Joanna Błaszczak https://orcid.org/0000-0002-8332-2827

University of Wrocław

Juliane Domke

Max Planck Society – Social Neuroscience Lab

Can Tense Be Subject to Grammatical Illusion? Part 2: Evidence from an ERP Study on the Processing of Tense and Aspect Mismatches in Compound Future Constructions in Polish

Abstract

In this part of the paper we report the results of an ERP study on the processing of two types of compound future in Polish, with infinitival and participial complements. In the study we monitored the EEG correlates of two types of temporal mismatches. Tense mismatches between the future auxiliary and the past tense modifier *wczoraj* ('yesterday') relative to the *jutro* ('tomorrow') baseline resulted in a biphasic (LAN + P600) signature. Aspect mismatches between the future auxiliary and the perfective aspect of the lexical complement (relative to the imperfective baseline) triggered a widely distributed positivity with a posterior maximum (P600). In addition, we wanted to assess whether matching tense specifications in different words of a sentence can cause grammatical illusions. To this aim, we tested whether the presence of the adverb *wczoraj* ('yesterday') (specified for [past]) could give rise to an illusion of grammaticality for perfectives as *l*-participles (allegedly [past] marked), but not as infinitives (not having any [past] specification). Neither behavioral nor electrophysiological results of the present study provided support for this hypothesis. Rather, the findings seem to suggest that TENSE might not belong to the features that are relevant for grammatical illusions, unlike NEGATION, COMPARATIVE, CASE, NUMBER, GENDER, which have been shown to be suspectible to grammatical illusions. We conclude with a discussion of possible underlying reasons for this negative result.

Keywords

processing of compound future constructions, tense and aspect mismatches, grammatical illusion, tense illusion, ERP, Polish

Streszczenie

W tej części artykułu przedstawiamy wyniki badania ERP dotyczącego przetwarzania dwóch rodzajów złożonych konstrukcji czasu przyszłego w języku polskim (z czasownikiem leksykalnym w formie imiesłowu i bezokolicznika). Badanie polegało na monitorowaniu korelatów EEG dwóch rodzajów temporalnych niezgodności. Niezgodności pod względem czasu pomiędzy czasownikiem posiłkowym czasu przyszłego a modyfikatorem czasu przeszłego *wczoraj* (w porównaniu z semantycznie zgodnym modyfikatorem *jutro*) skutkowały pojawieniem się dwufazowego komponentu LAN + P600. Niezgodność pod względem aspektu między czasownikiem posiłkowym a formą dokonaną czasownika leksykalnego (w porównaniu z wymaganą w tym kontekście formą niedokonaną) wywołała obszerną pozytywność z maksimum w tylnych obszarach mózgu (P600). Ponadto celem badania było sprawdzenie, czy (nie)dopasowanie pod względem kategorii czasu między określonymi słowami w zdaniu może spowodować iluzje gramatyczności poprzez określenie, czy obecność przysłówka *wczoraj* wywołuje złudzenie gramatyczności w przypadku dokonanych dopełnień imiesłowowych, ale nie bezokolicznikowych. Tylko pierwsze z nich są powierzchownie identyczne z formami czasu przeszłego, potencjalnie więc tylko one mają cechę [past]. Ta hipoteza nie zyskała potwierdzenia w dostarczonych przez badanie danych behawioralnych ani elektrofizjologicznych. Wyniki wydają się raczej sugerować, że CZAS, inaczej niż NEGACJA, KOMPARATYW, PRZYPADEK, LICZBA, RODZAJ, może nie należeć do cech mogących być przedmiotem złudzeń gramatyczności. Artykuł kończy się omówieniem różnych możliwych przyczyn tego negatywnego wyniku.

Słowa kluczowe

przetwarzanie złożonych konstrukcji czasu przyszłego, niezgodności czasu i aspektu, iluzja gramatyczności, iluzja czasu (iluzja temporalna), ERP, język polski

4. The present ERP study

To test the predictions formulated in Part I of this paper (see SPL 14(4)), two ERP experiments were conducted that were identical with respect to tasks and recording procedures. The only difference concerns the material. Critical sentences contained imperfective/perfective participles in Experiment 1, but in Experiment 2 imperfective/perfective infinitives were used instead (see Table 1 in Part I). This had to be done to avoid abnormal extension of length and duration of an ERP testing session. Data from both experiments were included in the same statistical analysis (see below), when necessary for direct comparison.
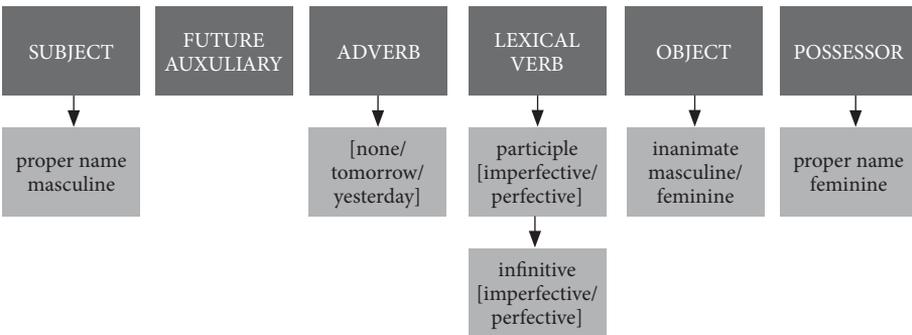
4.1. Participants

Forty four Polish native speakers aged between 21 and 29 years from the University of Wrocław (students from the Institute of English Studies) participated in the reported ERP study for partial course credit (in Experiment 1, there were 21 subjects (16 female), mean age 23.3 years (SD = 2.15); in Experiment 2 there were 23 subjects (20 female), mean age 21.5 years; SD = 0.93). All of them were right-handed according to the Edinburgh Handedness Inventory

(Oldfield 1971). The EEG data of 9 additional subjects (Experiment 1: n=7, Experiment 2: n=2) were excluded from the analysis due to excessive noise artifacts. All participants had normal or corrected vision. None had neurological or psychiatric disorders or reported neurological traumas.

4.2. Material

In each experiment there were 300 stimulus items (sentences): 50 sentences per condition (50 x 6). All of them had the same syntactic structure across all experimental conditions and, as presented in Figure 1, consisted of a subject (masculine proper name, 2 syllables), a future auxiliary *będzie*, optionally a compatible 'tomorrow'/incompatible 'yesterday' temporal adverbial, an imperfective or perfective lexical verb[1] (in the form of an *l*-participle in Experiment 1 and in the form of an infinitive in Experiment 2), and an inanimate object (1–4 syllables, 50% masculine, 50% feminine) with a possessor (femine proper name, 2–3 syllables). The schematic structure of sentences in different conditions is presented in Figure 1.

SUBJECT	FUTURE AUXULIARY	ADVERB	LEXICAL VERB	OBJECT	POSSESSOR
proper name masculine		[none/ tomorrow/ yesterday]	participle [imperfective/ perfective]	inanimate masculine/ feminine	proper name feminine
			infinitive [imperfective/ perfective]		

Figure 1. General schema of the experimental sentences

The single words in each sentence were not controlled for frequency. However, the content of the sentence material relied on daily actions and were informally tested for plausability prior to the experimental setup. Furthermore, it should be noted that exactly the same words were used across all experimental conditions, that is, the conditions did not differ in this respect.[2]

---

[1] The mean length (in terms of the number of syllables) was 3.01 syllables in the case of imperfective verbs and 3.26 syllables in the case of perfective verbs. Importantly, note that there were no differences between participles and infinitives in this respect, that is, all imperfective/ perfective participles had exactly the same length as the corresponding imperfective/perfective infinitives.

[2] It has been claimed in the literature that the participial complement is preferably used for masculine singular forms of the compound future, while the infinitival complement is used

All the stimuli were pseudo-randomly ordered and distributed over six blocks, each containing 50 sentences. After each sentence there was a grammaticality judgement question. To control the level of attention and the degree of processing difficulty an additional probe detection task was used (see also footnote 3). The probe words were equally distributed across conditions. There was an equal number of probes semantically or phonologically corresponding to subjects, objects and possessors. The probes were balanced for the expected YES and NO answers regarding the entire experimental run since the grammatical judgement task promoted imbalanced answers (75% NO, 25% YES).[3]

The whole set of 300 sentences was presented to all participants, however, in each experiment the material was presented in one of two versions: version A – in a descending order and version B – in an ascending order. Each version was presented to half of the subjects. Additionally, each version A and B was further subdivided into two variants differing in the coding for YES and NO buttons to avoid any potential effects of lateralized readiness potential.

4.3. Procedure

In each experiment participants were tested individually in one session, which lasted approximately 90 minutes including EEG preparation. Following the application of the EEG electrodes, subjects were seated approximately one meter in front of a Samsung 22" computer LCD computer screen in an electrically and acoustically shielded EEG chamber. All stimuli were presented in the center of the screen in a white courier font, size 48, on a black background using Presentation software package 16.3 12.20.12 (Neurobehavioral System Inc., 2012).

The experimental session was preceded by instructions and a trial session. As part of the instructions the participants were asked not to move or blink while a sentence was displayed. They were informed that the sentences would be presented segment by segment and that each sentence would be followed by a grammaticality judgment question as well as by a question whether

mostly with plural subjects but also for other nonmasculine forms (see, e.g., Mikoś 1985). It is also suggested that the sex of the speaker might play an important role in regard to the question of which type of compound future construction is chosen: male speakers seem to preferably use participial futures in first person sentences (Łaziński 2006: 319–320). The question of the frequency of use of participial and infinitival complements cannot be discussed here in detail for reasons of space. Importantly, as explained above, participial and infinitival futures (only with singular third person subjects) were tested in two separate experiments (with two different groups of participants). In other words, no participant of the study saw both variants of compound future so that the question of preference did not arise.

[3] The probe detection task was used instead of fillers. Note that due to the way in which the probes were used (as explained in the text), the participants' attention could be diverted from the critical items (adverb and verb). This strategy was sucessful as was additionally confirmed in debriefing interviews.

a particular probe word occurred in the currently presented sentence. The participants were instructed to provide their judgments as fast as possible. They were also instructed which button on the Razor keyboard corresponded to which answer.

After reading written instructions, the participants received a practice block with 10 sentences related to the experiment in order to familiarize each of them with the task.[4] After the practice session, the participants received explicit feedback about the errors they made. The practice session was followed by six experimental blocks containing 50 sentences each. After each block there was a break.

Each trial consisted of the following events: a fixation cross appeared in the center of the screen for 1000 ms, after which a stimulus sentence was presented word-by-word. Each word appeared in the center of the screen for 500 ms, followed by a short 100 ms long blank screen interval. Sentence-ending words appeared with a full stop and were followed by a 100 ms blank screen. After each presented sentence the question POPRAWNE? ("correct?") appeared on the screen for 500ms. After that, the words TAK ("yes") and NIE ("no") were presented on the screen, as a prompt for the grammaticality judgement task. Participants had 3000 ms to chose the right answer. Probes were presented for 500 ms as well. After that, the words TAK ("yes") and NIE ("no") were presented on the screen for 3000 ms, as a prompt for the probe detection task. After that time the next trial started automatically with the presentation of the new asterisk.

4.4. EEG recordings

The EEG-activity was measured with 24 Ag/AgCl-electrodes which were attached to the scalp using the Easycap system at Fz, FCz, Cz, CPz, Pz, POz, FC1, F3, C3, P3, O1, FC5, CP5, F7, P7, FC2, F4, C4, P4, O2, FC6, CP6, F8, P8. The ground electrode was positioned at AFz. Electrode positions were chosen in accordance with the international 10/20 system (Jasper 1958). Signals were referenced to the A1 electrode (left mastoid) during recording and re-referenced to the average of left (A1) and right (A2) mastoids before data processing. Horizontal eye activity was measured by placing two electrodes 2 cm lateral to the right (EOGR) and the left (EOGL) canthus. Vertical eye activity was measured by placing two electrodes 3 cm above (EOGU) and below (EOGD) the pupil of the right eye. Electrode impedances were kept below 5 kΩ. All electrophysiological signals were digitized with a frequency of 250 Hz.

---

[4] None of the sentences used in the practice session appeared later in the experiment proper.

4.5. Data processing and analysis

4.5.1. Behavioral data processing and analyses

In total, 1594 data points or 8.9% of all data were removed (because of missing data, bad performance, outlier character, or coding errors, among other things). After this, 16406 data points out of the original 18000 were left and subjected to further analysis. All final statistical analyses were conducted using R (R Core Team 2019) on a linux machine.[5] The main statistical method of choice considered linear mixed model as implemented in the lme4 package (Bates et al. 2015).[6]

For the analysis of reaction times (RT) a mixed model was used with full fixed effects structure in the three independent variables (ADVERB, ASPECT, VERB), taking into account a trial effect (learning effect represented by a decrease in participants' RT) as a third order polynomial and with a rather flexible random effects structure. The following random parameters were modeled: Intercepts for participant and item, and random slopes for adverb and aspect. Correlations between random terms were excluded. They lead to notorious computational problems and, probably, overfitting.

For the analysis of accuracy a similar model was used, including a full fixed effects structure in the three independent variables (ADVERB, ASPECT, VERB) and a rather flexible random effects structure. Random effects (error term in model equation) included random slopes for adverb and aspect. Correlations between random terms were excluded.

In further analysis the models for RT and accuracy were fitted to the best models by removing outliers and insignificant factors and/or their interactions and adjusting the initial random effects structure. In the final analysis, in both cases (RT and accuracy) we have a model with a simple fixed effects structure and a rather complicated random effects structure.

4.5.2. EEG data processing and analyses

EEG data were preprocessed using the Brain Vision Analyzer 2 software (Brain Products, Gilching). ERPs were time-locked to the critical adverb at each condition and had a duration of 2500 ms. All segments were calculated to a 200 ms pre-stream baseline yielding a total length of 2700 ms. ERPs were filtered (low cutoff 0.05 Hz) and corrected for artifacts (e.g., blinks, muscle and facial movements, as well as alpha waves). Segments including strong, visible artifacts (like pauses or periods of strong movement) were manually removed before proceeding. An ICA blink correction was performed for the remaining data using the Slope Algorithm for blink detection. After the blink correction,

---

[5] Version 3.5.1.
[6] Package version 1.1-21.

data were again inspected semi-automatically to monitor successful blink correction. Altogether, trials characterized by eye blinks, excessive muscle artifacts, or amplifier blocking were excluded. Artifacts were defined according to the following parameters: The maximal allowed voltage step per sampling point was 50 µV. The maximal allowed absolute difference of two values in a segment was 300 µV. The minimal allowed amplitude was −200 µV, and the maximal allowed amplitude was 200 µV. The lowest allowed activity (max–min) was 0.5 µV in an interval of 100 ms. In total, 20.5% of segments were rejected in Experiment 1 (out of 3160 segments altogether), and 18.3% of segments were rejected in Experiment 2 (out of 4134 segments altogether). This means that in total, 2512 artifact-free segments in Experiment 1 and 3380 artifact-free segments in Experiment 2 entered the averaging process. The segments correspond to the critical words, which were: temporal adverb, verb (an imperfective/perfective *l*-participle or infinitive), and object. ERPs were averaged offline within each experimental sentence type at the critical words for each subject at each electrode site. For visual purposes only, grand averages were smoothed with a 10 Hz low-pass filter. Statistical analyses concerning the relevant time windows (see below) were computed on nonfiltered data.

The following regions of interest (ROIs) were defined: right-anterior (C4, F4, F8, FC2, FC6), left-anterior (C3, F3, F7, FC1, FC5), midline-anterior (Fz, FCz, Cz), midline-posterior (CPz, Pz, POz), left-posterior (CP5, O1, P3, P7), right-posterior (CP6, O2, P4, P8). These regions of interest were chosen based on visual data inspection and previous studies on tense violations (see Part I (SPL 14(4)), section 2.1 and also section 2.3).

Statistical analysis was performed using R (R Core Team 2019) on a linux-machine.[7] For EEG data, linear mixed model was used as implemented in the lme4 package[8] (Bates et al. 2015). The models for the ERP data were fitted in two steps. First, a model for the data relevant for the comparison was fitted. This model contained the full fixed effects structure and random slopes for all predictors. We excluded correlations of random effects from the model. These parameters are notoriously hard to fit, prone to overfitting and rarely realistically interpretable. The second step (using the package LMERConvenienceFunctions (Trembley and Ransijn 2015)[9]) consisted in removal of outliers (function trim.data.frame()) and the refitting of the model for both random and fixed effects. Residuals of all cases of refitted models were approximately normally distributed. The planned comparisons were analysed with the *emmeans* package (Lenth 2019).[10]

---

[7] Version 3.6.1.
[8] Version 1.1-21.
[9] Version 2.10.
[10] Version 1.4.2.

4.6. Results

4.6.1. Behavioral results

The behavioral results (reaction times and accuracies) obtained for participle (Experiment 1) and infinitive (Experiment 2) conditions are structurally very similar altogether (see Figure 2 and 3). Importantly, it should be noted that VERB and both of its interactions are removed from the best model formula for accuracy as they turned out not to be significant (i.e., not to give a significant contribution to the model). Regarding the RT data, here also no modulation by VERB was found.
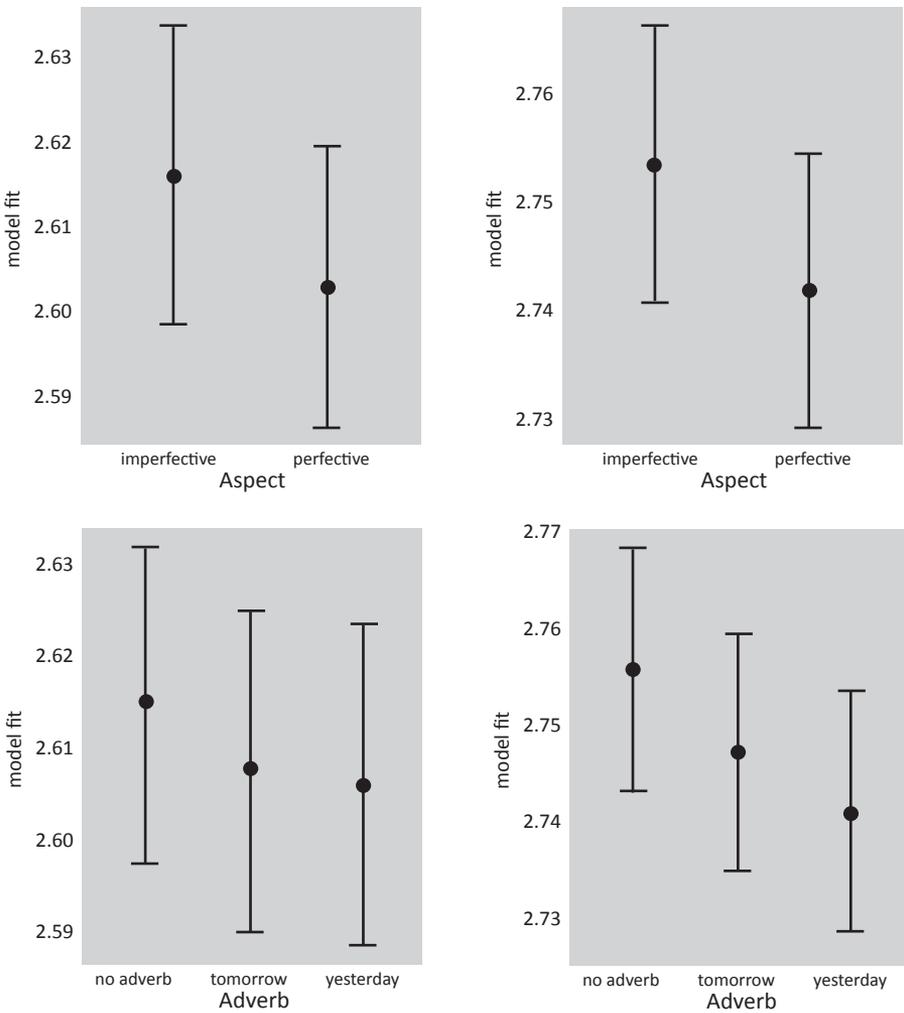
Figure 2. Behavioral data – reaction times. Experiment 1 (VERB: participle) – on the left side; Experiment 2 (VERB: infinitive) – on the right side
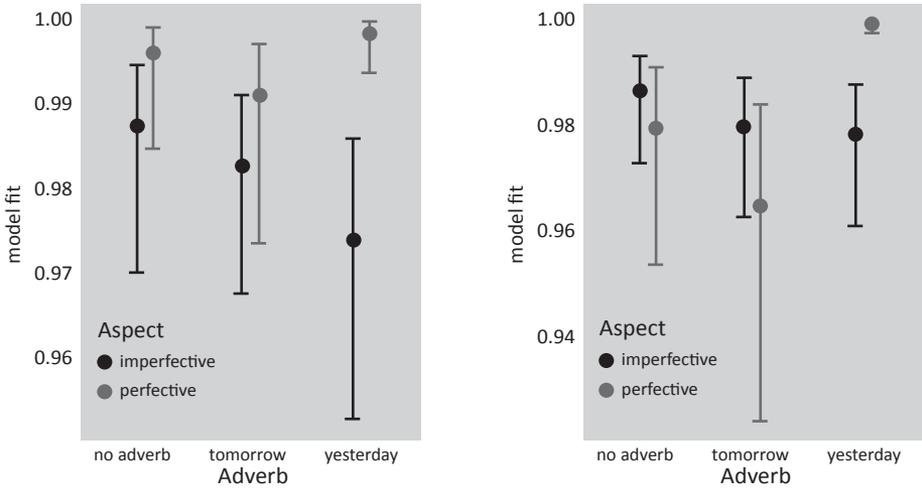
Figure 3. Behavioral data – accuracy. Experiment 1 (VERB: participle) – on the left side; Experiment 2 (VERB: infinitive) – on the right side

4.6.1.1. Reaction times

The following tables provide an overview of mean reaction (response) times for each condition separately for Experiment 1 (see Table 2) and Experiment 2 (see Table 3).

Table 2 Experiment 1 (VERB: participle): Descriptive statistics

Means and standard deviations (sd) over reaction times (RT) in milliseconds within each condition after averaging over participant first

Condition No.	ADVERB	ASPECT	mean.RT	sd.RT	N
1	none	imperfective	748 ms	276	19
2	none	perfective	685 ms	236	19
3	tomorrow	imperfective	759 ms	232	19
4	tomorrow	perfective	628 ms	192	19
5	yesterday	imperfective	745 ms	300	19
6	yesterday	perfective	594 ms	195	19

Table 3. Experiment 2 (VERB: infinitive): Descriptive statistics

Means and standard deviations (sd) over reaction times (RT) in milliseconds within each condition after averaging over participant first

Condition No.	ADVERB	ASPECT	mean.RT	sd.RT	N
1	none	imperfective	862 ms	262	22
2	none	perfective	765 ms	177	22
3	tomorrow	imperfective	842 ms	219	22
4	tomorrow	perfective	722 ms	194	22
5	yesterday	imperfective	771 ms	205	22
6	yesterday	perfective	656 ms	133	22

The results of the planned comparisons (relavant only for 'yesterday' conditions) are given in Table 4.

Table 4. Planned comparisons (only 'yesterday' conditions): Reaction times

	Contrast	estimate	SE	z.ratio	p.value
Comparison 1	'perfective_**participle**' vs. 'perfective_**infinitive**'	−0.0240	0.0152	−1.581	0.7
Comparison 2	'imperfective_**participle**' vs. 'imperfective_**infinitive**'	−0.0186	0.0152	−1.221	1.0
Comparison 3	'**imperfective**_participle' vs. '**perfective**_participle'	0.0215	0.0061	3.526	0.004
Comparison 4	'**imperfective**_infinitive' vs. '**perfective**_infinitive'	0.0160	0.0056	2.859	0.04

4.6.1.2. Accuracy

The following tables provide an overview of accuracies for each condition separately for Experiment 1 (see Table 5) and Experiment 2 (see Table 6).

Table 5. Experiment 1 (VERB: participle): Descriptive statistics. Accuracies. R = proportion of right answers, F = proportion of false answers

Condition No.	ADVERB	ASPECT	R (correct)	F (incorrect)	N
1	none	imperfective	91%	9%	909
2	none	perfective	97%	3%	910
3	tomorrow	imperfective	97%	3%	909
4	tomorrow	perfective	96%	4%	909
5	yesterday	imperfective	96%	4%	910
6	yesterday	perfective	99%	1%	910

Table 6. Experiment 2 (VERB: infinitive): Descriptive statistics. Accuracies. R = proportion of right answers, F = proportion of false answers

Condition No.	ADVERB	ASPECT	R (correct)	F (incorrect)	N
1	none	imperfective	97%	3%	1049
2	none	perfective	93%	7%	1055
3	tomorrow	imperfective	96%	4%	1054
4	tomorrow	perfective	90%	10%	1052
5	yesterday	imperfective	97%	3%	1054
6	yesterday	perfective	100%	0%	1055

The results of the planned comparisons (relavant only for 'yesterday' conditions) are given in Table 7.

Table 7. Planned comparisons (only 'yesterday' conditions): Accuracies

|	**Contrast**	*estimate*	*SE*	*z.ratio*	*p.value*
Comparison 1	'perfective_**participle**' vs. 'perfective_**infinitive**'	−0.8798	0.7292	−1.206	1.0000
Comparison 2	'imperfective_**participle**' vs. 'imperfective_**infinitive**'	−0.2857	0.4522	−0.632	1.0000
Comparison 3	'**imperfective**_participle' vs. '**perfective**_participle'	−1.9766	0.5118	−3.862	0.0009
Comparison 4	'**imperfective**_infinitive' vs. '**perfective**_infinitive'	−2.5707	0.6007	−4.280	0.0002

4.6.2. Electrophysiological results

The following section provides an overview of the results (main effects and interactions) of the ERP analysis conducted in four time windows (see also section 4.5.2). The time windows for the analysis were selected on the basis of visual inspection of grand average wave forms, and with reference to the previous relevant literature, in particular, literature related to the processing of tense and aspect violations (refer to section 2.1 in Part I (SPL 14(4)) as well as to the electrophysiological processing of ungrammatical sentences with two anomalies (refer to section 2.3 in Part I (SPL 14(4)). The critical items (on which effects were expected) are: **ADVERB**, **VERB**, and **OBJECT**. Below, we will discuss the outcomes of the planned comparisons separately for each relevant critical position (see section 3 in Part I (SPL 14(4)).

4.6.2.1. ADVERB position, 'yesterday' and 'tomorrow' conditions

At the ADVERB position, in the time window 300–500 ms post stimulus onset the mean potentials in the 'yesterday' conditions were more negative than

those in the 'tomorrow' conditions. In a later time window, 700–1000 ms post stimulus, we observe an opposite pattern: the mean potentials in the 'yesterday' conditions were more positive than those in the 'tomorrow' conditions (see Figure 4 and Figure 5).
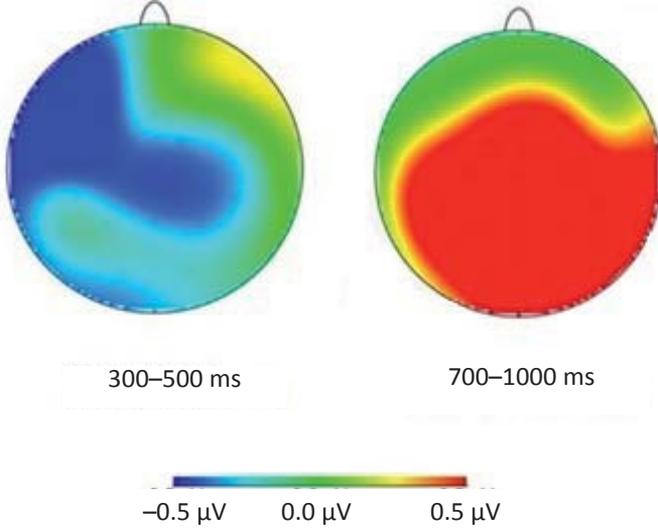

Figure 4. Headmaps depicting mean differences in voltage between 'yesterday' and 'tomorrow' conditions in the 300–500 ms and 700–1000 ms time windows post ADVERB onset
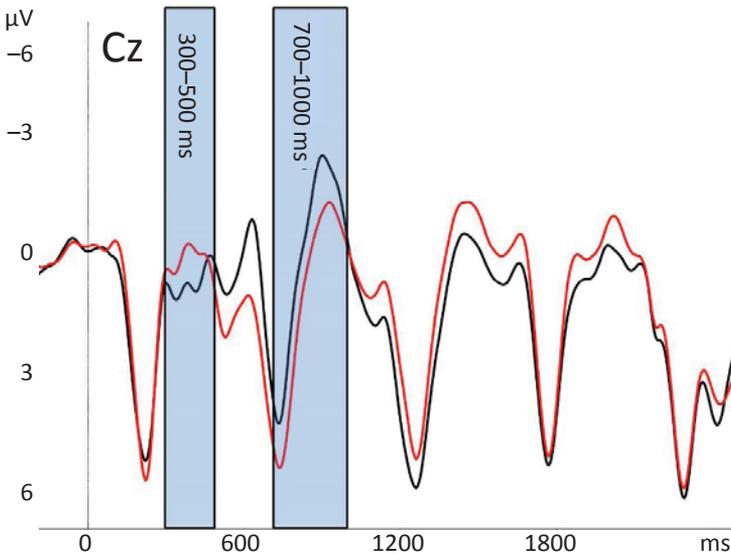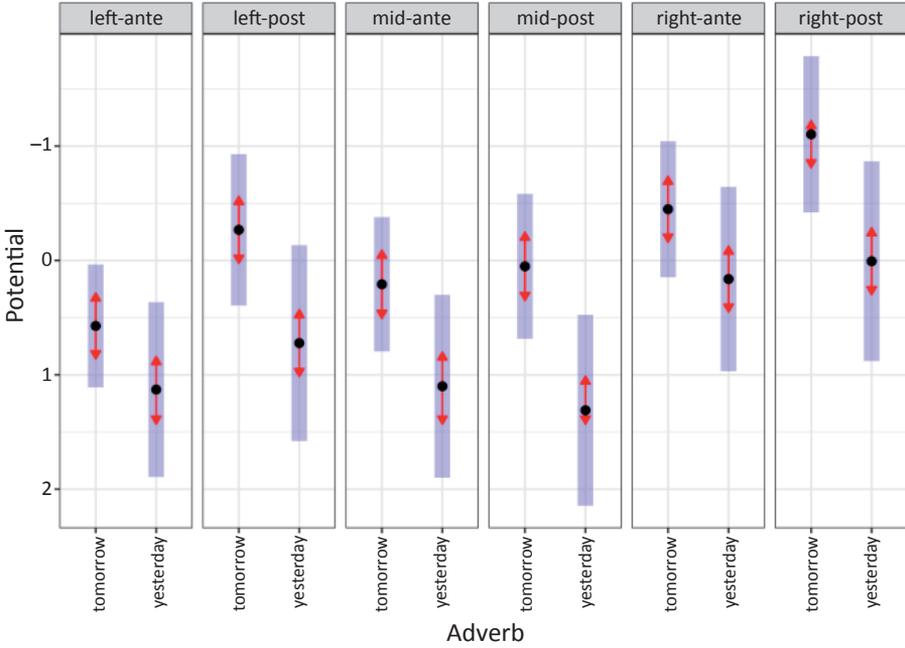

Figure 5. Comparison: 'yesterday' vs. 'tomorrow' conditions: Grand average pattern at one selected electrode site (Cz) in the analyzed time windows. Black line – congruent adverb ('tomorrow'), red line – incongruent adverb ('yesterday'). ADVERB onset at 0 ms

A single comparison between all 'yesterday' and 'tomorrow' conditions (where conditions are collapsed across VERB and ASPECT) conducted in these two time windows indeed revealed a significant effect of ADVERB. In the ealier time window (300–500 ms post ADVERB onset) this effect was significant only in two ROIs: left-anterior ($z = 3.056$, $p = 0.002$) and midline-anterior ($z = 2.008$, $p = 0.04$); see Figure 6. In the later time window (700–1000 ms post ADVERB onset) the effect of ADVERB was the strongest in all posterior ROIs (left-posterior: $z = -3.320$, $p = 0.0009$; midline-posterior: $z = -4.131$, $p < 0.0001$; right-posterior: $z = -3.721$, $p = 0.0002$), though the contrasts between 'tomorrow' and 'yesterday' conditions were also significant in the midline-anterior ($z = -2.929$, $p = 0.003$) and right-anterior ($z = -2.076$, $p = 0.04$) ROIs; see Figure 7.
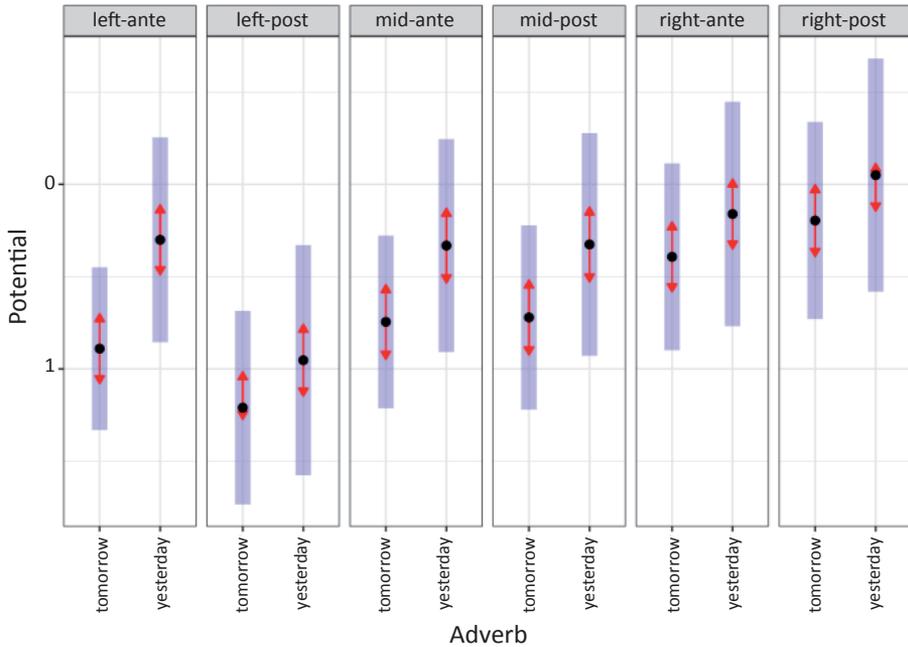

Figure 6. Comparison: 'yesterday' vs. 'tomorrow' conditions in the time window 300–500 ms post ADVERB onset. The black dots show the point estimates while the shaded areas are confidence intervals. The red arrows are constructed in such a way that overlapping arrows indicate a non-significant difference

Figure 7. Comparison: 'yesterday' vs. 'tomorrow' conditions in the time window 700–1000 ms post ADVERB onset. See Figure 6 for the description

4.6.2.2. VERB position, 'no adverb' conditions

At the VERB position in the time window 100–400 ms post VERB onset, the mean potentials for the perfective conditions were more negative than those for the imperfective conditions at the right-posterior, left-posterior and also weakly at the midline-posterior electrode sites. In a later time window, 700–900 ms post onset of the critical verb, the mean potentials in the perfective conditions were largely more positive compared to those in the imperfective conditions (see Figure 8 and Figure 9).
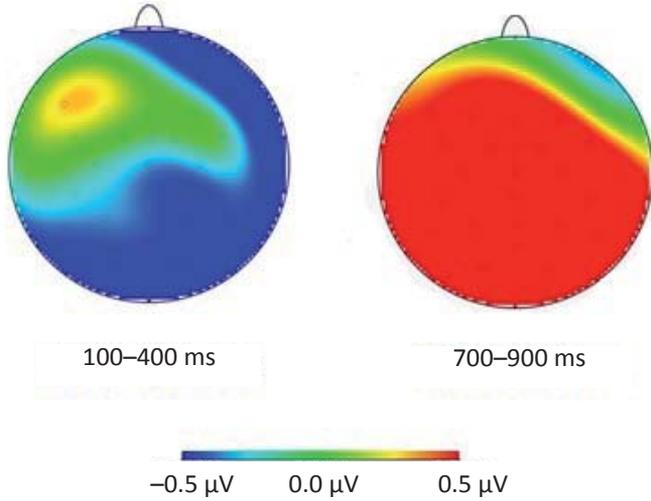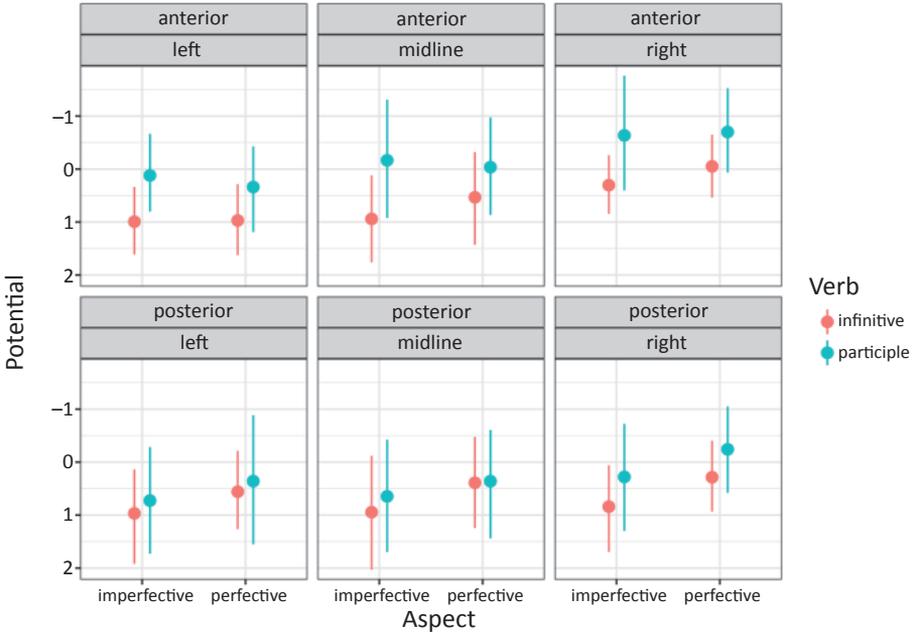
Figure 8. Headmaps depicting mean differences in voltage between perfective and imperfective conditions in the 100–400 ms and 700–900 ms time windows post VERB onset. Data restricted to 'no adverb' conditions
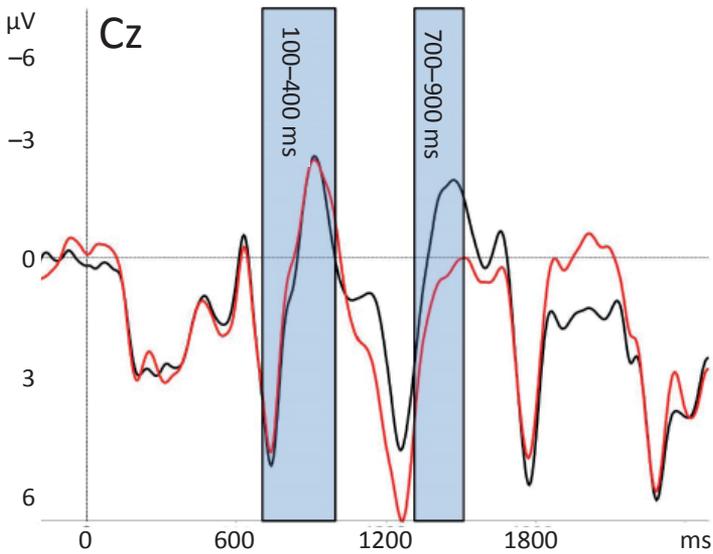

Figure 9. Comparison: perfective vs. imperfective conditions: Grand average pattern at one selected electrode site (Cz) in the analyzed time windows. Black line – imperfective, red line – perfective. Data restricted to 'no adverb' conditions. VERB onset at 600 ms

A 2×2 comparison of ASPECT*VERB was applied to all 'no adverb' conditions in these two time windows. In the earlier time window, 100–400 ms post VERB onset the differences in the voltages between perfective and imperfective conditions, which were detectable upon visual inspection, were not confirmed by the statistical analysis. The effect of ASPECT turned out to be not significant. Nor was there a significant ASPECT*VERB interaction in any of the ROIs. See Figure 10 for illustration.


Figure 10. Comparison: ASPECT*VERB in the 100–400 ms time window post VERB onset. Data restricted to 'no adverb' conditions. Per person averages over all electrodes are computed, from which bootstrapped confidence intervals are drawn visualizing the data with their between person variance

In the later time, 700–900 ms post VERB onset, the analysis revealed no ASPECT*VERB interaction either (see Figure 11), but there was a significant effect of ASPECT within ROIs. The difference between perfective and imperfective aspect was significant in all ROIs, indicating that the observed positivity was widely distributed but it had its maximum over posterior sites (cf. left-anterior: $t(49.9) = -2.936$, $p = 0.005$; midline-anterior: $t(56.6) = -3.653$, $p = 0.0006$; right-posterior: $t(50.0) = -2.471$, $p = 0.02$; left-posterior: $t(52.3) = -4.750$, $p < 0.0001$; midline-posterior: $t(56.7) = -5.940$, $p < 0.0001$; right-posterior: $t(52.6) = -3.744$, $p = 0.0004$); see Figure 12.
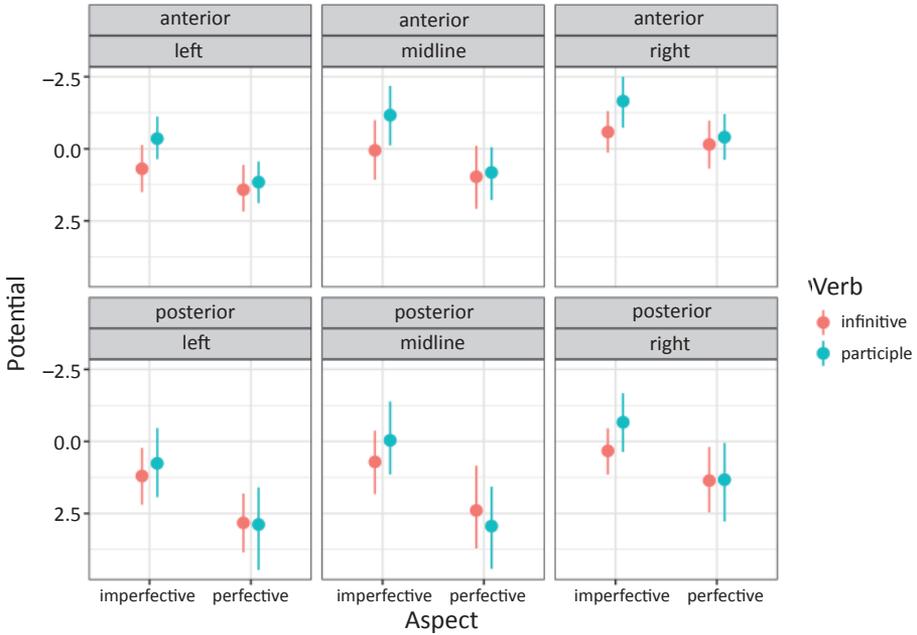
Figure 11. Comparison: ASPECT*VERB in the 700–900 ms time window post VERB onset. Data restricted to 'no adverb' conditions. See Figure 10 for the description
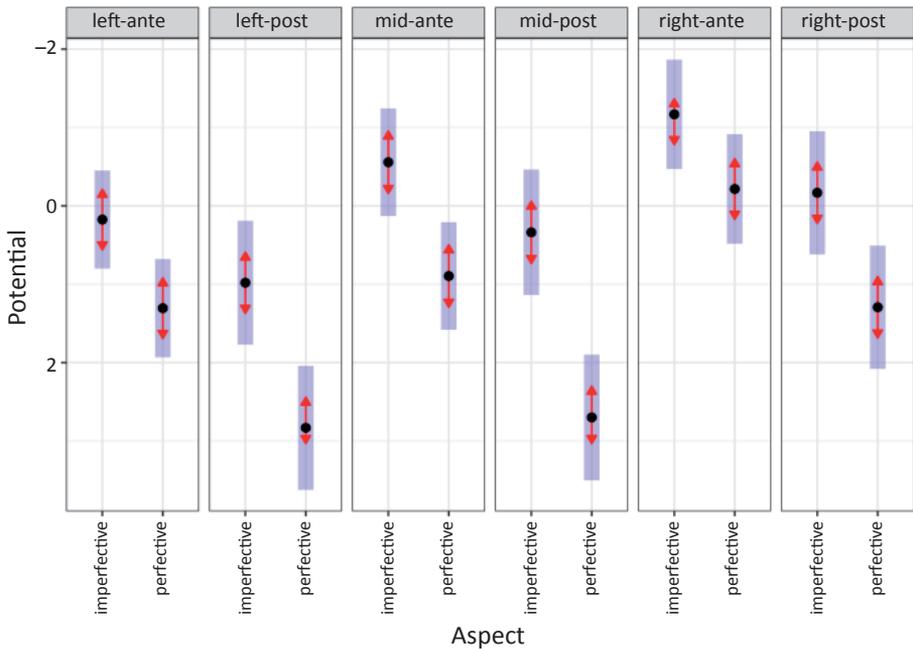
Figure 12. Depiction of the main effect of ASPECT in the 700–900 ms time window post VERB onset. Data restricted to 'no adverb' conditions. See Figure 6 for the description

4.6.2.3. VERB position, 'yesterday' conditions

A 2×2 comparison of ASPECT*VERB was applied to all 'yesterday' conditions at the VERB position (time window: 700–900 ms post VERB onset). The analysis revealed no significant main effects for ASPECT and VERB. Nor was the interaction ASPECT*VERB significant. See Figure 13 for illustration. Figure 14a and 14b depict grand average patterns in the relevant conditions on one selected electrode (Cz) in the analyzed time window. For reasons of space and because there are no statistically significant differences, voltage difference maps are not shown.
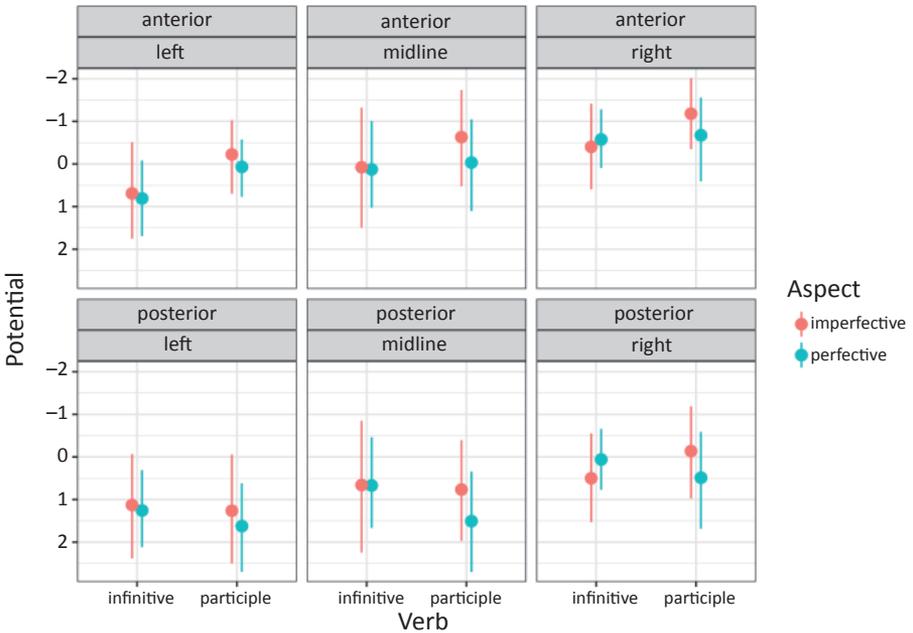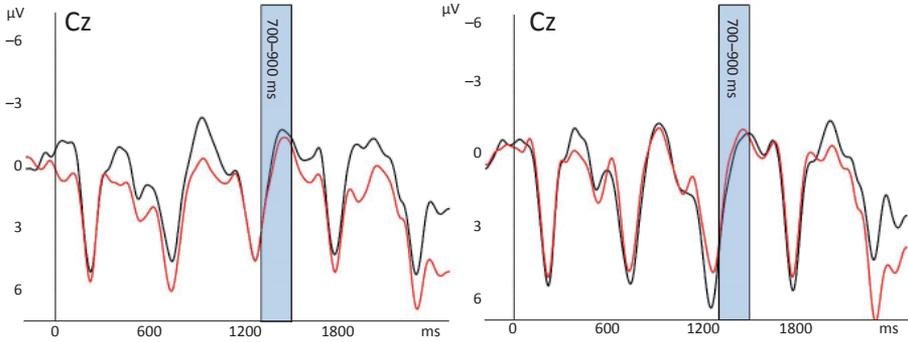

Figure 13. Depiction of the data for ADVERB = 'yesterday' in the 700–900 ms time window post VERB onset. See Figure 10 for the description

Figure 14a. Grand average patterns at one selected electrode site (Cz) in the analyzed time window 700–900 ms post VERB onset. VERB onset at 600 ms. Data restricted to 'yesterday' conditions. Imperfective conditions – on the left side; perfective conditions – on the right side. Black line – participle, red line – infinitive
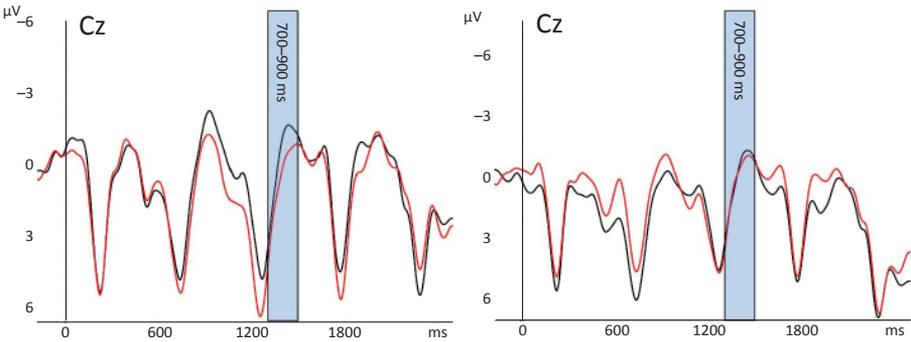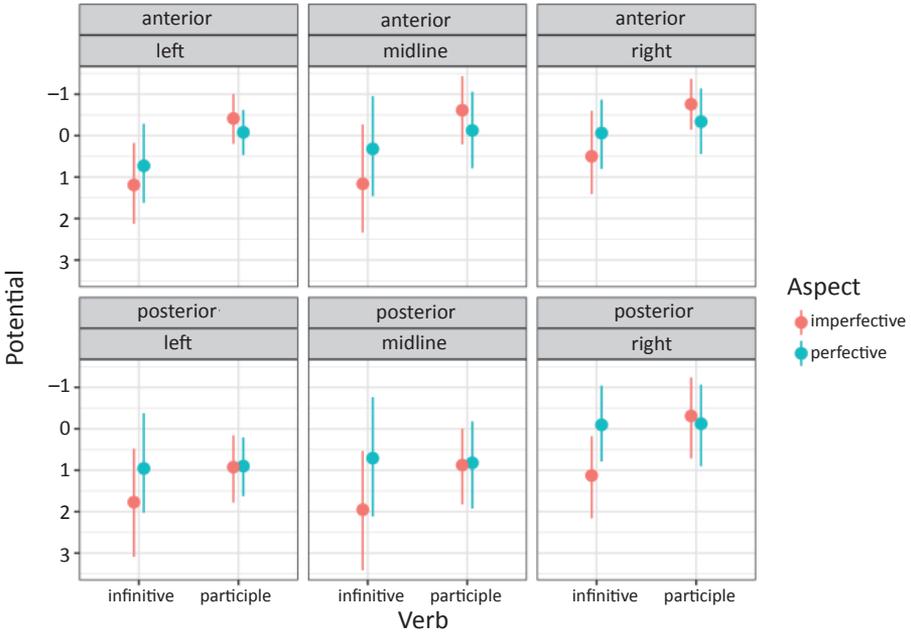

Figure 14b. Grand average patterns at one selected electrode site (Cz) in the analyzed time window 700–900 ms post VERB onset. VERB onset at 600 ms. Data restricted to 'yesterday' conditions. Participle conditions – on the left side; infinitive conditions – on the right side. Black line – imperfective, red line – perfective
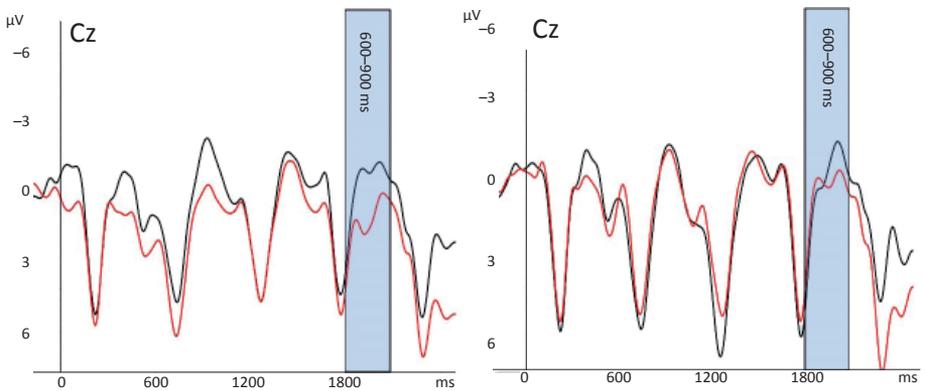
4.6.2.4. OBJECT position, 'yesterday' conditions

A 2×2 comparison of ASPECT∗VERB was applied to all 'yesterday' conditions at the OBJECT position (time window: 600–900 ms post OBJECT onset). The analysis revealed no significant main effects, neither for ASPECT nor for VERB. Furthermore, in the time window under consideration, threre was no significant interaction ASPECT∗VERB, though single comparisons within ROIs showed two cases of (marginally) significant ASPECT∗VERB pairwise interactions, namely in the midline-posterior ROI ($t(67.2) = 1.917$, $p = 0.0596$) and right-posterior ROI ($t(59.5) = 2.020$, $p = 0.0479$). See Figure 15 for illustration. Figure 16a and 16b depict grand average patterns in the relevant conditions on

one selected electrode (Cz) in the analyzed time window. For reasons of space and because there are no statistically significant differences, voltage difference maps are not shown.

Figure 15. Depiction of the data for ADVERB = 'yesterday' in the 600–900 ms time window post OBJECT onset. See Figure 10 for the description

Figure 16a. Grand average patterns at one selected electrode site (Cz) in the analyzed time window 600–900 ms post OBJECT onset. OBJECT onset at 1200 ms. Data restricted to 'yesterday' conditions. Imperfective conditions – on the left side; perfective conditions – on the right side. Black line – participle, red line – infinitive
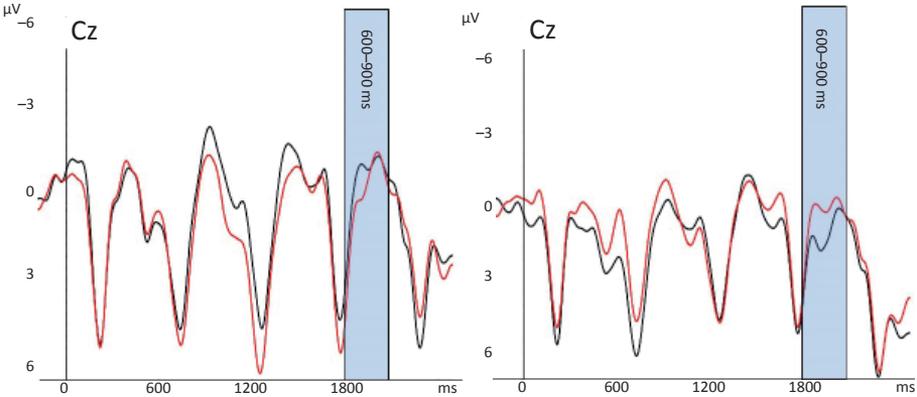
Figure 16b. Grand average patterns at one selected electrode site (Cz) in the analyzed time window 600–900 ms post OBJECT onset. OBJECT onset at 1200 ms. Data restricted to 'yesterday' conditions. Participle conditions – on the left side; infinitive conditions – on the right side. Black line – imperfective, red line – perfective

4.7. Discussion

4.7.1. Behavioral data

Under **the recency account** (see section 3.2 in Part I (SPL 14(4)), grammatical illusions were expected in ungrammatical sentences with the past tense modifier 'yesterday' and perfective participles but not in those with perfective infinitives. Accordingly, the expectation was to find longer response times and higher error rates (increase in acceptability due to grammatical illusion) for ungrammatical sentences with 'yesterday' and perfective participles than for ungrammatical sentences with 'yesterday' and perfective infinitives (no grammatical illusion). This expectation is clearly not fulfilled: in 'yesterday' conditions, perfective infinitives show longer response times (656 ms) than perfective participles (594 ms), though this difference is not significant (**Comparison 1**). As far as error rates are concerned in 'yesterday' conditions, in fact both in the case of perfective participles and perfective infinitives the participants were highly accurate (99% and 100%, respectively). Again the difference between these two conditions is not significant (**Comparison 1**).

As far as ungrammatical sentences with the past tense modifier 'yesterday' and imperfective complements are concerned, no grammatical illusions were expected, neither for participle nor for infinitive conditions. Accordingly, the expectation was that in 'yesterday' conditions, there should be no considerable differences in response times and error rates between participles and infinitives in the case at hand. Moreover, doubly anomalous ungrammatical sentences with perfective participles were expected to be perceived (due to a potential grammatical illusion) as at least as grammatical/acceptable as ungrammatical sentences with the mismatching past tense adverb 'yesterday'

and imperfective participles (no grammatical illusion). Accordingly, longer response times and higher error rates were expected for perfective participles relative to imperfective ones. By contrast, as no grammatical illusions were expected for ungrammatical sentences with the past tense modifier 'yesterday' and infinitives, the expection was that both perfective and imperfective infinitives would be correctly judged as ungrammatical and thus they would potentially result in comparable response times and accuracies.

The obtained results are not fully in line with these predictions. On the one hand, in 'yesterday' conditions no significant differences in response times and error rates between imperfective participles and imperfective infinitives were found (**Comparison 2**). However, contrary to expectation, in 'yesterday' conditions, perfective participles in fact show significantly shorter response times (594 vs. 745 ms) and significantly higher accuracy (99% vs. 96%) than imperfective participles (**Comparison 3**). Moreover, contrary to initial expectation, in fact exactly the same pattern of results was obtained for imperfective and perfective infinitives in 'yesterday' conditions. Also in this case perfective infinitives show significantly shorter response times (656 ms vs. 771 ms) and significantly higher accuracy (100% vs. 97%) (**Comparison 4**).

Note that the obtained findings are not compatible with the predictions formulated under **the coercion account** either (see section 3.2 in Part I (SPL 14(4)). Under this account, in 'yesterday' conditions grammatical illusions were expected to arise both for perfective and imperfective participles. By contrast, no grammatical illusions were expected to arise in infinitive conditions, neither for perfective nor imperfectives. Accordingly, in 'yesterday' conditions longer response times and higher error rates were expected for participles than infinitives, both in perfective and imperfective conditions. However, response times and error rates were expected not to differ significantly between both aspects, neither for participles nor for infinitives. Contrary to these predictions, as we have seen above, response times and accuracies did not significantly differ between participle and infinitive conditions, neither for perfectives (**Comparison 1**) nor for imperfectives (**Comparison 2**). However, contrary to expectation, there were in fact significant differences in response times and accuracies between perfectives and imperfectives, both in the participle (**Comparison 3**) and in the infinitive conditions (**Comparison 4**).

Taken together, the findings of the analysis of behavioral data seem to suggest that no "tense illusion" is at work in the case of compound futures with participial complements (potentially specified as [past]), as in ungrammatical sentences with the past tense modifier 'yesterday', they did give rise to the same behavior as compound futures with infinitival complements, which do not bear any [past] specification.[11]

---

[11] Recall from section 4.6.1 that the form of the verb turned out not to be a significant factor in the assumed statistical model.

4.7.2. Electrophysiological data

Like the behavioral data, also the ERP data have been analyzed to assess whether the presence of the adverb 'yesterday' (specified for [+past]) could give rise to an illusion of grammaticality for perfectives as *l*-participles, but not as infinitives. To this aim, two different hypotheses as formulated under the recency and the coercion account have been tested, separately for the VERB and the OBJECT position (see section 3.1 in Part I (SPL 14(4)). Regarding the VERB position, under **the recency account**, the predictions were that (a) infinitives should have an enhanced P600 relative to participles with perfectives, but not with imperfectives, and (b) perfectives should have an enhanced P600 relative to imperfectives with infinitives, but not with participles. Under **the coercion account**, the predictions were that (a) infinitives should have an enhanced P600 relative to participles, both for perfectives and imperfectives, and (b) perfectives should not have an enhanced P600 relative to imperfectives, neither for infinitives nor for participles. Our data does not allow us to choose between these hypotheses as neither of these predictions have been confirmed in the present study. A 2×2 comparison of ASPECT*VERB applied to all 'yesterday' conditions at the VERB position (time window: 700–900 ms post after VERB onset) revealed no significant effects at all, neither main effects nor interactions of ASPECT and VERB.

As far as the OBJECT position is concerned, under **the recency account**, the predictions were that (a) infinitives should have an enhanced negativity compared to participles for perfectives, but not imperfectives, and (b) perfectives should have a stronger negativity than imperfectives for infinitives, but not for participles. Under **the coercion account**, the predictions were that (a) infinitives should have a stronger negativity than participles for perfectives and imperfectives, and (b) there should be no difference between both aspects, neither for participles nor for infinitives. Also in this case our data does not allow us to choose between these two hypotheses. The predictions formulated under the recency account would be simply refuted by the lack of any significant interactions of ASPECT and VERB. The results do not support the predictions formulated under the coercion account either. If anything, the difference between infinitives and participles visible in the left-anterior ROI (see Figure 15), would in fact speak against this hypothesis. Note that – contrary to prediction – it is in fact the participle condition that is more negative for both aspects than the infinitive condition. Note furthermore that as mentioned in section 4.6.2.4, single comparisons within ROIs (ASPECT pairwise, VERB pairwise), conducted in the time window under consideration, revealed marginally significant differences in the midline-posterior and right-posterior ROIs. However, as the contrasts depicted in Figure 15 show – contrary to the expectation – perfective and imperfective conditions in fact differ for infinitives (but not for participles) in these two ROIs.

To conclude, also the electrophysiological results do not provide support for the assumption that matching TENSE specifications in different words of a sentence could cause grammatical illusions, similar to what has been described for other grammatical illusion phenomena.

Finally, recall that the other aim of the present study was to monitor the EEG correlates of mismatches between the future tense auxiliary and perfective aspect relative to the imperfective baseline as well as mismatches between the future auxiliary and the past tense modifier 'yesterday' relative to the 'tomorrow' baseline (see section 3.1 in Part I (SPL 14(4)).

Regarding the tense mismatch under consideration, we expected that in all 'tomorrow' vs. 'yesterday' comparisons the incongruent adverbial 'yesterday' would elicit a negativity (LAN) followed by a posterior positivity (P600). This prediction is confirmed in that we indeed found a biphasic pattern: an early negativity in the 300–500 ms time window post critical stimulus onset, which had a left- and midline-anterior distribution, followed by a late positivity in the 700–1000 ms time window post critical stimulus onset with a maximum in posterior areas. Given this, the observed negativity would resemble a LAN, which can be taken to reflect the actual detection of the anomaly (see, e.g., Münte et al. 1997) or the failure of a parsing operation (Hagoort 2003: 896). In our case, this would mean that the LAN reflects the failure in the tense agreement check between the future auxiliary and the past temporal adverbial, as there is obviously a conflict in the features: [+FUTURE] vs. [+PAST]. Or, as Baggio (2008: 51) suggests, since the parser cannot simultaneously solve the contradictory temporal constraints set up by the adverb and the auxiliary, the LAN could thus be taken to reflect the detection of an inconsistency in the set of temporal constraints.

The positivity in a later time window – 700–1000 ms post onset of the adverb might be taken to reflect enhanced processing costs related to a repair and/or reanalysis mechanism – the attempt of the parser at resolving the inconsistency between the future auxiliary and the past temporal adverb and constructing a representation of an ungrammatical sentence (see Molinaro et al. 2008: 964–965).

Regarding the aspect mismatch under consideration, we expected that perfectives should differ from imperfectives both in the participle and infinitive conditions, without there being significant differences between participles and infinitives. More specifically, we expected that the mismatching perfective aspect violating the selectional requirements of the future auxiliary, which is only compatible with imperfective complements, would elicit a P600, possibly preceded by an early negativity, both in the case of participles and infinitives. These predictions could be confirmed. No ASPECT*VERB interaction effect was found, in neither of the investigated time windows: 100–400 ms and 700–900 ms post onset of the verb. In the earlier time window, although the mean

potentials for the mismatching perfective conditions were more negative than those for the matching imperfective conditions, especially over posterior sites, these differences were statistically not significant. However, it is interesting to note that – as far as earlier time windows are concerned – aspectual mismatches seem to elicit negativities with a different distribution from that found in the case of tense mismatches. Whereas the latter normally have a left-anterior scalp distribution, the former tend to have a more posterior distribution (see Dillon et al. 2012; Zhang and Zhang 2008). However, more research is needed to confirm this initial observation (see also Dillon et al. 2012: 337–338).

In the later time window, 700–900 ms after onset of the verb, the difference between mismatching (perfective) and matching (imperfective) aspect was significant at all electrode sites, though the observed effect was the strongest in the posterior regions. Altogether, the results indicate that the aspect violation is very salient for the processor and the anomalous perfective aspect elicits a broadly distributed positivity response. Note that, as has been pointed out in the literature, the P600 might indeed be sensitive to the saliency of a violation (see, e.g., Coulson et al. 1998; Drenhaus et al. 2005; Nevins et al. 2007).

5. Conclusion and outlook

The present study had two aims: (i) to monitor the EEG correlates of mismatches between future tense auxiliary and perfective aspect relative to the imperfective baseline as well as mismatches between the future auxiliary and the past tense modifier 'yesterday' relative to the 'tomorrow' baseline, and (ii) to assess whether the presence of the adverb "yesterday" (specified for [+past]) could give rise to an illusion of grammaticality for perfectives as *l*-participles, but not as infinitives. Regarding the first goal, the present study showed that aspectual mismatches lead to a widely distributed positivity with a posterior maximum (P600) while tense mismatches resulted in biphasic (LAN + P600) signature. These findings are in line with previous research on tense/aspect mismatches.

Regarding the second goal and the central question of the present paper, the reported ERP study of the processing of compound (infinitival and participial) future constructions in Polish does not provide evidence for the hypothesis that matching TENSE specifications in different words of a sentence can cause grammatical illusions, similar to what has been described for other grammatical illusion phenomena. This conclusion is consistent with the claim found in the literature that the *l*-participle does not have the past tense specification (see, e.g., Błaszczak et al. 2014; Dornisch 1997; Spencer 2001; Witkoś 1998, contra, Fisiak et al. 1978 and Tajsner 1999). In these analyses, future or past time reference are derived for the whole periphrastic form, i.e., the

lexical verb and an auxiliary (which is covert (null) in the case of past tense; see Dornisch 1997: 188). This argumentation and the findings of the reported study in this paper are also consistent with the result of the ERP study by Bos et al. (2013), who examined violations of a past tense context (*zonet* 'a moment ago') with a noncongruent nonpast periphrastic verb form (e.g., *gaat malen* 'will grind') as compared to a congruent past periphrastic verb form (e.g., *heeft gemalen* 'has ground'). Importantly, though both periphrastic verb forms contained a present tense auxiliary: *gaat* 'will' and *heeft* 'has' respectively, only in the former case a present tense auxiliary evoked a positivity in the past tense context. Based on this observation, Bos et al. (2013) argue that "[t]ense violations only cause a positivity if they lead to an incongruent time reference" (p. 296). In other words, what matters is not just the tense form of the auxiliary as such but rather "the time reference of the complete verb forms" (Bos et al. 2013: 283). If correct, in the context of the present ERP study, this could be taken to mean that the time reference of the complete periphrastic verb form was future and that the superficial morphological similarity of the participial complement of the future auxiliary to a past tense form on its own is not enough to cause any tense related grammatical illusions. This conclusion would be in line with the observation from the literature on agreement attraction (see Bock and Eberhard 1993) that pseudo-plurals, that is, distractors that sound or look like plural nouns but are in fact singular nouns, do not elicite attraction errors, which also suggests that "a superficially plural-like appearance is not sufficient to cause interference during the computation of agreement" (Häussler 2012: 84).

If it is true – as the findings of the present study suggest – that TENSE does not indeed belong to the features that are relevant for grammatical illusions, the question is what differentiated features or grammatical constraints that are subject to grammatical illusions from those that are not (refer to the discussion in section 2.2 in Part I (SPL 14(4)). One factor that seems to matter is the question of how a feature is acquired by a given element: is it acquired in a purely syntactic way (by syntactic mechanisms), as is the case with number on verbs or structural case on nouns, or is it acquired lexically, as is the case with number on pronouns or lexical case on nouns (see Bader and Meng 1999; Bock et al. 2004; Eberhard et al. 2005, among others, for the relevant discussion). Tense seems to be lexically specified on auxiliaries but the time reference is compositionally derived for the whole periphrastic verb form.

Another important factor is the mechanism by which a given information is retrieved. Phillips et al. (2011) discuss two types of retrieval: by search and by content-addressable access. The first type of retrieval involves "pairwise comparisons […] between the information desired and candidate encodings in memory" (p. 149). By this type of search information can be selectively targeted in specific structural locations, thus it should be possible "to avoid

interference from similar information in strucuturally irrelevant locations"
(Phillips et al. 2011: 150). By constrast, the second type of retrieval involves
a simultaneous probing of the entire memory "with a set of cues to the desired
information" (p. 149). In this type of cue-based retrieval, grammatically illicit
candidates only partially matching the relevant retrieval cues can be activated
and "even partially activated constituents may mislead the comprehenders into
a false impression of grammaticality" (Phillips et al. 2011: 150).

In view of the negative findings of the present paper as to possible tense-
related grammatical illusion and given the assumption – as mentioned above
– that the time reference in compound verb forms is compositionally derived,
it seems reasonable to assume that in the case of tense processing the first type
of retrieval, that is, a more constrained and selective retrieval by search, is in-
volved. However, our study can be understood only as a starting point towards
the question of whether TENSE is one of the features that are relevant for
grammatical illusions rather than providing a conclusive answer. More stud-
ies would be needed to confirm or disconfirm the negative findings of the ERP
study reported in this paper. For example, one should examine whether the
discussed tense and aspect anomalies would elicit ERP responses similar to
those obtained in the present study if the past tense adverb instead of follow-
ing the future auxiliary preceded it, as indicated in (19).

(19) a. ***Wczoraj**   Janek   *będzie*   <u>malował</u>   pokój   Zosi.
         yesterday   Janek   will   paint$_{IPFV.PTCP.SG.M}$ room   (of) Zosia
         ('*Yesterday Janek will be painting Zosia's room.')
   b. ***Wczoraj**   Janek   *będzie*   <u>pomalował</u>   pokój   Zosi.
         yesterday   Janek   will   paint$_{PFV.PTCP.SG.M}$ room   (of) Zosia
         ('*Yesterday Janek will paint Zosia's room.')

One potential problem with the present study is that in experimental sen-
tences the critical words appear next to each other. It could thus be that some
potential effects were masked. In a future study more lexical material could be
inserted between the critical words to increase the distance between them (see
(20)) and to see whether any additional effects become visible.

(20) *Janek   *będzie*   <u>z Marysią</u>   **wczoraj**   <u>po południu</u>   *przeczytał*
     Janek   will   with Mary   yesterday   in afternoon   read$_{PFV.PTCP. SG.M}$
     książkę.
     book
     ('*Janek will yesterday in the afternoon read a book with Mary.')

Also it might have been the case that the participants have developed spe-
cific processing strategies, though this seems rather unlikely. As revealed by
debriefing interviews, the participants were mostly concerned with remem-
bering words (which was needed to successfully complete the probe detection

task) as they thought that the primary aim of the experiment was to check their multitasking abilities and short term memory (see also section 4.2, footnote 3).[12] In a future experiment, instead of presenting the whole experimental material to each participant, different versions (also comparing different word orders) could be used, so that each participant would see only a portion of the relevant material. Also a higher number of different types of fillers with a different structure from the experimental ones should be prefereably used. This would help to avoid syntactic priming across experimental sentences, to avoid boredom, and to avoid strategy development by the participants. Also in future research, in order to find out about a possible tense illusion one could contrast ratings elicited under time pressure and ratings elicited without time pressure. This would only work if filler sentences were included, ideally stimuli from an experiment that already led to established linguistic illusions.[13]

Acknowledgements

References

Bader Markus, Meng Michael (1999). Case attraction phenomena in German. Manuscript, University of Jena. (http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.112.2979&rep=rep1&type=pdf (accessed: 28 August 2018)).

Bader Markus, Meng Michael, Bayer Josef (2000). Case and reanalysis. *Journal of Psycholinguistic Research* 29, 37–52.

Baggio Giosuè (2008). Processing temporal constraints: An ERP study. *Language Learning* 58, 35–55.

---

[12] An anonymous reviewer draws our attention to yet another problematic issue. The task of probe detection has an influence on participants' reading strategies and sentence processing, and can lead to different EEG outcomes compared to other tasks. The grammaticality judgment task, on the other hand, might have made the participants parse the sentences rather more indepth than they would have under natural conditions, making a linguistic illusion more unlikely. All of these choices would heighten the risk of weakening existing effects, thus making false negatives more likely. In future experiments these aspects should be taken into consideration.

[13] We thank an anonymous reviewer for suggesting this to us.

BATES Douglas, MÄCHLER Martin, BOLKER Ben, WALKER Steve (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1), 1–48. R package lme4 version 1.1-17. DOI: 10.18637/jss.v067.i01.

BŁASZCZAK Joanna, JABŁOŃSKA Patrycja, KLIMEK-JANKOWSKA Dorota, MIGDALSKI Krzysztof (2014). The riddle of future tense in Polish: How much 'future' is there in 'future tense'? In *Future Tense(s) / Future Time(s)*, Philippe DE BRABANTER, Mikhail KISSINE, Saghie SHARIFZADEH (eds.), 165–204. Oxford: Oxford University Press.

BOCK Kathryn, EBERHARD Kathleen M. (1993). Meaning, sound and syntax in English number agreement. *Language and Cognitive Processes* 8, 57–99.

BOCK Kathryn, EBERHARD Kathleen M., CUTTING J. Cooper (2004). Producing number agreement: How pronouns equal verbs. *Journal of Memory and Language* 51, 251–278.

BOS Laura S., DRAGOY Olga, STOWE Laurie A., BASTIAANSE Roelien (2013). Time reference teased apart from tense: Thinking beyond the present. *Journal of Neurolinguistics* 26, 283–297.

COULSON Seana, KING Jonathan W., KUTAS Marta (1998). Expect the unexpected: Event related brain response to morphosyntactic violations. *Language and Cognitive Processes* 13, 21–58.

DILLON Brian, NEVINS Andrew, AUSTIN Alison C., PHILLIPS Colin (2012). Syntactic and semantic predictors of tense in Hindi: An ERP investigation. *Language and Cognitive Processes* 27(3), 313–344.

DORNISCH Ewa (1997). Auxiliaries and functional projections in Polish. In *Formal Approaches to Slavic Linguistics (FASL) IV: The Cornell Meeting 1995*. (= *Michigan Slavic Materials* 39), Wayles BROWNE, Ewa DORNISCH, Natasha KONDRASHOVA, Draga ZEC (eds.), 183–209. Ann Arbor: Michigan Slavic Publishers.

DRENHAUS Heiner, FRISCH Stefan, SADDY Douglas (2005). Processing negative polarity items: When negation comes through the backdoor. In *Linguistic Evidence: Empirical, Theoretical, and Computational Perspectives* (= *Studies in Generative Grammar* 85). Stephan KEPSER and Marga REIS (eds.), 145–165. Berlin: de Gruyter.

EBERHARD Kathleen M., CUTTING J. Cooper, BOCK Kathryn (2005). Making syntax of sense: Number agreement in sentence production. *Psychological Review* 112, 531–559.

FISIAK Jacek, LIPIŃSKA-GRZEGOREK Maria, ZABROCKI Tadeusz (1978). *An Introductory English-Polish Contrastive Grammar*. Warszawa: Państwowe Wydawnictwo Naukowe.

JASPER Herbert H. (1958). The ten-twenty electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology* 10, 371–375.

HÄUSSLER Jana (2012). *The Emergence of Attraction Errors during Sentence Comprehension*. PhD thesis, University of Konstanz (revised version for publication).

HAGOORT Peter (2003). Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *Journal of Cognitive Neuroscience* 15, 883–899.

LENTH Russell (2019). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.4.2. https://CRAN.R-project.org/package=emmeans.

ŁAZIŃSKI Marek (2006). *O panach i paniach (Polskie rzeczowniki tytularne i ich asymetria rodzajowopłciowa)* [About gentlemen and ladies (Polish nouns of address and their asymmetry in gender and sex)]. Warszawa: Wydawnictwo Naukowe PWN.

Mikoś Michael J. (1985). Alternative forms of the future imperfective tense in Polish. *Slavic and East European Journal* 29, 448–460.

Molinaro Nicola, Kim Albert, Vespignani Francesco, Job Remo (2008). Anaphoric agreement violation: An ERP analysis of its interpretation. *Cognition* 106, 963–974.

Münte Thomas F., Matzke Mike, Johannes Sönke (1997). Brain activity associated with syntactic incongruencies in words and pseudo-words. *Journal of Cognitive Neuroscience* 9, 318–329.

Nevins Andrew, Dillon Brian, Malhotra Shiti, Phillips Colin (2007). The role of feature-number and feature-type in processing Hindi verb agreement violations. *Brain Research* 116(4), 481–494.

Oldfield R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* 9, 97–113.

Phillips Colin, Wagers Matthew W., Lau Ellen F. (2011). Grammatical illusions and selective fallibility in real-time language comprehension. In *Experiments at the Interfaces* (= *Syntax and Semantics* 37), Jeffrey T. Runner (ed.), 153–186. Bingley: Emerald.

R Core Team. 2019. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Spencer Andrew (2001). The paradigm-based model of morphosyntax. *Transactions of the Philological Society* 99(2), 279–313.

Tajsner Przemyslaw (1999). A minimalist analysis of compound future in Polish. *Poznań Studies in Contemporary Linguistics* 33, 119–131.

Tremblay Antoine, Ransijn Johannes (2015). LMERConvenienceFunctions: Model Selection and Post-hoc Analysis for (G)LMER Models. R package version 2.10. https://CRAN.R-project.org/package=LMERConvenienceFunctions

Witkoś Jacek (1998). *The Syntax of Clitics: Steps towards a Minimalist Account*. Poznań: motivex.

Zhang Yaxu, Zhang Jingting (2008). Brain responses to agreement violations of Chinese grammatical aspect. *NeuroReport* 19(10), 1039–1043.

Joanna Błaszczak
Uniwersytet Wrocławski
Wydział Filologiczny
Instytut Filologii Angielskiej
ul. Kuźnicza 22
50-138 Wrocław
joanna.blaszczak(at)uwr.edu.pl

Juliane Domke
Forschungsgruppe Soziale Neurowissenschaften
Max-Planck-Gesellschaft
Campus Nord, Haus 5
Humboldt-Universität zu Berlin
Philippstrasse 13
10099 Berlin
juliane.domke(at)gmail.com