EUR Research Information Portal

Picturing meaning: An ERP study on the integration of left or right-handed first-person perspective pictures into a sentence context

Published in:
Language, Cognition and Neuroscience

Publication status and date:
E-pub ahead of print: 11/08/2015

DOI (link to publisher):
10.1080/23273798.2015.1062901

Document Version
Publisher's PDF, also known as Version of record

Document License/Available under:
Article 25fa Dutch Copyright Act

Citation for the published version (APA):
de Nooijer, JA., Gootjes, L., van Gog, T., Paas, F., & Zwaan, R. (2015). Picturing meaning: An ERP study on the integration of left or right-handed first-person perspective pictures into a sentence context. *Language, Cognition and Neuroscience*, *31*(2), 167-178. Advance online publication. https://doi.org/10.1080/23273798.2015.1062901

Link to publication on the EUR Research Information Portal

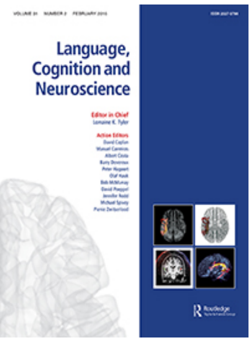
Terms and Conditions of Use
Except as permitted by the applicable copyright law, you may not reproduce or make this material available to any third party without the prior written permission from the copyright holder(s). Copyright law allows the following uses of this material without prior permission:

- you may download, save and print a copy of this material for your personal use only;
- you may share the EUR portal link to this material.

In case the material is published with an open access license (e.g. a Creative Commons (CC) license), other uses may be allowed. Please check the terms and conditions of the specific license.

Take-down policy
If you believe that this material infringes your copyright and/or any other intellectual property rights, you may request its removal by contacting us at the following email address: openaccess.library@eur.nl. Please provide us with all the relevant information, including the reasons why you believe any of your rights have been infringed. In case of a legitimate complaint, we will make the material inaccessible and/or remove it from the website.

Erasmus University Rotterdam
Making Minds Matter



Language, Cognition and Neuroscience

ISSN: 2327-3798 (Print) 2327-3801 (Online) Journal homepage: http://www.tandfonline.com/loi/plcp21

Picturing meaning: an ERP study on the integration of left or right-handed first-person perspective pictures into a sentence context

Jacqueline A. de Nooijer, Liselotte Gootjes, Tamara van Gog, Fred Paas & Rolf A. Zwaan

To cite this article: Jacqueline A. de Nooijer, Liselotte Gootjes, Tamara van Gog, Fred Paas & Rolf A. Zwaan (2016) Picturing meaning: an ERP study on the integration of left or right-handed first-person perspective pictures into a sentence context, Language, Cognition and Neuroscience, 31:2, 167-178, DOI: 10.1080/23273798.2015.1062901

To link to this article: http://dx.doi.org/10.1080/23273798.2015.1062901

LANGUAGE, COGNITION AND NEUROSCIENCE, 2016
VOL. 31, NO. 2, 167–178
http://dx.doi.org/10.1080/23273798.2015.1062901

Routledge
Taylor & Francis Group

Picturing meaning: an ERP study on the integration of left or right-handed first-person perspective pictures into a sentence context

Jacqueline A. de Nooijer[a], Liselotte Gootjes[a], Tamara van Gog[a], Fred Paas[a,b] and Rolf A. Zwaan[a]

[a]Institute of Psychology, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands; [b]Early Start Research Institute, University of Wollongong, Wollongong, Australia

ABSTRACT

Verbal and pictorial information are often processed together. Therefore, knowing how and when information from these modalities is integrated is important. In this ERP study we investigated integration of pictorial information into a sentence context. Right-handed participants heard sentences containing manual action verbs (e.g. "You are slicing the tomato"), while seeing a picture of a manual action. Pictures matched or mismatched the sentence content and the participants' handedness (i.e. pictures showed a left or right-handed perspective). Results showed a larger N400-amplitude for content-mismatching than for content-matching sentence-picture pairs. The N400-amplitude was not larger when the picture mismatched the participants' handedness. However, participants responded faster to right than to left-handed perspective pictures. This study suggests that with a sentence context, pictures are integrated with verbal information, but mental simulations either do not play a role in this process or this role might be too small to be visualised in the N400.

ARTICLE HISTORY
Received 3 July 2014
Accepted 11 June 2015

KEYWORDS
Sentence-picture integration; action simulation; ERP; N400; handedness

Deriving meaning from the integration of multiple sources of information is an essential part of language comprehension. Many researchers have searched for ways in which integration of verbal elements can be investigated. In the 1980s it was first shown that the difficulty to integrate two semantic constructions, becomes manifest in the ERP waveform by a negative going peak around 400 ms after a mismatch occurs (Kutas & Hillyard, 1980). For example, the word "cry" in the sentence "the pizza is too hot to cry", is difficult to integrate with the rest of the sentence, which results in an N400-effect with a centro-parietal maximum. Although there is some debate as to what stage in language processing the N400 reflects, lexical selection (i.e. word level) or lexical integration (i.e. discourse level), the account of lexical integration seems to be favoured (Salisbury, 2004). According to this account the N400 is an index of semantic processing that reflects the neural mechanisms of semantic integration into a context (Brown & Hagoort, 1993). Initially, studies on this semantic integration and the N400 have investigated integration of *verbal* elements into a sentence context (e.g. Connolly & Phillips, 1994; Hagoort, Hald, Bastiaansen, & Petersson, 2004; Kutas & Hillyard, 1980) or broader discourse (e.g. Nieuwland & Van Berkum, 2006; Salmon & Pratt, 2002; Van Berkum, Hagoort, & Brown, 1999). More recently, however, researchers have started to focus on the question of how information from other modalities, such as visual information (e.g. pictures), is integrated into a verbal context. It is very common to simultaneously encounter pictorial and verbal information, for example when reading a magazine or browsing the internet. Therefore, an interesting question is how information from these two sources is integrated. Concerning this type of integration, it was suggested that the language system incorporates semantic information coming from linguistic and extralinguistic domains over a similar neural time course and by recruitment of overlapping brain areas (Willems, Özyürek, & Hagoort, 2008). Thus, when we refer to integration of verbal and pictorial information, we mean the processing of information from two modalities into one coherent whole. This view of how integration works is in line with the one-step model of language comprehension. According to this model every source of information, whether linguistic or extralinguistic, immediately (i.e. as soon as the information from two modalities becomes available) constrains the interpretation of an utterance (e.g. Hagoort & Van Berkum, 2007; Tanenhaus & Trueswell, 1995). The visual (pictorial) context can, therefore, influence word recognition at the earliest moments during language processing (e.g. Spivey-Knowlton, & Sedivy, 1995; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995).

CONTACT Jacqueline A. de Nooijer ✉ j.a.denooijer@fsw.eur.nl

This integration of verbal and pictorial information was investigated at the word level. For example, when participants were first presented with a verbal object name at the categorical (e.g. dog) or specific (e.g. collie) level and then with a black and white picture that either matched or mismatched the object name, an N400-effect was found for both a basic and a subordinate mismatch (Hamm, Johnson, & Kirk, 2002). In a similar vein, when acoustically presenting 19-month olds and adults with words at the categorical level that were either congruent or incongruent with a picture content, both groups showed an N400-effect (Friedrich & Friederici, 2004; see also: D'Arcy & Connolly, 1999). However, a mismatch between verbal and pictorial information not only leads to difficulties with the integration of information at the word level but such difficulties might also occur at the sentence or discourse level. Several studies have investigated the integration of pictorial information into a sentence context. In some studies on sentence-picture integration, pictures were not presented simultaneously with the sentence, but with the picture as a replacement of a word in the sentence (e.g. Ganis, Kutas, & Sereno, 1996; Nigam, Hoffman, & Simons, 1992), where an N400-effect was found when a sentence was completed with a context inappropriate picture, implying that when the picture was congruent, the information from both modalities could be integrated (For similar results see: Knoeferle, Urbach, & Kutas, 2011; Wassenaar & Hagoort, 2007). Behavioural results have also shown that replacing words with pictures does not disrupt sentence processing (Potter, Kroll, Yachzel, Carpenter, & Sherman, 1986). However, also when the pictures only offer additional information, for example when participants were presented with short stories while simultaneously seeing a line drawing of an object that is congruent or incongruent with the context, an N400-effect for congruency was found (Willems et al., 2008). The finding that the N400-amplitude was larger for the incongruent items suggests that semantic integration was more difficult for the incongruent than for the congruent items. This shows that people even attempt to integrate information gleaned from pictures that offer additional information into the larger sentential context.

In addition to the study of integration of verbal and pictorial information, several studies investigated the integration of two sources of pictorial information. Interestingly, even without the presentation of verbal information, pictures themselves can lead to mismatches, resulting in integration difficulties. When the content of pictures is incongruent with world knowledge, an N400-effect arises, for example, when presenting participants with coloured pictures of people performing simple incongruent actions, (e.g. a woman cutting bread with a saw; Proverbio & Riva, 2009), tools shown in a false orientation (e.g. a screwdriver held horizontally where the screw is shown vertically; Bach, Gunter, Knoblich, Prinz, & Friederici, 2009), incorrect tool use (e.g. using a screwdriver to open a lock; Bach et al., 2009), inappropriate passing-receiving pictures (e.g. seeing a hand in a grasping position, for both the passing and receiving action; Shibata, Gyoba, & Suzuki, 2009) or when stories in the form of a series of pictures have an incongruent picture ending (West & Holcomb, 2002; see also: Mudrik, Lamy, & Deouell, 2010). These action-elicited N400 waves resemble the shape and timing of linguistic N400 waves, suggesting that the same neural mechanisms are involved in linguistic integration, as well as integration of information from the visual modality (Amoruso et al., 2013).

The research described above seems to suggest that verbal information and visual information can easily be integrated with one another, on a word, sentence or discourse level and even when two sources of information from the visual modality have to be integrated with each other, this is easily done. Even though language and sensorimotor processes seem to be integrated during the comprehension of everyday actions, it is still unclear how this happens in the brain (Amoruso et al., 2103). One factor that might be involved in the integration of information is the mental simulation created when processing information, as is stipulated in theories of embodied, or grounded, cognition. It has been suggested that the N400 can be understood within an embodied cognition framework (e.g. Hald, Marshall, Janssen, & Garnham, 2011), meaning that motor information can modulate meaning-related processes indicated by the N400 (Amoruso et al., 2013). Mental simulations also elicit (pre)motor activity (e.g. Aziz-Zadeh, Wilson, Rizzolatti, & Iacoboni, 2006) and might therefore influence the integration process. These mental simulations are evoked by both verbal and visual information. Studies have shown that motoric information is automatically activated when seeing an object or a picture of an object (Borghi et al., 2007; Sumner et al., 2007). Also, when reading an action word, the motor system becomes active (Hauk, Johnsrude, & Pulvermüller, 2004). Reading and seeing the same (manual) action should, therefore, in part lead to the same activation patterns in the brain which might facilitate integration.

If the mental simulation, that is, "the reenactment of perceptual, motor, and introspective states acquired during interaction with the world, body, and mind" (Barsalou, 2008, p. 618), evoked by the verbal information and the picture are congruent, this might facilitate the integration of verbal and pictorial information. However, pictures mismatching the mental simulation might hinder the integration of the picture in the sentence context. According to the body-specificity hypothesis hearing about an action leads to the creation of a body-specific mental simulation of that action. For example, when hearing the sentence: "You are stirring in the pot" right-handers would make a right-handed mental simulation (Casasanto, 2009). Therefore, seeing a left-handed picture perspective might provide a mismatch with the mental simulation and hinder integration of the picture in the sentence context, which could be reflected in the N400-component. Behavioural studies have shown that both memory and learning can be hindered when a perspective shown mismatched the participants' perspective. For instance, a recent study showed that motor activation induced by seeing pictures influenced memory performance (Apel, Cangelosi, Ellis, Goslin, & Fischer, 2012). Right-handers could remember more instructions when an object's handle was oriented to the right and actions also had to be performed with the right-hand. Also, research has shown that when learning new words coupled to a picture of a left or a right-handed picture perspective, right-handers recall fewer word definitions, when the picture seen during learning mismatches the right-handed mental simulation evoked by the verbal definition (De Nooijer, Van Gog, Paas, & Zwaan, 2013). Behavioural results can, however, reflect later processes than those reported from ERP data. If the effects of handedness on memory and learning tasks as reported above result from early processing, then a mismatch between the (right-handed) mental simulation evoked by the action verb and the left-handed picture will be reflected in the N400. If it is the result of slower-acting processes, it might be reflected in longer post-sentential reaction times to these items.

To summarise, we asked two questions. First, can we replicate the finding that pictorial information can be integrated with verbal information that is conveyed in a single sentence? We hypothesised on the basis of the current literature that pictures that mismatch the content of the sentence would evoke a larger N400-amplitude than pictures that match the sentence content, given that the mental simulations created by verbal and pictorial information are integrated with each other into a single sentence context.

Second, if pictorial information can be integrated with verbal information, is integration then facilitated when the handedness perspective of the picture matches that of the participant? Here we hypothesised that the N400-effect would be modulated by the hand perspective, where the largest N400 should occur for sentence-picture pairs that mismatch both in content and hand perspective (i.e. a left-handed picture). When both content information and handedness perspective mismatch, the integration process might be hindered. This effect might be strengthened by using sentences that are formulated in the second-person perspective, such as "you are stirring in the pot", (e.g. Sato & Bergen, 2013), combined with using first-person picture perspectives, which prime an actor's perspective.

Finally, on a sentence-picture verification task, longer response times are expected on the left-handed perspective items. In line with previous research (e.g. Apel et al., 2012; De Nooijer et al., 2013) we only expected this effect for right-handers, which is why we conducted the EEG experiment only with right-handers. To foreshadow, we did not find a perspective effect on ERPs, only on behavioural data. Based on a suggestion by an anonymous reviewer, however, we decided to test whether indeed only right-handers were influenced by handedness perspective in our sentence-verification task (which is different from tasks used in prior research on handedness effects). Therefore, this task was subsequently also investigated with left-handers, for whom we did not expect to find any differences in reaction times as a function of left- or right-handed perspective pictures (e.g. Apel et al., 2012; De Nooijer et al., 2013).

By investigating these questions, this study might contribute to answering the larger question of how language and sensorimotor processes are integrated during the comprehension of everyday actions, given that it is still unclear how this happens in the brain (Amoruso et al., 2013). With this study, we try to unravel one factor that might be relevant in this issue, namely the creation of mental simulations, by investigating whether pictures that mismatch the viewer's mental simulation hinder the integration process. Moreover, to move theories on grounded cognition forward, it is necessary to focus more on when we need or use grounded symbols and mental simulations and to try to understand the nature of these mental simulations (Dale, Dietrich, & Chemero, 2009; Dove, 2009; Zwaan, 2014). This study might contribute to answering such questions, as it can provide insight into whether mental simulations influence the integration of verbal and pictorial information, on which no information is available thus far.

Method

Participants

Twenty-five (16 female) undergraduate psychology students with a mean age of 23.2 years (SD = 3.2) participated in this study as part of an EEG tutorial. Participants were naïve to the experimental questions.

The experiment lasted approximately 10–15 min. All participants were native speakers of Dutch, had normal or corrected-to-normal vision and no neurological disorders. To avoid influencing participants during the experiment, we did not ask them about their handedness until after the experiment. One participant turned out to be left-handed and could therefore not be included in the analysis, leaving 24 participants. Data of another 3 participants were not of good enough quality (i.e. fewer than 20 segments remained in each condition after artefact rejection) leaving 21 participants for the final analyses.

In addition, 25 left-handers (16 female), of whom 24 were confirmed to be left-handers according to the Edinburgh Handedness inventory (meaning that one was ambidextrous and was therefore not included in the sample), participated in the sentence-picture verification task. These 24 participants had a mean age of 20.5 years (SD = 2.0) and received either course credit or a small monetary reward for participation.

Materials

Material consisted of 40 Dutch sentences and pictures. Sentences always consisted of the same four elements; the second-person pronoun "you" followed by a manual action verb, a definite article and an object (e.g. "Jij snijdt de tomaat" meaning "*You are slicing the tomato*"). Given that Dutch is an SVO language, the verb always appeared in the second position. Verbs had an average of 5.1 phonemes (SD = 1.39) and a mean log frequency of 1.0 (SD = 0.7). All sentences were recorded in a sound-attenuated room, spoken at a normal rate by a native Dutch female speaker. For all sentences we created a picture that matched the meaning of the sentence. In the picture a hand was shown that was executing the denoted action. The pictures were all taken from a first-person perspective and afterwards mirrored horizontally to create an otherwise identical right- and left-handed perspective. In addition, we made sure that all the hands depicted in the pictures were as neutral as possible (no rings, long nails, nail polish, etc.). The hands could, therefore, not be easily identified as being either male or female. This was done because handedness effects could possibly be greater in cases where the depicted hand has greater resemblance to the participant's hand. For an example of the stimuli see Figure 1.

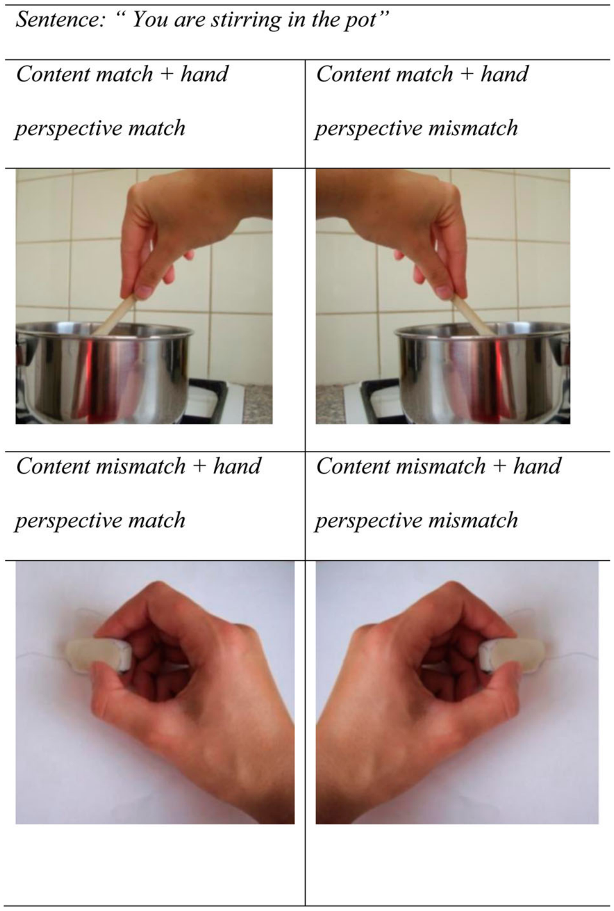


Figure 1. Example of the materials for four conditions.

Design and procedure

In a within-subjects design, participants were presented with spoken sentences combined with pictures that could (mis)match the action verb in two ways: (1) the content of the verb matched the content of the picture (i.e. hearing "you are slicing the tomato" and seeing a hand with a knife "slicing" a tomato) or mismatched the content (i.e. hearing "you are slicing the tomato" but seeing a hand "writing" something). (2) The hand perspective could match or mismatch the right-handed participants' hand perspective (i.e. seeing a right-hand slicing a tomato, vs. a left-hand slicing a tomato). This resulted in four conditions: (1) content and hand perspective match, (2) content match, hand perspective mismatch, (3) content mismatch, hand perspective match, and (4) content and hand perspective mismatch.

During the EEG experiment, which was implemented in E-Prime (Schneider, Eschman, & Zuccolotto, 2002) participants were seated in a comfortable chair and were instructed to minimise movement. They heard the sentences in four blocks, with each sentence being presented in each block, but coupled with a different picture, so that each sentence appeared once in each condition. (This procedure is similar to the one used in e.g. Bach et al., 2009; Friedrich & Friederici, 2004.) To prevent participants from guessing what the mismatching picture would be, we made sure picture-sentence pairs were never exchanged (e.g. if the mismatching picture for the sentence: "You are writing the report" was "cutting a tomato", then the mismatching picture for the sentence "You are cutting the tomato" would not be "writing a report"). Blocks were counterbalanced across lists, resulting in the use of four lists. Sentences were presented in a randomised order within blocks. Pictures were displayed for two seconds at the onset of the verb. All pictures were presented in the centre of the screen and were 400 × 400 pixels. The audio file lasted for three seconds after which the question: "Does the picture match the content of the sentence?" appeared. Participants were instructed to indicate as fast as possible whether the content of the sentence matched the picture by pressing the "J" (for "ja" = yes) or "N" (for "nee" = no) button on the keyboard with their dominant hand (the right or left middle finger and the right or left index finger respectively). The "J" and "N" are placed above each other on the keyboard, which should prevent an interference effect of button placement with the left and right-handed perspective pictures. Reaction times were calculated counting from the end of the sentence. Triggers were placed in order to mark the audio onset of the verb. Before the start of the experiment two practice trials were presented which contained different critical words than used in the main part of the experiment. Participants were told to attentively listen to and watch the stimuli and to blink their eyes only in between the sentences when a fixation cross was presented for 500 ms in the centre of the screen, at the beginning of each trial.

EEG recording

EEG was recorded from 32 electrode sites across the scalp using active Ag/AgCl electrodes (BioSemi, Amsterdam, the Netherlands, ActiveTwo amplifier system) placed in an elastic cap according to the 10/20 system. Electrodes were placed on standard sites (Fz, Cz, Pz, Oz, AF3, AF4, Fp1, Fp2, F3, F4, F7, F8, FC1, FC2, FC5, FC6, T7, T8, C3, C4, CP1, CP2, CP5, CP6, P3, P4, P7, P8, O1, O2, PO3, and PO4). Eye movements and blinks were monitored by four additional electrodes. Horizontal eye movements (HEOG) were measured by placing electrodes on the left and right outer canthi of the eye. Vertical eye movements (VEOG) were measured with electrodes above and below the left eye. Lastly, reference electrodes were placed on the left and right mastoid. Recordings were amplified using an ActiveTwo amplifier system and sampled at 512 Hz.

EEG analysis

EEG data were re-referenced offline to the linked mastoids. Segments of 1100 ms were created, including 100 ms before the target word onset (i.e. the action verb). Epochs were filtered with a 0.01–40 Hz band filter, and corrected for eye movements using the algorithm of Gratton, Coles, and Donchin (1983). Segments were only analysed when the correct answer was provided by the participant. In each condition the maximum number of data segments per participant was 40. If fewer than 20 segments remained in each condition after artefact rejection the participant was excluded from further analysis (3 participants), leaving 21 participants. From the remaining 21 participants on average 9% of the segments were rejected because of artefacts (12% in the content match, hand perspective mismatch, 12% in the content and hand perspective match, 7% in the content mismatch, hand perspective match and 6% in the content and hand perspective mismatch condition). Segments were normalised on the basis of the 100 ms pre-stimulus baseline. ERPs were calculated for each participant by averaging trials for each electrode and condition separately.

We used two time windows, an N400 time window (300–550 ms) and a late time window (600–900 ms). A late time window was included because some studies have found a broad negativity following the N400 in a similar design (e.g. Mudrik et al., 2010). We, therefore, included this time window, to investigate whether our results were in line with such a finding. On the basis of previous studies (e.g. Federmeier & Kutas, 2001; Holcomb & McPherson, 1994; Mudrik et al., 2010) the N400-effect for sentence-picture integration could be expected to have a more frontal maximum than the centro-parietal maximum that is usually reported (e.g. Kutas & Hillyard, 1980), however other studies found a centro-posterior N400 distribution for picture-sentence integration (e.g. Friedrich & Friederici, 2004; Knoeferle et al., 2011). To be able to detect where the N400-effect is strongest, we created four quadrants plus a midline section to examine the time windows. All recorded channels are included in these sections: left anterior (F3, F7, FC1, FC5, C3 AF3, FP1, and T7), right anterior (F4, F8, FC2, FC6, C4, AF4, FP2, and T8), left posterior (CP1, CP5, P3, P7, O1, and PO3) right posterior (CP2, CP6, P4, P8, O2, and PO4), midline (Fz, Cz, Pz, and Oz).

Statistical analyses

For the behavioural data of the right-handers two Repeated-Measures analysis of variance (ANOVA) were conducted; one based on participant variability and one based on item variability. Both analyses were conducted with content (picture-sentence [mis]match) and perspective (picture perspective [mis]match) as within-subjects factors for the reaction times on the sentence-picture verification task. Similar analyses were separately conducted for the data of the left-handers. Only reaction times to correct answers were analysed. In addition, reaction times faster than 100 ms or slower than 3000 ms were excluded from analyses. As a result 7.7% of the reaction time data from the right-handers and 8.3% of the reaction time data from the left-handers was not analysed.

For the EEG data of the right-handers a $5 \times 2 \times 2$ Repeated-Measures ANOVA was conducted with region (quadrants plus midline), content and perspective as within-subject factors for both the N400 and the late time window. In addition, to explore whether there are any differences in how the handedness perspective is processed by the two hemispheres, we performed an additional analysis with hemisphere (left, right), content (congruent, incongruent) and perspective (left, right) as within-subjects factors. When assumptions of sphericity are violated, results of Multivariate analyses are reported, given that these are not limited by the sphericity assumption (Jennings, Cohen, Ruchkin, & Fridlund, 1987). Only results with a *p*-value $< .05$ are interpreted as significant. When multiple comparisons are made, results were only considered as significant when they are Bonferroni-corrected for multiple comparisons.

In addition to classical frequentist Repeated-Measures ANOVAs, we also used Bayesian Repeated-Measures ANOVAs. Bayesian statistics allows us to calculate the probability a hypothesis is true, given the data. Bayes Factors for the separate effects in the Repeated-Measures ANOVA were computed with JASP 0.6 (http://jasp-stats.org). The Bayes Factor used is the inclusion Bayes Factor ($\mathrm{BF}_{\mathrm{inc}}$), which is an average of the likelihood of models that include the effect. It, therefore, compares all models with a certain factor against all the models without that factor. BFinc can be interpreted as follows concerning the evidence for the alternative hypothesis: $-\infty < B \leq 0.1$ is considered strong evidence against, $0.1 < B \leq (1/3)$ is substantial against, $(1/3) < B < 1$ barely worth mentioning against, $1 \leq B < 3$ barely worth mentioning for, $3 \leq B \ll 10$ substantial for, $10 \leq B < \infty$ strong for (Jeffreys, 1961). For example, when the BFinc is .02 this means that given the data, it is 50 times more likely that the null hypothesis is true, compared to the alternative hypothesis (i.e. the inverse of the inclusion Bayes Factor ($1/\mathrm{BF}_{\mathrm{inc}}$)).

Results

Behavioural data

Mean accuracy and reaction times on the accurate items are given in Table 1. Because of the high accuracy scores (Table 1) we only analysed the reaction times. The ANOVA for the right-handers showed no significant effect of content, $F(1, 23) = .56$, $p = .462$, $\eta_p^2 = .16$, BFinc $= .23$, on the reaction times on the accurate items. However, there was a significant effect of perspective, $F(1, 23) = 4.46$, $p = .046$, $\eta_p^2 = .16$, BFinc $= .41$, indicating that participants were faster to react to the right-handed than to the left-handed pictures. No significant interaction between content and perspective was found, $F(1, 23) = .02$, $p = .881$, $\eta_p^2 = .001$, BFinc $= .097$. Although we did expect to find an interaction between these factors, the null hypothesis is, given these data, over 10 times more likely than the alternative hypothesis (i.e. the inverse BF), which provides strong evidence for the null hypothesis. In addition, the item analyses showed no effect of content, $F(1, 39) < 1$, $p = .687$, $\eta_p^2 = .004$, BFinc $= .14$. There was also no significant interaction between content and perspective, $F(1, 39) < 1$, $p = .688$, $\eta_p^2 = .004$, BFinc $= .06$. Lastly, the effect of perspective was not significant, $F(1, 39) = 2.90$, $p = .096$, $\eta_p^2 = .069$, BFinc $= .42$.

The ANOVA for the left-handers showed a significant effect of content, $F(1, 23) = 4.59$, $p = .043$, $\eta_p^2 = .166$, BFinc $= 4.64$, indicating that the participants took longer to respond to the congruent items, compared to the incongruent items, which is in line with the accuracy scores for both left and right-handers, where it seemed that more correct answers were given on the incongruent items. There was no significant effect of perspective, $F(1, 23) < 1$, $p = .596$, $\eta_p^2 = .012$, BFinc $= .19$ nor a significant interaction between content and perspective, $F(1, 23) < 1$, $p = .752$, $\eta_p^2 = .004$, BFinc $= .21$. As the analysis for the left-handers showed no effect of perspective, only an analysis on participant's variability and not on item variability is reported here (Table 1).

ERP data

In Figure 2 the grand-averaged ERPs for four electrodes (two frontal and two central electrodes) are shown. We chose to show these electrodes because, as mentioned earlier, some studies investigating sentence-picture integration have found N400-effects with a centro-frontal maximum (e.g. Federmeier & Kutas, 2001; Holcomb & McPherson, 1994; Mudrik et al., 2010).

N400 time window (300–550 ms)

In Figure 3, the grand-averaged ERPs are shown for the four conditions in the four quadrants used. The ANOVA showed a significant main effect of region, $F(4, 17) = 19.71$, $p < .001$, $\eta_p^2 = .082$, BFinc $= \infty$, meaning that there is a significant difference between the frontal regions and the posterior and midline regions

Table 1. Means (and standard deviations) of accuracy (in %) and reaction times (in ms).

	Condition			
	Content + perspective match	Content match + perspective mismatch	Content mismatch + perspective match	Content + perspective mismatch
Accuracy				
Right-handers	94.3 (10.2)	94.0 (7.6)	99.2 (1.4)	99.2 (2.2)
Left-handers	97.5 (2.7)	97.3 (3.7)	99.5 (1.0)	99.0 (1.5)
Reaction time				
Right-handers	498 (136)	519 (175)	488 (157)	505 (156)
Left-handers	424 (153)	426 (145)	397 (118)	404 (103)

($p < .001$). This is, however, not relevant for our hypotheses, so we will not elaborate on these findings. There was also a significant effect of content, $F(1, 20) = 4.49$, $p = .047$, $\eta_p^2 = .18$, BFinc = 4.37, where we found a larger N400-amplitude for sentence-picture mismatching than for matching items. There was no effect of perspective, $F(1, 20) < 1$, $p = .951$, $\eta_p^2 = .001$, BFinc = .048. Given these data, the null hypothesis of no differences in performance on the perspective factor was about 21 times (inverse of the inclusion Bayes Factor) more likely than the alternative hypothesis, which could be described as strong evidence for the null hypothesis. Lastly, there was no interaction between content and perspective, $F(1, 20) = 1.49$, $p = .236$, $\eta_p^2 = .07$, BFinc = .057. Given these data, it is about 18 times more likely that the null hypothesis is true. It is, therefore, likely that there was truly no effect of perspective in the ERP data. Lastly, there was a significant interaction between region and

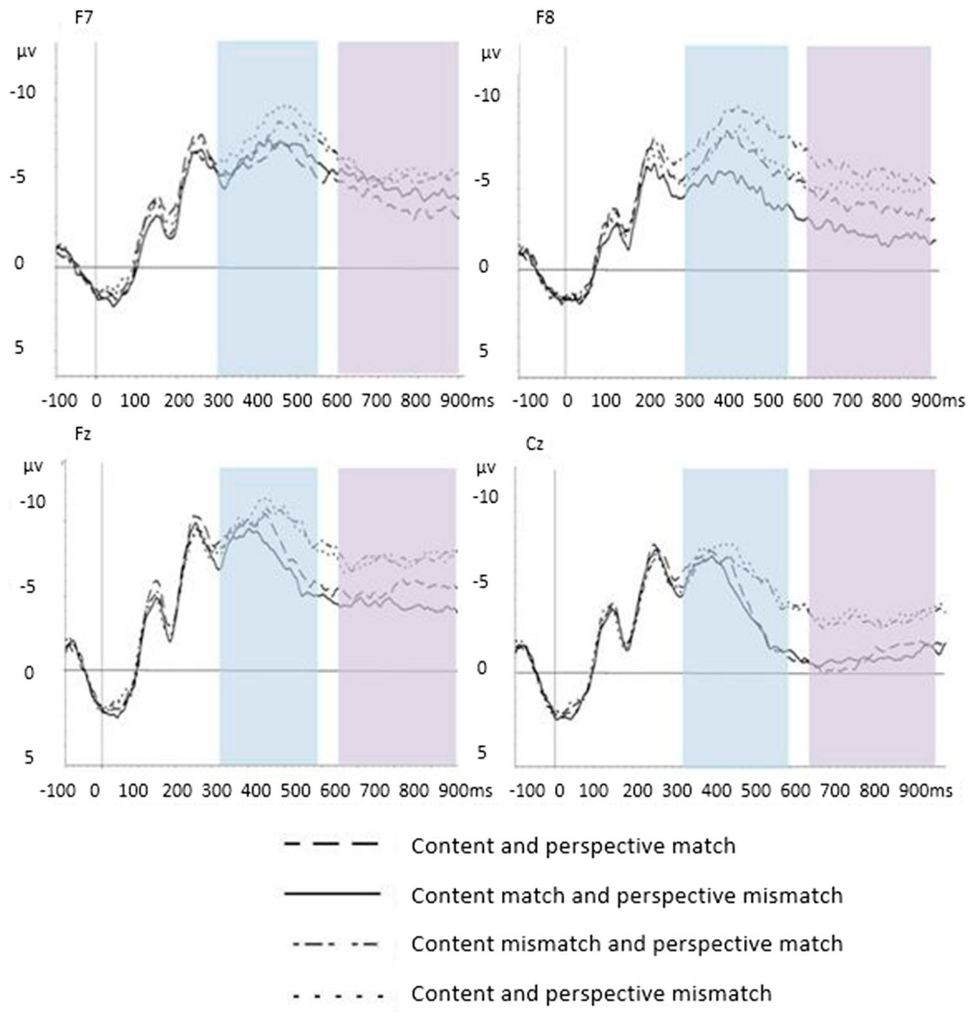


Figure 2. Grand-average ERPs for the four conditions at electrodes, F7, Fz, F8, Cz.
Note: Negativity is plotted upwards.

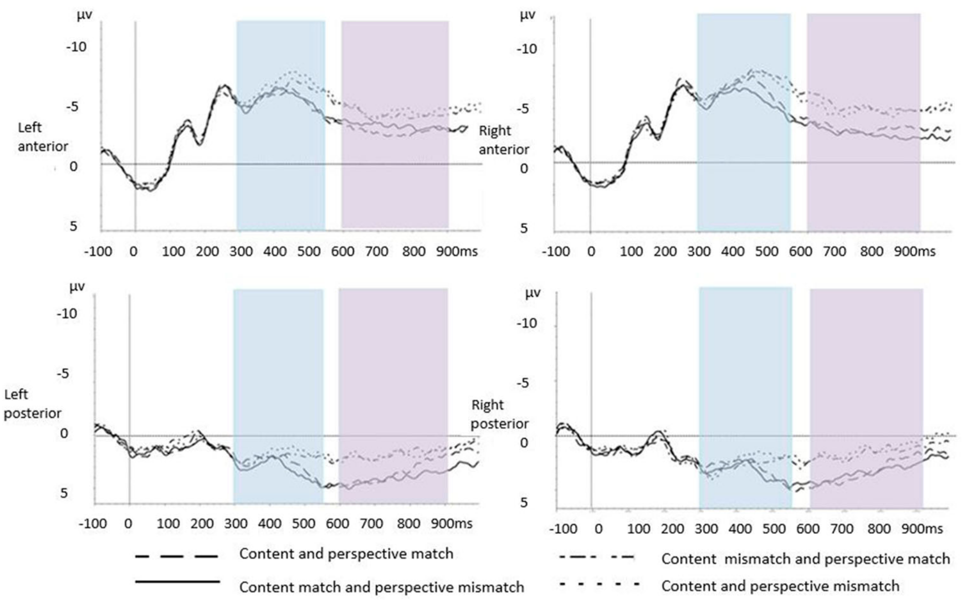


Figure 3. Grand-average ERPs for the four conditions for the four quadrants used.
Note: Negativity is plotted upwards.

perspective, $F(4,\ 17) = 4.28$, $p = .014$, $\eta_p^2 = .50$, BFinc $= .005$, although Bonferroni-corrected post hoc tests showed no significant differences between the left and right-handed perspective picture conditions in the different regions (left anterior: $F(1,\ 20) = 1.06$, $p = .315$, BFinc $= .33$; right anterior: $F(1,\ 20) = 1.25$, $p = .28$, BFinc $= .28$; left posterior: $F(1,\ 20) < 1$, $p = .839$, BFinc $= .22$; right posterior: $F(1,\ 20) < 1$ $p = .893$, BFinc $= .18$; midline: $F(1,\ 20) < 1$, $p = .737$, BFinc $= .26$).

In the design of this study we used a certain amount of repetition of stimuli, which was necessary, given that there are only a limited number of manual action verbs. However, repetition of stimuli can influence the size of the N400 (Rugg, 1985), therefore we performed a Repeated-Measures ANOVA on only the first block of stimuli before any repetition occurred to test for the effect of repetition. Even though the power of this analysis is much lower than the overall analysis, it does not seem to be the case that there is an effect of perspective anywhere, $F(1,\ 19) < 1$, $p = .438$, $\eta_p^2 = .03$, BFinc $= .13$). Given these data, the null hypothesis of no differences on the perspective factor was about 8 times more likely than the alternative hypothesis, meaning there is substantial evidence for the null hypothesis.

Lastly, we conducted a Repeated-Measures ANOVA with hemisphere (left, right), content (congruent-incongruent) and perspective (left-right) as within-subject factors, as we were interested in any interactions of content or perspective with hemisphere. We found a significant interaction between hemisphere and perspective, $F(1,\ 20) = 6.14$, $p = .022$, $\eta_p^2 = .235$, BFinc $= .02$. Although it appeared there was a larger negativity for the left-handed pictures processed in the left hemisphere and for the right-handed pictures processed in the right hemisphere, than for the left-handed pictures processed in the right hemisphere and the right-handed pictures processed in the left hemisphere, both the Bayes factor (which suggests that the null hypothesis is 50 times more likely) and follow-up post hoc tests (left hemisphere: $t(1,\ 20) < 1$, $p = .546$, BFinc $= .34$; right hemisphere: $t(1,\ 20) < 1$, $p = .549$, BFinc $= .35$), suggest there is no effect of perspective in either the left or right hemisphere.

Late time window (600–900 ms)

The ANOVA showed a main effect of region, $F(4,\ 17) = 36.20$, $p < .001$, $\eta_p^2 = .90$, BFinc $= 3.217 \times 10^{15}$, and a main effect of content, with a larger negative amplitude for the sentence-picture mismatching than for the matching items, $F(1,\ 20) = 13.72$, $p = .001$, $\eta_p^2 = .41$, BFinc $= 3.269 \times 10^8$. There was no effect of perspective, $F(1,\ 20) < 1$, $p = .991$, $\eta_p^2 = .001$, BFinc $= .049$, nor an interaction between content and perspective, $F(1,\ 20) < 1$, $p = .760$, $\eta_p^2 = .005$, BFinc $= .035$. Similar to the N400 time window, there was an interaction between region and perspective, $F(4,\ 17) = 5.14$, $p = .007$, $\eta_p^2 = .55$, BFinc $= .005$. As in the N400 time window, however, Bonferroni-corrected post hoc tests showed no significant differences between the left and right-handed picture conditions in the different regions (left anterior: $F(1,\ 20) < 1$, $p = .393$, BFinc $= .30$; right anterior: $F(1,\ 20) < 1$, $p = .542$, BFinc $= .23$; left posterior: $F(1,\ 20) < 1$, $p = .752$, BFinc $= .23$; right posterior: $F(1,20) < 1$,

$p = .763$, BFinc = .21; midline: $F(1, 20) < 1$, $p = .778$, BFinc = .22).

Lastly, like in the N400 time window the interaction between hemisphere and perspective was significant in the late time window. This analysis showed again a larger negativity for left-handed perspective pictures processed in the left hemisphere, and right-handed perspective pictures processed in the right hemisphere, than for right-handed pictures processed in the left hemisphere or left-handed perspective pictures processed in the right hemisphere. Like in the N400 time window, follow-up *t*-tests were not significant (left hemisphere: $t(1, 20) < 1$, $p = .667$, BFinc = .40; right hemisphere: $t(1, 20) < 1$, $p = .805$, BFinc = .30). No other interactions with the factor hemisphere were significant.

Discussion

In this study we investigated the integration of semantic information conveyed through sentences and pictures and the influence of mental simulations herein. Our results show a larger N400-amplitude for the sentence-picture mismatching than for the matching items in the N400 time window. This effect was relatively broad in its scalp distribution, given that there was no interaction between the factors content and region. Although some studies using daily actions found a more frontally distributed bias (e.g. Federmeier & Kutas, 2001; Holcomb & McPherson, 1994; Mudrik et al., 2010; West & Holcomb, 2002), others do not (Friedrich & Friederici, 2004; Knoeferle et al., 2011), which seems to be more in line with our findings. Given that the N400-component occurs in response to violations of semantic expectancy, this study suggests that people attempt to integrate pictorial information presented concurrently with verbal information presented in a single sentence context, where they have more difficulty with integration when the sentence content does not match the picture content. This finding is in line with other studies in the field that have investigated integration of verbal and pictorial information (e.g. Ganis et al., 1996; Knoeferle et al., 2011; Nigam et al., 1992; Willems et al., 2008) and with the one-step model of language processing according to which every source of information, whether linguistic or extralinguistic can immediately constrain the interpretation of an utterance (e.g. Spivey-Knowlton, & Sedivy, 1995; Tanenhaus et al., 1995).

Concerning the results for the late time window (600–900 ms) we found a larger negativity when the content of the sentence and picture mismatched. Previous research showed an N400-effect when presenting participants with congruent or incongruent visual scenes, but also a more broadly pronounced distributed negativity which the authors attributed to late processes of semantic evaluation and response preparation (Mudrik et al., 2010). This would also explain our results, given that our participants were required to give a response on whether picture and sentence matched in content. In other studies on this topic, a response was not always required (e.g. Willems et al., 2008) and therefore this late effect might have been less strong in these studies.

So what can these results tell us about how verbal and visual information are integrated, what is the underlying mechanism? One possibility, which we mentioned before, is that pictures elicit a mental simulation, while the sentence itself also evokes a mental simulation. These perceptual representations might be integrated into a coherent message. The creation of two matching mental simulations could then facilitate the integration of pictures in a single sentence context. If mental simulations facilitate the integration of information from two modalities, we would however expect the integration to be hindered by pictures that mismatch the observer's mental simulation (in this case the observer's hand preference), but this was not evidenced by our results in the N400. Although we used stimuli containing the pronoun "you" and first-person perspective pictures which both could have stimulated taking the actor's perspective in the simulation (Sato & Bergen, 2013), and even though the N400-effect seemed to be the strongest for the condition where the content and the hand perspective mismatched, there were no significant effects of hand perspective. Given that we did not instruct the participants on how to interpret the pronoun "you", participants could have taken the perspective of a bystander, which might have affected their expectations about the handedness perspective. This seems unlikely, however, because it has been shown that the use of the second-person pronoun leads to the adoption of the actor's perspective, without any instruction on how to interpret the pronoun (Sato & Bergen, 2013).

While the ERP results did not differ for perspective matches and mismatches, reaction times on the sentence-picture verification task did. This is in line with other behavioural tasks (i.e. learning and memory tasks), where an effect of these mental simulations was found (e.g. Apel et al., 2012; De Nooijer et al., 2013). Several studies have shown that right-handers' recall can have negative effects on memorisation (e.g. Linkenauger, Witt, Stefanucci, Bakdash, & Proffitt, 2009) and learning tasks (e.g. De Nooijer et al., 2013) when seeing the left-handed perspective, which may suggest that participants are influenced by the mismatching perspective in a conscious task, but not in automatic, fast, subconscious processes (as reflected in the N400). However, it could be the case that the faster reaction times of the right-handers on the right-handed perspective pictures are due to a type of motor priming, where responses with the right-hand were faster to right-handed perspective pictures. Another factor that could have played a role here is familiarity with the right-handed perspective, which might have caused faster processing of this perspective. Considering familiarity with the perspective, both left- and right-handers can be assumed to be more familiar with right-hander's actions as we most often observe other people's actions with a right-handed perspective (as 90% of the population is right-handed). However, with pictures from a first-person perspective, left-handers should be more familiar with the left-handed first-person perspective than with the right-handed perspective whereas they would be more familiar with the right-handed perspective when seeing pictures from the third-person perspective. If the results on the reaction time task are, therefore, due to motor priming or familiarity with the perspective, we would have expected left-handers (responding with their left-hand) to react faster to the first-person left-handed perspective pictures. There was, however, no difference in reaction time between the left-handed and right-handed perspective pictures, in left-handers. However, given that we do not have data on what perspective left-handers are most familiar with, we cannot entirely rule out that familiarity had (some) effect in these results.

Lastly, could it be the case that mental simulations (in the form of a handedness mismatch) do have an influence in the integration process, but was undetected in this study? Could, for example, the lack of finding an effect of perspective be due to the implicit nature of the task concerning perspective? No explicit mention was given to the perspective manipulation while participants did have to judge whether the content of the picture matched the sentence, which was the other manipulation. It seems that mental simulations in essence occur outside of awareness. Any effect of mismatching mental simulation on the integration of information would, therefore, have been expected when no attention is drawn to the handedness perspective. Because of this reason it is unlikely that any handedness effects were overshadowed by the effects of congruency. To provide more evidence for this statement, we conducted a Bayesian analysis that suggested that, given our data, the null hypothesis is 21 times more likely than the alternative hypothesis, which provides strong evidence for that hypothesis (Jeffreys, 1961).

Effects of mental simulations on language processes, as described earlier, were found without explicit attention to the left-right difference. Therefore, if there was an effect of perspective, we would indeed especially have expected this to occur when no attention was drawn to the perspective manipulation. Another factor one might argue could have influenced the results is the use of a design in which the sentences and pictures were presented more than once. This repetition was necessary because we wanted to investigate one type of verb, namely object-manipulation verbs that elicit motor activation. Given the limited number of these types of verbs, a certain amount of repetition was necessary to obtain enough power. Although repetition can influence the size of the N400-effect (Rugg, 1985), this influence has been found in some tasks (e.g. lexical decision) but not in others (e.g. word/number discrimination) (Bentin & McCarthy, 1994). Also, although the sentences and pictures were repeated, they were never repeated in the same combination, for which reason the correspondence of the sentence and picture, still had to be evaluated for each stimulus, which, therefore differs from when the exact same word in isolation is repeatedly presented and the same evaluation has to be given by the participant (as was the case in for example, Rugg, 1990). Lastly, an analysis of the first block of the experiment, before any repetition occurred, did not give any indication that an effect of hand perspective might have gone unnoticed. We, therefore, think it is implausible that the null-effect of perspective is due to a type II error.

To summarise, this study suggests that photographs of manual actions can quickly be integrated into a single sentence context. Concerning the hand perspective of the pictures, there is no neurophysiological evidence that hand perspective seems to matter for the ease of integration into the sentence, although it might have effects on behavioural tasks, involving learning, memory or comprehension. This study contributes to the literature on how and when information from different modalities is integrated. Even in a single context verbal and visual information are easily integrated, but in this integration process mental simulations in the form of a mismatching hand perspective, might not play a role based on the neurophysiological evidence presented in this study or this role might be too small to be visualised in the N400.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was funded by the Netherlands Organisation for Scientific Research and the National Initiative Brain & Cognition (project number: 056-33-016).

References

Amoruso, L., Gelormini, C., Aboitiz, F., González, M. A., Manes, F., Cardona, J. F., & Ibanez, A. (2013). N400 ERPs for actions: Building meaning in context. *Frontiers in Human Neuroscience*, *7*, 57. doi:10.3389/fnhum.2013.00057

Apel, J. K., Cangelosi, A., Ellis, R., Goslin, J., & Fischer, M. H. (2012). Object affordance influences instruction span. *Experimental Brain Research*, *223*, 199–206. doi:10.1007/s00221-012-3251-0

Aziz-Zadeh, L., Wilson, S. M., Rizzolatti, G., & Iacoboni, M. (2006). Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Current Biology*, *16*, 1818–1823. doi:10.1016/j.cub.2006.07.060

Bach, P., Gunter, T. C., Knoblich, G., Prinz, W., & Friederici, A. D. (2009). N400-like negativities in action perception reflect the activation of two components of an action representation. *Social Neuroscience*, *4*, 212–232. doi:10.1080/17470910802362546

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617–645. doi:10.1146/annurev.psych.59.103006.093639

Bentin, S., & McCarthy, G. (1994). The effects of immediate stimulus repetition on reaction time and event-related potentials in tasks of different complexity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 130–149. doi:10.1037/0278-7393.20.1.130

Borghi, A. M., Bonfiglioli, C., Lugli, L., Ricciardelli, P., Rubichi, S., & Nicoletti, R. (2007). Are visual stimuli sufficient to evoke motor information? Studies with hand primes. *Neuroscience Letters*, *411*, 17–21. doi:10.1016/j.neulet.2006.10.003

Brown, C., & Hagoort, P. (1993). The processing nature of the N400: Evidence from masked priming. *Journal of Cognitive Neuroscience*, *5*, 34–44. doi:10.1162/jocn.1993.5.1.34

Casasanto, D. (2009). Embodiment of abstract concepts: Good and bad in right- and left-handers. *Journal of Experimental Psychology: General*, *138*, 351–367. doi:10.1037/a0015854

Connolly, J. F., & Phillips, N. A. (1994). Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. *Journal of Cognitive Neuroscience*, *6*, 256–266. doi:10.1162/jocn.1994.6.3.256

Dale, R., Dietrich, E., & Chemero, A. (2009). Explanatory pluralism in cognitive science. *Cognitive Science*, *33*, 739–742. doi:10.1111/j.1551-6709.2009.01042.x

D'Arcy, R. C. N., & Connolly, J. F. (1999). An event-related brain potential study of receptive speech comprehension using a modified Token Test. *Neuropsychologia*, *37*, 1477–1489. doi:10.1016/S0028-3932(99)00057-3

De Nooijer, J. A., Van Gog, T., Paas, F., & Zwaan, R. A. (2013). When left is not right: Handedness effects on learning object-manipulation words using pictures with left or right-handed first-person perspectives. *Psychological Science*, *24*, 2515–2521. doi:10.1177/0956797613498908

Dove, G. O. (2009). Beyond perceptual symbols: A call for representational pluralism. *Cognition*, *110*, 412–431. doi:10.1016/j.cognition.2008.11.016

Federmeier, K. D., & Kutas, M. (2001). Meaning and modality: Influences of context, semantic memory organization, and perceptual predictability on picture processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 202–224. doi:10.1037/0278-7393.27.1.202

Friedrich, M., & Friederici, A. D. (2004). N400-like semantic incongruity effect in 19-month-olds: Processing known words in picture contexts. *Journal of Cognitive Neuroscience*, *16*, 1465–1477. doi:10.1162/0898929042304705

Ganis, G., Kutas, M., & Sereno, M. I. (1996). The search for "common sense": An electrophysiological study of the comprehension of words and pictures in reading. *Journal of Cognitive Neuroscience*, *8*, 89–106. doi:10.1162/jocn.1996.8.2.89

Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, *55*, 468–484. doi:10.1016/0013-4694(83)90135-9

Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science*, *304*, 438–441. doi:10.1126/science.1095455

Hagoort, P., & Van Berkum, J. (2007). Beyond the sentence given. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*, 801–811. doi:10.1098/rstb.2007.2089

Hald, L., Marshall, J., Janssen, D. P., & Garnham, A. (2011). Switching modalities in a sentence verification task: ERP evidence for embodied language processing. *Frontiers in Psychology: Cognition, Special issue on Embodied and Grounded Cognition*, *2*, 45. doi:10.3389/fpsyg.2011.00045

Hamm, J. P., Johnson, B. W., & Kirk, I. J. (2002). Comparison of the N300 and N400 ERPs to picture stimuli in congruent and incongruent contexts. *Clinical Neurophysiology*, *113*, 1339–1350. doi:10.1016/S1388-2457(02)00161-X

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, *41*, 301–307. doi:10.1016/S0896-6273(03)00838-9

Holcomb, P. J., & McPherson, W. B. (1994). Event-related brain potentials reflect semantic priming in an object decision task. *Brain and Cognition*, *24*, 259–276. doi:10.1006/brcg.1994.1014

Jeffreys, H. (1961). *The theory of probability*. Oxford: The Clarendon Press.

Jennings, J. R., Cohen, M. J., Ruchkin, D. S., & Fridlund, A. J. (1987). Editorial policy on analyses of variance with repeated measures. *Psychophysiology*, *24*, 474–478.

Knoeferle, P., Urbach, T. P., & Kutas, M. (2011). Comprehending how visual context influences incremental sentence processing: Insights from ERPs and picture-sentence verification. *Psychophysiology*, *48*, 495–506. doi:10.1111/j.1469-8986.2010.01080.x

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*, 203–205. doi:10.1126/science.7350657

Linkenauger, S. A., Witt, J. K., Stefanucci, J. K., Bakdash, J. Z., & Proffitt, D. R. (2009). The effects of handedness and reachability on perceived distance. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 1649–1660. doi:10.1037/a0016875

Mudrik, L., Lamy, D., & Deouell, L. Y. (2010). ERP evidence for context congruity effects during simultaneous object–scene processing. *Neuropsychologia*, *48*, 507–517. doi:10.1016/j.neuropsychologia.2009.10.011

Nieuwland, M. S., & Van Berkum, J. J. (2006). When peanuts fall in love: N400 evidence for the power of discourse. *Journal of Cognitive Neuroscience*, *18*, 1098–1111. doi:10.1162/jocn.2006.18.7.1098

Nigam, A., Hoffman, J. E., & Simons, R. F. (1992). N400 to semantically anomalous pictures and words. *Journal of Cognitive Neuroscience*, *4*, 15–22. doi:10.1162/jocn.1992.4.1.15

Potter, M. C., Kroll, J. F., Yachzel, B., Carpenter, E., & Sherman, J. (1986). Pictures in sentences: Understanding without words. *Journal of Experimental Psychology: General*, *115*, 281–294. doi:10.1037/0096-3445.115.3.281

Proverbio, A. M., & Riva, F. (2009). RP and N400 ERP components reflect semantic violations in visual processing of human actions. *Neuroscience Letters*, *459*, 142–146. doi:10.1016/j.neulet.2009.05.012

Rugg, M. D. (1985). The effects of semantic priming and word repetition on event-related potentials. *Psychophysiology*, *22*, 642–647. doi:10.1111/j.1469-8986.1985.tb01661.x

Rugg, M. D. (1990). Event-related brain potentials dissociate repetition effects of high-and low-frequency words. *Memory & Cognition*, *18*, 367–379. doi:10.3758/BF03197126

Salisbury, D. F. (2004). Semantic memory and verbal working memory correlates of N400 to subordinate homographs. *Brain and Cognition*, *55*, 396–399. doi:10.1016/j.bandc.2004.02.057

Salmon, N., & Pratt, H. (2002). A comparison of sentence-and discourse-level semantic processing: An ERP study. *Brain and Language*, *83*, 367–383. doi.10.1016/S0093-934X(02)00507-2

Sato, M., & Bergen, B. K. (2013). The case of the missing pronouns: Does mentally simulated perspective play a functional role in the comprehension of person? *Cognition*, *127*, 361–374. doi:10.1016/j.cognition.2013.02.004

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime: User's guide*. Pittsburgh: Psychology Software Incorporated.

Shibata, H., Gyoba, J., & Suzuki, Y. (2009). Event-related potentials during the evaluation of the appropriateness of cooperative actions. *Neuroscience Letters*, *452*, 189–193. doi:10.1016/j.neulet.2009.01.042

Spivey-Knowlton, M., & Sedivy, J. C. (1995). Resolving attachment ambiguities with multiple constraints. *Cognition*, *55*, 227–267. doi:10.1016/0010-0277(94)00647-4

Sumner, P., Nachev, P., Morris, P., Jackson, S. R., Kennard, C., & Husain, M. (2007). Human medial frontal cortex mediates unconscious inhibition of voluntary action. *Neuron*, *54*, 697–711. doi:10.1016/j.neuron.2007.05.016

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*, 1632–1634. doi:10.1126/science.7777863

Tanenhaus, M. K., & Trueswell, J. C. (1995). Sentence comprehension. In J. L. Miller & P. D. Eimas (Eds.), *Speech, language, and communication* (pp. 217–262). San Diego, CA: Academic Press.

Van Berkum, J. J., Hagoort, P., & Brown, C. M. (1999). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of Cognitive Neuroscience*, *11*, 657–671. doi:10.1162/089892999563724

Wassenaar, M., & Hagoort, P. (2007). Thematic role assignment in patients with Broca's aphasia: Sentence-picture matching electrified. *Neuropsychologia*, *45*, 716–740. doi:10.1016/j.neuropsychologia.2006.08.016

West, W. C., & Holcomb, P. J. (2002). Event-related potentials during discourse-level semantic integration of complex pictures. *Cognitive Brain Research*, *13*, 363–375. doi:10.1016/S0926-6410(01)00129-X

Willems, R. M., Özyürek, A., & Hagoort, P. (2008). Seeing and hearing meaning: ERP and fMRI evidence of word versus picture integration into a sentence context. *Journal of Cognitive Neuroscience*, *20*, 1235–1249. doi:10.1162/jocn.2008.20085

Zwaan, R. A. (2014). Embodiment and language comprehension: Reframing the discussion. *Trends in Cognitive Sciences*, *18*, 229–234. doi:10.1016/j.tics.2014.02.008