



## The difference between “giving a rose” and “giving a kiss”: Sustained neural activity to the light verb construction



Eva Wittenberg<sup>a,b,\*</sup>, Martin Paczynski<sup>a</sup>, Heike Wiese<sup>b</sup>, Ray Jackendoff<sup>a</sup>, Gina Kuperberg<sup>a,c</sup>

<sup>a</sup> Tufts University, United States

<sup>b</sup> Potsdam University, Germany

<sup>c</sup> Massachusetts General Hospital, United States

### ARTICLE INFO

#### Article history:

Received 17 January 2012

revision received 6 February 2014

#### Keywords:

Event-related potential  
Sentence processing  
Light verb constructions  
Argument structure  
Syntax-semantics interface  
Sustained negativity

### ABSTRACT

We used event-related potentials (ERPs) to investigate the neurocognitive mechanisms associated with processing light verb constructions such as “give a kiss”. These constructions consist of a semantically underspecified light verb (“give”) and an event nominal that contributes most of the meaning and also activates an argument structure of its own (“kiss”). This creates a mismatch between the syntactic constituents and the semantic roles of a sentence. Native speakers read German verb-final sentences that contained light verb constructions (e.g., “Julius gave Anne a kiss”), non-light constructions (e.g., “Julius gave Anne a rose”), and semantically anomalous constructions (e.g., “Julius gave Anne a conversation”). ERPs were measured at the critical verb, which appeared after all its arguments. Compared to non-light constructions, the light verb constructions evoked a widely distributed, frontally focused, sustained negative-going effect between 500 and 900 ms after verb onset. We interpret this effect as reflecting working memory costs associated with complex semantic processes that establish a shared argument structure in the light verb constructions.

© 2014 Elsevier Inc. All rights reserved.

### Introduction

Most theories of argument structure assume a tight coupling between syntactic and semantic structure, such that each noun phrase maps onto a single semantic role. And, indeed, in most sentences, this is the case. For example, in a sentence like “Julius gave Anne a rose”, the giver (the Agent) is associated with the subject of the verb (“Julius”), the givee (the Recipient) is expressed as the indirect object (“Anne”), and the gift (the Theme) is expressed as the direct object (“rose”).

However, consider a sentence like “Julius gave Anne a kiss”—the so-called light verb construction. These constructions are complex predicates whose verbs are said

to be semantically “light”, communicating only lexical and grammatical aspect, and the directionality of the action; the bulk of the predicative meaning stems from the event nominal within the construction (Butt, 2010; Wiese, 2006). While in a non-light construction such as “give someone a rose”, the verb “give” means “to hand over”, in “give someone a kiss”, the verb “give” only denotes a general sense of transfer and the event nominal “kiss” conveys the action itself. Thus, Julius acts not only as the Agent of the verb “give”, but also as the Agent of the direct object “kiss”, while Anne is both the Recipient of “give” and the Patient of “kiss”. This phenomenon is known as “argument sharing” (Baker, 1989; Butt, 2010; Durie, 1988; Jackendoff, 1974; Müller, 2010), and it violates the tight coupling of semantic and syntactic structure.

There have been several theoretical attempts to reconcile the lack of a direct correspondence between semantic

\* Corresponding author at: Center for Cognitive Studies, Tufts University, 11 Miner Hall, Medford, MA 02155, United States.

E-mail address: [eva.wittenberg@tufts.edu](mailto:eva.wittenberg@tufts.edu) (E. Wittenberg).

and syntactic argument structure in the light verb construction (Hale & Keyser, 1993, 2002; Goldberg, 2003). In this paper, we follow the Parallel Architecture framework (Culicover & Jackendoff, 2005; Wittenberg & Piñango, 2011), which allows both syntactic and semantic structure to be built independently, though the two are linked through a grammatical function tier (for further discussion, see Müller & Wechsler, *in press*; Wittenberg et al., *in press*). According to this theory, when a verb appears in a light verb construction with certain event nominals, the process of argument sharing is triggered: the arguments provided by the verb (in the case of “give”, the Agent, Patient, and Theme), and the arguments provided by the noun (in the case of “kiss”, Agent and Patient) need to be aligned.

As a result of the mechanisms that, according to the Parallel Architecture, are engaged during argument sharing (Culicover & Jackendoff, 2005, pp. 222–225), we predict that light verb constructions should incur processing costs during comprehension. Note that this hypothesis goes against what might be predicted on the basis of the frequency of light verb constructions, which, despite their complexity, are commonly encountered in everyday language. For example, according to the PropBank corpus (Palmer, Gildea, & Kingsbury, 2005), the most common English verbs appearing within light verb constructions, such as *take*, *have*, *make*, *do*, and *give*, are among the twenty most frequent verbs in English. More importantly, these verbs are more frequently encountered within light than non-light verb constructions (Wittenberg & Piñango, 2011). Thus, in the absence of other factors, these frequency data alone would predict reduced processing costs in association with the more frequent light verb construction than the less frequent non-light construction.

There have been only a few behavioral experiments examining light verb constructions. First, in a recent study, Wittenberg and Piñango (2011) asked participants to listen to German light verb constructions (e.g. “Weil der Student seiner Kommilitonin vor dem Seminar eine Zusammenfassung gab, spendierte sie ihm letzte Woche einen Kaffee”; English literal translation: “Because the student to his fellow student before class a summary gave, she bought him coffee last week.”). These light verb constructions were compared with non-light constructions using the same verbs (e.g. German: “Weil der Student seiner Kommilitonin vor dem Seminar einen Kugelschreiber gab, spendierte sie ihm letzte Woche einen Kaffee”; English literal translation: “Because the student to his fellow student before class a pen gave, she bought him coffee last week.”), or the same nouns (e.g. German: “Weil der Student seiner Kommilitonin vor dem Seminar eine Zusammenfassung kopierte, spendierte sie ihm letzte Woche einen Kaffee”; English literal translation: “Because the student for his fellow student before class a summary copied, she bought him coffee last week.”). The default Subject–Object–Verb (SOV) word order in German allowed the authors to probe processing costs at the critical verb where they predicted the effects of argument sharing would be most prominent. After these critical verbs, letter-string probes appeared on a screen and participants made a lexical decision about these probes. Participants were slower to respond to

probes appearing 300 ms after the offset of the verbs in the light verb constructions, compared to the two non-light constructions. The authors interpreted this as evidence for an increased processing load in computing light verb constructions (see also Piñango, Mack, & Jackendoff, *in press*, for similar findings in English).

In another recent behavioral study, Wittenberg and Snedeker (*in press*) used a conceptual sorting task to explore the argument structure of light verb constructions in English. During a training phase, participants were trained to sort pictorial depictions of events by the number of thematic roles they encoded (e.g. a picture of man giving a woman some flowers would be classified as a ‘three role’ event: man, woman, flowers). They were then asked to sort a mix of pictures and written sentences into these different types of event structures (with different numbers of thematic roles). Despite the fact that they have three syntactic arguments, events described by light verb constructions (e.g. “The teenager is giving his rival a kick”) were most frequently grouped with event structures with two semantic roles (Agent–Patient events, e.g. “The cowboy is taming the pony”). This suggests that light verb constructions do, indeed, typically involve a non-canonical mapping between semantic and syntactic event structure. However, in about a quarter of cases, the light verb constructions were grouped with three-role event structures (e.g. Source–Theme–Goal Events, like “The businessman is passing pamphlets to the pedestrians”). This in-between pattern provided indirect evidence for argument sharing; that is, light verb constructions may be intrinsically associated with two different argument structures that can be active at the same time: an Agent–Patient non-canonical argument structure in which the number of semantic and syntactic arguments mismatch, and a Source–Theme–Goal canonical structure in which the number of semantic and syntactic arguments match.

Together, these behavioral studies provide some evidence that both the processing and final interpretation of light verb constructions involve argument sharing. Nevertheless, there are some limitations in the interpretation of the results. First, Wittenberg and Piñango (2011) used a cross-modal lexical decision task, which imposes dual task demands, potentially altering participants’ processing of the sentences (see Pickering, McElree, Frisson, Chen, & Traxler, 2006, for a critique of this method). Second, Wittenberg and Snedeker (*in press*) probed participants’ final interpretation of these constructions, rather the time course of their online neural processing.

There has only been one study investigating neural activity associated with light verb constructions. In an MEG study, Briem et al. (2010) carried out three experiments in German. They contrasted light verbs like “geben” (“give”) with non-light verbs like “erwarten” (“expect”), either by themselves (Experiment 1), presented together with a subject pronoun (Experiment 2), or in object-verb-subject order (Experiment 3). In all experiments, light verbs (e.g. “geben”/“give”) evoked less activity than non-light verbs (e.g. “erwarten”/“expect”).

The authors interpreted these findings as reflecting reduced lexical processing due to the semantic underspecification of light verbs. At first glance, these findings appear

to contradict the behavioral findings described above, namely more cost associated with light verb constructions than non-light constructions. However, there are several confounds that limit the interpretation of Briem and colleagues' study. First, the non-light verbs were both longer and more morphologically complex than the light verbs (the stimuli can be found in the appendix of Briem, 2009); both of these differences can explain the early effects in visual regions observed in all three experiments. Also, in the second experiment, several of the pairings between the non-light verbs and subject pronouns resulted in ungrammatical phrases, which is likely to have contributed to increased neural activity (see Friederici & Frisch, 2000). Finally, the object–verb–subject word order in the third experiment is very uncommon in German, so it is difficult to generalize the results to more naturalistic language comprehension. Moreover, because the authors did not analyze activity after the verb, any effects of argument sharing associated with light verb constructions may have been missed altogether – for example, under the Parallel Architecture, one would expect effects from argument sharing on the post-verbal subject, which would receive a semantic role from the object.

#### *Event-related potentials and the present study*

In the present study, we used a different technique – event-related potentials (ERPs) – to explore the time-course and nature of processing light verb constructions. Of most relevance to this study are three groups of ERP components, summarized below.

First, the N400 is a negative-going potential peaking at approximately 400 ms post stimulus onset, and which is thought to reflect the retrieval or access to semantic features associated with an incoming word (Kutas & Federmeier, 2011). In sentence and discourse contexts, the amplitude of the N400 reflects the match or mismatch between the semantic features associated with this word and those activated by context. The N400 therefore tends to be smaller to words that are lexically predictable versus unpredictable in relation to their context, with predictability usually operationalized using cloze probability (Federmeier, Wlotko, De Ochoa-Dewald, & Kutas, 2007; Kutas & Hillyard, 1984).

Second, the P600 is a positive-going potential, which usually appears between 500 and 900 ms post stimulus onset. While this waveform was originally characterized as being most closely linked to syntactic violations and ambiguities (Hagoort, Brown, & Groothusen, 1993; Osterhout & Holcomb, 1992), it is also seen when the linguistic input is highly semantically implausible or incoherent (Kuperberg, Caplan, Sitnikova, Eddy, & Holcomb, 2006; Kuperberg, Kreher, Sitnikova, Caplan, & Holcomb, 2007; Kuperberg, Sitnikova, Caplan, & Holcomb, 2003; for reviews, see Kuperberg, 2007; Van de Meerendonk, Kolk, Chwilla, & Vissers, 2009 and Bornkessel-Schlesewsky & Schlesewsky, 2008, and see Kuperberg, 2013, for a recent discussion).

Finally, a set of anteriorly-distributed negativities have been associated with working memory operations required to maintain, link and select sentential and discourse

constituents during online processing (e.g. King & Kutas, 1995; Mueller, King, & Kutas, 1997; Nieuwland & Van Berkum, 2008a). Of most relevance to this study is emerging evidence that a set of prolonged negativity effects, starting at approximately 500 ms post stimulus onset and lasting for several hundred milliseconds, may be associated with the computation of complex but plausible event representations. For example, Baggio, Van Lambalgen, and Hagoort (2008) reported a sustained anteriorly distributed negativity beginning at approximately 450 ms after the onset of sentence-final verbs that signaled that an anticipated ongoing event (writing a letter) was prematurely terminated (by spilling coffee on the paper). A similar pattern was observed by Bott (2010) who reported a sustained anterior negativity starting at 500 ms after the onset of critical words (e.g. discovered) that signaled an additional event had taken place (a searching event). And, in a recent study, we observed a late-onset sustained negativity effect in association with iterative aspectual coercion, in which a punctive action (e.g. pounce) must be repeated for a period of time, specified by the preceding context (Paczynski, Jackendoff, & Kuperberg, *in press*). We will return to a more complete analysis of these studies in the Discussion section. For now, we note that in all three studies, the meaning of the events described cannot be derived solely from combining the meanings of the individual words with the surface syntactic structure; rather, the elements in the sentence need to be (re)-combined in a non-canonical way in order to arrive at a meaningful final sentence interpretation.

In the present study, we measured ERPs as participants read three types of sentences: light verb constructions, non-light constructions, and anomalous constructions (see Table 2). Similar to the behavioral study by Wittenberg and Piñango (2011), we used German sentences with verb-final ordering, which allowed us to examine processing on the verb, after the presentation of all arguments. In the light verb constructions, a light verb was combined with an eventive noun (“eine Ansage machte”, i.e., “made an announcement”). The non-light sentences used the same light verb, but paired it with a non-eventive noun (“einen Kaffee machte”, i.e., “made a coffee”), which resulted in a non-light construction. Finally, the anomalous constructions used the same verb, but paired it with an abstract noun that, in most cases, could be combined with a *different* light verb, but that rendered the overall construction ungrammatical (“ein Gespräch machte”, i.e., “made a conversation”; note that this construction is unacceptable in German; an English example would be “make a nap”). We examined ERPs evoked by the verb in each type of sentence. As the verb was identical across the three experimental conditions, any effects would necessarily arise from the combination of the verb with its arguments, rather than the lexical properties of the verb itself.

Our main focus was the contrast between the light and non-light constructions. As noted above, light verb constructions tend to be more frequent and thus more predictable than non-light constructions, which could potentially make them easier to process. The critical question was whether, despite their higher frequency and predictability, we would see evidence of argument sharing in the light verb

**Table 1**

Illustration of argument sharing in a light verb construction, as opposed to no argument sharing in a non-light sentence. Glossary: ag = Agent, pred = predicate, rec = Recipient, and pat = Patient.

	Non-Light construction	Light Verb construction
Example	<i>Henry gave a rose to Elsa</i>	<i>Henry gave a kiss to Elsa</i>
Syntax:	in both cases: SUBJ – VERB – OBJ 1 – OBJ 2	
Correspondence between syntactic element and semantic role		
SUBJ:	Agent of <i>give</i>	Agent of <i>give</i> and <i>kiss</i>
OBJ 1:	Theme of <i>give</i>	Theme of <i>give</i> + part of complex predicate
OBJ 2:	Recipient of <i>give</i>	Recipient of <i>give</i> and <i>kiss</i>
Syntax-Semantics Mapping	<i>Henry gave a rose to Elsa</i> ag <sub>1</sub> pred-1 theme <sub>1</sub> rec <sub>1</sub> [DP ] [VP ] [DP ] [PP ] <sub>1</sub>	<i>Henry gave a kiss to Elsa</i> ag <sub>1,2</sub> pred-1 pred-2/theme <sub>1</sub> pat <sub>2</sub> /rec <sub>1</sub> [DP ] [VP ] [DP ] [PP ] <sub>1,2</sub>

**Table 2**

Example sentences in German, English literal translation, and English SVO word order translation. Note that (3) is anomalous and unacceptable in German, due to semantic restrictions. The sentences within one set only differed at the object noun (italicized here). The critical verb (underlined here), to which ERPs were measured, did not differ between the three types of construction.

Context sentence:	
German:	Das Flugzeug war bereits hoch über den Wolken
English (literal):	The airplane was already high in the sky
(1) Non-Light construction:	
German:	Als die Stewardess einen Kaffee <u>machte</u>
English (literal):	When the stewardess a coffee <u>made</u>
(2) Light Verb construction:	
German:	Als die Stewardess eine Ansage <u>machte</u>
English (literal):	When the stewardess an announcement <u>made</u>
(3) Anomalous construction:	
German:	*Als die Stewardess ein Gespräch <u>machte</u>
English (literal):	*When the stewardess a conversation <u>made</u>
Continuation of each sentence:	
German:	... ging gerade die Sonne auf
English (literal):	... was just rising the sun
English word order:	"When the stewardess made a(n) coffee/announcement/*conversation, the sun was just rising."

constructions, as predicted by the Parallel Architecture (Culicover & Jackendoff, 2005). This might manifest as a late sustained negativity effect, similar to that previously observed in association with complex semantic operations involving the maintenance, computation and/or selection of non-canonical event structure representations within working memory (e.g. Baggio, Van Lambalgen, & Hagoort, 2008; Bott, 2010; Paczynski, Jackendoff, & Kuperberg, in press).

With regards to the contrast between light verb constructions and anomalous constructions, we predicted that the anomalous verbs would produce a P600 effect, similar to that previously observed in association with the detection of other violations of overall propositional coherence (see Kuperberg, 2007, 2013 for reviews), which may or may not be accompanied by an N400 effect.

## Methods

### Materials

One-hundred-and-twenty German scenarios were constructed, each with three conditions, as shown in Table 2. In each scenario, the first sentence provided some context

and was the same across the three conditions. The second sentence began with a subordinate clause, allowing for the verb-last word order. These subordinate clauses contained either (a) a light verb construction (light), (b) a non-light construction using the same verb (non-light), (c) or a noun and a light verb in a combination that yielded an anomalous interpretation (anomalous). Each of the sentences began with a subordinate conjunction, followed by the subject, optionally an adverbial phrase and an indirect object, and finally the critical noun and verb ("announcement made"/"coffee made"/"conversation made"). After the critical verb, the main clause began with its main verb, followed by the rest of the sentence. Thus, the three conditions only varied at the noun; all other words were held constant (see Table 2 for example sentences, and <http://www.nmr.mgh.harvard.edu/kuperberglab/materials.htm> for a full list of stimuli). Since the number of verbs that can enter light verb constructions is limited, we repeated the verbs up to eight times within a list, in all three conditions.

### Norming of stimuli

*Frequency.* The frequencies of nouns and verbs were retrieved from the DWDS corpus, which consists of a total of

100 million words (Geyken, 2007). On average, the nouns in light verb constructions appeared 1,314 times ( $SD = 1709$ ), the nouns in non-light constructions 1,360 times ( $SD = 1749$ ), and the nouns in anomalous constructions 1,396 times per million tokens ( $SD = 1710$ ). There were no significant differences in noun frequency between conditions (all comparisons:  $t < .28$ , all  $p > .78$ ).

*Noun-verb co-occurrence frequency.* We also used the same database to retrieve the co-occurrence frequency of the nouns and the verbs, that is, how often the nouns and verbs occurred together in the same syntactic configuration as in our stimuli: for light verb constructions, this was 1.33 ( $SD = .45$ ); for non-light constructions, it was .1 ( $SD = .21$ ); and for anomalous constructions, it was .18 per million tokens ( $SD = .65$ ). The nouns and verbs used in the light verb constructions thus occurred more often together than the nouns and verbs used for anomalous and non-light constructions ( $ts > .25$ ,  $ps < .02$ ); there was no difference in noun-verb co-occurrence frequency between the non-light and anomalous conditions ( $t = .77$ ,  $p > .44$ ).

*Cloze.* A cloze probability study was carried out with 162 German native speakers (58 males; average age: 28) who did not take part in the ERP experiment and who gave informed consent. The scenario stems (the context sentence and the main sentence up to and including the object) were randomized across three lists using a Latin Square design. The lists were then subdivided into three, and every participant was presented with one of the nine lists. Each sentence stem ended with an ellipsis (“...”) indicating that the sentence continued, and participants were asked to write the most likely next word. Cloze probabilities for each of the three conditions were calculated based on the percentage of respondents who produced a word that matched the light, non-light, or anomalous verb exactly. As expected, cloze was highest for the light condition, 53(35)%, lower for the non-light condition, 20(28)%, and zero (0) for the anomalous condition. Pairwise contrasts revealed significant differences in cloze probability between each of the three conditions (all  $ts > 6$ , all  $ps < 0.0001$ ).

*Lexical constraint.* From the cloze data, we also calculated the lexical constraint of the contexts in each construction. This was computed as the number of instances of the most frequent response over all responses (see Federmeier et al., 2007). As discussed by Federmeier et al. (2007), lexical can be partially dissociated from cloze itself: for example, a critical word can be zero cloze when its context is highly lexically constraining for another word, or when it is non-lexically constraining. Lexical constraint differed significantly across the three constructions ( $F(2, 357) = 16.07$ ,  $p < .001$ ), because as expected, the contexts of the light verb constructions were significantly more lexically constraining, 62(25)%, than the contexts of both the non-light constructions, 42(24)%, and the anomalous constructions, 45(23)%,  $ts > 2.9$ ,  $ps < 0.0001$ . There were no significant differences in contextual constraint between the non-light and anomalous constructions ( $t(119) = 1.4$ ,  $p > .15$ ).

*Plausibility ratings.* Plausibility ratings were collected from a different set of 58 German native speakers (18 M;

mean age: 26.7 years), who also did not take part in the ERP experiment and gave informed consent. List creation and randomization was the same as for the cloze ratings. Each sentence stem ended with the verb, followed by “...” to indicate that the sentence continued, and participants were asked to rate the scenarios up to and including the verb for plausibility on a scale from 1 (implausible) to 7 (plausible). Sentences with light verb constructions received an average rating of 5.8 (0.98), non-light constructions received an average rating of 5.6 (1.15), and anomalous sentences received an average rating of 2.8 (1.33). There were significant differences across the three conditions ( $F(2, 357) = 254.67$ ,  $p < .001$ ), but this was driven by the implausibility of the anomalous sentences relative to both the light sentences ( $t(119) = 20.6$ ,  $p < .0001$ ) and non-light sentences ( $t(119) = -17.2$ ,  $p < .0001$ ); there was no difference in plausibility between light and non-light sentences ( $t(119) = 1.8$ ,  $p > .5$ ). Thus, the light and non-light sentences were matched on plausibility.

In the ERP experiment, the 360 scenarios were counter-balanced across three lists, using a Latin Square design. Each participant thus saw 40 instances of each of the three constructions (120 scenarios in total), but never encountered a given scenario more than once. However, across all participants, each scenario was seen in all three conditions.

Eighty filler scenarios were then added to each list: 20 were semantically and syntactically normal, and 60 introduced a semantic anomaly in the final part of the second sentence. Thus, in each list, half of the scenarios were normal and half of the scenarios contained a semantic anomaly, the majority of which occurred towards the end of the second sentence, ensuring that participants read to the end of each two-sentence scenario.

### Participants

Twenty native German-speaking participants (8 M; mean age: 27.3 years) participated in the ERP experiment. All participants were right-handed and had normal or corrected-to-normal vision, were not taking any medication, and were screened to exclude a history of psychiatric or neurological disorders. They provided written consent before participating, as specified by the guidelines of the Tufts University Institutional Review Board, and received \$20 for their participation.

### ERP procedure

Each participant was randomly assigned to one of three lists and was given six practice trials at the beginning of the experiment. All stimuli were presented visually. Each two-sentence trial began with presentation of a fixation cross at the center of the screen for 450 ms, followed by a 100 ms blank screen, followed by the context sentence, presented as a whole (displayed for 800–1600 ms, depending on length). After the context sentence, a fixation cross was again presented for 450 ms, followed by a 100 ms blank screen, and then the second sentence was presented word-by-word (450 ms per word; interstimulus interval: 150 ms), followed by a question mark. At this point,

participants judged whether or not the scenario they had just read was a natural sentence of German. They were told that sentences may not be natural for different reasons, and if sentences seemed “odd” in any way, they should indicate that it was not natural. For the light and non-light constructions and for the normal filler sentences, the expected answer was “yes”; for the semantically anomalous constructions and for the anomalous fillers, the expected answer was “no”, so that in total, every participant ideally judged half the sentences to be acceptable, and half to be unacceptable. After making their judgments, participants pressed a button to move onto the next trial. Each list contained four blocks.

### EEG acquisition

Twenty-nine tin electrodes were held in place on the scalp by an elastic cap, see Fig. 1. Electrodes were also placed below the left eye and at the outer canthus of the right eye to monitor vertical and horizontal eye movements, and on the left and right mastoids. Impedance was kept below 5 k $\Omega$  for all scalp electrode sites, 2.5 k $\Omega$  for mastoid electrode sites and 10 k $\Omega$  for the two eye channels. The EEG signal was amplified by an Isolated Bioelectric Amplifier System Model HandW-32/BA (SA Instrumentation Co., San Diego, CA) with a bandpass of 0.01–40 Hz, and was continuously sampled at 200 Hz by an analogue-to-digital converter. A digitizing computer simultaneously monitored the stimuli and behavioral responses.

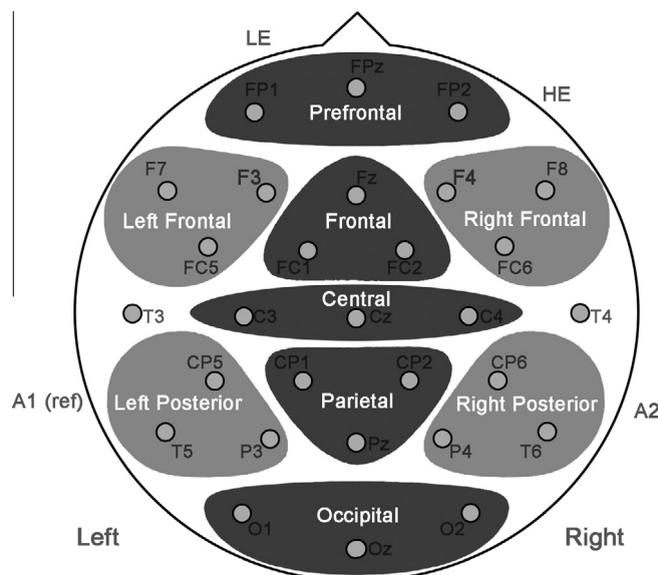
### Data analysis

Response congruency was computed as the percentage of responses that matched our prior classifications. A

congruent response was considered a judgment of the light verb and non-light constructions as “acceptable”, and the anomalous constructions and the violated fillers as “unacceptable”.

Averaged ERPs were time-locked to the onset of the critical verbs. They were formed off-line from trials free of ocular and muscular artifact, and were quantified by calculating the mean amplitude (relative to a 50 ms pre-stimulus baseline) in time windows of interest. On the verb, two main time windows were chosen: 300–500 ms to capture the N400 component and perhaps the beginning of an anterior negativity effect, and 500–900 ms (the relevant time frame for sustained negativities and the P600). We also ran analyses for the same time windows from the onset of the preceding noun, to determine whether there were any carryover effects to the critical verbs, and for the 50–200 ms and 200–300 ms time windows from verb onset to determine whether there were any early effects.

The scalp surface was subdivided into regions along the anterior–posterior distribution, at both mid and peripheral sites (each region contained 3 electrode sites, see Fig. 1). Two omnibus analyses of variance (ANOVAs), one covering mid regions and another covering peripheral regions across the scalp, were conducted in each time window. In the mid-regions ANOVA, Construction and Region were within-subjects factors; in the peripheral regions ANOVA, Hemisphere was an additional within-subjects factor. Significant main effects of Construction or interactions between Construction and Region (and/or Hemisphere) were followed up by carrying out ANOVAs comparing each type of Construction with one another. Any further interactions between Construction and Region were then followed-up by examining the effects of Construction in each three-electrode Region individually. In all analyses,



**Fig. 1.** Electrode montage with regions used for analysis. Regions in dark grey were part of the mid-regions omnibus ANOVA and regions in light grey were part of the peripheral regions omnibus ANOVA. The Left Frontal and Right Frontal regions together constituted the Frontal Peripheral regions, and the Left Posterior and Right Posterior Regions constituted the Parietal Peripheral regions.

the Geisser–Greenhouse correction was used in cases with more than one degree of freedom in the numerator (Greenhouse & Geisser, 1959) to protect against Type I error resulting from violations of sphericity. In these cases, we report the original degrees of freedom with the corrected  $p$  value, using a significance level of  $\alpha = 0.05$ .

## Results

### Behavioral data

Response congruency with our a priori classifications did not differ significantly between light verb constructions, 72(4.1)%, non-light constructions, 69(3.9)%, and anomalous constructions, 75(6.2)%,  $F(2, 51) = 1.54$ ,  $p = .22$ . Similar response congruency was found for the filler scenarios that contained violations, 88(5.5)%, and those that did not contain violations, 76(1.4)%.

### ERP data

Approximately 8.8% of the critical trials were rejected due to artifact. There was no significant difference in rejection rates across conditions,  $F(2, 51) = 0.04$ ,  $p = .97$ . All ERP analyses reported are based on correctly classified trials only. Similar results were obtained when analyses were repeated including all trials. Additionally, we also carried out ERP analyses on a subset of scenarios in which the level of contextual constraint was matched across the three condition; again, these analyses yielded results similar to the main analyses, and are thus not reported separately.

### Nouns

ERPs evoked by the object nouns were similar across the three construction types. Omnibus mid-regions and peripheral regions ANOVAs revealed no significant main effects of Construction and no interactions between Construction and Region and/or Hemisphere in either the 300–500 ms window (all  $F_s < 0.44$ , all  $p_s > .63$ ) or the 500–900 ms time window (all  $F_s < 0.88$ , all  $p_s > 0.42$ ).

### Verbs

#### Early effects

There were no significant main effects of Construction and no interactions between Construction, Region and/or Hemisphere in either the 50–200 ms (all  $F_s < 0.47$ , all  $p_s > 0.62$ ) or the 200–300 (all  $F_s < 1.33$ , all  $p_s > 0.27$ ) time windows.

#### 300–500 ms

As shown in Fig. 2, there was little divergence between the waveforms evoked by the verbs within the N400 time window, and indeed there were no significant main effects or interactions (see Table 3). Because previous work on enriched composition has sometimes reported shorter-lived effects within the N400 time window (e.g. Brennan & Pykkänen, 2008; Kuperberg, Choi, Cohn, Paczynski, & Jackendoff, 2010), we also carried out analyses in

successive 50 ms time windows between 300 and 500 ms, but found no significant main effects or interactions.

#### 500–900 ms

In this time window, the waveforms to verbs in the three conditions diverged from one another (see Fig. 2). These differences were reflected by main effects of Construction (mid-regions:  $F(2, 34) = 6.79$ ,  $p = .01$ ; peripheral regions:  $F(2, 34) = 5.76$ ,  $p = .01$ ), as well as by interactions between Construction and Region (mid-regions:  $F(8, 136) = 7.18$ ,  $p < .001$ ; peripheral regions:  $F(2, 34) = 12.5$ ,  $p < .001$ ). We followed up these omnibus effects by carrying out pair-wise ANOVAs comparing each type of construction with one other.

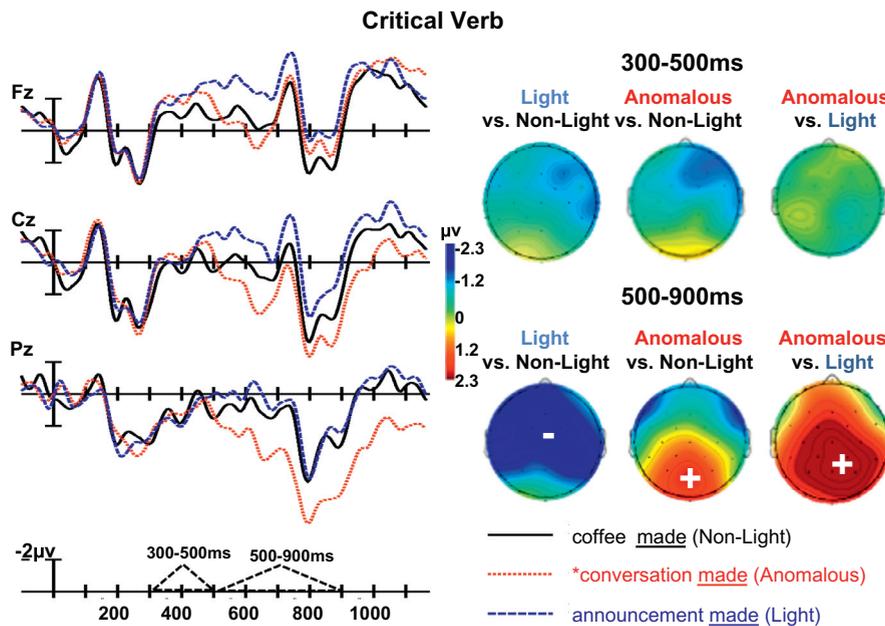
The waveform to verbs in the light verb constructions was more negative-going than to verbs in the non-light constructions. This effect was widespread, as reflected by significant main effects of Construction in both the mid-regions and peripheral regions, but its magnitude varied across scalp region, as reflected by an interaction between Construction and Region in the mid-regions. Follow-ups in individual regions showed significant effects in the frontal and central regions, a near-significant effect in the parietal region, but no significant effect in the occipital region (see Table 3 for statistics).

The anomalous constructions evoked a larger positive deflection than the non-light constructions. The magnitude of this positivity effect also varied across the scalp surface as reflected by significant interactions between Construction and Region in both the mid-regions and peripheral regions. Follow-ups showed significant effects of Construction in the parietal and occipital regions, a near-significant effect in the peripheral parietal regions, but no effect in any other region (see Table 3 for statistics).

As expected, the divergence of the positive-going waveform in the anomalous constructions from the negative-going waveform in the light verb constructions was statistically robust: there were main effects of Construction in both the mid-regions and peripheral regions, as well as a significant interactions between Construction and Region; the effect of Construction was significant in all but the prefrontal and frontal peripheral regions.

## Discussion

In this study, we examined the time-course and nature of processing light verb constructions using ERPs. Using German subordinate sentences with verb-final word order enabled us to directly examine neural activity at the verb – the point at which semantic roles are usually assigned and hence where we predicted the effects of argument sharing in light verb constructions to be most prominent. We contrasted light verb constructions with non-light and anomalous constructions (see Table 1). In the N400 time window, there were no significant differences in the waveforms observed to the three types of construction. Within the 500–900 ms time window, however, there was clear divergence in the waveforms evoked in the light verb constructions, relative to the other two conditions: the ERP to



**Fig. 2.** Left: Grand-averaged waveforms to critical verbs in all three sentence types at Fz, Cz, and Pz. Solid black lines indicate non-light sentences; dotted red lines indicate anomalous sentences; dashed blue lines indicate light verb construction sentences. The plots are shown using a 50 ms peri-stimulus baseline. Right: Voltage maps show differences between ERPs to critical verbs between 300 and 500 ms and between 500 and 900 ms. The scale is valid for all voltage maps; a minus or plus sign on the voltage map indicates significant differences (see Results section for details). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the light verb constructions was more negative/less positive than that to the other two constructions; this effect was widespread but with a frontal focus. The anomalous constructions elicited a posteriorly distributed positivity effect.

#### Light verb constructions

The divergence in neural activity to verbs in the light (versus the non-light) constructions cannot be attributed to lexical differences because identical verbs were seen in both construction types. It also cannot be an effect of differences in plausibility, as the two sentence types did not differ from each other in that measure. Spillover effects from the preceding object argument also cannot explain this effect, since there was no divergence between conditions before the verb appeared. The effect also cannot be reduced to differences in the lexical predictability of the verbs across the two constructions; less predictable words usually evoke a larger negativity than more predictable words within the earlier N400 time window (Kutas & Hilliard, 1984). In this study, however, the light verbs were actually more lexically predictable than the non-light verbs (as assessed through our cloze ratings), and the divergence in waveforms was seen in a later 500–900 ms time window. Finally, these findings are not easily explained by differences in lexical constraint between the light and non-light contexts prior to the verb: words that violate highly lexically constraining contexts can evoke anteriorly-distributed late positivity effects (Federmeier et al., 2007). In this study, however, the lexical constraint of both constructions (49%) was much lower than the

average lexical constraint of the sentences used by Federmeier, Wlotko, De Ochoa-Dewald and Kutas (2007: 85%). Moreover, when we repeated all analyses in a subset of stimuli in which the light and non-light contexts were matched for lexical constraint, the effect remained significant.

One question that arises is whether the ERP effect observed to the light (versus non-light) constructions reflects a larger late anterior negativity effect to the light verb constructions or a larger late anterior positivity effect to the non-light constructions. This study alone cannot distinguish between these two possibilities. However, there are several reasons why we think that it is more likely to reflect a late anterior negativity effect to the light verb constructions. First, as discussed above, late anterior positivity effects are typically produced by words that violate highly lexically constraining contexts (Federmeier et al., 2007), which was not the case here. Second, in the linguistic literature, light verb constructions are usually seen as the special cases, while the full, non-light versions of those verbs are the point of reference (Butt, 2003; Jespersen, 1965). We follow this theoretical assumption by treating the non-light constructions as the baseline condition. Third, our interpretation of this ERP effect as a late sustained anterior negativity effect is in line with an emerging ERP literature associating similar effects with the processing of plausible but non-canonical event structures, as we discuss next.

There are now several studies reporting late sustained negativity effects in association with processing non-canonical, complex event structures. First, Baggio, Van Lambalgen, and Hagoort (2008) reported a sustained

**Table 3**

ANOVAs comparing ERPs evoked by the critical verb between 300–500 ms and 500–900 ms after their onset. C = main effect of Construction, C × AP = Construction × Anterior–Posterior distribution interaction. Significant effects and effects approaching significance are in bold.

Effect	300–500 ms		500–900 ms		
	F (df)	p	F (df)	p	
<i>A. Light vs. Non-Light</i>					
Midline Omnibus	C	.64 (1,17)	.43	6.69 (1,17)	<b>.02</b>
	C × AP	2.03 (4,68)	.16	4.22 (4,68)	<b>.03</b>
Peripheral Omnibus	C	1.62 (1,17)	.22	10.27 (1,17)	<b>.00</b>
	C × AP	1.24 (3,51)	.30	.85 (3,51)	.43
Anterior Frontal	C	.85 (1,17)	.37	4.83 (1,17)	<b>.04</b>
Frontal	C	1.56 (1,17)	.22	8.55 (1,17)	<b>.01</b>
Central	C	1.59 (1,17)	.22	8.22 (1,17)	<b>.01</b>
Parietal	C	.11 (1,17)	.75	3.27 (1,17)	<b>.08</b>
Occipital	C	2.13 (1,17)	.16	.37 (1,17)	.55
Frontal Peripheral	C	2.01 (1,17)	.17	8.23 (1,17)	<b>.01</b>
Parietal Peripheral	C	.66 (1,17)	.43	7.19 (1,17)	<b>.01</b>
<i>B. Non-Light vs. Anomalous</i>					
Midline Omnibus	C	.88 (1,17)	.36	1.71 (1,17)	.21
	C × AP	2.01 (4,68)	.15	13.78 (4,68)	<b>.00</b>
Peripheral Omnibus	C	2.24 (1,17)	.15	.11 (1,17)	.75
	C × AP	1.94 (3,51)	.15	16.01 (3,51)	<b>.00</b>
Anterior Frontal	C	1.45 (1,17)	.24	2.05 (1,17)	.17
Frontal	C	1.52 (1,17)	.23	.06 (1,17)	.81
Central	C	1.36 (1,17)	.26	1.62 (1,17)	.24
Parietal	C	.21 (1,17)	.65	4.84 (1,17)	<b>.04</b>
Occipital	C	.83 (1,17)	.37	11.61 (1,17)	<b>.00</b>
Frontal Peripheral	C	2.72 (1,17)	.12	2.45 (1,17)	.14
Parietal Peripheral	C	1.06 (1,17)	.32	4.16 (1,17)	<b>.05</b>
<i>C. Light vs. Anomalous</i>					
Midline Omnibus	C	.04 (1,17)	.84	16.13 (1,17)	<b>.00</b>
	C × AP	.23 (4,68)	.72	5.67 (4,68)	<b>.01</b>
Peripheral Omnibus	C	.04 (1,17)	.84	9.96 (1,17)	<b>.00</b>
	C × AP	.32 (3,51)	.74	7.49 (3,51)	<b>.00</b>
Anterior Frontal	C	.00 (1,17)	.94	1.27 (1,17)	.27
Frontal	C	.00 (1,17)	.96	11.9 (1,17)	<b>.00</b>
Central	C	.00 (1,17)	.96	18.59 (1,17)	<b>.00</b>
Parietal	C	.04 (1,17)	.85	17.69 (1,17)	<b>.00</b>
Occipital	C	.72 (1,17)	.41	9.17 (1,17)	<b>.01</b>
Frontal Peripheral	C	.03 (1,17)	.86	1.46 (1,17)	.24
Parietal Peripheral	C	.04 (1,17)	.84	19.02 (1,17)	<b>.00</b>

anteriorly distributed negativity effect, beginning at approximately 450 ms after the onset of a verb that implied the interruption of an ongoing event, e.g. original Dutch: “Het meisje was een brief aan het schrijven toen haar vriendin koffie op het papier (vs. tafelkleed) morste”; English translation: “The girl was writing a letter when her friend spilled coffee on the paper (vs. tablecloth)” (Paczynski, Jackendoff, & Kuperberg, *in press*). In this sentence, full integration of the verb establishes that the ongoing event (writing a letter) was not completed because coffee was spilled on the paper (rather than on the tablecloth). Second, Bott (2010) observed a similar effect beginning at 500 ms after the onset of verbs that implied an additional event to the one explicitly stated—so-called *additive aspectual coercion* (e.g. original German: “In zwei Stunden hatte der Förster die Falle entdeckt”; English translation: “Within two hours, the ranger had discovered the trap”), compared to the non-coerced control condition (e.g. original German: “Nach zwei Stunden hatte der Förster die Falle entdeckt”; English translation: “After two hours, the

ranger had discovered the trap”). In the sentence containing an aspectual coercion, full integration of the verb establishes not only that the trap has been discovered, but that it was also being searched for. Finally, in a recent study, we observed a prolonged widely distributed negativity effect (also with an anterior focus), beginning at 500 ms after the onset of verbs that implied multiple iterations of a punctive event—so-called *iterative aspectual coercion*, e.g. “For several minutes (vs. After several minutes) the cat pounced on the rubber mouse.” In this sentence, full integration of the verb establishes that the cat did not just carry out a single pounce, but rather several pounces lasting for several minutes.

What all these constructions have in common—and what they also have in common with light verb constructions—is that the interpretation of the event cannot be simply derived from the meaning of the individual lexical items and the syntactic structure of the sentences: rather, the elements in the sentence need to be combined a non-canonical way in order to arrive at a meaningful final inter-

pretation of the event. In the case of interrupted accomplishments, there is a need to change the default event representation from a completed event to an aborted event, based on subsequent discourse information (Baggio et al., 2008); in the case of additive coercion, the event representation needs to be modified based on temporal information outside the verb's semantics (Bott, 2010); in the case of aspectual coercion, the aspect of the event is again derived from information outside of the verb's lexical semantics, that is, from the prepositional phrase (Paczynski, Jackendoff, & Kuperberg, in press). And, finally, in the present study of light verb constructions, the event type and main meaning are not provided by the verb, but rather by the event nominal.

We suggest that, in all these cases, the sustained negativity effect reflects an extended process of integrating the incoming word in order to derive these non-canonical event structure representations. In the present study, we suggest that in a sentence like "Julius gives Anne a kiss", it involves reconciling the mismatch between semantic and syntactic argument structure so that argument structure of the noun ("kiss") is layered onto that of the verb ("give") and Julius acts both as the Agent of "give" and "kiss", and Anne acts both as the Recipient of "give" and the Patient of "kiss"—a process of argument sharing (see Wittenberg & Snedeker, in press, for further discussion).

There are several possible mechanisms by which such argument sharing might proceed. One possibility is that it reflects a process of constructing a complex event structure only after attempting to construct a canonical event structure – similarly to encountering a semantic garden path. Another is that both a canonical and non-canonical event structure are active for a short time, and that comprehenders then *select* the non-canonical light verb construction by enhancing its activation and suppressing the canonical event structure. Both accounts would be consistent with the findings by Wittenberg and Snedeker (in press) that light verb constructions can be associated with two different event structures. The latter account would also be consistent with previous studies that have associated prolonged anterior negativity effects with a process of selecting between alternative event structures (e.g. non-literal over a literal event structures in comprehending novel metaphors (e.g. Coulson & Van Petten, 2007), cartoon stories (Nieuwland & Van Berkum, 2006, Experiment 1) and jokes (e.g. Coulson, 2001), or between alternative predicted specific events (Wlotko & Federmeier, 2012). On either account, the anteriorly-distributed sustained negativity effect is likely to reflect the additional working memory demands of maintaining and manipulating these event structures to arrive at the final sentence meaning (see King & Kutas, 1995; Mueller, King, & Kutas, 1997; Nieuwland & Van Berkum, 2008a for other examples of anteriorly-distributed sustained negativity effects associated with working memory costs).

Importantly, we are not claiming that the type of sustained anterior negativity observed here, and in these previous studies, is *specific* to complex semantic operations in coming to event representation. Sustained anteriorly distributed negativities have also been observed in situations where other types of constituents must be held within

working memory to link elements within and across sentences (King & Kutas, 1995; Kluender & Kutas, 1993; Nieuwland & Van Berkum, 2008a, 2008b; Van Berkum, Brown, & Hagoort, 1999; Van Berkum, Zwitserlood, Hagoort, & Brown, 2003). By the same token, we also do not think that all complex semantic compositional processes are necessarily associated with anteriorly distributed sustained negativity effects. For example, we and others have reported modulation on the N400 component in association with complement coercion (Baggio, Choma, Van Lambalgen, & Hagoort, 2010; Kuperberg et al., 2010): complement noun-phrases in coerced sentences (e.g. "The man began the book...") evoked a larger negativity between 300–500 ms than in non-coerced sentences (e.g. "The man read the book..."), which we interpreted as primarily reflecting the semantic mismatch between an action-requiring verb like "begin" and the semantic features of the entity "book" at the point of the complement noun phrase.

#### *Anomalous constructions*

The anomalous constructions were created by pairing light verbs with abstract nouns that are usually associated with different light verbs, such as "give a nap" (compare "take a nap"). In contrast to acceptable light verb constructions, which evoked an anteriorly distributed negativity, anomalous light verb constructions evoked a posteriorly distributed positivity—the P600. This P600 was similar to that previously reported by our group (Kuperberg et al., 2003, 2006, 2007) as well as by others (e.g. Hoeks, Stowe, & Doedens, 2004; Kim & Osterhout, 2005; Van de Meerendonk, Kolk, Chwilla, & Vissers, 2009; Van Herten, Kolk, & Chwilla, 2005) to selection restriction violations occurring on verbs. A P600 effect is also seen on nouns that violate the selection restrictions of their preceding verbs, particularly in judgment tasks (e.g. Kuperberg et al., 2010; Paczynski & Kuperberg, 2011, 2012). Although there is debate about the precise functional significance of this effect, there is some general agreement that it reflects additional analysis or reanalysis that is triggered when the parser computes a proposition, using all available linguistic information, and this is classified as being impossible or incoherent at the critical word (see Bornkessel-Schlesewsky & Schlewsky, 2008; Kuperberg, 2007; Paczynski & Kuperberg, 2012; Van de Meerendonk et al., 2009). More specifically, we have suggested that it is triggered by a conflict between a strong, high-certainty prediction for one specific type of event structure, and an incoming word whose integration violates this strong prediction, and that it reflects a process of updating the contents of working memory through bottom-up attempts to establish a new event structure (see Kuperberg, 2013, for a recent discussion of the posterior late positivity/P600 effect as being triggered by an event structure prediction error). In the present study, we take the presence of the P600 to the light verbs paired with the wrong eventive nouns as evidence that the parser tried, but failed, to generate a coherent complex predicate. In other words, when successful, argument sharing was associated with an anterior negativity effect. However, when initial integration of the input yielded an event

structure that conflicted with what was anticipated, this triggered additional analysis or reanalysis processes, which manifested as a posterior P600 effect.

### Implications

This paper contributes to both the psycholinguistic and theoretical linguistic literatures. With respect to the ERP psycholinguistic literature, our finding that light verb constructions evoke a widely distributed negativity effect with an anterior focus between 500–900 ms adds to a growing body of evidence that certain complex semantic operations can engage quite different neurocognitive operations from the types of semantic feature matching that are thought to be reflected by modulation of the N400 component. More specifically, we have suggested that it may reflect a process of computing semantically complex events, possibly involving the top-down selection of a non-canonical semantic–syntactic mapping (event structure). This is distinguished both from the N400 component, which is thought to reflect facilitated access or retrieval of the semantic features of a specific word that have been activated by the context (see Paczynski & Kuperberg, 2012, for recent discussion), as well as from the P600 effect which may reflect a bottom-up attempt to come up with a novel event structure when a strongly predicted event structure is violated by initial attempts to integrate the input (Kuperberg, 2013).

At a theoretical level, our findings may help arbitrate between competing linguistic accounts of how light verb constructions are represented. This picture is inconsistent with theories which model light verb constructions as syntactically less complex (Folli, Harley, & Karimi, 2004; Hale & Keyser, 1993, 2002; Jung, 2002; see Wittenberg et al., in press, for detailed discussion): a syntactically less complex structure should result in reduced processing demands, even if one follows a more modern form of the Derivational Theory of Complexity (Phillips, in press). In contrast, our data does support models that consider both syntactic and semantic compositionality as contributing to the language architecture (Butt, 2010; Culicover, 2013; Culicover & Jackendoff, 2005; Jackendoff, 2002; Müller, 2010).

In sum, we have shown that light verb constructions evoke a widely distributed, but frontally focused, sustained negativity effect between 500–900 ms after verb onset, despite being more frequent and predictable than non-light constructions. These data show how the study of a syntactically simple and common phenomenon can reveal complex underlying neural processes, yielding insights that are relevant for both linguistic theory and for understanding mechanisms of language comprehension.

### Acknowledgments

This work was supported by NIMH (R01 MH071635) and the Brain and Behavior Research Foundation (NARSAD, with the Sidney Baer Trust) to Gina R. Kuperberg, as well as by a “European Recovery Program” scholarship to Eva Wittenberg, provided by the German Federal Ministry of Economics and Technology. We also thank Claire Bonial for help with the English corpus data, the conference audi-

ences at CUNY 2011, AMLaP 2011, and SPR 2011, Hugh Rabagliati, as well as three anonymous reviewers and the editor, Martin Pickering, for many helpful comments.

### References

- Baggio, G., Choma, T., Van Lambalgen, M., & Hagoort, P. (2010). Coercion and compositionality. *Journal of Cognitive Neuroscience*, 22, 2131–2140.
- Baggio, G., Van Lambalgen, M., & Hagoort, P. (2008). Computing and recomputing discourse models: An ERP study. *Journal of Memory and Language*, 59, 36–53.
- Baker, M. C. (1989). Object sharing and projection in serial verb constructions. *Linguistic Inquiry*, 20, 513–553.
- Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2008). An alternative perspective on “semantic P600” effects in language comprehension. *Brain Research Reviews*, 59, 55–73.
- Bott, O. (2010). *The processing of events*. Philadelphia, PA: John Benjamins Publishing Company.
- Brennan, J., & Pyllkänen, L. (2008). Processing events: Behavioral and neuromagnetic correlates of aspectual coercion. *Brain and Language*, 106, 132–143.
- Briem, D. (2009). *Neurotopographie der Speicherung und Verarbeitung von lexiko-semantic und syntaktisch unterschiedlichen Verb-Klassen* (pp. 1–169). Unpublished doctoral dissertation. Universität Konstanz.
- Briem, D., Balliel, B., Rockstroh, B., Butt, M., Walde, S. S., & Assadollahi, R. (2010). Distinct processing of function verb categories in the human brain. *Brain Research*, 1249, 173–180.
- Butt, M. (2003). The light verb jungle. *Harvard Working Papers in Linguistics*, 9, 1–49.
- Butt, M. (2010). The light verb jungle: Still hacking away. In M. Amberber, M. Harvey, & B. Baker (Eds.), *Complex predicates in cross-linguistic perspective* (pp. 48–78). Cambridge University Press.
- Coulson, S. (2001). *Semantic leaps: Frame-shifting and conceptual blending in meaning construction*. Cambridge University Press.
- Coulson, S., & Van Petten, C. (2007). A special role for the right hemisphere in metaphor comprehension? ERP evidence from hemifield presentation. *Brain Research*, 1146, 128–145.
- Culicover, P. (2013). *Grammar and complexity*. Oxford: Oxford University Press.
- Culicover, P. W., & Jackendoff, R. (2005). *Simpler syntax*. Oxford; New York: Oxford University Press.
- Durie, M. (1988). Verb serialization and verbal-prepositions. *Oceanic Linguistics*, 27, 1–23.
- Federmeier, K. D., Wlotko, E. W., De Ochoa-Dewald, E., & Kutas, M. (2007). Multiple effects of sentential constraint on word processing. *Brain Research*, 1146, 75–84.
- Folli, R., Harley, H., & Karimi, S. (2004). Determinants of event type in Persian complex predicates. *Lingua*, 115, 1365–1401.
- Friederici, A. D., & Frisch, S. (2000). Verb argument structure processing: The role of verb-specific and argument-specific information. *Journal of Memory and Language*, 43, 476–507.
- Geyken, A. (2007). The DWDS corpus: A reference corpus for the German language of the 20th century. In *Collocations and Idioms* (pp. 23–40), London.
- Goldberg, A. E. (2003). Constructions: a new theoretical approach to language. *Trends in Cognitive Sciences*, 7, 219–224.
- Greenhouse, S., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, 24, 95–112.
- Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes*, 8, 439.
- Hale, K., & Keyser, S. J. (1993). On argument structure and the lexical expression of syntactic relations. *The View from Building*, 20, 53–100.
- Hale, K. L., & Keyser, S. J. (2002). *Prolegomenon to a theory of argument structure* (No. 39). Cambridge, MA: MIT Press.
- Hoeks, J. C. J., Stowe, L. A., & Doedens, G. (2004). Seeing words in context: The interaction of lexical and sentence level information during reading. *Cognitive Brain Research*, 19, 59–73.
- Jackendoff, R. (1974). A deep structure projection rule. *Linguistic Inquiry*, 5, 481–505.
- Jackendoff, R. (2002). *Foundations of language*. Oxford University Press.
- Jespersen, O. (1965). *A modern English grammar on historical principles, Part VI. Morphology*. London: George Allen and Unwin Ltd.
- Jung, D. (2002). Light verb just as a little v. *Kansas working papers in linguistics* (pp. 59–74).

- Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language*, 52, 205–225.
- King, J., & Kutas, M. (1995). Who did what and when? Using word-and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience*, 7, 376–395.
- Kluender, R., & Kutas, M. (1993). Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience*, 5, 196–214.
- Kuperberg, G. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research*, 1146, 23–49.
- Kuperberg, G., Caplan, D., Sitnikova, T., Eddy, M., & Holcomb, P. (2006). Neural correlates of processing syntactic, semantic, and thematic relationships in sentences. *Language and Cognitive Processes*, 21, 489–530.
- Kuperberg, G., Choi, A., Cohn, N., Paczynski, M., & Jackendoff, R. (2010). Electrophysiological correlates of complement coercion. *Journal of Cognitive Neuroscience*, 22, 2685–2701.
- Kuperberg, G., Kreher, D., Sitnikova, T., Caplan, D., & Holcomb, P. J. (2007). The role of animacy and thematic relationships in processing active English sentences: Evidence from event-related potentials. *Brain and Language*, 100, 223–238.
- Kuperberg, G. (2013). The pro-active comprehender: What event-related potentials tell us about the dynamics of reading comprehension. In B. Miller, L. Cutting, & P. McCardle (Eds.), *Unraveling the behavioral, neurobiological, and genetic components of reading comprehension* (pp. 176–192). Baltimore: Paul Brookes Publishing.
- Kuperberg, G., Sitnikova, T., Caplan, D., & Holcomb, P. J. (2003). Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research*, 217, 117–129.
- Kutas, M., & Federmeier, K. (2011). Thirty years and counting: Finding meaning in the N400 component of the Event-Related Brain Potential (ERP). *Annual Review of Psychology*, 62, 621–647.
- Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, 307, 161–163.
- Müller, S. (2010). Persian complex predicates and the limits of inheritance-based analyses. *Journal of Linguistics*, 46, 601–655.
- Mueller, H. M., King, J. W., & Kutas, M. (1997). Event related potentials elicited by spoken relative clauses. *Cognitive Brain Research*, 5(3), 193–203.
- Müller, S., & Wechsler, S. (in press). Lexical approaches to argument structure. *Theoretical Linguistics*, 40(1–2).
- Nieuwland, M. S., & Van Berkum, J. J. A. (2006). When peanuts fall in love: N400 evidence for the power of discourse. *Journal of Cognitive Neuroscience*, 18, 1098–1111.
- Nieuwland, M. S., & Van Berkum, J. J. A. (2008a). The interplay between semantic and referential aspects of anaphoric noun phrase resolution: Evidence from ERPs. *Brain and Language*, 106, 119–131.
- Nieuwland, M. S., & Van Berkum, J. J. A. (2008b). The neurocognition of referential ambiguity in language comprehension. *Language and Linguistics Compass*, 2, 603–630.
- Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, 31, 785–806.
- Paczynski, M., Jackendoff, R., & Kuperberg, G. (in press). When events change their nature: The neurocognitive mechanisms underlying aspectual coercion. *Journal of Cognitive Neuroscience*.
- Paczynski, M., & Kuperberg, G. (2011). Electrophysiological evidence for use of the animacy hierarchy, but not thematic role assignment, during verb argument processing. *Language and Cognitive Processes*, 26, 1402–1456.
- Paczynski, M., & Kuperberg, G. R. (2012). Multiple influences of semantic memory on sentence processing: Distinct effects of semantic relatedness on violations of real-world event/state knowledge and animacy selection restrictions. *Journal of Memory and Language*, 67, 426–448.
- Palmer, M., Gildea, D., & Kingsbury, P. (2005). The proposition bank: An annotated corpus of semantic roles. *Computational Linguistics*, 31, 71–206.
- Phillips, C. (in press). Parser-grammar relations: We don't understand everything twice. In M. Sanz, I. Laka, & M. Tanenhaus (Eds.), *Language down the garden path: The cognitive and biological basis for linguistic structure*. Oxford University Press.
- Pickering, M. J., McElree, B., Frisson, S., Chen, L., & Traxler, M. J. (2006). Underspecification and aspectual coercion. *Discourse Processes*, 42, 131–155.
- Piñango, M. M., Mack, J., & Jackendoff, R. (2014). Semantic combinatorial processes in argument structure: Evidence from light verbs. In *Proceedings of the annual meeting of the Berkeley Linguistics Society, 2006* (in press). <<http://ase.tufts.edu/cogstud/jackendoff/papers/PinangoMack&Jackendoff2006BLS.pdf>>.
- Van Berkum, J. J. A., Brown, C. M., & Hagoort, P. (1999). Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of Memory and Language*, 41, 147–182.
- Van Berkum, J. J. A., Zwitserlood, P., Hagoort, P., & Brown, C. (2003). When and how do listeners relate a sentence to the wider discourse? Evidence from the N400 effect. *Cognitive Brain Research*, 17, 701–718.
- Van de Meerendonk, N., Kolk, H. H., Chwilla, D. J., & Vissers, C. T. W. (2009). Monitoring in language perception. *Language and Linguistics Compass*, 3, 1211–1224.
- Van Herten, M., Kolk, H. H. J., & Chwilla, D. J. (2005). An ERP study of P600 effects elicited by semantic anomalies. *Cognitive Brain Research*, 22, 241–255.
- Wiese, H. (2006). "Ich mach dich Messer" – Grammatische Produktivität in Kiez-Sprache (Kanak Sprach'). *Linguistische Berichte*, 207, 245–273.
- Wittenberg, E., & Snedeker, J. (in press). It takes two to kiss – But does it take three to give a kiss? Conceptual sorting based on thematic roles. *Language and Cognitive Processes*.
- Wittenberg, E., Jackendoff, R., Kuperberg, G., Paczynski, M., Snedeker, J., & Wiese, H. (in press). The processing and representation of light verb constructions. In A. Bachrach, I. Roy, & L. Stockall (Eds.), *Structuring the argument*. John Benjamins.
- Wittenberg, E., & Piñango, M. M. (2011). Processing light verb constructions. *The Mental Lexicon*, 6, 393–413.
- Wlotko, E. W., & Federmeier, K. D. (2012). So that's what you meant! Event-related potentials reveal multiple aspects of context use during construction of message-level meaning. *NeuroImage*, 62, 356–366.