

Prosodic clues to syntactic processing—a PET and ERP study

K.N. Strelnikov,^{a,d,*} V.A. Vorobyev,^{a,c} T.V. Chernigovskaya,^b and S.V. Medvedev^a

^a*Institute of the Human Brain RAS, Saint-Petersburg, Russia*

^b*Saint-Petersburg State University, Russia*

^c*Centre for Cognitive Neuroscience, Department of Psychology, University of Turku, Turku, Finland*

^d*Cognitive Brain Research Unit, University of Helsinki, Helsinki, Finland*

Received 29 April 2005; revised 5 August 2005; accepted 25 August 2005

Available online 26 September 2005

Syntactic processing of spoken speech often involves prosodic clues processing. In the present PET and ERP study, subjects listened to phrases in which different prosodic segmentation dramatically changed the meaning of the phrase. In the contrast of segmented vs. non-segmented phrases, PET data revealed activation in the right dorso-lateral prefrontal cortex and in the right cerebellum. These brain structures, therefore, might be part of the syntactic analysis network involved in prosodic segmentation and pitch processing. ERP results revealed frontal negativity that was sensitive to the position of the segmenting pause, possibly reflecting prosody-based semantic prediction. The present results are discussed in the context of their relation to brain networks of emotions, prosody, and syntax perception.

© 2005 Elsevier Inc. All rights reserved.

Keywords: Syntactic; Prosodic; Language; Brain imaging; Speech perception; PET; ERP

Introduction

The way in which words are thematically related in a sentence is conveyed by syntactic structure. The neuroanatomical basis of syntactic parsing is believed to include the left perisylvian associative cortex, with possible contribution of the homologous contralateral cortex, as suggested by brain lesion studies (Grodzinsky, 1990, 1995; Berndt et al., 1996; Caplan et al., 1996). Broca's area seems to be the most frequently mentioned candidate for the key brain structure of syntactic analysis (Swinney and Zurif, 1995; Swinney et al., 1996). The activation level of Broca's area correlated with syntactic complexity in some PET studies for both visual (Just et al., 1996; Stromswold et al., 1996) and auditory (Caplan et al., 1998; Caplan et al., 1999) sentence presentation. It is also interesting that areas adjacent to Broca's area were shown to

subserve syntactic errors detection (Indefrey et al., 2001; Friederici et al., 2003). In numerous studies, the processing of syntactic structures was also associated with some ERP components—e.g., P600 and the so-called left anterior negativity (LAN), which has its maximum above Broca's area (e.g., Neville et al., 1991; Kluender and Kutas, 1993). It should be noted that the literature includes discussion on the nature of these ERP components (e.g., Friederici, 2002; Ullman, 2004).

However, the regions in the right hemisphere, homologous to Broca's and Wernicke's areas, were also implicated during syntactic processing (Just et al., 1996). Humphries et al. (2001) found bilateral activations in the anterior temporal cortex during speech sounds, as opposed to non-speech sounds. In their PET study, Mazoyer et al. (1993) showed that the bilateral anterior temporal cortex was the only structure specifically activated by listening to sentences, whereas Broca's area was activated not only by sentence listening but also by hearing separate words. Friederici et al. (2000) also showed that syntactic processing of speech bilaterally influenced anterior temporal cortex activation. Thus, syntactic processing seems to be subserved by the cross-hemispheric neural networks.

According to Homae et al. (2002), activation in the left inferior frontal gyrus during sentence processing does not depend on sensory modality. Still, the brain processing of oral speech peculiarities remains mostly unclear. Natural language perception and oral speech syntax are based primarily on the auditory modality—in particular, on such prosodic features as changes in pace, tone, and loudness (Shattuck-Hufnagel and Turk, 1996). Further understanding of oral speech processing requires an investigation of its prosodic mediation (Friederici, 2002). Studies of prosodic processing have suggested the involvement of either the right hemisphere or both hemispheres (Baum and Pell, 1999; Meyer et al., 2002, 2004; Kotz et al., 2003), whereas modality-independent syntactic processing is usually associated with the left hemisphere (e.g., Chernigovskaya and Deglin, 1986; Caplan et al., 1998, 1999; Indefrey et al., 2001; Röder et al., 2002). However, an interaction between the structural conditions and prosodic conditions was observed bilaterally in the anterior temporal lobe along the superior temporal gyrus (Humphries et al., in press). ERP studies have revealed some commonalities in prosody and syntax

* Corresponding author. Cognitive Brain Research Unit, Department of Psychology, P.O. Box 9, University of Helsinki, FIN-00014 Helsinki, Finland.

E-mail addresses: kuzma.strelnikov@cbru.helsinki.fi, strelkuz@yandex.ru (K.N. Strelnikov).

Available online on ScienceDirect (www.sciencedirect.com).

processing: the same electrophysiological phenomenon accompanied the processing of prosodic boundaries in spoken speech and the processing of commas during silent reading (Steinhauer, 2003; Steinhauer et al., 1999). Generally, though the direct interaction of prosody and syntax is proved, the existing data do not allow judgment on the specific brain mechanisms of the prosody/syntax interface. Despite the theoretical and methodological complexities of tackling this problem, it is an intriguing question for further neurolinguistic studies.

The main aspect of this problem is how the existing data on syntax processing correspond to the necessity of using prosodic clues for spoken speech syntactic parsing. One prosodic cue, important for syntactic analysis of spoken phrases, is the position of a segmenting pause (accompanied by the appropriate pitch and syllable duration changes), which corresponds to the comma in writing. The aim of the present study was to investigate the brain mechanisms of such segmentation of spoken phrases, with the hypothesis that they activate the prosody/syntax interface neural networks. Russian appears to be an appropriate language to use in this study because it allows a much higher degree of freedom in the order of words than English or the Roman languages. Word order is frequently not informative in Russian; thus, prosody in spoken language (pauses, stress, pitch, duration) and punctuation in written language remain the only means to distinguish between the completely different meanings of two otherwise identical sentences.

Our stimuli were designed so that the perceived syntactic structures depended exclusively on the prosodic cues. Thus, the prosody/syntax interface was expected to be activated at most while comparing segmented phrases with non-segmented. The positron emission tomography (PET) technique was used to determine the brain areas recruited in this mechanism, and the event-related potentials (ERP) technique was implemented to specify the timing of this recruitment.

Methods

Subjects

Twelve volunteers (men, aged 18–35 years) participated in the PET study. Eighteen volunteers (10 men, 8 women, aged 18–30 years) participated in the ERP study. None had a past history of psychiatric or neurological diseases, and none was on any medication. All subjects were native Russian speakers, with at least a secondary school education, and all were right-handed. Written informed consent was obtained from each subject in a form approved by the Ethical Committee of the IHB and in agreement with the Declaration of Helsinki (1975).

PET and ERP data acquisition and analysis

PET data were acquired with a Scanditronix PC2048-15B tomograph providing 15 slices in 9.7 cm axial field of view with 5–6 mm inplane resolution (Holte et al., 1989). A transmission scan with rotating pin source was performed for each subject right after head positioning and immobilization in order to correct emission PET data for attenuation. An individually fitted thermoplastic mask was used to maintain the same position of subject's head inside of the tomograph. A catheter for tracer administration was placed in the subjects' right antecubital vein. Prior each emission scan, a bolus injection of 40 mCi of [^{15}O]-water was administered through

the catheter. Four 60-s scans were performed for each subject (2 scans per condition). Emission data acquisition started automatically at the moment of enough events rate, detected by the PET camera, that is approximately 15–20 s after the injection. The minimal time interval between scans was 15 min. Stimulation started 10 s before the injection and lasted for 100 s enough to cover the scanning period.

The emission PET data were reconstructed into images using 7 mm Hanning filter. Then, a statistical parametric mapping (SPM99, Wellcome Department of Cognitive Neurology, London, UK) software was used for further data analysis including realignment, spatial normalization to the Montreal Neurological Institute (MNI) brain template, smoothing with 16 mm isotropic Gaussian kernel, and statistical assessment (Friston et al., 1994). Differences in global flow were removed by proportional scaling of value in each gray matter voxel of an image to image's mean. In order to test for inter-condition activation differences, pair-wise subtractions (contrasts) between corresponding images were performed using *t* statistics. The ensuing images (SPM-*t* maps) were estimated using a voxel-level threshold $t = 4.83$ that corresponded to $P < 0.05$ false positive probability with correction for multiple non-independent comparisons (Worsley et al., 1996). Anatomical localization of cluster maxima was performed using the Talairach and Tournoux (1988) brain atlas after accounting for a spatial difference between the MNI brain template and the atlas brain by nonlinear coordinate transformation, as provided by the MRC Cognition and Brain Sciences Unit (<http://www.mrc-cbu.cam.ac.uk/Imaging/Common/mnispace.html>).

EEG data were recorded with 19 Ag/AgCl electrodes, mounted according to the “10–20” system, referenced to the ears. The signals (0.03–100 Hz) were continuously recorded and sampled at 250 Hz. Trials with eye artifacts ($>75 \mu\text{V}$) were rejected (20–30% of epochs), the threshold was set in each channel. The signals were digitally filtered (band pass 2–30 Hz), baseline correction was applied, using the 100 ms prestimulus interval. ERPs were averaged for each presented sentence. The analysis epoch was 1600 ms, including the 100 ms prestimulus interval. However, because the pause positions differed in the compared phrases, the compared phrases had different timing structures. Since the effect of the different pause positions in a phrase was the object of the study, a direct comparison of the ERPs for the whole phrases was methodologically inappropriate. To solve this problem, parts of the resulting epochs, corresponding to a particular word, differently located in a sentence, were compared. An epoch length for the compared words was limited to 400 ms, so that the differing parts of stimuli following these words were not included in the comparison. Differences in the ERP amplitudes were assessed using two-way MANOVA (electrode, the word position, $P < 0.05$) for the 20 ms time intervals around the maximum of the observed differences; *P* values underwent a modified Bonferroni correction to reduce the risk of Type 1 errors.

Stimuli

The speech stimuli were 60 phrases, recorded by a female professional speaker who was instructed to read them neutrally (without emotional load) at a pace of about 100 words/min. Half of the phrases were segmented by the intonational pause (accompanied by the appropriate pitch and syllable duration changes) in the speaker's voice. Depending on its position, the intonational pause changed the meaning of a sentence (e.g., “The teacher said, ‘A

student” and “‘The teacher,’ said a student”). The remaining phrases were not segmented by the pause (e.g., “Father bought him a coat”). These non-segmented sentences were lexically different from the segmented ones. Phrases recorded by the speaker were then equalized by their averaged RMS power, duration of the semantic pauses (110 ± 10 ms), and full time of a phrase (2.5 ± 0.5 s), preserving the natural sound of the phrases. During the experimental sessions, the phrases were delivered through loudspeakers at a volume chosen by the subject to obtain maximum speech perception comfort.

Visual stimuli were word strings presented in white letters on a black background, using a 17-in. computer screen in the center of the subject’s field of view, at a distance of 1.5 m from the subject’s eyes. The visual and acoustic stimulations were performed using “Presentation” software (Version 0.40, <http://www.neuro-bs.com>).

Stimulation time structure

At the beginning of each trial, a question and two variants of response appeared, as illustrated by Fig. 1. Then, after a random interval in the range of 200–600 ms, a phrase started to sound. As soon as an answer was chosen, the subject was to press either the left or the right response button. The response changed the screen to display a white central fixation cross on the black background (Fig. 1). Then, the next trial started, and the same process continued for all trials. This paradigm was the same for both of the conditions used in the study.

Preliminary testing of the experimental conditions

The above-described timing structure of the present PET experiment was developed in a preliminary behavioral study on a separate group of 12 subjects and aimed to include the maximum number of phrases in a scan. Though the subjects qualified all the phrases as natural sounding, only phrases with the mean error rate of less than 25% were selected for the actual PET study. Eleven segmented phrases were rejected: although these phrases sounded natural, the speaker did not articulate the required prosodic clues distinctly. There was no significant difference in error rate and reaction time between the opposite variants of phrases (e.g., “The teacher said, ‘A student’” and “‘The teacher,’ said a student”).

In order to check for the possible emotional reactions to changes in the complexity of tasks, galvanic skin response (GSR) and heart-beating rate (HBR) were registered during the preliminary study.

Experimental conditions

In the PET study, each of the two experimental conditions was repeated twice in a series of four scans for each subject. The order of conditions was counterbalanced between subjects.

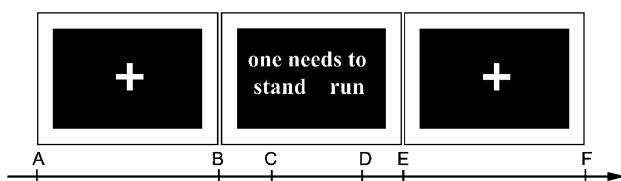


Fig. 1. Timing of the auditory and visual stimulation. Time intervals: AB and EF—a cross on the black background: 300–600 ms; BC—the task presentation before the phrase sounding 200–600 ms; CD—the phrase sounding; DE—choosing a correct answer by a subject.

“Segmented” condition

Phrases with segmentation were presented binaurally. Subjects were to choose the correct meaning of a phrase from the two given on the screen and to press the relevant key with their right hand. For example, the phrase sounded “To run, not to stand”; on the screen was written “One needs to: stand/run” or the phrase was “Peter, said Ivan”; on the screen was written “It was said by: Ivan/Peter.”

“Non-segmented” condition (control to the “Segmented” condition)

Phrases without segmentation were presented binaurally. Subjects were to choose the correct meaning of a phrase from the two given on the screen and to press the relevant key with their right hand. For example, the phrase sounded “Father bought him a coat”; on the screen was written “Father bought him: a coat/a watch.”

All the Russian phrases and questions presented in the study were grammatically and syntactically sound. Prior to PET scanning, each subject was trained to perform the above-described tasks, using an extra set of phrases that were not used in the actual PET experiment.

It was supposed that the pair comparisons of the corresponding test and control conditions would extract brain areas involved in analysis of the prosody-related syntactic structure of the utterances.

The ERP study involved presentation of four phrases in which the placement of the pause and the corresponding comma in writing dramatically changed the meaning:

- To chop not, to saw. (in Russian: “Rubit nelzya, pilit.”)
- To chop, not to saw. (“Rubit, nelzya pilit.”)
- To saw not, to chop. (“Nelzya pilit, rubit.”)
- To saw, not to chop. (“Pilit, nelzya rubit.”).

These four phrases were selected from 60 phrases presented in the PET study. The criterion of selection was that the words contain only stop consonants. It was subjectively confirmed by reaction time and by error rate that all of the presented phrases’ syntactic structures were equally natural in Russian. The sounding of a phrase started at the beginning of each trial, while the central fixation cross on the black background was presented visually. When the sounding of the phrase was over, a question and two variants of response (one on the left side and one on the right side of the screen) appeared for 2 s. A subject was to press either the left or the right response button, according to the chosen variant on the screen. The four phrases were presented at random, 150 times each; the whole ERP study lasted 40 min.

Results

Behavioral results

No difference in error rate was revealed between the “Segmented” and “Non-segmented” conditions ($P > 0.05$, t test), although the reaction time was significantly longer for the condition “Segmented” (Table 1, confidence interval indicated for the means at $P < 0.05$). Neither the GSR nor the HBR

Table 1
The reaction time (RT) and error rate (ER) for the presented types of stimuli

Condition	Mean RT, s ($P < 0.05$)	Error rate, % ($P < 0.05$)
Segmented	3.8 ± 0.1	16 ± 4
Non-segmented	2.8 ± 0.07	15 ± 3

demonstrated a significant difference between the test conditions and the control conditions ($P > 0.05$, t test).

PET results

Table 2 summarizes the obtained results in terms of peak values and anatomic location.

The “Segmented” condition elicited significantly higher activation than the “Non-segmented” condition in the right dorsolateral prefrontal cortex (DLPFC), with its peak approximately at the junction of Brodmann areas (BAs) 44, 45, and 9. Another activation was found in the medial posterior part of the right cerebellum (Fig. 2).

Areas of CBF increase in the “Non-segmented” as compared to the “Segmented” condition were found in the deep posterior Sylvian cortex, bilaterally close to the primary auditory cortex, with extension to BA 40 on the left side (Fig. 3).

ERP results

Significant negativity in the frontal and central electrodes was observed for the word “saw” (“pilit”) when it preceded the intonational pause (i.e., “To saw, not to chop”), as compared to the same word when it ended the sentence (i.e., “To chop not, to saw”). In a comparison of the symmetrical electrodes, this negativity differed for only the inferior frontal electrodes: it was significantly higher in the right inferior frontal (F8) electrode than in the left inferior frontal (F7) electrode ($P < 0.05$). The latency of the negativity was 250–350 ms, corresponding to the offset (at 310 ms) of the word “saw” (“pilit”) (Fig. 4). A comparison of this negativity for the first and the second half of the presented stimuli (75 first and 75 last stimuli) did not reveal a significant difference ($P > 0.05$). Although there was a tendency for a similar negativity in response to the word “chop” (“rubit”) in the position before the pause as compared to the end of the sentence, it did not reach the level of significance ($P > 0.05$). The amplitude of the N100 (measured in the interval of 100–120 ms) and P200 (measured in the interval of 190–210 ms) components was significantly larger

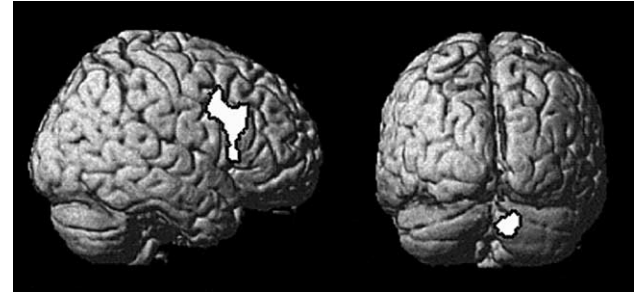


Fig. 2. Activation areas obtained when the segmented phrases were actively analyzed (rCBF increase in the “Segmented” condition vs. the “Non-segmented” condition). The right prefrontal and cerebellar activations are shown as projections to the right and posterior brain surfaces, respectively. For the purpose of better illustration, an uncorrected $P < 0.0001$ voxel-level threshold was used for this figure.

in the central electrodes for the words “saw” (“pilit”) and “chop” (“rubit”) when they preceded the intonational pause, compared to the same words when they ended the sentence ($P < 0.05$).

Discussion

The present study used phrases segmented into syntagmas (groups of words united by a uniform intonational contour). This segmenting was based on two main cues—the pause itself and pause-related pitch changes in the syntagmas. Several sentences of the same kind were presented in the PET experiment, but only one set of four sentences was presented many times in the ERP experiment (due to the timing limitations of the ERP methodology, each of the phrases used for PET could not be repeated 150 times to get ERPs). In this study, phrases with prosody-based syntactic complexity (“Segmented”) elicited more activation in the right dorsolateral frontal cortex and the right cerebellum, as compared with less complicated phrases (“Non-segmented”). In the ERPs, the word immediately preceding the critical pause, which indicated the meaning of the sentence, yielded more ERP negativity with a maximum in the right frontal electrodes, as compared with the same word at the end of the sentence (i.e., when the meaning of the sentence was already clear).

Our ERP results suggest the key role of the syntagma when it precedes the pause in the brain analysis of such segmentation. Since in the ERP design, the same stimuli must be repeated many times, ERP studies of speech perception, of course, do not reflect natural

Table 2
Peak values and anatomic locations of the observed PET activations

Contrast [brain region, Brodmann area (BA)]	P corr.	t value	Coordinates (mm)		
			x	y	z
<i>“Segmented” vs. “Non-segmented”</i>					
R. mid. frontal gyrus/inf. frontal sulcus, BA 9/46/44	0.013	5.24	48	22	28
R. cerebellum	0.037	4.92	10	−84	−36
<i>“Non-segmented” vs. “Segmented”</i>					
L. Sylvian sulcus, BA 42/40/13	0.002	5.79	−50	−22	20
L. Heschl gyrus BA 41/42	0.010	5.30	−50	−22	10
R. Heschl gyrus, BA 41	0.016	5.17	42	−26	8

The corrected voxel-level P values, t values, and the location of the local maxima in terms of the brain structures, Brodmann areas, and stereotactic coordinates are presented. Abbreviations: L = the left hemisphere; R = the right hemisphere; inf. = inferior; mid = middle.

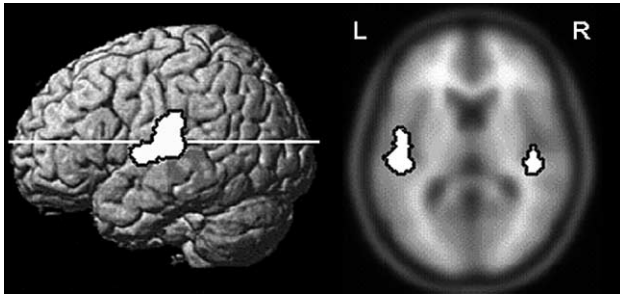


Fig. 3. Areas of rCBF increase in the “Non-segmented” condition as compared to the “Segmented” condition. The white line on the rendered right hemisphere surface depicts the level of the horizontal slice shown on the right side of the figure. The letters L and R indicate the left and right sides of the slice, respectively. As for Fig. 1, the uncorrected voxel $P < 0.0001$ threshold was used here for illustration purposes.

speech processing; they present only a model that emphasizes some aspects of speech processing. In the present study, one phrase was repeated many times, forming the short-term and activating the long-term memory prosody-related traces. This repetition helped decrease the activity of the online brain lexical analysis, revealing or emphasizing the anticipatory function of the brain, based on the remembered prosodic features of the presented phrases. Using prosodic anticipation, a subject who had heard the first syntagma could already predict the second one. This phenomenon can be defined as prosody-based semantic prediction. Our assumptions about the predictive role of prosody in speech perception are based on the recent cognitive model of Hawkins (2003) where the author considers the multi-purpose function of phonetic information and the polysystemic nature of speech within language, so that the meaning of an utterance may be interpreted directly from the global sound pattern. Since prosody in our study is a clue to syntactic

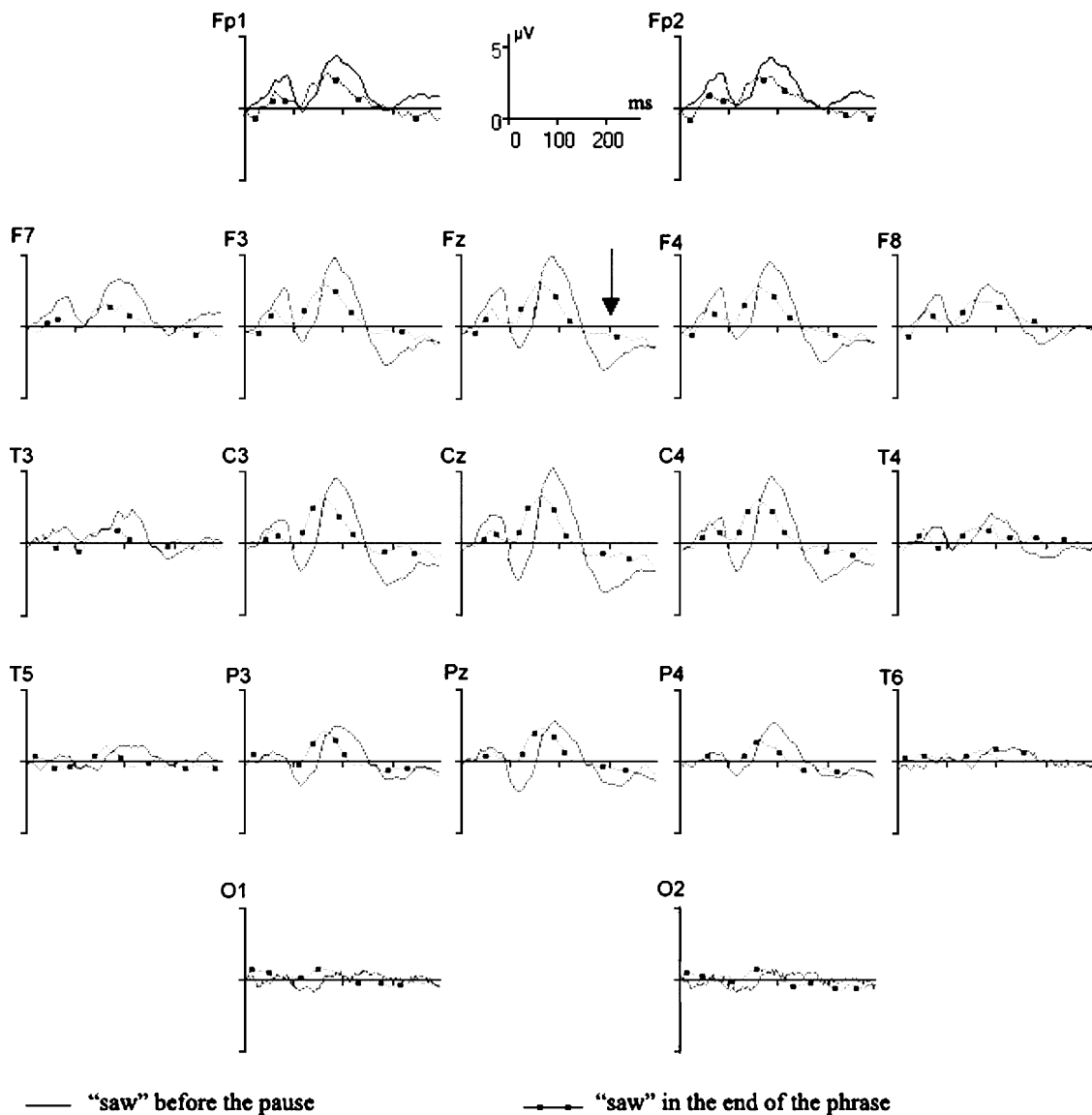


Fig. 4. ERPs to the word “saw” (“pilit”) before the segmenting pause and at the end of the sentence. The initial time point corresponds to the beginning of the word. Negative amplitude is plotted downward. In the position before the pause, the word “saw” (“pilit”) elicited negativity with the latency of 250–350 ms (marked with arrow in Fz).

structuring, we apply here the term “semantic” to syntactic structures, basing it on a number of linguistic and psycholinguistic studies on syntactic semantics (see [Kako and Wagner, 2001](#) for the review). The observed ERP negativity ([Fig. 4](#)) might be an indicator of this prosody-based semantic prediction at the neural level. The negativity before the critical pause may be interpreted as preparatory activity of the brain in anticipation of important prosodic information. This preparatory activity occurs in response to the change of pitch that always precedes a pause in natural speech. The nature of the increased N100 and P200 components for the words before the pause can be the same, but still another interpretation is possible. When the word before the pause is in the middle of the sentence, the word-onset components for it could be higher than for the same word at the end of the sentence. However, this could be part of the same psycholinguistic mechanism because a word at the end of the sentence is not so useful for linguistic anticipation. Though no distinct Closure Positive Shift (CPS) ([Steinhauer, 2003](#)) was observed in this study, the ascending part of the negativity preceding the pause may be considered part of the CPS mechanism. The relationship of the negativity preceding the pause to CPS seems to be an interesting issue for further research.

Since prosody in this study was regarded as a clue to understanding syntactic structure, an increase in prosodic complexity was accompanied by an increase in syntactic complexity. Such close interaction of prosody and syntax in our study was supposed to activate the brain mechanisms of the prosody/syntax interface, which can differ to some extent from the known data on purely syntactic and purely prosodic processing. Thus, the relation of our results to the literature data should be considered in an effort to understand whether there is such a difference for the interface mechanism and to what extent.

The right-sided DLPFC was previously shown to participate in facial expression identification ([Nakamura et al., 1999](#); [Kesler-West et al., 2001](#)), affective prosody perception ([Buchanan et al., 2000](#)), pitch discrimination ([Zatorre et al., 1992](#); [Hsieh et al., 2001](#); [Klein et al., 2001](#)), prosodic modulations perception ([Meyer et al., 2002, 2004](#)), and syntactic analysis ([Just et al., 1996](#); [Embick et al., 2000](#)). Such functional diversity, observed in a single area, seems to be quite natural from the developmental point of view: the child begins to categorize facial emotions very early ([Kotsoni et al., 2001](#)); then, obviously, the child compares these facial expressions with speech sounds, gradually learning to understand some aspects of emotional prosody for what pitch analysis is obligatory. The acquired skills for prosodic clues perception can be used later to judge prosodic clues for sentence parsing ([Shattuck-Hufnagel and Turk, 1996](#)). The development of these abilities takes place in the highly overlapping brain areas. The known principle of correlational learning ([Hebb, 1949](#)) suggests that synchronously activated neurons should link into cell assemblies and brain areas with similar functions are often neighboring or symmetrical. Even while syntactic structures are visually presented, the corresponding intonational patterns are reproduced in inner speech, and the prosodic analysis of such “self-hearing” might contribute to the activation of the same areas (e.g., the right DLPFC, as in our study). We suppose that a substantial functional overlap and only relative specialization (or functionally mosaic structure) of neural populations take place in this multi-functional brain area.

Unfortunately, the poor temporal resolution of experiments with a blocked design (which is still the only one possible for PET) prevents the distinction between activation related to early phases and that related to late phases during a trial. In the present

experiment, the activation differences among conditions could possibly be due to the sentence processing, to the task performance, or both. Considering these methodological limitations, [Plante et al. \(2002\)](#) argued that frontal activations in language tasks are often related to memory and planning demands. Applying this idea to the right DLPFC area activated in our study may lead to a better understanding of its functioning. Identification of the perceived prosody patterns is based on comparing them with the long-term memory prosody patterns for different types of sentences. It might be suggested that, in an active task, attention enhances this process of comparison, increasing the extent of the specific memory usage and, consequently, the level of activation of the observed areas. We believe that prosody/syntax interface functioning is highly related to the interaction of the different memory types, helping to map what is heard to what is known, and can be understood with the use of previous linguistic experience as a basis. This idea comes close to the [Ullman \(2004\)](#) model of complimentary interactions between declarative and procedural memory as underlying language abilities. Prosody processing is not explicitly treated in this model; but, following the general logic of the model, prosody may be mainly subserved by the procedural memory system.

The observed right cerebellar activation might be related to the perception of speech timing. Indeed, the cerebellum is often believed to be a key structure for timing estimation, which is important for many sensory, motor, and cognitive activities ([Ackermann et al., 1999](#); [Ivry and Richardson, 2002](#); [Salman, 2002](#)). It is also possible that the right cerebellar activation is related to estimating phonetic and semantic borders of syntagmas or to keeping the phrase structure in the working memory during processing (see [Marien et al., 2001](#) for a review).

Reverse contrasts are often difficult to interpret. However, an increase of CBF in the left perisylvian cortex in the “Non-segmented” control condition, compared with the “Segmented” condition in our study, might be explained by the prefrontal modulation (suppression) of the sensory cortex activation, shown in some PET and brain lesion studies ([Frith et al., 1991](#); [Knight et al., 1999](#)).

The possible differences in difficulty level and emotional state between the test and control conditions were addressed by the analysis of the behavioral (RT) and physiological (GSR and HBR) parameters, registered during the preliminary session. Although a significant difference in RT existed between the “Segmented” condition and the “Non-segmented” condition, this difference was unlikely to be an objective criterion of the non-specific task difficulty and might easily have been caused by additional cognitive components, which involve more neural circuits and thus demand more time for their processing. The error rate seems to be a more reliable indicator of the non-specific difficulty in neurocomputational systems ([Basheer and Hajmeer, 2000](#)). Since there was no significant difference in the error rate between the “Segmented” and “Non-segmented” conditions, the PET data were unlikely to be influenced by the change in task difficulty. The preliminary study also showed no significant inter-conditional differences in emotional state, as indicated by the GSR and the HBR recordings.

Conclusions

The right posterior prefrontal cortex and the right medial posterior cerebellar area participate in the brain network of spoken speech syntactic parsing, being involved in the prosody/syntax

interface. The acquired ERP data support the idea that prosody-based semantic prediction is important for such processing. Furthermore, comparing our results with other brain mapping studies, we conclude that the right posterior prefrontal cortex might represent the functional overlap of brain networks of emotions, prosody, and syntax perception.

Acknowledgments

The authors are grateful to Professor N.D. Svetozarova for her expertise in linguistic issues. The authors also thank the staff of the PET and radiochemistry laboratories of IHB RAS for their help with the PET procedure. The research was supported by Grants No. 03-06-80068 from the Russian Foundation for Basic Research (RFBR) and Nos. 04-04-00083a and 03-06-00206 from the Russian Foundation for the Humanities (RFH).

References

- Ackermann, H., Graber, S., Hertrich, I., Daum, I., 1999. Cerebellar contributions to the perception of temporal cues within the speech and nonspeech domain. *Brain Lang.* 67, 228–241.
- Basheer, I.A., Hajmeer, M., 2000. Artificial neural networks: fundamentals, computing, design, and application. *J. Microbiol. Methods* 43, 3–31.
- Baum, S.R., Pell, M.D., 1999. The neural bases of prosody: insights from lesion studies and neuroimaging. *Aphasiology* 13, 581–608.
- Berndt, R.S., Mitchum, C.C., Haendiges, A.N., 1996. Comprehension of reversible sentences in “agrammatism”: a meta-analysis. *Cognition* 58, 289–308.
- Buchanan, T.W., Lutz, K., Mirzazade, S., Specht, K., Shah, N.J., Zilles, K., Jancke, L., 2000. Recognition of emotional prosody and verbal components of spoken language: an fMRI study. *Brain Res. Cogn. Brain Res.* 9, 227–238.
- Caplan, D., Hildebrandt, N., Makris, N., 1996. Location of lesions in stroke patients with deficits in syntactic processing in sentence comprehension. *Brain* 119, 933–949.
- Caplan, D., Alpert, N., Waters, G., 1998. Effects of syntactic structure and propositional number on patterns of regional cerebral blood flow. *J. Cogn. Neurosci.* 10, 541–552.
- Caplan, D., Alpert, N., Waters, G., 1999. PET studies of syntactic processing with auditory sentence presentation. *NeuroImage* 9, 343–351.
- Chernigovskaya, T.V., Deglin, V.L., 1986. Brain functional asymmetry and neural organization of linguistic competence. *Brain Lang.* 29, 141–153.
- Embick, D., Marantz, A., Miyashita, Y., O’Neil, W., Sakai, K.L., 2000. A syntactic specialization for Broca’s area. *Proc. Natl. Acad. Sci. U. S. A.* 97, 6150–6154.
- Friederici, A.D., 2002. Towards a neural basis of auditory sentence processing. *Trends Cogn. Sci.* 6, 78–84.
- Friederici, A.D., Meyer, M., von Cramon, D.Y., 2000. Auditory language comprehension: an event-related fMRI study on the processing of syntactic and lexical information. *Brain Lang.* 74, 289–300.
- Friederici, A.D., Ruschmeyer, S.A., Hahne, A., Fiebach, C.J., 2003. The role of left inferior frontal and superior temporal cortex in sentence comprehension: localizing syntactic and semantic processes. *Cereb. Cortex* 13, 170–177.
- Friston, K.J., Worsley, K.J., Frackowiak, R.S.J., Mazziotta, J.C., Evans, A.C., 1994. Assessing the significance of focal activations using their spatial extent. *Hum. Brain Mapp.* 1, 214–220.
- Frith, C.D., Friston, K.J., Liddle, P.F., Frackowiak, R.S., 1991. A PET study of word finding. *Neuropsychologia* 29, 1137–1148.
- Grodzinsky, Y., 1990. *Theoretical Perspectives on Language Deficits*. MIT Press, Cambridge.
- Grodzinsky, Y., 1995. A restrictive theory of agrammatic comprehension. *Brain Lang.* 50, 27–51.
- Hawkins, S., 2003. Roles and representations of systematic fine phonetic detail in speech understanding. *J. Phon.* 31, 373–405.
- Hebb, D.O., 1949. *The Organization of Behavior*. A Neuropsychological Theory. Wiley, New York.
- Holte, S., Eriksson, L., Dahlbom, M., 1989. A preliminary evaluation of the Scanditronix PC2048-15B brain scanner. *Eur. J. Nucl. Med.* 15, 719–721.
- Homae, F., Hashimoto, R., Nakajima, K., Miyashita, Y., Sakai, K.L., 2002. From perception to sentence comprehension: the convergence of auditory and visual information of language in the left inferior frontal cortex. *NeuroImage* 16, 883–900.
- Hsieh, L., Gandour, J., Wong, D., Hutchins, G.D., 2001. Functional heterogeneity of inferior frontal gyrus is shaped by linguistic experience. *Brain Lang.* 76, 227–252.
- Humphries, C., Willard, K., Buchsbaum, B., Hickok, G., 2001. Role of anterior temporal cortex in auditory sentence comprehension: an fMRI study. *NeuroReport* 12, 1749–1752.
- Humphries, C., Love, T., Swinney, D., Hickok, G., in press. Response of anterior temporal cortex to syntactic and prosodic manipulations during sentence processing. *Hum. Brain Mapp.*
- Indefrey, P., Hagoort, P., Herzog, H., Seitz, R.J., Brown, C.M., 2001. Syntactic processing in left prefrontal cortex is independent of lexical meaning. *NeuroImage* 14, 546–555.
- Ivry, R.B., Richardson, T.C., 2002. Temporal control and coordination: the multiple timer model. *Brain Cogn.* 48, 117–132.
- Just, M.A., Carpenter, P.A., Keller, T.A., Eddy, W.F., Thulborn, K.R., 1996. Brain activation modulated by sentence comprehension. *Science* 274, 114–116.
- Kako, E., Wagner, L., 2001. The semantics of syntactic structures. *Trends Cogn. Sci.* 5, 102–108.
- Kesler-West, M.L., Andersen, A.H., Smith, C.D., Avison, M.J., Davis, C.E., Kryscio, R.J., Blonder, L.X., 2001. Cerebral substrates of facial emotion processing using fMRI. *Cogn. Brain Res.* 11, 213–226.
- Klein, D., Zatorre, R.J., Milner, B., Zhao, V., 2001. A cross-linguistic PET study of tone perception in mandarin Chinese and English speakers. *NeuroImage* 13, 646–653.
- Kluender, R., Kutas, M., 1993. Bridging the gap: evidence from ERPs on the processing of unbounded dependencies. *J. Cogn. Neurosci.* 5, 196–214.
- Knight, R.T., Staines, W.R., Swick, D., Chao, L.L., 1999. Prefrontal cortex regulates inhibition and excitation in distributed neural networks. *Acta Psychol. (Amst.)* 101, 159–178.
- Kotsoni, E., de Haan, M., Johnson, M.H., 2001. Categorical perception of facial expressions by 7-month-old infants. *Perception* 30, 1115–1125.
- Kotz, S.A., Meyer, M., Alter, K., Besson, M., von Cramon, D.Y., Friederici, A.D., 2003. On the lateralization of emotional prosody: an event-related functional MR investigation. *Brain Lang.* 86, 366–376.
- Marien, P., Engelborghs, S., Fabbro, F., De Deyn, P.P., 2001. The lateralized linguistic cerebellum: a review and a new hypothesis. *Brain Lang.* 79, 580–600.
- Mazoyer, B.M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrrier, O., 1993. The cortical representation of speech. *J. Cogn. Neurosci.* 5, 467–479.
- Meyer, M., Alter, K., Friederici, A.D., Lohmann, G., von Cramon, D.Y., 2002. fMRI reveals brain regions mediating slow prosodic modulations in spoken sentences. *Hum. Brain Mapp.* 17, 73–88.
- Meyer, M., Steinhauer, K., Alter, K., Friederici, A.D., von Cramon, D.Y., 2004. Brain activity varies with modulation of dynamic pitch variance in sentence melody. *Brain Lang.* 89, 277–289.
- Nakamura, K., Kawashima, R., Ito, K., Sugiura, M., Kato, T., Nakamura, A., Hatano, K., Nagumo, S., Kubota, K., Fukuda, H., Kojima, S., 1999. Activation of the right inferior frontal cortex during assessment of facial emotion. *J. Neurophysiol.* 82, 1610–1614.
- Neville, H.J., Nicol, J.L., Barss, A., Forster, K.I., Garrett, M.F., 1991. Syntactically based sentence processing classes: evidence from event-related brain potentials. *J. Cogn. Neurosci.* 3, 151–165.
- Plante, E., Creusere, M., Sabin, C., 2002. Dissociating sentential prosody

- from sentence processing: activation interacts with task demands. *NeuroImage* 17, 401–410.
- Röder, B., Stock, O., Neville, H.J., 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.
- Salman, M.S., 2002. The cerebellum: it's about time! But timing is not everything—New insights into the role of the cerebellum in timing motor and cognitive tasks. *J. Child Neurol.* 17, 1–9.
- Shattuck-Hufnagel, S., Turk, A.E., 1996. A prosody tutorial for investigators of auditory sentence processing. *J. Psycholinguist. Res.* 25, 193–247.
- Steinhauer, K., 2003. Electrophysiological correlates of prosody and punctuation. *Brain Lang.* 86, 142–164.
- Steinhauer, K., Alter, K., Friederici, A.D., 1999. Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nat. Neurosci.* 2, 191–196.
- Stromswold, K., Caplan, D., Alpert, N., Rauch, S., 1996. Localization of syntactic comprehension by positron emission tomography. *Brain Lang.* 52, 452–473.
- Swinney, D., Zurif, E., 1995. Syntactic processing in aphasia. *Brain Lang.* 50, 225–239.
- Swinney, D., Zurif, E., Prather, P., Love, T., 1996. Neurological distribution of processing resources underlying language comprehension. *J. Cogn. Neurosci.* 8, 174–184.
- Talairach, P., Tournoux, J., 1988. *A Stereotactic Coplanar Atlas of the Human Brain*. Thieme, Stuttgart.
- Ullman, M.T., 2004. Contributions of memory circuits to language: the declarative/procedural model. *Cognition* 92, 231–270.
- Worsley, K.J., Marrett, S., Neelin, P., Vandal, A.C., Friston, K.J., Evans, A.C., 1996. A unified statistical approach for determining significant signals in images of cerebral activation. *Hum. Brain Mapp.* 4, 58–73.
- Zatorre, R.J., Evans, A.C., Meyer, E., Gjedde, A., 1992. Lateralization of phonetic and pitch discrimination in speech processing. *Science* 256, 846–849.