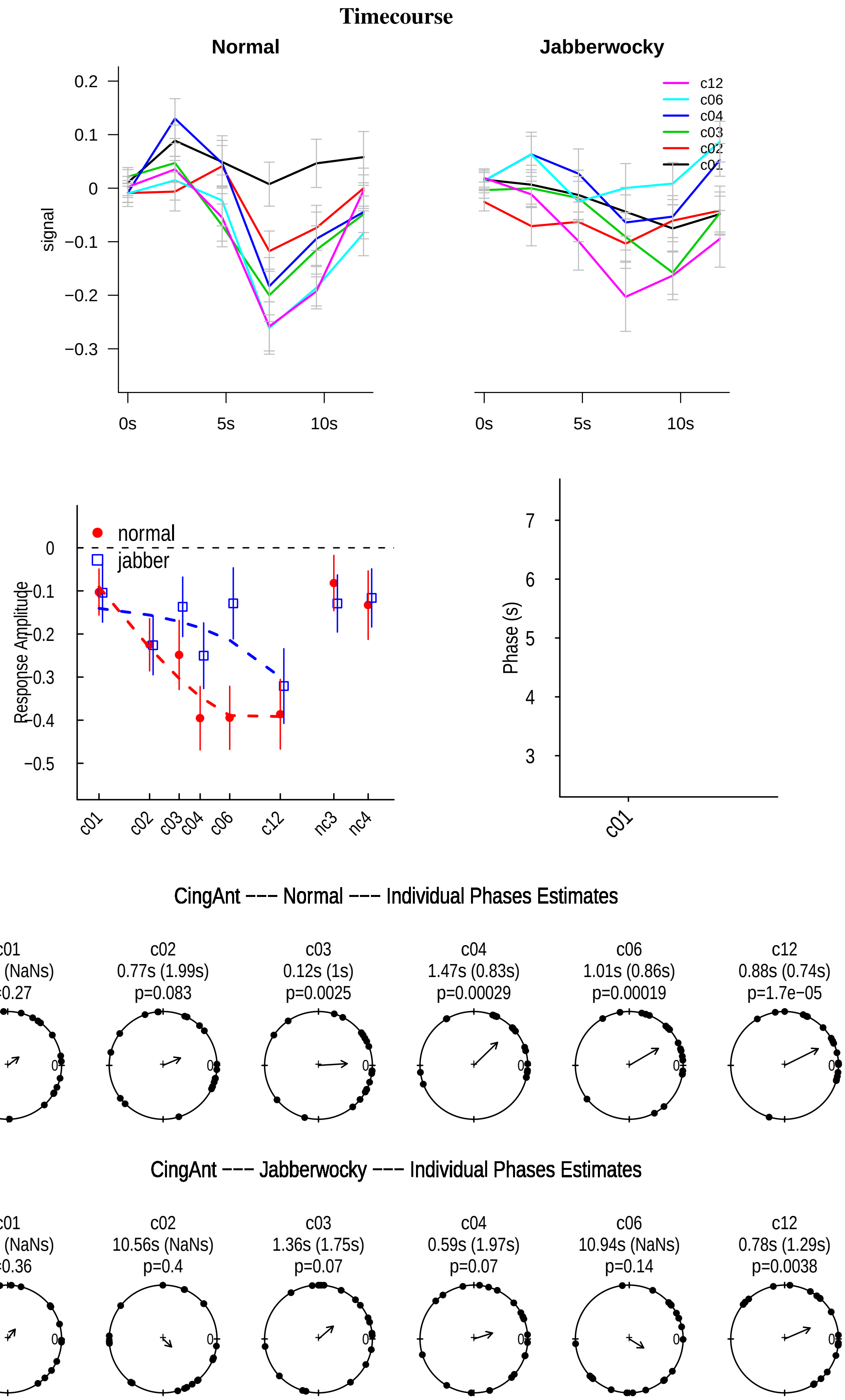


Fig. S14: Response in Middle Cingulate (3 -27 39)

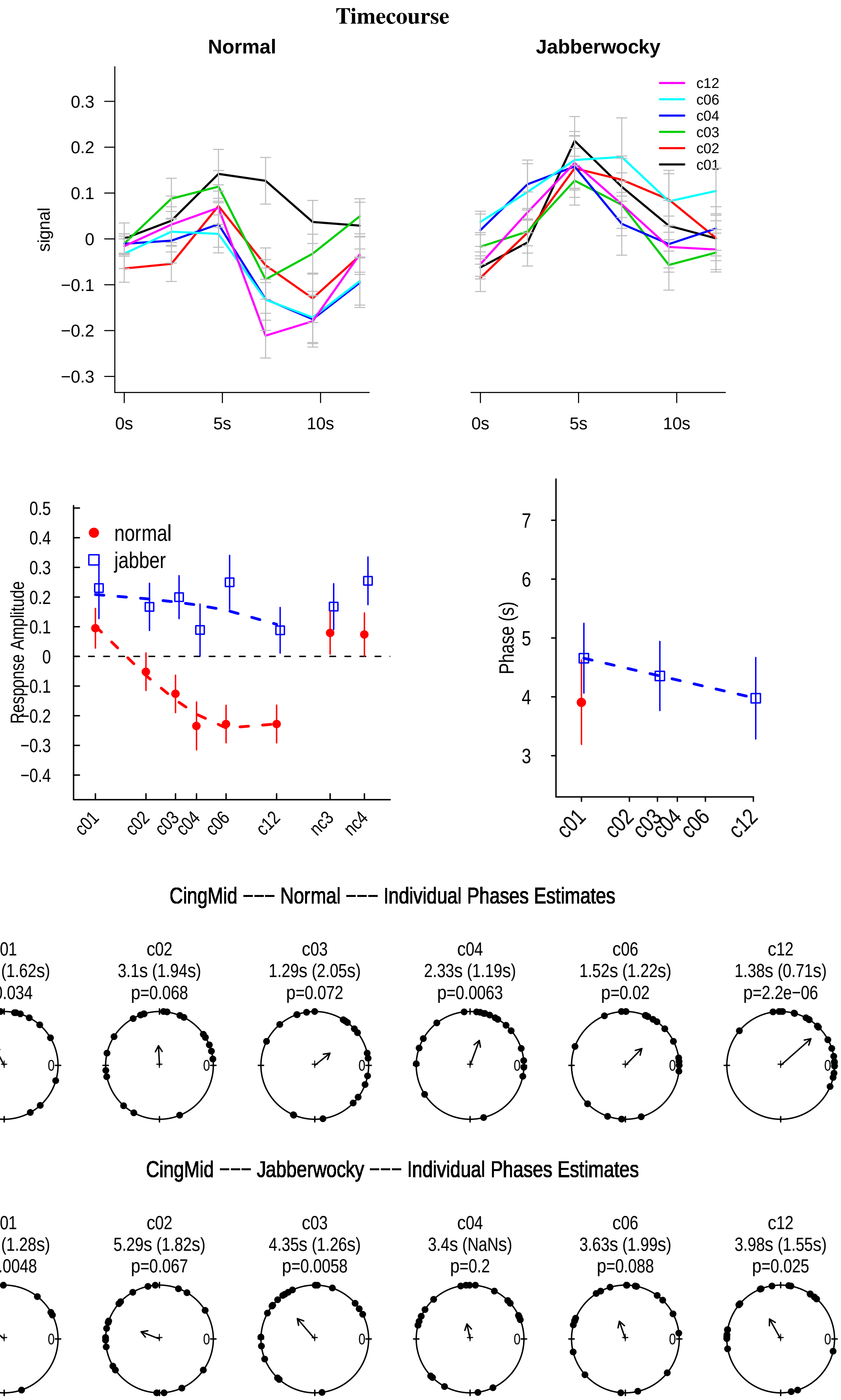
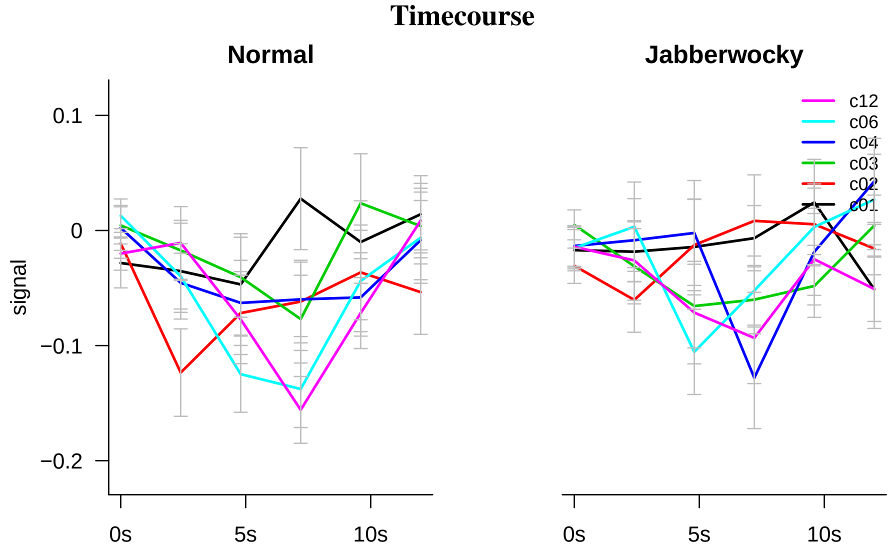
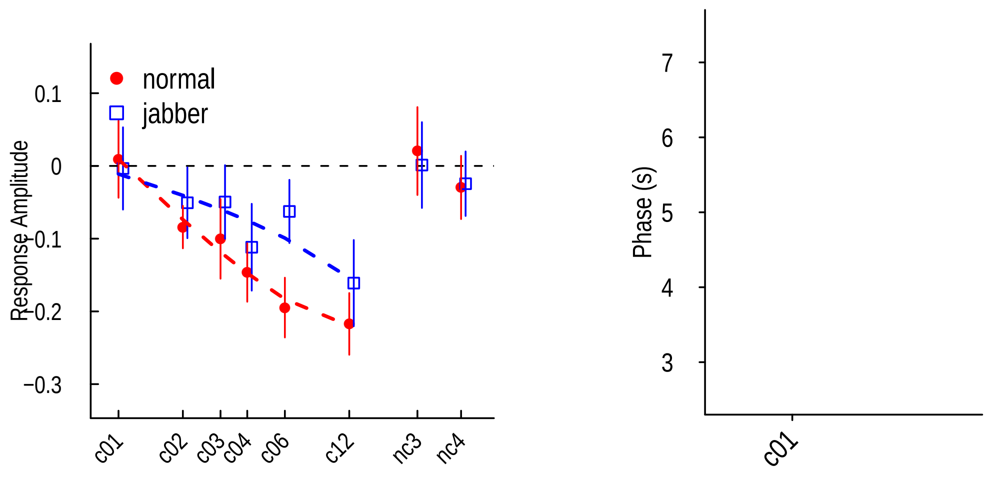


Fig. S15: Response in Right Frontal Middle Gyrus (36 30 45)





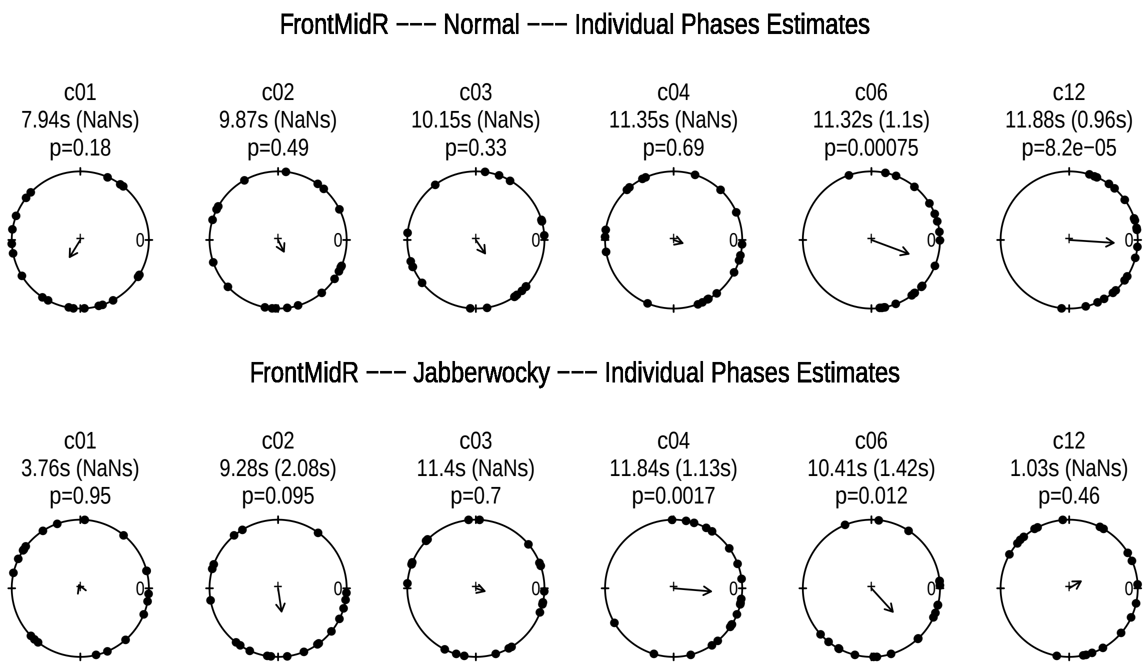
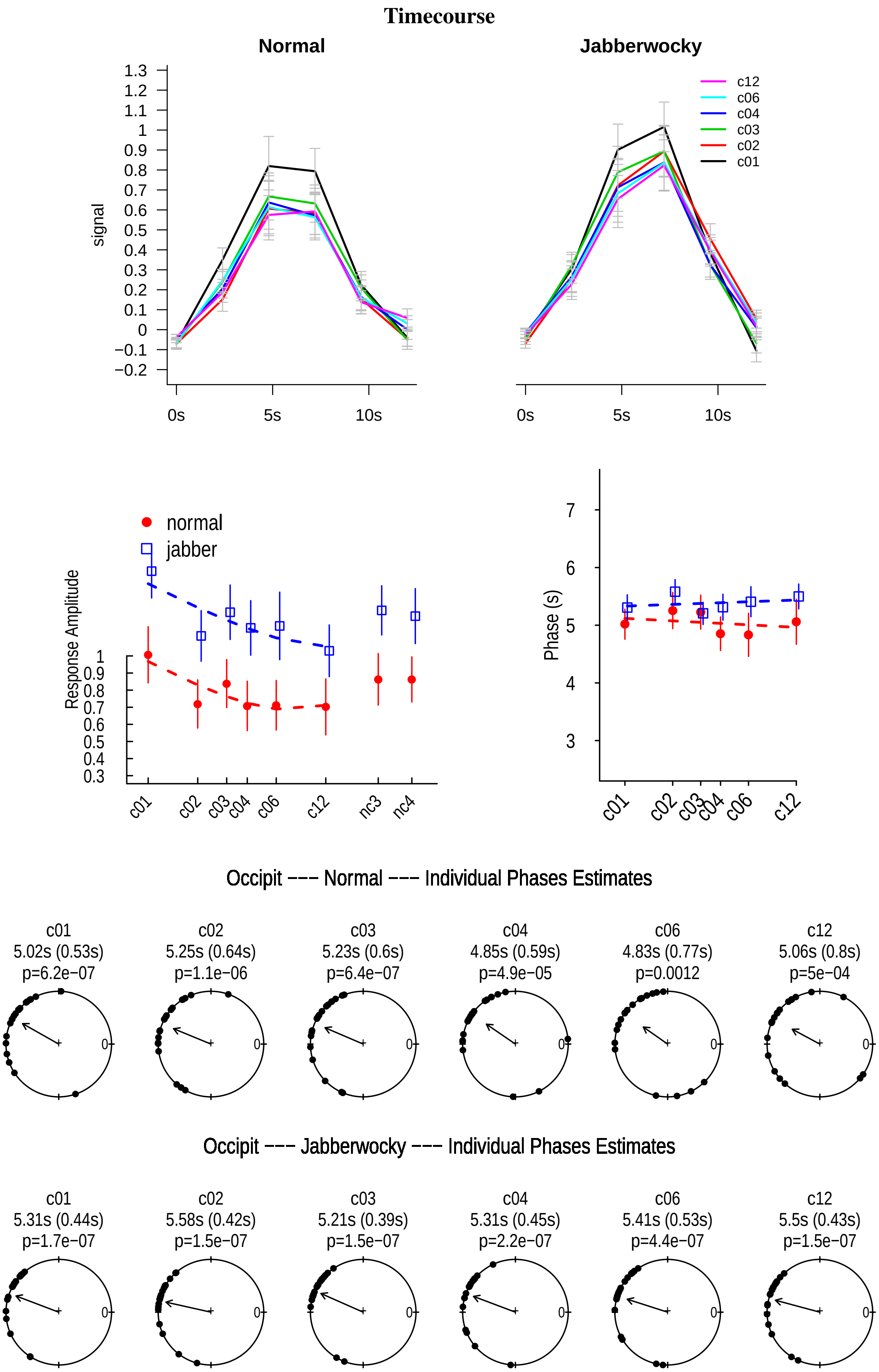


Fig. S16: Response in Left Occipital Middle Gyrus (-15 -90 -9)



5 Sentences from which constituents were extracted

1. Il reçoit un sac de céréales cultivées dans le nord du pays.
2. Elle regarde les éclairs qui traversent le ciel couvert de nuages gris.
3. Le boucher veut savoir si son père veille sur toute sa famille.
4. Il cherche le nom de la fille qui a appelé ses parents.
5. Les élèves veulent résoudre un problème qui demande des connaissances de physique.
6. Le prêtre salue le menuisier qui répare le plafond de la chapelle.
7. Le bijoutier a perdu la boîte de la chaîne en argent massif.
8. Le balayeur respire un air chargé de fines particules de poussière grise.
9. Elle désigne son bras pris dans un gros bloc de plâtre blanc.
10. La mère a peur que son enfant tombe dans la rivière glacée.
11. Il sursaute en entendant le grincement de la lourde porte en bois.
12. Le père quitte la maison pleine de souvenirs de son triste passé.
13. Le comptable vide toutes les pièces de son appartement de jeune célibataire.
14. Il ouvre la portière de sa voiture à la carrosserie très abîmée.
15. Il croit lire son nom écrit sur une des lignes presque effacées.
16. Ils construisent un haut mur de pierres extraites de la carrière voisine.
17. Nous regrettons que vous ayez décidé de quitter notre maison de production.
18. Il déteste la couleur des carreaux de sa deuxième salle de bain.
19. Nous envisageons de bâtir une maison qui dominera la vallée du fleuve.
20. Tes enfants refusent de lire le dernier chapitre de ce livre ennuyeux.
21. La crasse forme un mince film qui couvre leurs visages de travailleurs.
22. Il cuisine en suivant des recettes héritées de son vieil oncle Gustave.
23. Il a peur du sang qui pourrait couler de sa petite égratignure.
24. Il écoute le concierge raconter les aventures du locataire du premier étage.
25. Les étoiles brillent en émettant une lumière inconstante au cours du temps.
26. Il est mort des terribles suites de sa dernière chute de cheval.
27. Elle est passée prendre des nouvelles de la santé de sa nièce.
28. Le garçon déballe les légumes achetés sur la grande place du village.
29. Le jeune docteur rit en entendant son cousin inventer des anecdotes invraisemblables.
30. La bombe explose en faisant entendre une série de petits craquements brefs.
31. Il écoute les légendes sur la naissance de cet ordre de chevalerie.
32. Cet homme tient les comptes des gens qui détestent faire des additions.
33. Le libraire veut que les clients respectent son classement par ordre alphabétique.
34. Le lièvre remue ses longues oreilles couvertes de fin duvet brun clair.
35. Le boutiquier voudrait que tu achètes ces chaussures en peau de serpent.
36. Le voleur préfère attendre que la nuit tombe sur la ville endormie.
37. Il examine les minuscules fissures du bois du rebord de la fenêtre.
38. Le berger caresse la douce laine des moutons de son troupeau paisible.
39. Il sourit en découvrant un paquet tout emballé dans du papier doré.
40. La fillette cherche à comprendre le langage des pétales de pâquerettes roses.
41. Le plombier a renoncé à réparer la fuite du joint en plastique.
42. Elle aime caresser le revers de ses draps en satin de coton.
43. Il croit pouvoir revendre les diamants de la mine de ses parents.
44. Il a tordu les barreaux de sa cellule de la prison régionale.
45. Elle défait les noeuds de la corde qui retient le petit bateau.
46. Elle a accepté que son fils achète une télévision à écran plat.
47. Il dit que tu hésites à couper tes longs cheveux blond cendré.
48. Le vacancier écoute le bruit des vagues qui déferlent sur le rivage.
49. Il compte les mots qui couvrent la page imprimée en petits caractères.
50. Je crois que tu devrais accepter la proposition de ton nouvel associé.
51. Elle a oublié de prévenir la deuxième vendeuse du magasin de chaussures.
52. Il continue à nier que son fils a inventé toutes ces accusations.
53. Ils ont ciré leur magnifique parquet aux lattes en bois de chêne.
54. Le garagiste parvient à extraire le clou enfoncé dans le pneu avant.
55. Le passant examine le luxueux canapé abandonné sur le bord du trottoir.

56. La cuisinière fait griller des noisettes qui répandent une odeur très appétissante.
57. Elle essuie la sueur qui goutte de son front barré de rides.
58. Le banquier cherche la référence des derniers chèques de ce client dépensier.
59. Il découvre le trou de la souris qui a grignoté sa tartine.
60. Il va voir le bénévole qui taille de larges tranches de pain.
61. Le négociateur écoute le ronronnement du climatiseur de la salle de conférences.
62. Il abandonne la poursuite du petit éléphant aux défenses en ivoire éclatant.
63. Le jardinier redoute les brusques changements de la température de la serre.
64. Il parvient à déchiffrer les indices du visage de son redoutable adversaire.
65. Le joueur sort la carte dissimulée dans la doublure de sa manche.
66. Le soleil va accélérer la fonte de cette mince couche de glace.
67. Son doigt effleure les touches du grand piano décoré de bois précieux.
68. Le pompier sauve le bébé endormi au fond de son berceau blanc.
69. Le peintre nettoie le manche de ses brosses en poils de sanglier.
70. Le voyageur fait coulisser la porte de son compartiment de deuxième classe.
71. Il raye le fond de sa poêle au coûteux revêtement de téflon.
72. Il guette le léger glissement des chaussons de sa femme de ménage.
73. Mon ami aimerait comprendre les annotations de son professeur de grec ancien.
74. On raconte que cette femme collectionne des bouts de ficelle blanc cassé.
75. Il possède la patience nécessaire à un excellent vendeur de choses inutiles.
76. Le locataire va chercher les clés de son joli appartement si fonctionnel.
77. Le médecin examine les taches sur la peau de son patient terrifié.
78. Cette cliente veut prendre des pilules efficaces contre le mal de tête.
79. Le coureur fait le tour de la piste recouverte de bitume dur.
80. Le garçon choisit un costume de chasseur de grands crocodiles du Nil.
81. Il parvient à annoncer la nouvelle de la mort de son hamster.
82. Tu peux jeter les revues qui traînent sur la table du salon.
83. Je vais élaborer un plan qui permettra de redresser nos récentes pertes.
84. Notre poule pond des oeufs à la coquille dure comme du béton.
85. Le fermier dit que sa fille souhaite apprendre à traire les vaches.
86. Il a brisé le bouchon de la bouteille de jus de raisin.
87. Elle écoute la tranquille respiration de son amant couché sur la couette.
88. Il efface les traces de son passage dans le champ de blé.
89. La petite regarde comment son père grille les épis de maïs mur.
90. Il décrit les paysages qui défilent devant ses yeux de voyageur émerveillé.
91. La pluie frappe les ardoises des toits des immeubles de la ville.
92. Elle demande que son frère fasse des ronds de fumée de cigarette.
93. Le notaire consulte les lourds volumes empilés sur les étagères de métal.
94. Le garçon remonte ses lunettes à la monture en plastique bleu clair.
95. Elle a brûlé toutes les lettres de son ancien fiancé trop infidèle.
96. La balle reste coincée dans le fin canon cabossé en plusieurs endroits.
97. Le chasseur vise le cerf déjà blessé à la patte de devant.
98. Il redoute de devoir abandonner tout espoir de guérir avant le printemps.
99. Le libraire trébuche sur le rebord du trottoir couvert de verglas brillant.
100. Le chercheur croit que sa théorie va changer les façons de penser.
101. Elle écoute le feu qui crépite dans la cheminée de marbre gris.
102. Elle recompte les invités qui ont oublié de faire parvenir leur réponse.
103. Il croit pouvoir oublier les beaux moments passés auprès de sa maîtresse.
104. Ses enfants mettent des disques qui évoquent des jours du passé lointain.
105. Elle feuillette le dernier catalogue de la société de vente par correspondance.
106. Il rêvasse en rongeant un bout de son ongle déjà bien abîmé.
107. Elle cherche à séduire ce garçon très ami avec son amant regretté.
108. Les applaudissements retentissent dans le pieux silence de la salle de concert.
109. Elle sent que les larmes coulent sur ses joues rougies de fièvre.
110. Les prisonniers essaient de communiquer en utilisant la lumière de leurs lampes.
111. Sa collègue ramasse une écharpe tombée de la chaise à roulettes noires.
112. Le postier sourit en faisant tourner la lourde poignée de métal poli.
113. Le pianiste plaque le dernier accord de cette sonate difficile à jouer.

114. Le coureur masse sa cheville couverte de bandages de tissu jaune pâle.
115. Elle espère revoir son amant parti au chevet de sa femme malade.
116. La fée promet de faire apparaître un marteau au manche en cristal.
117. Il prétend que chaque participant pourra gagner une statue en chocolat blanc.
118. La stagiaire rêve de dormir sur un matelas de pétales de roses.
119. Il aborde les clients qui hésitent devant les rayonnages chargés de vêtements.
120. Cet élève a commencé à apprendre tous les mots de son dictionnaire.
121. Un frisson parcourt son dos offert aux regards de la foule hostile.
122. Le malade refuse de prendre les médicaments vendus dans les pharmacies ordinaires.
123. Le touriste déchiffre toutes les publicités des interminables couloirs du métro parisien.
124. Elle sursaute en entendant le grincement des planches du parquet mal entretenu.
125. La fourmi escalade des monceaux de pelures de pommes de plusieurs variétés.
126. La fermière trait la vache aux yeux brillants comme de petites billes.
127. Cet élève consulte un volume aux pages constellées de taches de café.
128. Il écoute le bruit de la pluie qui bat contre le carreau.
129. Le contrôleur regarde le voyageur fouiller dans son sac rempli de prospectus.
130. Le professeur interroge les élèves qui bavardent dans les rangs du fond.
131. Le chat joue avec la pelote qui vient de tomber du bureau.
132. Le peintre propose de refaire la décoration de la salle à manger.
133. Le touriste ouvre son portefeuille qui commence à déborder de papiers inutiles.
134. Le serveur prend la commande des clients assis près des toilettes publiques.
135. Il élabore un plan de restructuration de la compagnie qui fait faillite.
136. Il explique que son inspiration vient de ses voyages à la mer.
137. La sonnerie interrompt les méditations du sage assis sur un coussin gris.
138. Elle respire le parfum des roses qui ornent la table du salon.
139. Le chorégraphe réunit les danseurs qui veulent faire partie du prochain ballet.
140. La fille sourit en comprenant que son fils veut surprendre son père.
141. Il pâlit en écoutant les détails de la mort de son épouse.
142. Le cuisinier remue la sauce qui cuit dans une casserole en acier.
143. Le rescapé nettoie les plaies qui risquent de provoquer une infection grave.
144. Elle a égaré le cartable qui contenait les copies du dernier contrôle.
145. Il refuse de supporter la fumée des cigarettes de ses voisins impolis.
146. Il craint de manquer le train qui doit partir à quatre heures.
147. Sa main heurte la tasse posée sur le bras du fauteuil rouge.
148. Le garçon lance un caillou qui ricoche sur la surface du bassin.
149. Le romancier a fini le manuscrit promis à son éditeur si impatient.
150. Le médecin traverse la rivière qui coule devant sa maison de campagne.
151. Le surveillant éteint les lumières du couloir des internes du lycée français.
152. Le magicien fait disparaître les bonbons étalés sur le sol du théâtre.
153. Les larmes coulent sur les joues de la fille qui pleure doucement.
154. Elle coupe le fil qui pend de sa robe en soie noire.
155. Le vieux marin divague en buvant des verres de rhum des Antilles.
156. Le garçon répond que sa soeur cherche un tube de colle blanche.
157. La femme porte un sac alourdi de bouteilles de jus de fruit.
158. Paul tente de raisonner sa soeur qui se vexe sans bonne raison.
159. Les branches brûlent en lançant des étincelles qui illuminent la nuit sombre.
160. Elle cherche des gants qui puissent protéger ses mains de pianiste frileuse.
161. Elle marche en balançant ses hanches moulées dans une jupe de satin.
162. Le passant grave une croix qui entaille le tronc du peuplier imposant.
163. Les guerres volent la jeunesse de ces hommes qui devraient vivre heureux.
164. La reine boit un thé qui calme les irritations de sa gorge.
165. Pierre nie le manque de place dans cette chambre sous les toits.
166. La banque exige que le client ferme le compte de sa fille.
167. La balle tombe en éclaboussant la dame qui longe le trottoir mouillé.
168. Les meurtres effrayent les gens qui vivent près de ce bois obscur.
169. Le juge dit que Charles a le droit de voir son fils.
170. La fillette compte les gouttes qui tombent sur le sol de pierre.
171. La cire laisse des traces qui durcissent en formant des traînées blanches.

172. La femme porte le deuil de son mari mort à la guerre.
173. Le seau fuit en trempant la paille de la grange à foin.
174. Les traits forment des taches qui tranchent sur la blancheur du papier.
175. Le père jette le tas de lettres envoyées à sa jolie fille.
176. Le chanvre est une fibre qui donne de la toile trop rêche.
177. Le promeneur lâche la laisse de son chien qui aboie sans arrêt.
178. Le concierge déteste ce genre de livres qui lassent les lecteurs cultivés.
179. Il tape sur le clou qui fend le bois de la commode.
180. Une taupe creuse des trous qui défigurent le jardin de la voisine.
181. Il joue le rôle du comte qui meurt en maudissant ses enfants.
182. Le rat sort de son trou creusé dans le mur du salon.
183. Le fils crache en regardant sa mère qui garde un silence agacé.
184. Il cherche à joindre le curé de la paroisse de son parrain.
185. Le garçon refuse de manger de la viande pleine de nerfs blanchâtres.
186. Il écrit sur un bloc enfoncé dans la poche de sa veste.
187. Il regarde ses enfants choisir les jouets qui finiront dans la caisse.
188. La chanteuse jette le dé qui roule sur la table de jeu.
189. Le gel forme du givre qui luit sur les feuilles de laurier.
190. Elle dit que les saints vivent sur les nuages du ciel bleu.
191. Il entre sur la scène inondée de la lumière des projecteurs puissants.
192. Ils ont découvert la bactérie qui a causé des épidémies de grippe.
193. Ils ont construit une cabane de branchages couverts de feuilles encore vertes.
194. Son fils regarde une émission sur la migration des pigeons de Paris.
195. Il vérifie les installations qui assurent la sécurité des passagers du vol.
196. Le journaliste écrit un article qui dénonce les fraudes du maire sortant.
197. Le locataire raconte que la gardienne adore éplucher les pommes de terre.
198. Il rate le virage qui marque le début du circuit en montagne.
199. Il a peur que les talons rayent le parquet de bois tendre.
200. Le professeur dessine un diagramme incompréhensible pour ses élèves épuisés de fatigue.

6 Sentences from which non-constituents were extracted

1. Le moineau qui gazouille sur la branche du pommier a réveillé Paul.
2. Ce beau tapis tissé de longs poils de lama a coûté cher.
3. Le pélerin fatigué par son voyage sur les routes cherche un hôtel.
4. La femme du concierge de notre immeuble de prestige consulte des voyantes.
5. La récente perte de crédibilité de ses associés maladroits inquiète mon chef.
6. Le nouveau directeur de la banque de mes partenaires est très ambitieux.
7. Le sifflement de notre radiateur qui commence à fonctionner agace mon mari.
8. Notre voisin atteint par un cancer du poumon droit perd ses cheveux.
9. Le joli air qui retentit dans le matin silencieux plaît à tous.
10. La fête de notre bruyant locataire qui déménage demain provoque des disputes.
11. La chaude couleur du blé qui mûrit au soleil apaise les promeneurs.
12. Les tempêtes qui agitent la surface du lac bleu effraient mon amie.
13. Les livres rangés sur le meuble en fer rouillé ont coûté cher.
14. La révélation des secrets des habitants de ces quartiers a choqué marie.
15. La construction de ce monument consacré aux enfants aveugles a duré longtemps.
16. Les yeux du lion dessiné sur le tableau blanc sont très réussis.
17. La lumière des néons du plafond de ces couloirs modifie les couleurs.
18. Les vibrations de la vitre au fond du bureau dérangent la secrétaire.
19. La couleur de cette encre achetée le mois dernier fatigue les yeux.
20. Les réparations du robinet de la salle de bain ont été difficiles.
21. Le procès de ce bandit qui vole des montres passionne les gens.
22. Le soldat qui dort dans le lit de camp lavera vos chaussures.
23. Le lait de ces splendides vaches de race charolaise tourne plus vite.
24. Le nettoyage du grenier de notre maison de campagne occupera les enfants.

25. La fleuriste qui a préparé cet arrangement de roses appelle une collègue.
26. Le professeur qui a organisé cette conférence de médecine répond aux questions.
27. Les frites servies dans ce restaurant de la banlieue dégoulinent de graisse.
28. La boulangère qui mange un gâteau au chocolat blanc attend un visiteur.
29. Le professeur qui surveille toutes les épreuves du concours prend des notes.
30. Le jeune homme qui a apporté un énorme bouquet adore les fleurs .
31. Un prisonnier lassé de résister à ses médecins têtus avalerait les tranquilisants.
32. Le concierge curieux de connaître la maîtresse du locataire espionnent leurs rendez-vous.
33. Le garçon qui a ramené le courrier du jour ouvre les enveloppes.
34. La dame qui veut nettoyer les toilettes du grenier énerve mes parents.
35. Ces contrebandiers arrêtés par des policiers habillés en civils enfreignent la loi.
36. Les promeneurs qui ont poussé la porte du magasin achètent le livre.
37. Ce contrôleur qui dérange les voyageurs du train omnibus ruine mon trajet.
38. La ceinture autour des hanches du coureur de sprint gêne ses mouvements.
39. Le médecin qui a signé la feuille de soins recommande ces médicaments.
40. Les hésitations du chef de la meute de loups perturbent ses compagnons.
41. Cet avocat qui discute avec la maîtresse de maison rénove son cabinet.
42. Les retards des collaborateurs qui travaillent sur ce dossier décalent nos prévisions.
43. La boulangère qui tient la boutique de ma rue accumule les ennuis.
44. Le bétail du berger qui habite dans le quartier ravage nos pelouses.
45. Les convives invités par le mari de ma tante partiront plus tôt.
46. Ces émissions qui passent après le repas du soir intéressent les enfants.
47. Les élèves désireux de faire progresser leurs langues étrangères cherchent des correspondants.
48. Cet hôtelier perché sur un escabeau de métal rouillé fixe une pancarte.
49. Les neveux du buraliste qui vend de la réglisse apprennent le japonais.
50. Ce cavalier qui vient chaque jour de la semaine convoite la coupe.
51. Ce collaborateur au visage marqué de restes de varicelle joue au bridge.
52. Les joueurs qui ont gagné la partie de cartes désirent partir maintenant.
53. Ce cadenas fondu dans du métal difficile à trouver coûtera plus cher.
54. Les yeux de cette couturière fatiguée de trop travailler voient moins bien.
55. Le marchand avide de séduire des clients de qualité présente ses produits.
56. La voyageuse aux mains rougies par le froid coupant sort ses gants.
57. Cette personne responsable de maintenir le calme des enfants cherche une solution.
58. La girafe au cou constellé de taches marron foncé cligne des yeux.
59. Cet électricien auteur de travaux respectés dans le monde prépare une conférence.
60. Le jardinier qui a planté ces géraniums rouge vif regrette son choix.
61. Le serveur qui apporte les boissons destinées aux invités tombe par terre.
62. Cet enquêteur aux méthodes tirées de romans de gare résoudra nos problèmes.
63. Les conseillers du maire élu aux élections de mai avouent leurs erreurs.
64. Le gaz qui fuit de la bombonne de fonte a asphyxiée Marie.
65. Le chef de ce restaurant classé dans les guides rape des carottes.
66. Le cavalier imprégné du romantisme des livres de contes enfourche son cheval.
67. Cette araignée qui tisse des toiles de soie fine dégoûte ma mère.
68. Le présentateur de la chaîne de télévision par satellite annonce la nouvelle.
69. Cette demoiselle qui baisse ses yeux de biche apeurée raconte des mensonges.
70. Ce conducteur qui refuse de prendre plus de repos causera des accidents.
71. Cet ouvrier élu meilleur pâtissier de ce quartier animé refuse de parler.
72. Le notaire enfermé dans son étude sans aucune fenêtre perd ses couleurs.
73. Les talons des chaussures de ma femme si élégante rayent les parquets.
74. Le procureur qui instruit le procès des agressions récentes pose une question.
75. Les insectes qui vivent dans les recoins très sombres sortent la nuit.
76. La promenade le long de la falaise de craie a épuisé Jean.
77. Ce boutiquier qui détestait les hésitations des clients indécis a fait faillite.
78. Le plombier qui a installé les joints du lavabo forme un apprenti.
79. Le machiniste qui porte un pantalon couvert de boue travaille le jeudi.
80. Le greffier qui enregistre les plaintes des nouveaux arrivants a perdu patience.
81. Le livre qui sert à caler cette lourde armoire pèse deux kilogrammes.
82. Le pompier qui a sauvé le bébé de Paul accepte une bière.

83. La secrétaire qui a annulé les entrevues prévues demain doit rappeler Jean.
84. Ce criminel qui a braqué une banque de quartier protège ses complices.
85. Cette serveuse curieuse de connaître le passé du client prend sa commande.
86. Le clochard aux pieds fendus par des crevasses anciennes demande des chaussures.
87. Ce fleuriste spécialisé dans les compositions pour les mariages présente ses prix.
88. Ce meunier qui doit réparer les ailes du moulin consulte son entourage.
89. Le meurtrier accusé de plusieurs assassinats de vieilles dames étonne les enquêteurs.
90. Cet ustensile affuté avec une meule de pierre rugueuse coupe bien mieux.
91. Les documents qui encombrent ma valise à roulettes spéciales viennent du ministre.
92. Le bijoutier qui a confectionné ma bague de fiançailles réalise des miracles.
93. Le patient qui recrache toutes ses pilules de calmant contrarie les infirmières.
94. Ce comédien qui prend quelques minutes de repos mérité passe la serpillère.
95. Le crocodile qui vient de sortir du fleuve boueux terrorise les touristes.
96. Notre ami qui écrit des poèmes sur la nature cueuille des pâquerettes.
97. Le mari de la conseillère qui organise le dîner craque des allumettes.
98. La bactérie qui contamine les patients de cet hôpital perd sa virulence.
99. La pianiste qui va entrer sur la scène glissante peigne ses cheveux.
100. Le réceptionniste accusé de décourager les clients du soir défend ses intérêts.
101. Les bénévoles engagés pour construire un muret de béton dessinent des plans.
102. Ce chateau dressé sur une colline de cette région attire des visiteurs.
103. La maquilleuse qui a oublié de nettoyer ses pinceaux bafouille des excuses.
104. Le policier qui prend la déposition de ce plaignant lâche son crayon.
105. Le naufrage du paquebot vanté sur ces grandes publicités cause la stupeur.
106. Les écailles de ce serpent dissimulé dans les herbes brillent au soleil.
107. Le chirurgien qui a pratiqué cette opération si délicate revient du bloc.
108. Les invités qui ont cassé le plat de porcelaine proposent leur argent.
109. Le cuisinier qui prépare un entremet nappé de caramel invente ses recettes.
110. Le facteur sur sa bicyclette chargée de sacs postaux agite la main.
111. Le braconnier perdu dans les forêts de la montagne redoute les loups.
112. Le producteur qui a proposé cette émission de télévision remporte les votes.
113. Cet acteur qui part voir les hippopotames du zoo a quarante ans.
114. Les artistes qui consultent des voyantes aux tarifs astronomiques manquent de confiance.
115. Le pharmacien qui a refusé de vendre les cachets a eu tort.
116. Les prévisions de ce chercheur connu pour son honnêteté alarment la presse.
117. Le musicien qui a joué la sonate pour violon range son instrument.
118. Le concertiste qui a apporté la partition du menuet commence à jouer.
119. Ce candidat qui veut passer son permis de conduire accumule les fautes.
120. Ton collègue qui a fréquenté ces hôtels de luxe a quitté Paris.