Available online at www.sciencedirect.com
SCIENCE DIRECT®
Cognitive Brain Research 18 (2004) 130–141
COGNITIVE BRAIN RESEARCH
www.elsevier.com/locate/cogbrainres

Research report

P600 related to rule violation in an arithmetic task

M. Isabel Núñez-Peña*, M. Luisa Honrubia-Serrano

Department of Behavioral Sciences Methods, University of Barcelona Passeig Vall d'Hebron, 171 08035 Barcelona, Spain

Accepted 30 September 2003

Abstract

The goal of the present paper was to study if a similar neurophysiological process is required for treating violations of both arithmetical rules and linguistic syntactic structures. It has been shown that syntactic violations elicit the P600/syntactic positive shift (SPS) component, reflecting secondary parsing processes or repairing of an incorrect syntactic structure. However, late positivities, similar to the P600/SPS component, are also elicited by other types of violations (e.g. harmonic anomalies or violations in non-linguistic abstract rules), so this component is thought to be an index of detection for any anomaly in rule-governed sequences. We carried out an experiment where violations of arithmetic rules were presented. These violations were evident to a greater or lesser degree (a number very different or very similar to that which correctly completed a series of seven numbers). The type of rule was also manipulated, and increasing and decreasing series were presented. Results showed a late centro-parietal positivity related to arithmetic violations, whose amplitude was larger, the more evident the violation presented was, in both addition and subtraction. It is concluded that a similar neurophysiological process could be required for the processing of violations in numerical sequences and in linguistic syntactic structures. When the rule was broken, another component was present for the adding operation: an early negativity peaking between 250- and 300-ms post-stimulus. Regarding this negative peak, although some possible explanations are drawn, further research needs to be carried out in order to gather more knowledge about it.

Theme: Neural basis of behavior
Topic: Cognition

Keywords: P600; SPS; P3b; Syntax; Rule-governed anomalies; Arithmetic

1. Introduction

In recent years, many researchers have studied the linguistic specificity of some components of event related potentials (ERPs) (for a review, see Refs. [14,55,56]). This interest is due to the fact that these linguistic components can be used to gather more knowledge about language processing. Psychophysiological research has so far reported a component that seems to be specific for semantic processing[1], and other components related to syntactic processing. One of these syntactic components is the core of this paper.

Three ERP components have been associated with syntactic processing: the early left anterior negativity (ELAN), the left anterior negativity (LAN), and the syntactic positive shift (SPS)—this latter component was called P600 by Osterhout and Holcomb [46] and SPS by Hagoort et al. [21]. ELAN, LAN and P600/SPS components are elicited by a variety of syntactic and morphosyntactic violations. ELAN and LAN are two negative components that present similar scalp distribution but which differ in latency. ELAN is a negativity peaking approximately 150–200 ms after stimulus, which has its maximum amplitude in the left anterior region of the scalp, and that has been reported in phrase structure violations in both visual and auditory modalities [16,42]. LAN is also a negative peak with a left anterior distribution, but differs from ELAN in latency—it presents its maximum amplitude at about 400-ms post-stimulus. This component has been reported in inflectional verb violation [48,54]. The functional nature of these components is still

* Corresponding author. Tel.: +34-93-312-58-53; fax: +34-93-402-13-59.

E-mail address: inunez@ub.edu (M.I. Núñez-Peña).

[1] The N400 component [34] is a negative peak that shows maximum amplitude around 400 ms and a centro-parietal scalp distribution. It is a default response to every word and its amplitude is inversely proportional to word's expectancy. N400 has been said to reflect the activation of the semantic representations of a word (for a review of this component see [5,33,35,36,47]).

0926-6410/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/j.cogbrainres.2003.09.010

unclear, although they are believed to reflect the first-pass parsing stage, which is said to operate quite quickly and automatically—these components have been found to be independent of attentional factors [15,17,19,20,22]. Gunter et al. [20] refer to LAN as a "more primary/automatic reflection of syntactic violation than the P600" ([20], p. 672). The LAN component has also been associated with aspects of working memory [30,32] and with the processing of word class information [6,57].

The most extensively studied component related to syntactic processing is P600/SPS, a late positive component, peaking around 600 ms after stimulus onset, which has a centro-parietal distribution. P600/SPS is elicited by syntactic violations, and is considered to be a reflection of secondary parsing processes including re-analysis or repair of an incorrect structure [17,19,22]. Moreover, P600/SPS has a more controlled nature than either LAN or ELAN. Gunter and Friederici [19] found that P600 amplitude was reduced when they presented verb inflection violations to subjects who were required to judge if a word in a sentence was printed in upper case.

Another characteristic of the P600/SPS component is that its amplitude is also sensitive to the severity of syntactic anomalies. Osterhout et al. [49] presented syntactically correct sentences, violations of verb subcategorization and subcategorization biases. Whereas the first violations produce an irreversible ungrammaticality (thus making the sentence uninterpretable), violations of subcategorization biases merely require a structural re-analysis of the sentence (in this case, the subject is forced to re-analyze the sentence to turn to the less-preferred syntactic analysis for the sentence). In other words, sentences were manipulated in such a way that they were easy, difficult or impossible to integrate within the prior context. P600/SPS amplitude was modulated by integration difficulty, it being largest for the strongest violation. Osterhout et al. ([49], p. 798) concluded that the larger the P600/SPS, the more difficult it is to construct a grammatical representation. Two specific explanations were given for this finding. First, that the "P600 amplitude is sensitive to the syntactic fit between a sentence constituent and preceding sentence structure in a manner analogous to the way that N400 amplitude is a function of the semantic fit between a word and preceding context". Second, that "the amplitude variation in the P600 reintroduces the notion of cost of reprocessing" ([49] p. 798). The first explanation accounts for the P600/SPS amplitude as a function of the syntactic expectations, whereas the second explanation claims that P600/SPS amplitude depends on the ease with which an alternative analysis can be constructed.

Although P600/SPS could be considered to be a specific linguistic component, some authors question this. Firstly, it has been said that P600/SPS could be a P3b component. P3b is a centro-parietal component which is thought to reflect the resolution of prior uncertainty and the task-relevant surprise value of the stimulus: its amplitude is proportional to the probability of the stimulus and its latency varies with the difficulty of the task [52,53]. Some years ago, there was an interesting debate about P600/SPS-P3b similarities between Coulson et al. [9,10] and Osterhout and Haggort [45]. The question as to whether P600 is a specific syntactic index or a reflection of a more general mechanism is still unresolved.

A second group of authors argues against the language specificity of P600/SPS and claims that this component, rather than being syntactic, is related to any anomaly in rule-governed sequences. Patel, Gibson, Ratner, Besson and Holcomb [51] compared ERP patterns elicited by syntactic anomalies in sentences and harmonic anomalies in music. A positive peak similar in latency, polarity, amplitude and scalp distribution was found whenever an anomaly, syntactic or harmonic, was presented (late positivities to deviant notes in musical melodies have also been reported by Besson and Faïta [3] and Janata [25]). Patel et al. [51] also manipulated the integration difficulty in music and language and found similar results to those reported by Osterhout et al. [49]: the amplitude of the positive peak increased when the stimulus was more difficult to integrate. This amplitude variation was present in both language and music. Consequently, Patel et al. ([51], p. 726) stated that "the positivities to structurally incongruous elements in language and music do not appear to be distinguishable". Similarly, a late positivity comparable to the P600/SPS component has been reported when presenting orthographic anomalies (mis-spelled words) [41] or violations in non-linguistic abstract sequences [4,39,40]. These results once again question the syntactic specificity of the P600/SPS component.

Studies with arithmetical tasks have also reported late positive components (LPC) when an arithmetic rule is broken [43,44]. Niedeggen and Rösler [43] used an arithmetic task consisting of multiplication problems ($a \times b = c$). The solution to the product—number c—was manipulated by presenting the correct number or an incorrect number that differed either by a small, medium, or large numerical distance from the correct product. An LPC was elicited whenever an incorrect number was presented. This component peaked around 600-ms post-stimulus and had a posterior distribution; moreover, the amplitude of this component increased monotonically with increasing numerical distance from the correct solution. Nieddegen and Rösler [43] stated that "this pattern suggests that the LPC amplitude is a function of the implausibility of a presented solution, a possibility that fits with the well-established interpretation that the LPC amplitude is always inversely proportional to the subjective probability of its evoking event" ([43], p. 274). Although these authors related this LPC to the P3b, it is also similar to the component reported by Osterhout et al. [49] in syntactic violations, and that reported by Patel et al. [51] in musical violations. All three components match in terms of polarity, topography, latency, and sensitivity of the same experimental manipulation. As for this last point, reaction time studies have described the *split effect*: the more distant the proposed results from the correct result in an arithmetic operation, the faster subjects classify the result as incorrect [1,2]. The split effect is thought to reflect the comparison of the result computed by the subject with that proposed by the experimenter—in other words, it reflects the calculation verification. In this sense, split effect reflects a process similar to the syntactic verification reported in sentences [49,51] or the harmonic verification reported in music [51]. Recently, El Yagoubi, Lemaire and Besson [13] have reported differences in the ERP pattern when subjects have to verify small and large-split problems. It has been found that small-split problems are associated with less positivity than large-split problems.

Although the arithmetic LPC reported by Niedeggen and Rösler could be considered to be a general index of anomaly detection in rule-governed sequences, the number processing model proposed by Dehaene and Cohen [11] enables us to relate the arithmetic LPC to linguistic processing. Dehaene and Cohen's triple-code model postulates that arithmetic facts, such as multiplication and addition tables, are stored in a verbal word frame "in which numbers are represented as syntactically organized sequences of words" ([11], p. 85). This model also proposes that complex operations with numbers, such as subtraction problems or number comparison—which are not learnt by rote verbal learning—need to pass through the magnitude representation of the quantity in order to be solved. This magnitude representation is thought to be stored in a preverbal system of arithmetic reasoning. Single-case studies performed with brain-damaged patients supported the triple-code model (see, for example, Ref. [8]): "Some patients with left parietal lesions exhibit a loss of the sense of numerical quantity with a relative preservation of rote language-based arithmetic such as multiplication tables. Conversely, aphasia following left-hemispheric brain damage can be associated with a selective impairment of rote arithmetic and a preserved sense of quantity" ([12], p. 973). According to the Dehaene and Cohen model, the LPC reported by Niedeggen and Rösler [43] could be explained in terms of the linguistic P600/SPS component, because arithmetical fact retrieval depends on the general language processing system.

1.1. The present study

The aim of the present study was to find out if P600/SPS is a language-specific component or if it could be elicited by other violations of rule-governed sequences. We worked with series of seven numbers where a simple arithmetic rule could be easily discovered. The last number of the series could end it in three different ways: the correct number, an incorrect number very similar to the correct one, or an incorrect number very different from the correct one. Thus, we wanted to manipulate the integration difficulty of the number with the previous series. The arithmetic rule was also manipulated and consisted of adding or subtracting a constant number. Addition and subtraction were selected because evidence about some processing dissociation for these two arithmetical rules could be investigated. Moreover, neither processing increasing series nor processing decreasing series seem to rely on memorized syntactically organized sentences, according to Dehaene and Cohen's model [11]. As it has been mentioned before, this model argues that only addition and multiplication tables are stored as syntactically organized sequences of words. Therefore, increasing and decreasing series of seven numbers can be considered non-linguistic sequences governed by arithmetical rules.

According to the previous reasoning, if a violation of an arithmetical rule elicits a positivity peaking around 600 ms, with maximum amplitude at centro-parietal regions and whose amplitude varies depending on the difficulty of integration of the number in the previous series, it can be concluded that this positivity is functionally similar to the P600/SPS component. Consequently, similar neurophysiological processes could be required for the processing of numerical sequences and linguistic syntactic structures.

2. Materials and methods

2.1. Participants

Fourteen volunteers (11 women; age 20–33 years, mean = 22.1, standard deviation = 3.4; 12 right-handed) took part in the experiment. All participants were university students and had normal or corrected-to-normal visual acuity. Written informed consent was obtained from all participants prior to the start of the experiment.

2.2. Stimulus materials

The stimuli were series of seven Arabic numbers: the first six digits constituted the series and the seventh was the ending. Series were constructed in the following way: the numbers 1 to 16 were selected to begin each series, and the same constant quantity was consecutively added or subtracted[2]. Thus, 48 increasing and 48 decreasing series were generated. The series were followed by a seventh number that could be: (a) a number that continued the series in the correct way—from now on, this will be referred to as the *correct ending*; (b) an incorrect number very different from the correct one—from now on, this will be referred to as *clear incorrect ending*; and (c) an incorrect number very similar to the correct one[3]—from now on, this will be referred to as *confusing incorrect ending*.

[2] We used the following expression to generate the series: $x_{i+1} = x_i + c$ where c is a constant that could take the values -2, -3, -4, 2, 3 or 4.

[3] In order to make the second type of stimuli—clear incorrect endings—we added 25 points to the correct number; as for the third type of stimuli—confusing incorrect ending—we added or subtracted 1 point from the correct end. For example, in the ascending sequence *7, 10, 13, 16, 19, 22*, the correct ending was *25*, the clear incorrect ending was *50*, and the confusing incorrect ending was *24*. In the descending sequence *25, 22, 19, 16, 13, 10*, the correct ending was *7*, the clear incorrect ending was *32*, and the confusing incorrect ending was *8*.

The experiment was controlled by the STIM 2.0 program (NeuroScan, Herndon, VA). Numbers were shown in white color on a black background, and subtended a visual angle of 1.63° vertically and 0.8° (for one-digit stimuli) or 1.63° (for two-digit stimuli) horizontally.

2.3. Procedure

Participants came to the laboratory on two different days in order to avoid tiredness. Each session lasted about 2.5 h, with an inter-session period of 30 days. The main difference between the two sessions concerned the instructions given to the subjects: one day they were instructed to answer only if the ending was correct, whereas the other day they were told to answer only to incorrect endings. This variable was manipulated in order to control response-preparation potentials, so as to avoid overlapping of components (therefore, we asked subjects to give a motor answer to some trials but not to others). The order of the sessions was counterbalanced and each order was randomly assigned to every subject (half of the subjects started with the response to correct trials session, and the rest started with the response to incorrect trials session).

Participants sat in a comfortable armchair during the ERP recording. The session began with a training period consisting of the presentation of some trials similar to those used in the recording period. Because we wanted to be sure that subjects clearly understood the task to be performed, the training period finished when one of the following learning criteria was reached: (a) the participant had correctly answered the first 10 consecutive trials; or (b) the subject correctly answered 90% of trials. Feedback was given whenever a participant's response was wrong. There was a 30-s rest after every 20 trials.

When the training period was over, the recording period started. Subjects were instructed to be relaxed and to keep their eyes looking at the screen. They were also told that it was important to avoid blinking, and in case they needed to blink, they should try to wait until an asterisk or a rest message appeared on the screen. Each trial consisted of a series of seven Arabic numbers that were presented sequentially on the screen (Fig. 1 shows a trial scheme). Numbers remained on the screen for 1000 ms and the ISI was 1500 ms. After the ending, which was the last number of the series, an asterisk was shown for 500 ms. This asterisk had the function of giving the participant the following information: (1) the series had finished; (2) blinking was allowed; and (3) a motor response was required. As for this third point, the response required of the subject depended on the session: in one session, participants only had to respond to correct trials, whereas in the other session, they had to respond to incorrect trials. During the recording session, in contrast with the training session, no feedback over an incorrect response was given. The inter-trial interval was 1500 ms.

In each session, 12 blocks of 24 trials were presented to every participant. A message indicating a 30-s rest appeared on the screen after 12 trials, and there was a 5-min rest in the middle of the recording session. The type of trials and their order of appearance were controlled within each block. A block included four trials of each type (correct, clear incorrect, and confusing incorrect) randomly presented, but with the following restrictions: (a) no more than three correct trials or three incorrect trials could be consecutively presented; and (b) no more than three increasing series or three decreasing series could appear consecutively. All participants were tested on 576 trials, 48 for each experimental condition.

2.4. Electrophysiological recording

Recording and analysis of ERPs were performed with the SCAN 3.0 software (NeuroScan). EEG was recorded from seven tin electrodes mounted on a commercial electro-cap (Electro-Cap International, Eaton, OH) and positioned according to the 10–20 International System: C3, C4, P3, P4, Cz, Fz, and Pz. Connected ear lobes served as reference, and an equidistant point between Fpz and Fz was used as the location of the ground electrode. For monitoring eye movement and blinks, we used four Ag/AgCl electrodes: two for the VEOG recording, placed above and below the right eye, and two for the HEOG recording, placed at the two external canthi.

EEG and EOG channels were continuously digitized at a rate of 250 Hz by a SynAmp™ amplifier (5083 model, NeuroScan). A band-pass filter was set at 0.05 to 30 Hz, and a notch filter of 50 Hz was used. Electrode impedance was always kept below 5 kΩ.

2.5. Data analysis

Analysis of the electrophysiological response was carried out on the seventh number of the series. Firstly, epochs for

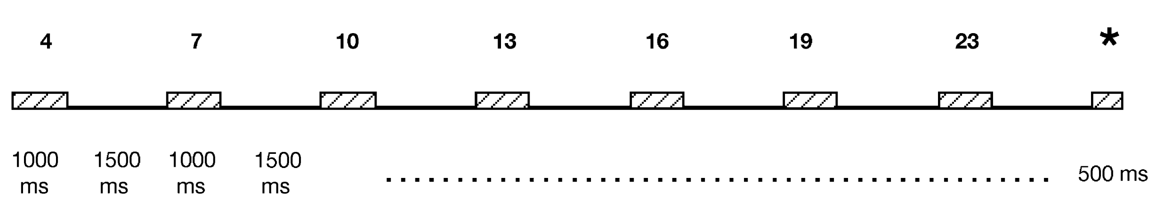

Fig. 1. Schematic representation of the structure of a trial.

every subject in each experimental condition were averaged, relative to a pre-stimulus baseline that was made up of the 100 ms of activity preceding the epoch of interest. Trials with artifacts (voltage exceeded ± 50 μV in any channel) and those with response errors were excluded from the ERP average. Secondly, ERP data were analyzed by computing the mean amplitude in the 0–200-ms window, in 50-ms windows from 200- to 500-ms post-stimuli, and in 100-ms windows from 500- to 800-ms post-stimuli. We analyzed 50-ms windows between 200 and 500 ms to identify early negativities, and 100-ms windows between 500 and 800 ms to look for late positivities. The purpose of working with these windows was to carry out a detailed analysis of the waves: we wanted to identify the temporal interval where the differences between experimental conditions began, where these differences finished, and where these differences reached their maximum amplitude.

A $2 \times 2 \times 3 \times 7$ repeated-measures analysis of variance (ANOVA) was performed, taking as factors *series type* (increasing or decreasing), *response type* (with or without motor response), *ending type* (correct, confusing incorrect and clear incorrect), and *location* (C3, C4, P3, P4, Fz, Cz, and Pz). A complementary ANOVA was performed to examine possible laterality effects at centro-parietal loca-

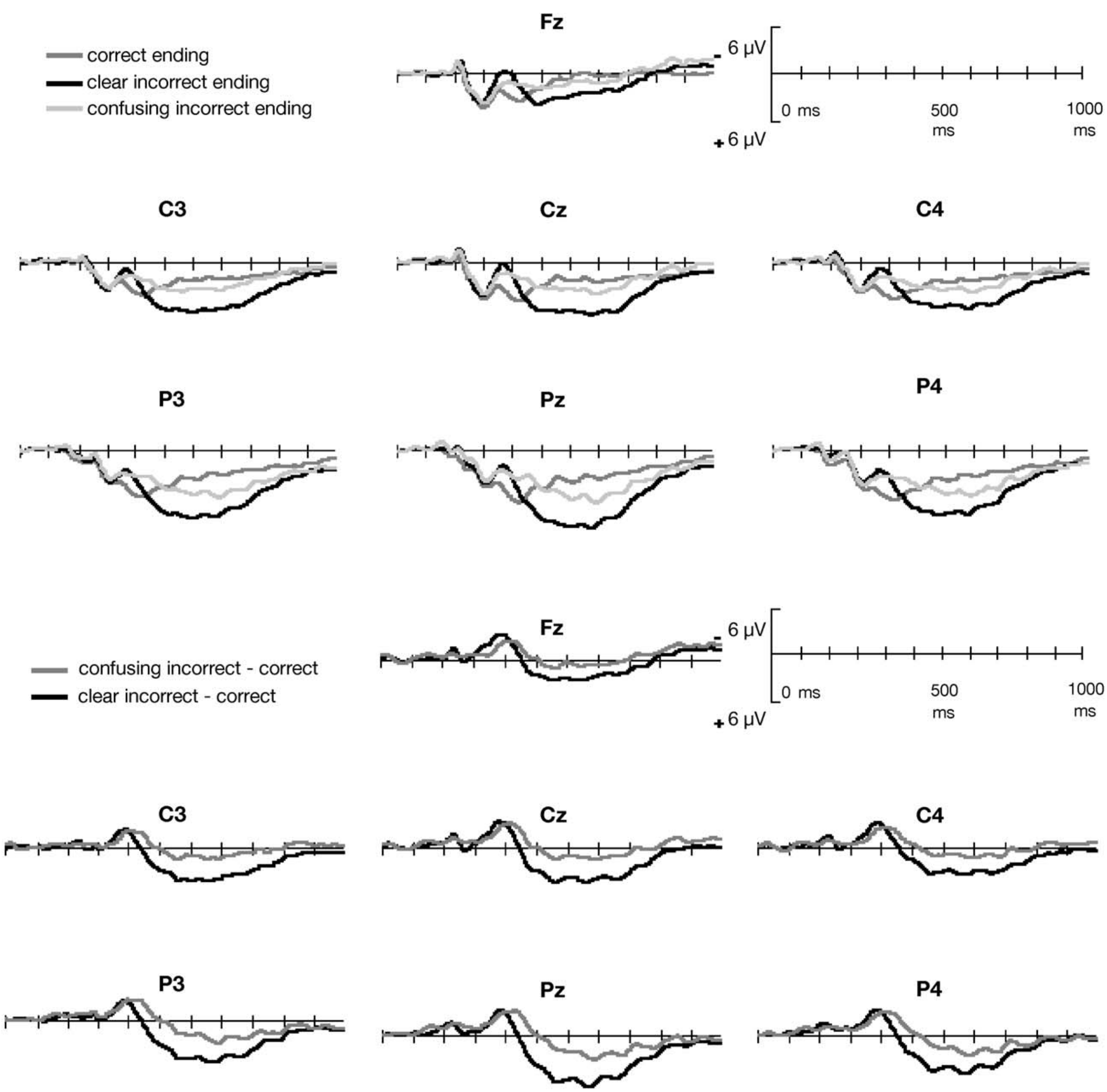

Fig. 2. Increasing series. Top: Grand-averages ($n = 14$) for the three types of ending. Bottom: Difference waves for increasing series: they were obtained after subtracting correct ending ERPs from clear and confusing incorrect ending ERPs.

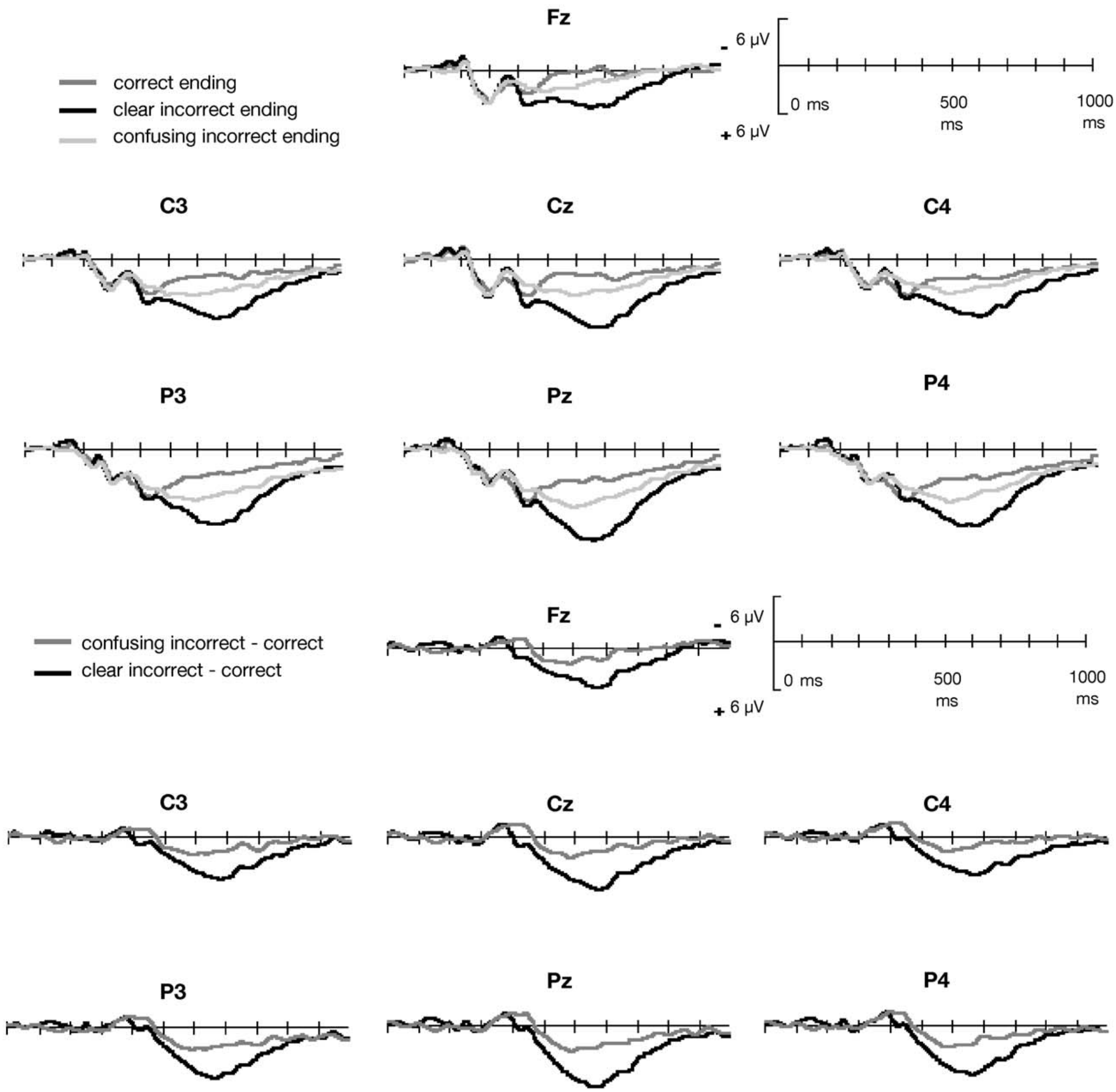

Fig. 3. Decreasing series. Top: Grand-averages ($n = 14$) for the three types of ending. Bottom: Difference waves for decreasing series: they were obtained after subtracting correct ending ERPs from clear and confusing incorrect ending ERPs.

tions by including *hemisphere* (left and right) as factor. Repeated-measures ANOVAs were carried out with the Greenhouse–Geisser correction [18] for sphericity departures, which was applied when appropriate: we reported the F value, the uncorrected degrees of freedom, probability level following correction, and the η^2 effect size index[4] [7,23,31]. Whenever a main effect reached significance, we calculated contrasts using a non-pooled error term [28,29]; in order to analyze significant interactions, simple effects were calculated. In both cases, the Hochberg approach was used to control for the increase in type I error [27].

3. Results

Figs. 2 and 3 show the grand average ERPs for the different final numbers in increasing and decreasing series, respectively. As for the increasing series—those consisting of adding a number—the presence of an early negativity, peaking between 250 and 300 ms, should be noted, followed by a positive wave (peaking around 600 ms) of large amplitude when a clear incorrect ending was presented. A similar pattern can be seen in confusing

[4] η^2 is a statistical index of the strength of the relationship between variables. It provides information about the proportion of variability in the observed data that is accounted for by treatment.

Table 1
ANOVA for the main effect of ending type, and the interactions ending type × series type, and ending type × electrode within each window

Time (ms)	Ending type			Ending × Series			Ending × Electrode		
	F	p	η^2	F	p	η^2	F	p	η^2
0–200	0.378[a]	0.689	0.028	1.920[a]	0.167	0.129	1.290[a]	0.229	0.090
200–250	0.671[b]	0.477	0.049	2.116[a]	0.141	0.140	1.528[a]	0.205	0.105
250–300	9.579[a]	0.001***	0.424	4.576[a]	0.020*	0.260	2.434[b]	0.062	0.158
300–350	10.478[a]	0.000***	0.446	6.555[b]	0.013*	0.335	2.634[b]	0.045*	0.168
350–400	13.260[a]	0.000***	0.505	3.736[b]	0.056	0.223	2.489[b]	0.061	0.161
400–450	29.483[a]	0.000***	0.694	3.316[a]	0.052	0.203	4.133[b]	0.007**	0.241
450–500	42.249[a]	0.000***	0.765	1.893[a]	0.171	0.127	7.386[b]	0.000***	0.362
500–600	43.245[a]	0.000***	0.769	3.030[a]	0.066	0.189	9.984[b]	0.000***	0.434
600–700	27.847[a]	0.000***	0.682	1.407[a]	0.263	0.098	7.509[b]	0.000***	0.366
700–800	8.796[a]	0.001***	0.404	0.412[a]	0.667	0.031	3.422[b]	0.026*	0.208

[a] Sphericity assumed.
[b] Greenhouse–Geisser correction.
* $0.05 \geq p > 0.01$.
** $0.01 \geq p > 0.001$.
*** $p \leq 0.001$.

incorrect endings, but, in this case, the late positivity was smaller. Finally, correct endings elicited a 250-ms negative peak followed by a quick positive wave (peaking around 300 ms), both showing less amplitude than those waves elicited by incorrect endings.

With regard to decreasing series—those consisting of subtracting a number—the pattern shown by the three types of ending was similar to those of increasing series. However, as can be seen in the difference waves in Figs. 2 and 3, there is an important difference regarding the early negativity: it seems that there is no difference in the amplitude of this negative peak between the three types of endings when the series is decreasing. As for the late positivity, the pattern of differences was similar to those described in the previous paragraph.

3.1. 0–200-ms interval

The ANOVA performed in the 0–200-ms window did not show any statistically significant effect, either main or interactive, over the voltage mean (some of the results are shown in Table 1). As shown in Figs. 2 and 3, an N1-P2 complex is associated with the presentation of the stimulus and it does not differ as a function of the type of stimulus. None of the manipulated factors (series type, response type, and ending type) elicits a differential ERP pattern in the 0–200-ms post-stimulus interval.

3.2. 200–500-ms interval

The interval from 200- to 500-ms post-stimulus was analyzed by performing ANOVAs, taking the mean amplitude in successive 50-ms windows as the dependent variable (the main results are shown in Table 1). The variable *type of response* did not have any effect, either principal or interactive, over the voltage; therefore, we only report results for the other variables.

The main effect of ending type reached statistical significance[5] in the 250- to 500-ms windows. The high effect sizes in the last two windows (400–500 ms) should be noted: 69% and 76% of variance being explained by the variable ending type, respectively. As we were interested in identifying in detail the differences between the three types of endings, paired contrasts were run using a non-pooled error term [24,25]. Although results showed similar differences between the three types of endings from 350 to 500 ms, the amplitude pattern was different in the interval 250–300.

Amplitude was more negative for both incorrect endings than for the correct endings in the 250–300-ms window: $F(1,13)=16.73$, $p=0.001$, $\eta^2=0.56$, when comparing clear incorrect with correct endings, and $F(1,13)=7.84$, $p=0.015$, $\eta^2=0.37$, in the confusing incorrect to correct ending contrast. There were not, however, significant differences between the two types of violations. As we have already mentioned, with respect to the 350- to 500-ms windows, there was a similar voltage pattern. In the 350–400-ms window, voltage was more positive for clear incorrect ending than for the other types of stimuli—$F(1,13)=14.31$, $p=0.002$, $\eta^2=0.52$, in the clear incorrect to correct ending contrast, and $F(1,13)=19.86$, $p=0.001$, $\eta^2=0.60$, in the clear to confusing incorrect ending contrast: there were no significant differences when comparing correct stimuli with confusing incorrect stimuli. However, for the other two windows—those between 400 and 500 ms—every paired contrast reached a p-value lower than 0.006: voltage was more positive for clear incorrect endings, medium for confusing incorrect endings, and more negative for correct endings. In summary, there are three neuroelectric patterns in the interval between 250 and 500 ms. First, the voltage is

[5] We considered that an effect reached significance when the p-value was less than or equal to 0.05.

Table 2
Mean differences (in μV) between the three types of ending in the 250–800-ms interval[a]

	250–300	300–350	350–400	400–450	450–500	500–600	600–700	700–800
Confusing incorrect–Correct	– 1.63*	– 2.32*	– 0.57	1.31*	2.39**	2.44**	1.92*	1.13
Clear incorrect–Correct	– 2.13**	– 0.66	1.94*	4.04**	5.49**	6.31**	5.34**	3.16*
Clear incorrect–Confusing incorrect	– 0.5	1.66	2.51**	2.73**	3.1**	3.87**	3.42**	2.03*

Overall means for increasing and decreasing series are presented.
[a] Increase in type I error was controlled for every contrast by means of the Hochberg approach.
*$0.016 \geq p > 0.001$.
**$p \leq 0.001$.

more negative for any violation than for a correct ending up to 300 ms. Second, a change can be seen in the interval between 300 and 350 ms, where the voltage for clear incorrect endings becomes more positive. Third, the voltage is more positive for clear incorrect than for confusing incorrect and correct stimuli between 350 and 500 ms. Amplitude mean differences for each window are shown in Table 2.

The negativity found in the 250–300-ms window was modulated by the *type of series*, as shown by the significant interaction between *type of ending* and type of series in this window; $F(2,26) = 4.57$, $p = 0.02$, $\eta^2 = 0.26$. Simple effect analyses showed that although voltage was more negative for incorrect than for correct endings in the adding series, there were no differences between the three types of stimuli in subtracting series. Table 3 shows the mean differences for the three types of ending in both types of series.

The interaction ending type × location reached significance in the 300–350, 400–450 and 450–500 ms windows. Simple effect analyses showed that voltage differences between the three types of ending appeared mainly at centro-parietal locations. Specifically, results were as follows: (1) in the 300–350-ms window, significant differences were found when comparing both types of violations at C3, P3, Cz, and Pz; when we compared correct and confusing incorrect endings, differences were shown at C4, P3, P4, Cz, and Pz (voltage was more negative for confusing incorrect ending than for the other types of stimuli—correct and clear incorrect stimuli, there being no significant difference between these last two types of stimuli). (2) In the 400–450-ms window, there were differences between correct and clear incorrect endings in all the locations, between correct and confusing incorrect ending at C3, P3, Cz, and Pz, and between both types of incorrect ending in all the locations, except Fz. In all cases, the voltage was more positive for a clear incorrect ending, medium for confusing incorrect ending, and more negative for correct endings. (3) In the 450–500-window, simple effect analyses detected differences between the three types of ending in all the locations; in this case, the voltage pattern was the same as that described for the previous window. In summary, the results show the following pattern: although differences between the three types of stimuli reached significance at every location in the 450–500-ms window, these differences were centro-parietally distributed in the interval 300–350 and 400–450 ms.

Finally, analysis to examine possible laterality effects at centro-parietal locations was performed and yielded a significant ending type × hemisphere effect at the 250–300-ms windows ($F(2,26) = 5.61$, $p = 0.009$, $\eta^2 = 0.30$), and the 300–350-ms window ($F(2,26) = 6.92$, $p = 0.004$, $\eta^2 = 0.25$). Simple effect analyses showed that there was not difference between left and right hemisphere for any type of ending (correct, confusing incorrect and clear incorrect), although differences between types of ending were a bit larger over right hemisphere (mean and standard error for the three types of ending in both hemispheres are reported in Table 4).

3.3. 500–800-ms interval

Mean amplitude in 100-ms windows from 500- to 800-ms post-stimulus was analyzed by ANOVA (the main results are shown in Table 1). Again, the variable type of response did not reach significance either in its principal effect or in its interactive effects with the other variables.

The type of ending was significant for all three windows analyzed. Once again, the high effect sizes in the two first windows (500–700 ms) should be noted: 76% and 68% of variance in the observed data being accounted for by this variable. Paired contrasts showed a similar voltage pattern for these two first windows: significant differences were found between correct and clear incorrect endings, correct and confusing incorrect endings, and confusing and clear incorrect endings (p values < 0.001 in the 500–600-ms window; p values < 0.014 in the 600–700-ms window). This pattern was the same as that reported within the 400–500-ms interval: more positive voltage for clear incorrect ending, medium for confusing incorrect ending, and more negative for correct ending. As regards the 700–800-ms window, differences were only significant in the correct

Table 3
Mean differences (in μV) between the three types of ending for increasing and decreasing series in the 250–300-ms window[a]

	Increasing series	Decreasing series
Confusing incorrect–Correct	– 2.17*	– 1.1
Clear incorrect–Correct	– 3.06*	– 1.2
Clear incorrect–Confusing incorrect	– 0.89	– 0.1

[a] Increase in type I error was controlled for every contrast by means of the Hochberg approach.
*$p \leq 0.002$.

Table 4
Mean and standard error (in μV) for the three types of ending in both hemispheres in the 250–300- and the 300–350-ms window.

	250–300-ms window			300–350-ms window		
	Correct	Confusing incorrect	Clear incorrect	Correct	Confusing incorrect	Clear incorrect
Left hemisphere	4.60 (0.63)	3.24 (0.42)	3.02 (0.66)	6.26 (0.84)	4.30 (0.50)	6.00 (0.73)
Right hemisphere	5.21 (0.60)	3.45 (0.58)	2.86 (0.86)	6.73 (0.76)	4.17 (0.63)	5.50 (0.77)
Mean difference	−0.61	−0.21	0.16	−0.47	0.13	0.5

Overall means for increasing and decreasing series are presented.

to clear incorrect ending contrast ($F(1,13)=14.35$, $p=0.002$, $\eta^2=0.52$), and in the clear to confusing incorrect ending contrast ($F(1,13)=6.06$, $p=0.025$, $\eta^2=0.32$). The voltage pattern was similar to that observed in the other windows. Table 2 shows mean amplitude differences for each window, and it can be seen that the maximum differences between conditions appeared within the 500–600-ms interval.

As for the type of ending × location interaction, significance was reached in all three windows analyzed. Simple effect analyses revealed a voltage pattern very similar to that reported within the 400–500-ms interval. Once more, the differences were mainly centro-parietal. The results were as follows: firstly, within the 500–600-ms window, clear incorrect endings showed a more positive voltage than did correct and confusing incorrect endings in all the locations; moreover, confusing incorrect differed from correct endings—higher positivity in the first—at C3, P3, P4, Cz, and Pz. Secondly, with respect to the 600–700-ms window, clear incorrect ending differed from the other two types of endings in all the locations; confusing and clear incorrect endings differing only at P3, P4, and Pz. Lastly, within the 700–800-ms window, correct and clear incorrect endings were different in all the locations, except for Fz; correct and confusing incorrect endings differed at P3 and P4; differences between both types of violation reached significance at C3, C4 and Cz. In summary, the differences between the three types of endings—larger positivity for clear incorrect ending, medium for confusing incorrect ending, and lesser for correct ending—were significant mainly at centro-parietal locations.

To conclude, neither type of series nor hemisphere has any effect over the voltage in the 500–800-ms interval, either in its main effect or in its interaction with the other variables.

4. Discussion

The main goal of the present study was to examine whether the P600/SPS component is a specific linguistic component or if, conversely, it could be considered to be a more general index related to violations in any rule-governed sequence. To this end, we presented anomalies in increasing and decreasing series, whose difficulty of integration was manipulated. We hypothesized that P600/SPS specificity would be brought into question if a late positivity, similar to this component, was elicited by violations in arithmetical series.

Our results support the hypothesis that both syntactic and arithmetical violations share a similar neurophysiologcial process. A positive peak was elicited whenever an incorrect number ending was presented, in both increasing and decreasing series. This positivity started in the 350–400-ms interval and showed its maximum amplitude in the 500–600-ms window. Moreover, it presented a centro-parietal distribution and its amplitude was sensitive to the type of violation that was presented. Concerning this last point, it was found that the amplitude varied depending on the difficulty of integration of the number in the series: the easier it was to refuse a number as a correct ending, the larger the amplitude of this late positivity. In this sense, we consider that whereas clear incorrect endings were impossible to integrate in the series, confusing incorrect endings could indeed produce some kind of ambiguity in the subject that made him/her try to integrate the number. Consequently, our results could be explained in terms of difficulty of integration.

These results are consistent with those previously reported in other papers. A P600 variation of amplitude has been reported both in syntactic and harmonic violations. As for the first—syntactic violation—Osterhout et al. [49] stated that P600/SPS amplitude is sensitive to the construction of a syntactically correct structure. They manipulated the severity of syntactic anomalies encountered during sentence processing, and found that P600 amplitude was a function of the degree of the syntactic anomaly; the more severe the violation presented, the greater the amplitude. These authors gave two possible explanations for this variation of amplitude. First, they claim that this variation could be explained in terms of the syntactic fit, or expectation, of a word with the preceding sentence structure—as what happens with the N400 amplitude in a cloze probability task [37]. Second, the amplitude variation can be explained in terms of the cost of syntactic reprocessing.

Patel et al. [51] reported similar results working with anomalies in sentences and music. Both syntactic and harmonic violations elicited a P600 component, whose amplitude varied depending on the degree of integration of the stimulus in the previous context: the larger the difficulty of integration, the larger the amplitude of the

P600. They concluded, "whatever process gives rise to the P600 it is unlikely to be language-specific" ([51], p. 726).

The present results also agree with those reported by Münte et al. [41], where a late positivity was elicited by orthographic anomalies, and those reported by Besson and Macar [4], who reported a similar positivity elicited by incongruous stimuli, in both geometric patterns and melodies. Similar results have been obtained by presenting abstract structures in sequences of visually presented stimuli [39,40]. In these experiments, Lelekov and Dominey [39] and Lelekov et al. [40] reported a late positivity around 500 ms that was evoked by violations of non-linguistic abstract structures. Again, this positivity was similar to the P600/SPS reported in studies of syntactic processing, so these authors claim that "abstract structural organization for both linguistic and non-linguistic sequences may rely on partially overlapping neurophysiological processes" ([39], p. 83).

In conclusion, the similarities between the positivity elicited by arithmetic violations—in both addition and subtraction—and the P600/SPS component allow us to conclude that they are comparable components. They share similar latency, topography and are sensitive to the same experimental manipulation, so the more parsimonious explanation is that they are reflecting a similar process. In fact, taking into account the results reported in other studies, P600 could be considered not as a specific index of breaking linguistic rules, but as a more general index of violation in any rule-governed sequence. In this sense, we agree with Patel et al. [51] when they state, "P600 reflects processes of knowledge-based structural integration" ([51], p. 727). For the moment, P600 has been related to syntactic, harmonic, abstract structure and arithmetic anomalies, and in these four cases, there was a rule that had been broken.

There is another argument that has to be considered concerning the nature of P600/SPS. We commented in the introduction that there has been a recent debate over whether the P600/SPS component could in fact be a P3b component, and the question remains unsolved. Some researchers defend the syntactic specificity of the P600/SPS component [45,50], whereas others cast doubt on its specificity and claims that it could be a P3b [9,10,20,41]. The defence of the P3b hypothesis is based on the argument that this component is similar to the P600/SPS in many aspects: (1) P3b is a positive peak; (2) it shows a centro-parietal distribution; (3) its latency depends on the difficulty of the task; (4) its amplitude is a function of stimulus probability[6]; and (5) P3b is elicited whenever an unexpected stimulus is presented, so it is said to be an index of context updating. In this sense, "the P600/SPS component could be considered a P3b elicited by encountering a rare linguistic event of ungrammaticality" ([9], p. 45). The defenders of P600/SPS specificity, on the other hand, argue that "the neural response to syntactic anomalies is, at least to an interesting degree, distinct from the neural response to non-linguistic anomalies" ([45], p. 8). Returning to our experiment, the results we obtained could also be explained in terms of P3b, because the component we found matched the P3b component in all the aspects mentioned above.

A second interesting result was found in the present experiment: an early negativity peaking between 250 and 300 ms was elicited by arithmetic violations in increasing but not in decreasing series. This finding suggests a neurophysiological difference between addition and subtraction processing, since differences between correct and incorrect endings were only present when the rule consisted of adding a number. One way of interpreting this difference is based on the claim that addition is a more automatic process that subtraction. LeFevre et al. [38] found that the mere presentation of two digits like "4 2" yielded an automatic activation of the addition result "6". In this experiment, subjects were presented with a number and were asked to say if it was in the previous list of two. Reaction time was larger when the target was the sum of the digits than when it was an unrelated digit.

The fact that addition is considered to be a more automatic process that subtraction allows us to speculate about the similarities between this early arithmetic negativity and the LAN component. The P600/SPS component is usually preceded by a negativity peak—the LAN—which is considered to be an index of automatic first parsing processing that operate quite quickly and automatically. It is possible that the arithmetic negativity for increasing series could be a similar component to this syntactic negativity. Hoen and Dominey [24] have reported a component similar to the LAN in a non-linguistic cognitive sequencing task. Therefore, it could be expected that the automatic LAN was elicited in this kind of arithmetic operation. However, there are two reasons why we have insufficient evidence to claim that this addition-related early negativity is the LAN component. First, early negativities related to syntactic anomalies have a left anterior distribution, whereas arithmetic negativity in additions did not have a clear distribution. Moreover, a seven-electrode array is insufficient to provide topographical conclusions. Second, LAN peaks around 400 ms and the negative component found in arithmetic sequences peaked in the 250–300-ms range. Therefore, further research needs to be carried out in order to gain more knowledge about the addition-related early negativity.

In summary, two main conclusions can be drawn from our study. Firstly, the P600/SPS component may not be linguistic and could be a more general index of violation of any rule-governed sequences, or even a P3b. Secondly, the question about the study of possible similarities of the LAN component and the addition-related early negativity is an interesting question that remains open to further investigation.

[6] P3b amplitude is inversely proportional to the stimulus probability, this probability being objective or subjective [26,53].

Acknowledgements

This research was supported by grant FI-PG/95-1.191 from the Generalitat de Catalunya. The authors thank Antonio Solanas, Carles Escera and Nuria Sebastian for their helpful comments on an previous draft of this paper. The authors also thank two anonymous reviewers for their helpful suggestions on an earlier version of this manuscript. A special acknowledgment to Isabel Palmerín, who gave us time to finish this paper.

References

[1] M.H. Ashcraft, J. Battaglia, Cognitive arithmetic: evidence for retrieval and decision processes in mental addition, J. Exp. Psychol. Hum. Learn. Mem. 4 (1978) 527–538.

[2] M.H Ashcraft, E.H. Stazyk, Mental addition: a test of three verification models, Mem. Cogn. 9 (1981) 185–196.

[3] M. Besson, F. Faïta, An event-related potential (ERP) study of musical expectancy: comparison of musicians with nonmusicians, J. Exp. Psychol. Learn. Mem. Cogn. 21 (1995) 1278–1296.

[4] M. Besson, F. Macar, An event-related potential analysis of incongruity in music and other non-linguistic contexts, Psychophysiology 24 (1987) 14–25.

[5] C. Brown, P. Hagoort, On the electrophysiology of language comprehension: implications for the human language system, in: M.W. Pickering, M. Pickering, C. Clifton (Eds.), Architectures and Mechanisms for Language Processing, Cambridge Univ. Press, Cambridge, 2000, pp. 213–237.

[6] C. Brown, P. Hagoort, M. ter Keurs, Electrophysiological signatures of visual lexical processing: open and closed-class words, J. Cogn. Neurosci. 11 (1999) 261–281.

[7] J. Cohen, Statistical Power Analysis for the Behavioral Sciences, Lawrence Erlbaum Associates, Hillsdale, NJ, 1988.

[8] L. Cohen, S. Dehaene, Calculating without reading: unsuspected residual abilities in pure alexia, Cogn. Neuropsychol. 17 (2000) 563–583.

[9] S. Coulson, J.W. King, M. Kutas, Expect the unexpected: event-related brain response to morphosyntactic violations, Lang. Cogn. Processes 13 (1998) 21–58.

[10] S. Coulson, J.W. King, M. Kutas, ERPs and domain specificity: beating a straw horse, Lang. Cogn. Processes 13 (1998) 653–672.

[11] S. Dehaene, L. Cohen, Towards an anatomical and functional model o number processing, Math. Cogn. 1 (1995) 83–120.

[12] S. Dehaene, E. Spelke, R. Pinel, R. Stanescu, S. Tsivkin, Sources of mathematical thinking: behavioral and brain-imaging evidence, Science 284 (1999) 970–974.

[13] R. El Yagoubi, P. Lemaire, M. Besson, Different brain mechanisms mediate two strategies in arithmetic: evidence from event-related brain potentials, Neuropsychologia 41 (2003) 855–862.

[14] A.D. Friederici, The neurobiology of language processing, in: A.D. Friederici (Ed.), Language Comprehension: A Biological Perspective, Springer, Berlin, 1998, pp. 265–292.

[15] A.D. Friederici, E. Pfeifer, A. Hahne, Event-related brain potentials during natural speech processing: effects of semantic, morphological and syntactic violations, Cogn. Brain Res. 1 (1993) 183–192.

[16] A.D. Friederici, A. Hahne, D.Y. von Cramon, Firs-pass versus second-pass parsing processes in a Wernicke's and a Broca's Aphasic: electrophysiological evidence for a double dissociation, Brain Lang. 62 (1998) 311–341.

[17] A.D. Friederici, D.Y. von Cramon, S.A. Kotz, Language related brain potentials in patients with cortical and subcortical left hemisphere lesions, Brain 122 (1999) 1033–1047.

[18] S. Geisser, S.W. Grenhouse, An extension of box's results on the use of the F distribution in multivariate analysis, Ann. Math. Stat. 29 (1958) 885–891.

[19] T.C. Gunter, A.D. Friederici, Concerning the automaticity of syntactic processing, Psychophysiology 36 (1999) 126–137.

[20] T.C. Gunter, L.A. Stowe, G. Mulder, When syntax meets semantics, Psychophysiology 34 (1997) 660–676.

[21] P. Hagoort, C. Brown, J. Groothusen, The syntactic positive shift (SPS) as an ERP measure of syntactic processing, Lang. Cogn. Processes 8 (1993) 439–483.

[22] A. Hahne, A.D. Friederici, Electrophysiological evidence for two steps in syntactic analysis: early automatic and late controlled processes, J. Cogn. Neurosci. 11 (1999) 194–205.

[23] W.L. Hays, Statistics, 5th ed., Holt, Rinehart and Winston, New York, 1994.

[24] M. Hoen, P.F. Dominey, ERP analysis of cognitive sequencing: a left anterior negativity related to structural transformation processing, NeuroReport 11 (2000) 3187–3191.

[25] P. Janata, ERP measures assay the degree of expectancy violation of harmonic contexts in music, J. Cogn. Neurosci. 7 (1995) 153–164.

[26] R. Johnson, A triarchic model of P300 amplitude, Psychophysiology 23 (1986) 367–384.

[27] H.J. Keselman, Testing treatment effects in repeated measures designs: an update for psychophysiological researchers, Psychophysiology 35 (1998) 470–478.

[28] H.J. Keselman, J.C. Keselman, Comparing repeated measures means in factorial designs, Psychophysiology 25 (1988) 612–618.

[29] H.J. Keselman, J.C. Keselman, Analysis of repeated measurements, in: L.K. Edwards (Ed.), Applied Analysis of Variance in Behavioral Science, Marcel Dekker, New York, 1993, pp. 105–145.

[30] J.W. King, M. Kutas, Who did what and when? using word- and clause-level: ERPs to monitor working memory usage in reading, J. Cogn. Neurosci. 1 (1995) 376–395.

[31] R.E. Kirk, Practical significance: a concept whose time has come, Educ. Psychol. Meas. 56 (1996) 746–759.

[32] R. Kluender, M. Kutas, Bridging the gap: evidence from ERPs on the processing of unbounded dependencies, J. Cogn. Neurosci. 5 (1993) 196–214.

[33] M. Kutas, K.D. Federmeier, Electropsysiology reveals semantic memory use in language comprehension, Trends Cogn. Sci. 12 (2000) 463–470.

[34] M. Kutas, S. Hillyard, Reading senseless sentences: brain potentials reflect semantic incongruity, Science 207 (1980) 203–205.

[35] M. Kutas, C. Van Petten, Event-related brain potentials studies of language, in: P.K. Ackles, J.R. Jennings, M.G.H. Coles (Eds.), Advances in Psychophysiology, JAI Press, Greenwich, CT, 1988, pp. 138–187.

[36] M. Kutas, C. Van Petten, Psycholinguistics electrified: event-related brain potential investigacions, in: M.A. Gernsbacher (Ed.), Handbook of Psycholinguistics, Academic Press, San Diego, CA, 1994, pp. 83–143.

[37] M. Kutas, T. Lindamood, S.A. Hillyard, Word expectancy and event-related brain potentials during sentence processing, in: S. Kornblum, J. Roquin (Eds.), Preparatory States and Processes, Erlbaum, Hillsdale, NJ, 1984, pp. 217–238.

[38] J. LeFevre, J. Bisanz, L. Mrkonjic, Cognitive arithmetic: evidence for obligatory activation of arithmetic facts, Mem. Cogn. 16 (1988) 45–53.

[39] T. Lelekov-Boissard, P.F. Dominey, Human brain potentials reveal similar processing of non-linguistic abstract structure and linguistic syntactic structure, Neurophysiol. Clin. 32 (2002) 72–84.

[40] T. Lelekov-Boissard, P.F. Dominey, L. Garcia-Larrea, Dissociable ERP profiles for processing rules vs. instances in a cognitive sequencing task, NeuroReport 11 (2000) 1129–1132.

[41] T.F. Münte, H. Heinze, M. Matzke, B.M. Wieringa, S. Johannes, Brain potentials and syntactic violations revisited: no evidence for specificity of the syntactic positive shift, Neuropsychologia 36 (1998) 217–226.

[42] H.J. Neville, J.L. Nicol, A. Brass, K.I. Forster, M.F. Garrett, Syntactically based sentence processing classes: evidence from event-related brain potentials, J. Cogn. Neurosci. 3 (1991) 151–165.

[43] M. Niedeggen, F. Rösler, N400 effects reflect activation spread during retrieval of arithmetic facts, Psychol. Sci. 1 (1999) 271–276.

[44] M. Niedeggen, F. Rösler, K. Jost, Processing of incongruous mental calculation problems: evidence for an arithmetic N400 effect, Psychophysiology 36 (1999) 307–324.

[45] L. Osterhout, P. Haggort, A superficial resemblance does not necessarily mean you are part of the family: counter-arguments to Coulson, King and Kutas (1998) in the P600/SPS-P300 debate, Lang. Cogn. Processes 14 (1999) 1–14.

[46] L. Osterhout, P. Holcomb, Event-related brain potentials elicited by syntactic anomaly, J. Mem. Lang. 31 (1992) 785–806.

[47] L. Osterhout, P.J. Holcomb, Event-related potentials and language comprehension, in: M.D. Rugg, M.G.H. Coles (Eds.), Electrophysiology of Mind: Event-Related Brain Potentials and Cognition, Oxford Univ. Press, Oxford, 1995, pp. 171–208.

[48] L. Osterhout, L.A. Mobley, Event-related brain potentials elicited by failure to agree, J. Mem. Lang. 34 (1995) 739–773.

[49] L. Osterhout, P.J. Holcomb, D.A. Swinney, Brain potentials elicited by garden-path sentences: evidence of the application of verb information during parsing, J. Exp. Psychol. Learn. Mem. Cogn. 20 (1994) 786–803.

[50] L. Osterhout, R. McKinnon, M. Bersick, V. Corey, On the language specificity of the brain response to syntactic anomalies: is the syntactic positive shift a member of the P300 family? J. Cogn. Neurosci. 8 (1996) 507–526.

[51] A.D. Patel, E. Gibson, J. Ratner, M. Besson, P.J. Holcomb, Processing syntactic relations in language and music: an event-related potential study, J. Cogn. Neurosci. 10 (1998) 717–733.

[52] J. Polich, A. Kok, Cognitive and biological determinants of P300: an integrative review, Biol. Psychol. 41 (1995) 103–146.

[53] W.S. Pritchard, Psychophysiology of P300, Psychol. Bull. 89 (1981) 506–540.

[54] F. Rösler, P. Pütz, A. Friederici, A. Hahne, Event-related brain potentials while encountering semantic and syntactic constraint violations, J. Cogn. Neurosci. 5 (1993) 345–362.

[55] S.J. Segalowitz, H. Chevalier, Event related potential (ERP) research in neurolinguistics: Part I. Techniques and applications to lexical access, in: B. Stemmer, H.A. Whitaker (Eds.), Handbook of Neurolinguistics, Academic Press, San Diego, CA, 1998, pp. 95–110.

[56] S.J. Segalowitz, H. Chevalier, Event related potential (ERP) research in neurolinguistics: Part II. Language processing and acquisition, in: B. Stemmer, H.A. Whitaker (Eds.), Handbook of Neurolinguistics, Academic Press, San Diego, CA, 1998, pp. 111–123.

[57] M. ter Keurs, C. Brown, P. Hagoort, Lexical processing of vocabulary class in patients with Broca's aphasia: an event-related brain potential study on a grammatical comprehension, Neuropsychologia 20 (2002) 1547–1561.