Abstract The process of describing an object’s location relative to another object results in ambiguity. How do people handle this ambiguity? The present studies examined spatial language processing when use of different reference frames results in ambiguity. We investigated whether electrophysiological (ERP) measures of cognitive processing may elucidate underlying reference frame processing; in particular, we were interested in semantic integration. ERP results showed a larger N400, peaking between 300 and 375 ms, when the intrinsic frame was not used. Behavioural results mirrored this finding, indicating a reduced cognitive processing requirement for the intrinsic reference frame. Previous work has not definitively tied spatial reference frame processing to specific ERP components and their associated cognitive processes. Although the N400 peak seen in this data is early, additional work supports the N400 interpretation, thereby linking spatial frame processing to semantic integration. Results are discussed within the larger context of spatial reference frame processing.

Résumé L’opération qui consiste à décrire l’endroit où se trouve un objet par rapport à un autre est source d’ambiguïté. Mais comment les gens composent-ils avec cette ambiguïté ? Les présentes études se sont penchées sur le traitement du langage spatial lorsque le recours à différents cadres de référence produit de l’ambiguïté. Nous avons tenté de déterminer si des mesures électrophysiologiques (ERP) permettent d’éclairer la nature du traitement du cadre de référence sous-jacent; nous nous sommes tout particulièrement intéressés à l’intégration sémantique. Ces mesures montrent une composante N400 plus importante, avec une pointe se situant entre 300 et 375 ms, dans les cas où le cadre intrinsèque n’est pas utilisé. Les résultats comportementaux concordent avec cette conclusion, puisqu’ils indiquent des exigences réduites dans le traitement cognitif appliqué au cadre de référence intrinsèque. Des travaux antérieurs n’ont pas encore établi de lien clair entre le traitement du cadre de référence spatial et des composantes particulières des potentiels évoqués, ainsi que les processus cognitifs qui leur sont associés. Malgré que la pointe que nous avons observée à l’onde N400 soit précoce, une autre étude appuie notre interprétation, reliant ainsi le traitement du cadre spatial à l’intégration sémantique. Ces résultats sont interprétés dans le contexte plus général du traitement du cadre de référence spatial.
same terminology. Other ambiguities arise when the vertical plane is considered.

Relative and intrinsic reference frames can either be aligned or misaligned. In the misaligned case, does one reference frame receive priority? Researchers have found mixed results in considering this question. Some argue that the relative frame receives priority because it is perceptually available. Other frames must be cognitively computed, through processes such as mental rotation. Other researchers have found little contribution of the relative frame, but this may be specific to the vertical plane where gravity plays a strong role.

Most research on spatial frame processing has used behavioural measures, and found interesting results, but has left questions unanswered. If multiple reference frames are initially activated and one is eventually selected, do behavioural measures reflect activation or selection? Carlson-Radvansky and Jiang (1998) cleverly showed multiple activation and inhibition/selection using negative priming. ERPs, because they can illustrate the time-course of cognitive processing, should provide converging evidence and illustrate both reference frame activation and selection processes. In fact, recent work by Carlson, West, Taylor, and Herndon (2000) showed such evidence using ERPs.

Our earlier work (Taylor, Naylor, Faust, & Holcomb, in press) used ERPs to examine spatial frame processing on the horizontal plane. Results showed greater negativity between 200 and 450 ms when the intrinsic frame was not used. However, the absence of other ERP work on spatial frame processing, the low relevance of spatial attention studies using ERPs, and the poor temporal resolution of neuroimaging studies of spatial processing made interpreting this result difficult. Based on psycholinguistic work, we hypothesized that this negativity reflected the N400 component.

The N400 was first described by Kutas and Hillyard (1980a, 1980b) in relation to anomalous sentence comprehension. ERPs time locked to the final word of a sentence showed a greater N400 for semantically anomalous words. More recent interpretations of the N400 attribute it to general semantic integration. Holcomb (1993) suggests that the N400 should be evident in any situation requiring semantic processing; information that is more difficult to integrate results in larger N400s. Semantic integration is integral to spatial frame processing, since one is interpreting or generating a spatial description based on a scenario. The semantics of the description must be applied to the semantics of the scenario. The ambiguity of multiple reference frames adds to the difficulty of semantic integration, just as semantic ambiguity does (van Petten & Kutas, 1987). As such, if one reference frame is prioritized, semantic integration for this frame may be easier.

Although we hypothesized that results of Taylor et al. (1999) reflected the N400, it is possible that the results reflected another component. The negative-going component in this study peaked at approximately 325 ms. The N400 more typically peaks at or even after 400 ms. The present studies attempt to show the utility of the ERP methodology for examining spatial frame processing by more closely tying processes underlying spatial frame interpretation to issues of semantic integration and the N400.

**Experiment 1**

To examine the link between spatial frame processing and semantic integration, participants viewed pictures of two objects, one with intrinsic sides and one without (a donut). Participants then saw a spatial term (right, left, front, or back) and decided whether it correctly described the donut's location, considering both the intrinsic and relative reference frames. Picture/locative term combinations resulted in four reference frame conditions, defined by the reference frame use by the locative term (intrinsic, relative, both, neither).

**METHOD**

**Participants.** Twelve undergraduates (11 male, 1 female) participated in partial fulfillment of a course credit. All participants were right handed, had normal or corrected-to-normal vision, and were native English speakers.

**Materials.** Pictorial stimuli consisted of 120 images depicting two objects. One object in each picture had intrinsic sides (e.g., car, chair). The second object in each picture was the same, donut-shaped object, and did not have intrinsic sides. Sixteen combinations...
resulted from four orientations of the intrinsic object (facing towards the participant, away from the participant, to the left, or to the right) and four donut locations (intrinsically in front of, in back of, to the left of, or to the right of the intrinsic object). See Figure 1 for an example picture. Word stimuli consisted of three locative terms, left, right, and front.

**Apparatus.** Tin electrodes (Electro-cap International) were placed at 13 scalp sites, including occipital left and right (O1 and O2), frontal left and right (F7, F8), frontal (FZ), central (CZ), and parietal (PZ) and over anterior temporal cortex (ATL, ATR), posterior temporal cortex (TL and TR), at temporo-parietal cortex (WL and WR). Two eye electrodes measured eye movements and blinks, one placed horizontally to the right eye and the other placed vertically under the left eye. An electrode placed at the left mastoid (A1) served as a reference and an electrode placed over the right mastoid (A2) measured any mastoid asymmetry. Impedance levels were kept below 5 kohms. EEGs were amplified by Grass Model 12 Neurodata Acquisition System, with gain set at 50 k. The band-pass levels were set between .01 and 30 hz, and the sample rate was at 200 hz. Trials with excessive eye movements, blinks, or blockage were rejected.

**Procedure.** Participants sat in a sound attenuated room. Presentation of stimuli subtended approximately four degrees of visual angle. Prior to the start of the experiment, participants filled out a handedness questionnaire. During application of the electrode cap, participants read an instruction sheet providing a general overview of the study. With the electrode cap in place, participants received a joystick with two buttons labeled “T” and “F.” The order of T and F placement was counterbalanced across participants.

Trials consisted of a “ready” signal, where participants pressed either button to begin the trial. Next, the picture of the two objects appeared for 1,000 ms, followed by a cross for 800 ms, then one of the locative terms for 1,000 ms, and finally a “T/F” prompt signaling the participant’s response. A 300-ms ISI appeared between each of these elements for experimental pacing. ERPs were time locked to the locative term. Participants were to respond “true” if the spatial term described the donut’s location, using either the intrinsic or the relative reference frame. For example, if the car were facing to the left and the donut appeared on the left side of the screen, the terms left and front both describe the donut’s location. The experimenter explained the difference between the two possible reference frames and told participants they should consider both. Participants received 10 practice trials.

In the experimental run, participants received 240 trials. Trials were defined by which frame the locative term described. Forty trials each represented scenes and locative terms where both reference frames were correct, where the intrinsic frame was correct, and where the relative frame was correct. Finally, to achieve a balance between true and false responses, 120 trials showed scenes and locative terms where neither frame was correct.

**RESULTS**

We formed averaged ERPs for each of the experimental conditions, excluding trials containing ocular or movement artifacts. To quantify the ERPs, mean amplitudes for the 200 ms to 425 ms epoch were calculated. Repeated measures ANOVAs, using Geisser/Greenhouse corrections for degrees of freedom greater than one, were used to examine experimental differences. Analyses looked at lateral and midline electrode sites separately. Lateral analyses examined factors of reference frame, hemisphere, and electrode site (five sites). Midline sites included reference frame and electrode site (three sites). Specific pairwise comparisons followed up the main analyses. Alpha levels were set at .05 level.

Analyses showed a main effect of reference frame (lateral: $F(3, 33) = 9.188$; midline: $F(3, 33) = 5.065$). Results also showed a reference frame by hemisphere interaction, $F(3, 33) = 4.959$. This greater negativity was larger for the right hemisphere.

Pairwise comparisons teased apart the reference frame effect. Overall, these analyses indicate that when the locative term did not describe the intrinsic frame, greater negativity appeared in the waveforms (see Figure 2). Comparison of trials using the relative frame and those using the intrinsic frame showed a significant reference frame effect (midline: $F(1, 11) = 7.664$) and a reference frame by hemisphere interaction at lateral sites, $F(1, 11) = 7.295$. A comparison of both frame trials to neither frame trials also showed a significant reference frame effect (lateral: $F(1, 11) = 21.729$; midline: $F(1, 11) = 8.335$). Again, reference frame interacted with hemisphere, $F(1, 11) = 8.453$. A comparison between both frame trials and relative frame trials showed a similar reference frame main effect (lateral: $F(1, 11), 13.878$; midline: $F(1, 11) = 7.281$) and a reference frame by hemisphere interaction, $F(1, 11) = 7.288$. Note that these three comparisons pit a condition where the intrinsic frame is used to one where it is not, all showing a greater negativity, peaking between 300 and 375 ms, when the intrinsic frame was not used. Additionally, reference frame differences tended to be larger over the right hemisphere. The greater right hemisphere difference appears to derive from a greater
left hemisphere positivity. A comparison pitting two conditions where the intrinsic frame was not used (neither frame vs. relative frame) showed no main effect, although a follow-up analysis showed a peak latency difference for lateral sites, \(F(1, 11) = 17.341\). No other comparisons showed significant differences.

**DISCUSSION**

Results showed a greater negativity, peaking between 300 and 375 ms after presentation of the locative term, when the intrinsic frame was not used. This replicates Taylor et al., (1999) providing converging evidence that spatial description/frame processing involves the N400 and semantic integration. The present study uses a simplified methodology from Taylor et al., requiring participants to respond to a single locative term instead of a descriptor phrase. Despite the lowered semantic demands of the spatial description, the N400-like effect remained. The present study also examines all four possible reference frame conditions, adding the condition where the intrinsic and relative frames align. One could argue that this condition could further ease semantic integration, since both reference frames afford the same description. This, however, does not appear to be the case, as the both reference frame condition did not differ significantly from the intrinsic frame condition.

Despite this replication, questions remain unanswered. Does the greater negativity truly reflect the N400 component and consequently semantic integration? Why does the frame that on the surface seems to require greater cognitive processing, the intrinsic frame, appear to require less semantic integration? The relative frame is perceptually available, whereas the intrinsic frame must be calculated through cognitive operations such as mental rotation.

We still do not have definitive evidence that the negativity seen in this study reflects the N400. The N400 does not generally peak as early as seen in this study, thus questioning whether it is an N400. Supporting the N400 is the fact that the distribution of the effect fits the typical N400 distribution, with the largest effects at midline central/parietal sites. The greater frontal distribution is also common with pictorial stimuli. Since the

---

*Figure 2. ERP results of reference frame conditions. The label “participant reference frame” corresponds to the relative reference frame. The label “object reference frame” corresponds to the intrinsic reference frame.*
early peak is problematic, Experiment 2 will use more typical N400 paradigms to determine whether earlier N400 peaks are possible. Experiment 3 will examine reference frame processing using behavioural measures for convergent evidence. Greater semantic integration generally leads to reduced accuracy and increased reaction times. Additionally, it will explore an explanation for the initial processing of the intrinsic frame.

Experiment 2
In this experiment, we sought to determine if the negativity seen in Experiment 1’s ERPs reflected a genuine N400, or whether it reflects another, as yet unidentified, component. Evidence in favour of a different component comes from the relatively early peak latency of the negativity (~325 ms). More typically, the N400 peaks at or even after 400 ms. The early peak, however, may be explained by the stimulus design of Experiment 1. Unlike most N400 studies, stimuli in Experiment 1 repeated many times. Participants may have quickly become familiar with the stimuli, thus saving some processing overhead on subsequent trials with repeated stimuli. However, the experimental conditions in Experiment 1 differed enough from traditional N400 paradigms to question the veracity of this position. Experiment 2 uses traditional semantic priming designs to examine whether repeated stimuli influence the peak latency of the N400.

METHOD: EXPERIMENT 2A
Participants. Sixteen right-handed native English-speakers took part in the study (8 male, 8 female). All were undergraduates at Tufts University taking Introductory Psychology courses.

Stimuli. Modified stimuli from Experiment 1 were used here. Pictures of the 10 objects, without the donut, appeared in three different orientations. Pictures were paired with target words front, left, and right. Of the 150 picture-word pairs, half had a match between the object orientation and the target word, and half did not.

Procedure. Each trial started with a word “READY,” and participants pressed a button to continue. After that, the screen was blank for 500 ms, then the picture appeared for 1,000 ms followed a fixation cross for 800 ms, then the target word appeared for 1,000 ms, and finally the “T/F” prompt appeared, signaling participants to respond. A 300-ms ISI occurred between each stimulus in a trial for experiment pacing. The positions of the true and false buttons were counterbalanced across participants.

For each picture-word pair, participants decided whether the side of the object closest to them matched the target word, using an intrinsic reference frame. Before the experiment, participants were familiarized with each object and participated in 10 practice trials.

ERP recording. The electroencephalogram (EEG) was recorded from 29 tin electrodes held in place by an elastic cap (Electrode-Cap International). Only data from the 13 sites near those used in Experiment 1 will be reported here (Fz, Cz, Pz, O1, O2, F7, F8, P3, P4, CP5, CP6, FC5, FC6). All other ERP recording procedures were the same as in Experiment 1.

METHOD: EXPERIMENT 2B
Participants. The volunteers from Experiment 2A also participated in this experiment.

Stimuli. Eighty colour drawings of animals and 80 colour drawings of tools were paired with target words “animal” or “tool.” Objects were shown on a white background. Half of the pictures in each object category matched the target word and half did not. Stimulus pairs were presented in a random order. Each participant saw each picture only once, but across the participant sample each picture occurred in the matching and mismatching conditions.

Procedure. This experiment was completed after Experiment 2A, following a short break. The procedure was the same as in Experiment 2A with the following exceptions. No fixation was shown after the picture presentation. Instead, directly after a 300-ms ISI, the target word appeared for 500 ms, followed by an 800-ms ISI before the response prompt. Participants pressed a “True” button if the picture matched the category name and a “False” button if it did not. Participants received 10 practice trials.

RESULTS: EXPERIMENT 2A
Mean amplitude analyses used the 200- to 425-ms time window.

Visual inspection. As can be seen in Figure 3, the ERPs from Experiment 2A were very similar in overall morphology to those from Experiment 1. Following the P2, there were several late ERP components including a negativity (N400) peaking at 300 ms (especially in the Mismatch condition), a positivity (P3) peaking at 350 ms, and either a late negativity around 500 ms or a slow return from baseline after the 350-ms positivity.

Epoch analyses. Target words that mismatched the preceding prime picture direction produced a larger negative-going waveform between 200 and 400 ms than did targets that matched (midline: \( t(1,15) = 30.24 \); lateral:
$F(1,15) = 23.32$). Over lateral sites this effect was somewhat larger over temporal and parietal than over more anterior sites (site by target type interaction, lateral: $F(4,60) = 5.03$, $p < .02$). Between 400 and 600 ms, matching targets produced slightly more negative ERPs than did mismatching targets (midline: $F(1,15) = 7.23$; lateral: $F(1,15) = 4.36$).

**RESULTS: EXPERIMENT 2B**

**Visual inspection.** Examination of Figure 4 reveals that the ERPs from this experiment were very similar to those in Experiments 1 and 2A. The major difference was that Experiment 2A produced a somewhat larger positivity at 350 ms.

**Epoch analyses.** The target words that mismatched the category of the prime picture produced a larger negative going waveform between 200 and 400 ms than did targets that matched (midline: $F(1,15) = 20.99$; lateral: $F(1,15) = 12.62$). There were no significant target type effects between 400 and 600 ms.

**DISCUSSION**

The results strongly suggest that the main findings in Experiment 1 illustrated an N400. Experiments 2A and 2B used standard N400 paradigms, but repeated stimuli many times, as was done in Experiment 1. The repeated stimuli appear to reduce cognitive preprocessing requirements associated with the stimuli, thus allowing for an earlier N400 peak. Therefore, it seems reasonable that the repetition of target items was also responsible for an earlier N400 peak in Experiment 1, rather than the negativity reflecting a different ERP component.

**Experiment 3**

Experiment 3 had two purposes. First, we hoped to provide behavioural data convergent with Experiment 1 results, thus aiding interpretation of the N400 finding. Behavioural data generally coincides with the N400 effect, [e.g., Bentin, McCarthy, & Wood] although this relationship is not absolute (e.g., Holcomb, 1993; Kounios & Holcomb, 1992). Second, we wished to further explore issues of semantic integration by varying the experimental procedures. Models of sentence/picture verification suggest the involvement of four processes: sentence encoding, picture encoding, comparison, and response. Further, these models suggest that sentence and picture information is compared serially (Carpenter & Just, 1975; Clark & Chase, 1972).
If so, comparisons between picture and verbal information should be more difficult when the picture has multiple interpretations and the picture information must be held in memory, both of which were true in Experiment 1. One way to strategically handle the memory load is to process information requiring greater cognitive computation, in this case the intrinsic frame, when the picture is perceptually available. If this strategy is used, semantic integration between the spatial term and the reference frame not initially processed should be more difficult. The present experiment explores this idea.

METHOD
Participants. Sixty undergraduates (39 female, 21 male) participated in partial fulfillment of a course credit. All participants were native English speakers.

Materials. Materials were identical to those used in Experiment 1, with the following exceptions. Six of the ten intrinsic objects were used. The telephone, television, and house were eliminated because further testing indicated that people interpret the right/left axis of these objects using a relative reference frame. The Sphynx was eliminated due to its unusual nature. This experiment also used four canonical locative terms (right, left, front, and back) for completeness. One hundred and ninety-six trials took into account object facing direction, donut location, and locative term.

Procedure. The procedure followed that of Experiment 1, with the following exceptions. First, we collected behavioural data, instead of ERPs. Second, participants responded when the locative term appeared, rather than waiting for the response signal. RT was measured from onset of the locative term to key press. Third, participants performed under one of three experimental conditions, defined by the timing between picture and locative term presentation. The delay condition matched Experiment 1. Participants saw the picture, followed by the locative term. In the simultaneous condition, the picture and locative term appeared at the same time. In the simultaneous-delay condition, participants saw the picture, then the locative term appeared while the picture remained.

RESULTS
Here we will report results paralleling Experiment 1
analyses, namely reference frame effects, and those related to the experimental condition. Other analyses are beyond the scope of this paper and can be found with related research in Taylor and Rapp (2001).

Accuracy. Results showed a main effect of reference frame, $F(3, 171) = 66.04$, $MSe = 0.009$. Participants were least accurate when the locative term used the relative frame ($M = 0.73$), second least accurate when the term used either the intrinsic frame ($M = 0.92$) or neither frame ($M = 0.93$), and most accurate when the term used both frames ($M = 0.96$).

This reference frame main effect was qualified by an interaction with condition, $F(6, 171) = 2.84$, $MSe = 0.009$. While the overall pattern follows the reference frame main effect, responses to relative frame differed as a function of condition. For relative frame trials, the delay condition resulted in the lower accuracy (see Figure 5).

Reaction time. Analyses for RT data used two of the experimental conditions, delay and simultaneous-delay. In these conditions, participants could process the picture prior to seeing the locative term. In contrast, RTs for the simultaneous condition include initial picture processing time, which is not of interest for the current study.

Results showed a reference frame main effect, $F(3, 114) = 22.91$, $MSe = 212,645$. Participants responded slowest to the relative frame ($M = 1,771$ ms) or to neither frame ($M = 1,850$), next slowest to the intrinsic frame ($M = 1,599$) and fastest when both frames were used ($M = 1,461$). Results also revealed a condition main effect, $F(1, 38) = 4.42$, $MSe = 2,936,500$. Participants in the delay condition ($M = 1,528$ ms) responded faster than those in the simultaneous-delay condition ($M = 1,813$).

Although, taken together, the RT and accuracy results might reflect a speed/accuracy trade-off, this seems unlikely. RT and accuracy results paralleled one another for the reference frame effects, and it is unlikely that a speed/accuracy trade-off would appear for one variable and not another.

DISCUSSION

Behavioural results mirrored the N400 results of Experiment 1. Participants responded slower and less accurately when the intrinsic frame was not used. As also reflected by the N400 in Experiment 1, participants appear to have more difficulty integrating the locative term to the relative frame interpretation of the picture.

The results also demonstrate that picture availability when processing the spatial term affects semantic integration. When the picture is not available, participants may strategically process some of the information, more likely information requiring more cognitive effort. Later, if the spatial term does not match this initial processing, additional semantic integration is needed and can presumably be extracted from memory. The greatest support for this argument comes from the interaction between reference frame and experimental condition for accuracy. This interaction primarily involved the relative frame trials, where participants responded less accurately in the delay condition (i.e., when they had to rely on memory of the picture).

General Discussion

Spatial directions are assigned using reference frames. A reference frame consists of orthogonal axes, defined by specific parameters, such as orientation, direction, and scale. From the reference frame, a spatial template for a specific spatial term can be established and used to determine appropriate usage of the term. Matching spatial terms to specific situations requires semantic interpretation of the term and the situation. The present studies examined the role of semantic integration in reference frame use, focusing on two reference frames, the intrinsic and relative.

We specifically examined the N400 in spatial reference frame processing, because of its link to semantic integration. Event-related potentials are sensitive to the time course of cognitive processing and ERP components have been linked to specific cognitive processes. Experiment 1 showed a larger negativity between 300 and 375 ms after the locative term when the term did not apply to the intrinsic reference frame. This finding replicated our earlier work, using a simplified, yet more
complete, design. Experiments 2A and 2B confirmed that the negativity in Experiment 1 was an N400, rather than another, unidentified component, by showing that the N400 can have an early peak when stimuli repeat multiple times.

Experiment 3 provided converging behavioural evidence for the N400 interpretation. ERPs showed a larger N400 when the intrinsic frame was not used; behavioural data showed slower RTs and lower accuracy in the same situation. Further, they are consistent with other recent ERP work on spatial reference frames, indicating early selection and preference of a reference frame.

Experiment 3 also indicated that the primacy of the intrinsic frame may be strategic. Participants showed increased accuracy for the relative frame when they could process it perceptually. Others have shown similar differences between sequential and simultaneous presentation and sentence verification (e.g., Glushko & Cooper, 1978). The comparison process involved in sentence/picture verification should be made easier when all information is not held in memory.

The process of semantically mapping a spatial term onto a scene is complicated when that scene has multiple interpretations. Since participants first see the scene, they must derive an interpretation of this scene to later match to the spatial term. The results here indicate that participants put priority on the intrinsic interpretation. Further semantic interpretation is needed if the spatial term does not match this intrinsic interpretation.

Taken together with Taylor et al. (1999), the results of the present studies are fairly unequivocal. For the experimental situation defined in these studies, participants require less semantic integration when spatial terms match the intrinsic frame interpretation than when they match the relative frame interpretation. Further research is needed to help define situational variables that may influence reference frame priority.

Experiment 1 was submitted to Tufts University by the second author in fulfillment of his Senior Honor’s Thesis. We would like to thank Lauren Brodsky, Justin Belmont, and Belkiss Murati for their help collecting data. Additionally, we would like to thank David Rapp, Laura Carlson, and an anonymous reviewer for comments on an earlier draft of this manuscript.

Requests for reprints should be addressed to Holly A. Taylor, Department of Psychology, Tufts University, Medford, MA 02155 (E-mail: htaylor@tufts.edu).

References


