Common ERP responses to narrative incoherence in sentence and picture pair comprehension

Anne-Lise Jouen, Nicolas Cazin, Sullivan Hidot, Carol Madden-Lombardi, Jocelyne Ventre-Dominey, Peter Ford Dominey

Brain and Cognition 153 (2021) 105775

Contents lists available at ScienceDirect

Brain and Cognition

journal homepage: www.elsevier.com/locate/b&c

Common ERP responses to narrative incoherence in sentence and picture pair comprehension

Anne-Lise Jouen [a], Nicolas Cazin [b], Sullivan Hidot [b], Carol Madden-Lombardi [b], Jocelyne Ventre-Dominey [b], Peter Ford Dominey [b,*]

[a] Department of Neuropsycholinguistics, Université de Genève, Geneva, Switzerland
[b] INSERM UMR1093-CAPS, Université Bourgogne Franche-Comté, UFR des Sciences du Sport, Dijon, France

ARTICLE INFO

Keywords:
ERP
Language
Late positivity
Narrative

ABSTRACT

Understanding the neural processes underlying the comprehension of visual images and sentences remains a major open challenge in cognitive neuroscience. We previously demonstrated with fMRI and DTI that comprehension of visual images and sentences describing human activities recruits a common extended parietal-temporal-frontal semantic system. The current research tests the hypothesis that this common semantic system will display similar ERP profiles during processing in these two modalities, providing further support for the common comprehension system. We recorded EEG from naïve subjects as they saw simple narratives made up of a first visual image depicting a human event, followed by a second image that was either a sequentially coherent narrative follow-up, or not, of the first. Incoherent second stimuli depict the same agents but shifted into a different situation. In separate blocks of trials the same protocol was presented using narrative sentence stimuli. Part of the novelty is the comparison of sentence and visual narrative responses. ERPs revealed common neural profiles for narrative processing across image and sentence modalities in the form of early and late central and frontal positivities in response to narrative incoherence. There was an additional posterior positivity only for sentences in a very late window. These results are discussed in the context of ERP signatures of narrative processing and meaning, and a current model of narrative comprehension.

1. Introduction

A major function of higher cognitive processing is making sense of the world around us based on accumulated experience that is organized in a narrative structure (Bruner, 1990). Binder et al. (2009) note that knowledge acquired from experience underlies our ability to understand, and forms the basis of the semantic system. In a meta-study of 120 PET and fMRI studies that access meaning from words, they identified a widely distributed network that suggests that semantic representations tap into sensory, motor, affective and cognitive systems recruited in human experience.

We hypothesized that such a broadly distributed semantic coding is not restricted only to verbal material, but rather that there will be a common network for representing meaning issued from all experience, including verbal and also visual image input. This idea is consistent with early work on language comprehension (Biederman, 1981; Friedman, 1979; Gernsbacher & Faust, 1991; Mandler & Johnson, 1976), and more recent work demonstrating common processes for integrating linguistic and visual picture information (Ganis et al., 1996; Shinkareva et al., 2011; Willems et al., 2008), and for integrating audio and video narrative (Baldassano et al., 2017; Chen et al., 2017; Zadbood et al., 2017). In this context, the current research continues our long term effort to better understand this network, and to test the hypothesis of a common neural system involved in understanding sentences and images. In our first test of this hypothesis, we determined that a common semantic network would be recruited in the comprehension of visual images, and sentences that depict human events (Jouen et al., 2015). fMRI and DTI revealed the fronto-temporo-parietal spatial organization of this network and aspects of its connectivity. Interestingly, the common network was quite similar to that identified by Binder et al. (2009), including major activation in the angular gyrus and temporo-parietal cortex. This was inspired by the groundbreaking work of Vandenberghe et al. (1996) in understanding the common semantic system, and extended their approach from simple images and single words, to rich

* Corresponding author.
E-mail address: peter.dominey@inserm.fr (P.F. Dominey).

https://doi.org/10.1016/j.bandc.2021.105775
Received 8 March 2021; Received in revised form 13 July 2021; Accepted 22 July 2021
Available online 29 July 2021

images and full sentences.

In further investigation of this hypothesis we performed detailed mapping of the white matter pathways and functional connectivity linking the cortical nodes of this distributed network, with major hubs in the anterior temporal cortex and the temporo-parietal cortex (Jouen et al., 2018). The dense connectivity of these areas highlights their roles in integrative processing. Such densely connected temporo-parietal regions could serve as anchor points for the convergence of multimodal cortical representations during experience. During comprehension, activation of such convergence zones could then allow for divergent reactivation of multimodal areas in reconstructing meaning (Lallee & Dominey, 2013; Meyer & Damasio, 2009). While fMRI and DTI thus provide a view of the spatially distributed network organization of the semantic system and its connectivity that will be engaged in making meaning from narrative, a different approach is required for characterization of the temporal dynamics of this integrative processing. The current research will examine the underlying spatio-temporal brain activation trajectories that compliment these functional and anatomical fMRI studies.

Given its temporal precision, electroencephalography (EEG) is a more suitable tool for investigating the temporal unfolding of processes in narrative integration. We thus set out to compare the temporal dynamics of the EEG signal in visual image and sentence processing using short (two-element) narratives that allowed us to manipulate the sequential narrative coherence of the successive stimuli within the narratives. We used sequential stimuli that are semantically and syntactically well formed, but with the second that is coherent, or not, in its narrative succession with respect to the first. The goal was to determine if there are ERP responses that are common to image and sentence processing, which would provide further evidence for common underlying neural mechanisms.

While it is not the focus of our study, to set the context it is worth noting that one of the most robust ERP effects in language processing is the N400, a negativity around 400 ms after words that are semantically anomalous, such as “socks” in “He spread his warm bread with socks,” vs. the more fitting “butter” (Kutas & Hillyard, 1980). Since its discovery, the N400 has been extensively investigated, and remains a key indicator of semantic processing (Kutas & Federmeier, 2011). N400 effects have been elicited not only by semantic anomalies in language, but by images, e.g. photos of human goal-directed activities that violate concepts in semantic memory (Sitnikova et al., 2008).

A second language-related ERP that has been extensively documented is the P600 or syntactic positive shift that was initially characterized as being elicited by syntactic anomalies (Hagoort et al., 1993; Kim & Osterhout, 2005; Osterhout & Hagoort, 1999; Osterhout & Holcomb, 1992). These early observations led to a general distinction between N400 reflecting semantic processing vs P600 reflecting syntactic processing. This view has become more refined, as these two can be evoked together with no syntactic anomaly, with the N400 followed by post N400 positivities (PNP) reflecting semantic predictability and plausibility (DeLong et al., 2014). This distinction has been further refined, in the context of the semantic illusion (Brouwer et al., 2012; Hoeks et al., 2004), where semantic anomalies that respect memory retrieval constraints but violate scene integration constraints can produce a P600 in the absence of N400 (Brouwer et al., 2012). Such a P600 in the absence of N400 was clearly evoked in a narrative discourse where two sentences were perfectly correct in terms of syntax and semantics, but where the second required a discourse-level inference to generate a coherent semantic representation (Burkhardt, 2007).

Written and visual narratives have been investigated separately using EEG in protocols where different dimensions of narrative coherence have been manipulated. Late ERP positivities with some variability in their localization and timing are frequently observed in these coherence manipulations. This can be seen during violation of the expectation or goal in visual narrative (Cohn et al., 2014; Sitnikova et al., 2008), and violation of semantic expectations built up earlier in a text in verbal narrative (Bornkessel-Schlesewsky & Schlesewsky, 2008; Brouwer et al., 2012; Paczynski & Kuperberg, 2012; Xiang & Kuperberg, 2015). Cohn and Foulsham (2020) recently observed late positivities extending up to 1100 ms during visual narratives where zoomed panels likely required more integration processing.

Kuperberg and colleagues (Brothers, Wlotko, Warnke, & Kuperberg, 2020) have recently focused on the processing characteristics associated with late frontal and posterior positivities during language comprehension. A common element for the late frontal and late posterior positivities is that a discourse context must first be established. When subsequent text requires an update of this rich situation model, the late frontal positivity is favored. In contrast, when subsequent text is anomalous in the context of the situation model, the late posterior positivity is favored.

In addition to these late positivities, earlier positivities in the 300 ms time-frame have been associated with semantic processing (Polich, 2007). The timing and distribution of these effects can be modulated by different dimensions of the stimuli. Early positive responses related to the P3b can have peak latencies well over 400 ms (Verleger et al., 2005). A long and rich history of research has investigated the links between these earlier positivities, later positivities and their relation with cognitive processes that that are sensitive incoherence in semantic predictability (Leckey & Federmeier, 2020). In the domain of visual narrative, positivities between 400 and 600 ms were evoked in narrative sequences that manipulated motion lines in a non-expected way (anomalous motion lines) (Cohn & Maher, 2015). It will thus be of interest to determine if there are corresponding positivities in our manipulations of visual and textual narrative.

In this overall context, we sought to identify ERP responses that can reliably reveal (or not) common processes for sentence and image processing, in the particular case where the individual stimuli are well-formed, and what is manipulated is the relation between these stimuli. The novelty of the current research is the focus on the detection of a form of narrative coherence in the comprehension of sentences and images in the same subjects. It will contribute to the literature in the field in two ways: first by extending the results on the common semantic system for visual and textual narrative from the functional (Jouen et al., 2015) and anatomical (Jouen et al., 2018) MRI domains into the more temporally resolved ERP domain. This allows a characterization of the spatio-temporal trajectories of brain activation implemented in the distributed networks identified in the fMRI and connectivity studies. Second, it will consider visual and textual narrative processing in the context of late frontal and posterior positivities that have been characterized in the language domain (Brothers, Wlotko, Warnke, & Kuperberg, 2020; Kuperberg, Brothers, & Wlotko, 2020).

To address this problem we adapted the protocol from Jouen et al. (2015) to the EEG domain, retaining the overall structure of the protocol that involved the presentation of natural visual images depicting human activity, and in separate blocks of trials, short whole sentences depicting the same type of human activity. We introduced an aspect of temporal succession corresponding to simple narratives. In both the sentence and the image domains, two successive stimuli were presented, separated by a variable delay. The first stimulus (either a picture or sentence) established a rich context or mental model. The second stimulus could be sequentially coherent or incoherent in a narrative context with respect to the first. Incoherent second stimuli depict the same agents but shifted into a different situation. In the context of the distinction made by Kuperberg and colleagues (Brothers, Wlotko, Warnke, & Kuperberg, 2020), this would correspond to an unexpected follow-up that requires update of a situation model, vs. anomalous follow-up that requires reanalysis or repair. Brain responses to these stimuli were analyzed to determine if there was evidence for a common processing of sequential narrative coherence in sentence and image modalities, in the form of common EEG responses to incoherence in these modalities.

We predict that in the incoherent conditions, for both sentences and images, there will be an increased late positivity, similar to those evoked

in the semantic illusion and discourse conditions (Brouwer et al., 2012; Burkhardt, 2007; Hoeks et al., 2004), and in visual narrative processing (Cohn & Foulsham, 2020; Cohn & Maher, 2015; Cohn et al., 2014; Sitnikova et al., 2008). The important point is that each of the two successive stimuli are syntactically and semantically well-formed with the incoherence at a higher level of narrative integration, typically associated with a late positivity (Brouwer et al., 2012; Burkhardt, 2007). In such cases of well-formedness of the separate sentences and images, we do not predict that the incoherent stimuli will produce N400 responses.

2. Methods

2.1. Participants

Eighteen healthy right-handed volunteers participated in the experiment (10 females, 8 males, native French speakers, without prior neurological history, 25.5 ± 4.1 years of age). The study was performed under approval (Authorization No. 10028) from the Rhône-Alpes Préfecture review board authorizing biomedical research. In accordance with the Declaration of Helsinki, all participants were advised of the physical details of the experiment, and gave their informed written consent. The experiment was conducted in French. Stimuli and instructions are translated here in English.

2.2. Stimuli

The temporal unfolding of the paradigm is presented in Fig. 1. For the image conditions one-hundred and twenty (120) images pairs were selected from the Getty photo database (http://www.gettyimages.fr/). A set of sixty pairs of images made up sequentially coherent narratives, and a second set of sixty pairs of images made up sequentially incoherent narratives. In a pretest phase, 10 participants evaluated our set of images such that both images were accurately described as the intended event, and the coherence (or lack of coherence) was correctly evaluated for 93% of pairs by all 10 subjects. The few pairs that were incorrectly evaluated for coherence were replaced.

In parallel we created a set of 60 pairs of sequentially coherent sentences, and a set of 60 pairs of sequentially incoherent sentences. The sentence pairs represented the same kinds of daily events that were depicted in the image pairs. Each sentence was also evaluated by 10 participants for comprehensibility (>90%) as well as imageability (how easy it is to form a mental image of the described event) on a scale of 1 (extremely difficult) to 5 (extremely easy), yielding a high imageability score of 4.53. In addition, the sentence pairs were correctly evaluated for 93% of pairs by 9 out of 10 participants, and the few sentences that generated incorrect coherence evaluations were replaced.

Image conditions were balanced in terms of portrait/landscape orientation. Both sentence and image conditions were balanced in terms of gender, age and number of people in the pictured/described event. Sentence conditions were controlled for number of letters in each sentence, grammatical constructions (reflexive verbs, relative propositions), and open-class word frequency using the Lexique database (New et al., 2001).

When a narrative was sequentially coherent, the images depicted the same people performing a logical sequence of activities (e.g. entering a bakery; buying bread). Sequentially incoherent pairs depicted the same

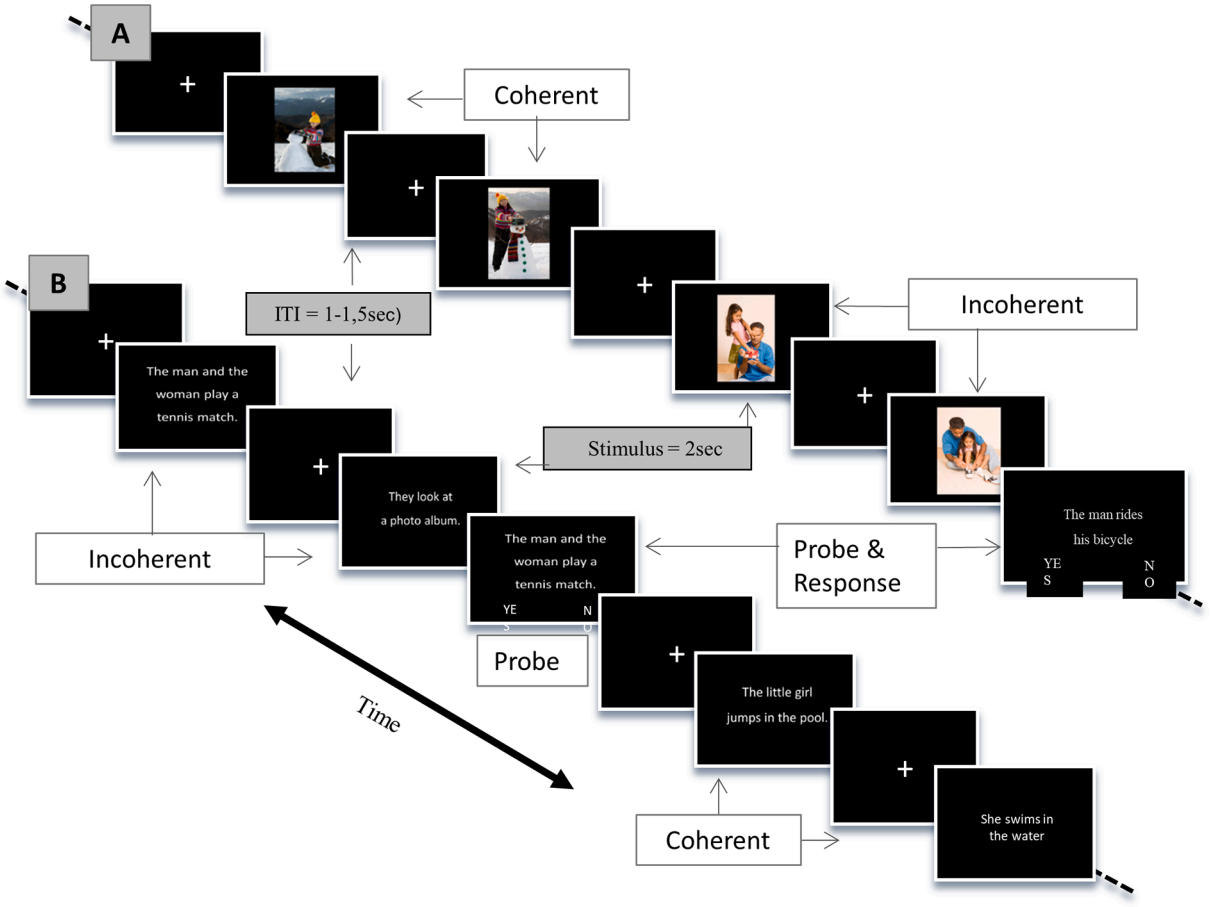

Fig. 1. Temporal unfolding of the experimental protocol for the image (A) and sentence (B) conditions. Trials begin with a fixation point for random delay between 1000 and 1500 ms. The first stimulus (sentence or image) is then presented for 2000 ms. After a delay of 1–1.5 s the second stimulus is presented for 2000 ms. On some trials, a probe question is presented and left visible until the subject responds yes or no.

people performing two unrelated activities. The images were controlled and counterbalanced for the number of people illustrated (1 or 2). Likewise, for the sentence conditions, sixty pairs of sentences were created that made up sequentially coherent narratives, and sixty pairs for incoherent narratives.

Example sentences

Sequentially Coherent

S1- The man opens the driver side door. S2 - He starts the car.
S1 – The woman and man play a game of cards. S2 – She wins the hand.

Sequentially Incoherent

S1 - The woman lounges by the pool. S2 - She irons the laundry.
S1 – The little girl tickles the little boy. S2 – They brush their teeth.

In the same way that images can implicate the same people in different situations over the presentation of two successive stimuli, the sentences describing people performing an action can follow-up (or not) in the second sentence. To reproduce a similar effect as with the images, that is, that certain information does not need to be re-analysed (e.g. the identity of the people, their different roles in the relations that link them) we chose to use complete nominal specification for stimulus 1 (e.g. "The man and the woman looked at the movie announcements") and replace the agents with pronouns to refer to the same protagonists in stimulus 2 (e.g. "They bought tickets at the counter."). This is consistent with pragmatic discourse rules and avoids the repeated name penalty (Gordon et al., 1993). While we do not use a serial word presentation, we ensured that the S2 sentences were of the form "Pronoun VERB ….", so that the verb which allows detection of narrative coherence was accessed immediately. This is confirmed in the ERP responses below.

2.3. Experimental paradigm

Subjects were seated in front of a visual display. Visual stimuli (sentences or images) were presented at the center of the screen, subtending a visual angle of approximately 5°. Subjects saw a visual image (or sentence) depicting a human event, and after a pause saw a second image (or sentence) that was either sequentially coherent, or not, with the first image (or sentence). Stimulus one and two were always in the same modality (sentence or visual image).

A trial started with a fixation point for a variable delay of 1-1.5 s, then the first stimulus was presented for 2 s. After a delay of 1–1.5 s the second stimulus was presented for 2 s followed either by a fixation point for 1–1.5 s or a question. Two thirds of the trials were followed by probe questions in order to maintain vigilance. The probe question was displayed until the subject responded yes or no to the question by pressing with their right hand respectively a right and left key on a button pad. The probe question was always related to the stimuli of the previous trial to avoid any memory confounds. Trials were blocked by modality (image or sentence), two blocks per modality, for a total of 4 blocks. Each block had 30 coherent, and 30 incoherent trials.

2.4. EEG acquisition and preprocessing

We acquired continuous neural activity with 64 channel EEG (Biosemi, ActiveTwo, version 5.36) sampled at 2 Khz while subjects performed the task. EEG data was processed using EEGLAB. Preprocessing was performed with the FASTER plugin for EEGLAB (Nolan, Whelan, & Reilly, 2010), with bandpass filter from 1 Hz to 95 Hz, notch filter (50 Hz), artifact rejection and epoching from −350 to 3000 ms relative to the stimulus onset. Thus, for sentences and images, EEG epochs were synchronized with the onset of the stimulus in the sentence and image conditions All electrodes were referenced with respect to the mastoids. Artifacts from eye movements, blinks and temporal muscle activity were identified and removed using a second-order blind identification (SOBI) algorithm implemented in the EEGLAB toolbox interface (Lio & Boulinguez, 2013).

2.5. ERP analysis

Average waveforms were computed across all trials per condition, using a baseline subtraction with the interval [−200 0 ms]. Multifactor ANOVAs and post-hoc comparisons with Greenhouse-Geisser correction for sphericity were performed using Statistica, on the factors: Modality (Image, Sentence), Narrative Coherence (Coherent, Incoherent) and Topography for different time windows. To account for scalp topography, we divided the electrodes in a 3 (left, middle, right) × 3 (anterior, central posterior) spatial topography as illustrated in Fig. 2C. In the three way ANOVAs, the Topography factor thus had 9 levels.

The analysis was performed on the critical second stimulus for three time windows chosen as a function of known stimulus driven responses and the morphology of the ERP responses illustrated in Fig. 2A and B. The first is a 400–500 ms window with a stimulus driven positivity corresponding to the classic context-processing P3 (Polich, 2007), which can have onset latency of ~400 ms (Verleger et al., 2005). The second is a 900–1150 ms window illustrated in Fig. 1B which overlaps with late positivity effects observed in sentence and discourse processing and visual narrative processing e.g. (Brouwer et al., 2012; Cohn & Foulsham, 2020; Cohn et al., 2014; Kim & Osterhout, 2005; Kuperberg, 2007). Finally, a late window at 1750–1950 ms corresponds to a previously little explored period where we observe task related effects. Again, for each of these time windows, ERPs were analyzed in a 3-way ANOVA with factors Coherence (Coherent, Incoherent), Modality (Sentence, Image), and Topography (9 regions as identified in Fig. 2).

3. Results

In Fig. 2A we see that, as expected, there are no effects of narrative coherence for stimulus 1, and these appear only in response to the second stimulus in a representative frontal central electrode FC1 as illustrated in Fig. 2B. Thus, in Fig. 2A we see temporally aligned stimulus-related responses for sentences and images. In Fig. 2B we can observe a task-related response in the positivity at 400 ms for sentences and images. We note that this alignment for sentence and image conditions indicates that the holistic stimulus presentation results in comparable temporal profiles, and that any temporal variability introduced by our method does not compromise the ERP response.

3.1. Early positivity effects

Fig. 3 displays the task-related effects in the 400–500 ms period. In the ERP plot in 3A we observe a clear separation between incoherent and coherent stimuli for images and sentences. In 3B we see that this separation is present across all topographic sites, with an increase in the positivity, particularly for responses to images in the posterior sites. 3C illustrates scalp topography for the coherence effect for images and sentences, displaying a central posterior topography.

The observation of increased amplitude in the incoherent condition is confirmed by the ANOVA main effect for Coherence ($F(1,220) = 109$, $p < 0.001$). The lack of effects for Modality ($F(1,220) = 1.46$, $p = 0.228$), and for Coherence × Modality interaction ($F(1,220) = 3.31$, $p = 0.07$) confirm that the Coherence effect is not different for sentences and images. The effects for the three posterior sites is confirmed by the main effect for Topography ($F(8,220) = 53.2$, $p < 0.001$). The Modality × Topography interaction ($F(8,220) = 27.3$, $p < 0.001$) corresponds to the observation that this effect is more pronounced for Images than Sentences. The three-way interaction ($F(8,200) = 0.314$), $p > 0.1$) was not significant.

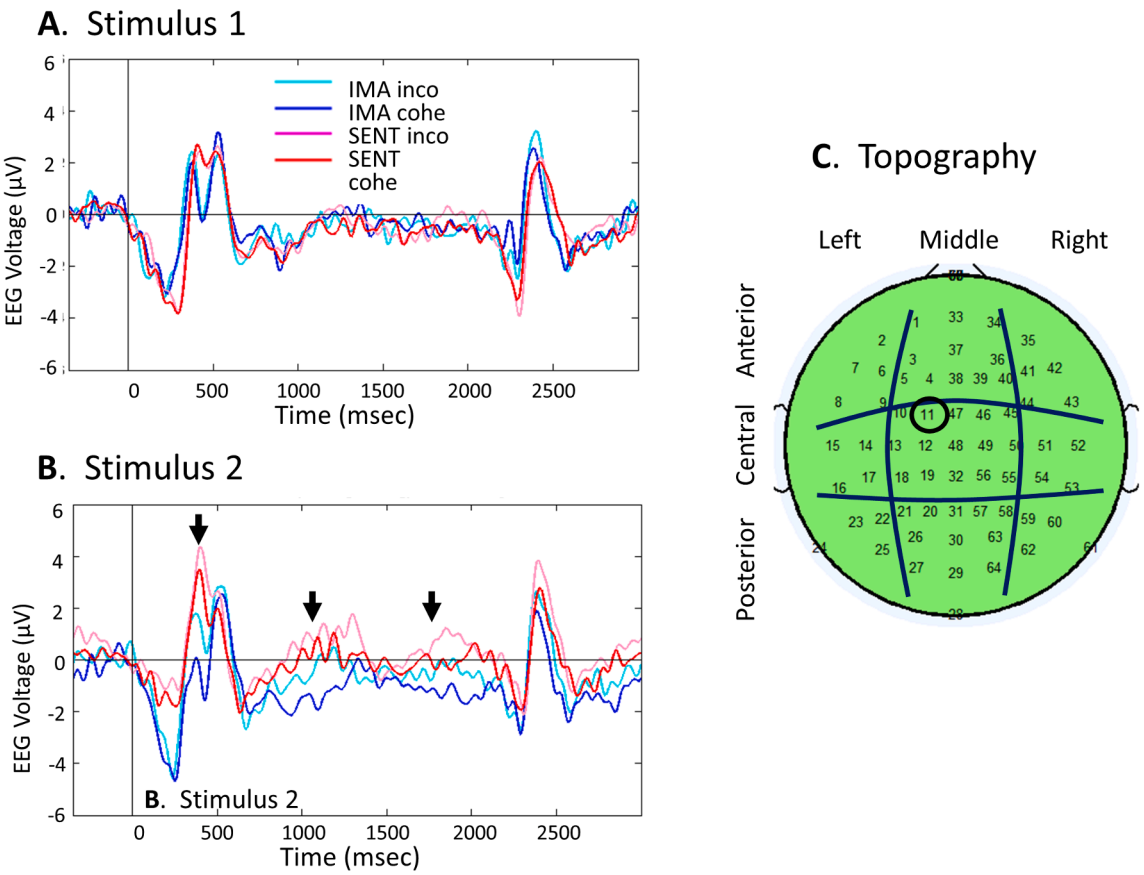

Fig. 2. ERP timing and topography. A-B Comparison of ERP responses to first and second stimuli. The large deflections seen at the start and end of the analysis epoch reflect visual onset and offset responses, respectively. A. Stimulus1 – small effects of modality are seen, but no effects of sequential coherence are visible (as expected). B. Stimulus 2 - Sequential coherence effects become visible. C. Spatial topography of electrodes, and division into nine topographic areas in a 3x3 grid. Example electrode displayed in A and B indicated with black circle on topographic plot in C.

3.2. *Later positivity effects*

Fig. 4 displays the task-related effects in the 900–1150 and 1750–1950 ms time frames. In the ERP plot in 4A we observe a clear separation between incoherent and coherent stimuli for images and sentences in these late time windows. Panel 4B illustrates this separation directly incoherent-coherent signal. In 4C and D we provide a view of the coherence effect at the distinct topographic locations. We see that this separation is present across multiple topographic sites, this time with a decrease for images in the posterior sites. Fig. 5 displays the topographic scalp distribution across the two late time windows and two different presentation modes. There we see the frontal distribution in all cases, and an extension to posterior sites for sentences in the 1750–1950 ms time frame.

The observation in the 900–1150 ms window of increased positive amplitude in the incoherent condition is confirmed by the ANOVA main effect for Coherence ($F(1,220) = 236$, $p < 0.001$). The observation of increased positivity for Sentences vs Images is confirmed by the main effect for Modality ($F(1,220) = 2332$, $p < 0.001$), The lack of Coherence × Modality interaction ($F(1,220) = 2.38$, $p = 0.124$) confirms that the Coherence effect is not different for sentences and images. The observation in Fig. 5 of a frontal topography is confirmed by the main effect for Topography ($F(8,220) = 69.9$, $p < 0.001$). The Modality × Topography interaction ($F(8,220) = 27.3$, $p < 0.001$) corresponds to the observation that responses diminish from anterior to posterior sites for Image but not Sentence (Figs. 4C and 5). The three-way interaction ($F(8,220) = 2.44$, $p < 0.05$) is marginally significant. Post-hoc (Scheffe) tests revealed a significant effect ($p < 0.05$) for coherence for Sentence and Images for the frontal zone, and central middle zone only for images, and no effects for the posterior middle zone.

Similarly, the observation in the 1750–1950 ms window of increased positive amplitude in the incoherent condition is confirmed by the ANOVA main effect for Coherence ($F(1,220) = 482$, $p < 0.001$). The observation of increased positivity for Sentences vs Images is confirmed by the main effect for Modality ($F(1,220) = 2202$, $p < 0.001$), The Coherence × Modality interaction ($F(1,220) = 47.3$, $p < 0.001$) corresponds to the observation in Fig. 4D of a more pronounced coherence effect for sentence vs. image. Post-hoc Scheffe tests confirm the significant effect of coherence for sentences ($p < 0.001$), and images ($p < 0.001$) The observation in Fig. 5 of a frontal topography is confirmed by the main effect for Topography ($F(8,220) = 67.5$, $p < 0.001$). The Modality × Topography interaction ($F(8,220) = 38.8$, $p < 0.001$) corresponds to the observation that responses diminish from anterior to posterior sites for Image but not Sentence (Fig. 4D and 5). The three-way interaction ($F(8,220) = 4.57$), $p < 0.001$) is significant. Post-hoc (Scheffe) tests revealed a significant effect ($p < 0.05$) for coherence for Sentence and Images for the frontal and central middle zones, and only for sentences in the posterior middle, as seen in Fig. 5D.

4. Discussion

In the current research we analyzed ERPs generated while subjects were exposed to complex human-activity-related stimuli presented in two-event narrative sequences. The sequences were presented in two distinct perceptual formats or dimensions: sentences and images. Our objective was to test the hypothesis that common neural processes are involved in making sense of these narrative sequences in the sentence and image modalities. We would consider the hypothesis validated if

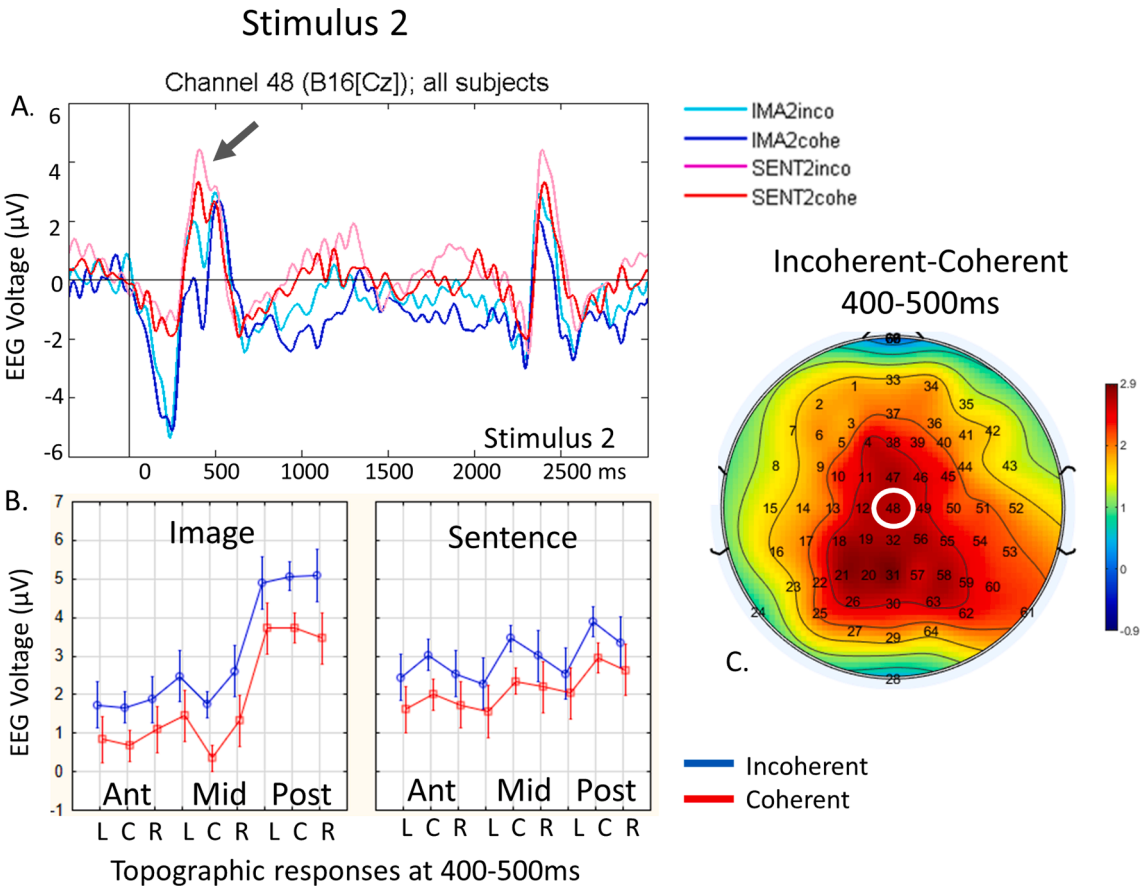

Fig. 3. Early Positivity 400–500 ms effect. A. ERP. This positivity makes the Incoherent vs. Coherent distinction, with an increased positivity for the sequentially incoherent stimuli. Illustrated with Electrode Cz (48). B. Topographic effect of positivity in 400–500 ms window. L – Left, C – Central, R – Right. Ant – Anterior, Mid – Middle, Post – Posterior. C. Topographic map of Incoherent-Coherent in 400–500 ms window. Electrode 48 in A indicated by white circle.

there is a principal common ERP response, without excluding the possibility of additional modality specific responses. Here we thus examine the neural dynamics associated with engaging the semantic system as required for understanding two-element narratives made up of images or sentences. We orient the discussion around the principal results, particularly the semantic processing associated with the late EEG responses, the corresponding behavior, and steps towards conceptual and neural implementations of explanatory models.

4.1. Early semantic effects

Already by 400 ms for both sentence and image modalities we observe common processing revealed as the sequential coherence effect. Such early responses have been observed in language processing for syntactic class expectation violations (Neville et al., 1991), and in image processing in response to comic strip panels in which the motion lines were in conflict with the depicted motion, or when they were absent (Cohn & Maher, 2015). This early positivity observed in our study and by Cohn & Maher has a central – posterior topography, similar to that of the P300b in response to unexpected stimuli (Polich, 2007). This suggests that this early positivity is a response to a violation of an expectation with respect to the narrative structure linking the two successive stimuli.

4.2. Late common semantic processing

We observed late frontal-centrally distributed positive ERP responses during the processing of sentences and images in the sequentially incoherent condition. The late positivity that we observe in the 900–1150 ms timeframe is similar to those that have been observed when sentences and images do not fit with the previously established context (Bornkessel-Schlesewsky & Schlesewsky, 2008; Brouwer et al., 2012; Burkhardt, 2007; Kuperberg, 2007; Xiang & Kuperberg, 2015), and in visual processing of comic strip sequences (Cohn & Foulsham, 2020).

While frontal positivities were initially produced in response to syntactic anomalies, they have since been identified in a number of situations that do not involve syntactic recovery. For example Kaan and Swaab (2003) observed a late (500–900 ms) frontal positivity associated with ambiguity resolution and/or increases in discourse complexity. Kuperberg (2007) reviews studies where late positivities are evoked, with and without syntactical anomalous sentences. Of particular interest, in the sentence "Every morning at breakfast the eggs would eat …", the response to eggs (which is semantically anomalous as inanimate eggs cannot eat) was a robust P600, in the absence of an N400 (Kuperberg et al., 2003). Kim and Osterhout (2005) likewise observed a very late positivity that extended beyond 900 ms during the reading of sentences with agency incoherence on the verb. Kuperberg suggests that some degree of semantic association between a verb and its arguments may trigger a P600. Similarly, Xiang and Kuperberg (2015) have argued that a late posterior positivity component is triggered when a near certain prediction is followed by an input that requires a switch to a new generative model representing relationships between events. Paczynski and Kuperberg (2012) advocate the P600 as reflecting a conflict between semantic memory-based predictions, and the detection of propositional incoherence. Brouwer et al. (2012) consider that late positivities are invoked by semantic integration processes. Late frontal positivities have been associated with cross-modal processing of irony

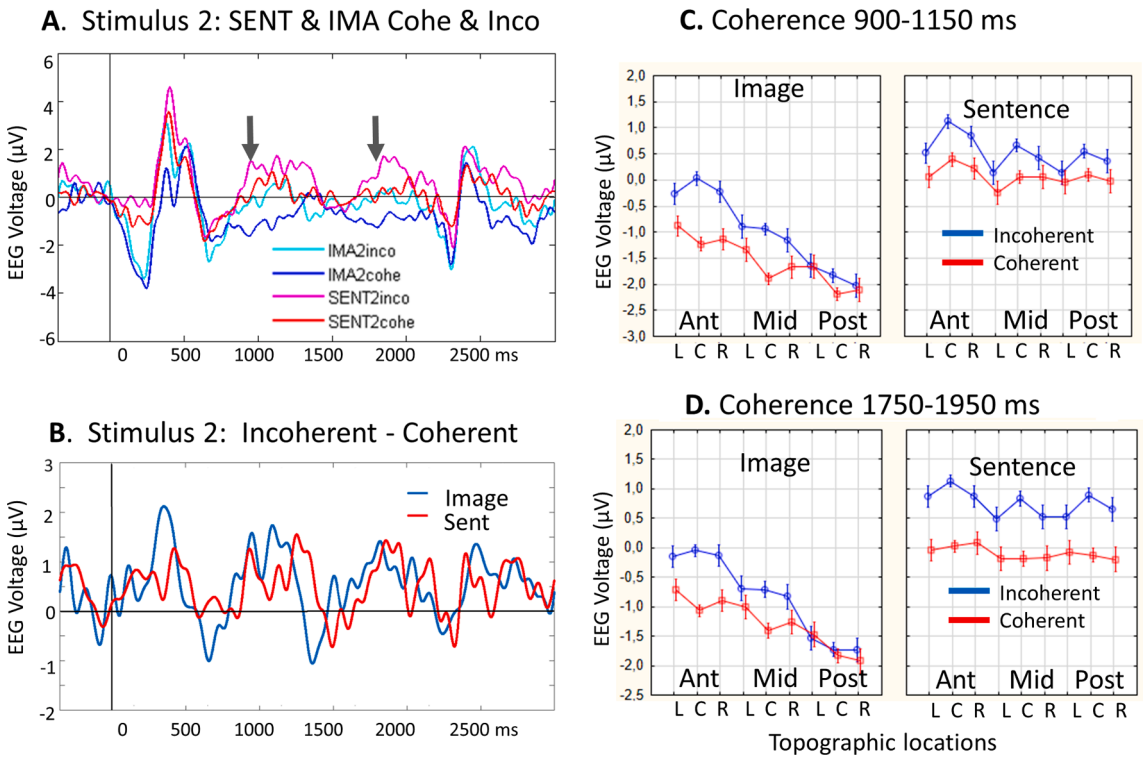

Fig. 4. Late positivities (900–1150, and 1750–1950 ms). A. ERP displays sensitivity to the Incoherent vs. Coherent distinction for sentences and images. B. Scalp map for sentences and images Incoherent-Coherent contrast for 900–1150, and 1750–1950 ms periods. Topographic responses at 900–1150 ms in C, and at 1750–1950 ms in D.

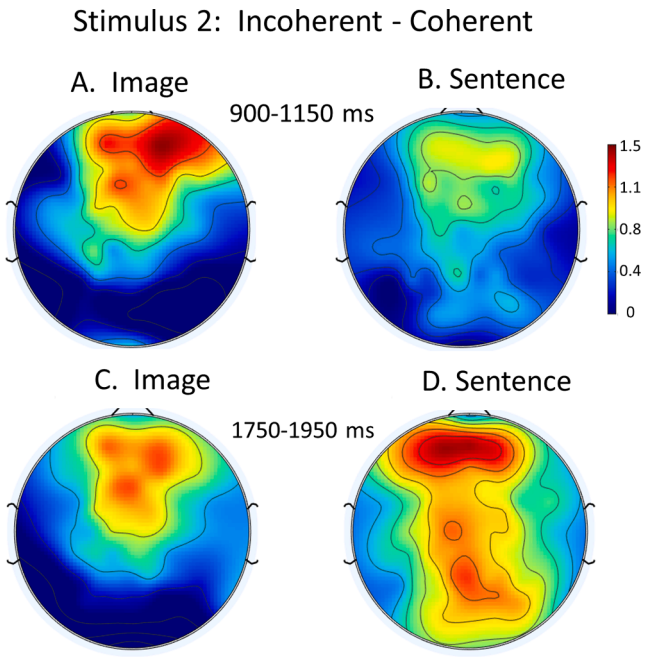

Fig. 5. Scalp distribution across 900–1150 ms and 1750–1950 ms time windows, and Image and Sentence presentation modes. In 900–1150 ms window, Image (A) and Sentence (B) have similar frontal distribution. In 1750–1950 ms time window, Image (C) retains the frontal distribution and Sentence (D) has a distribution than extends from frontal to posterior.

(Weissman & Tanner, 2018), and more generally with possible but unpredictable words (DeLong et al., 2014). Cohn and Foulsham (2020) observed a similar late positivity to comic strip panels that require processing for integration into the sequential context. Our sequentially incoherent condition would thus invoke such integration processes producing late central-frontal positivities, similar to those observed in discourse processing requiring inferencing (Burkhardt, 2007).

Given these observations, our results can be situated in the theoretical framework of language comprehension developed by Kuperberg and colleagues (Brothers, Wlotko, Warnke, & Kuperberg, 2020; Kuperberg, Brothers, & Wlotko, 2020). In their framework, during comprehension, the brain builds up a predictive situation model. When the situation model is sufficiently rich, the predictive model will respond in distinct manners to follow-ups that are either unexpected or anomalous. Unexpected follow-ups trigger a late frontal positivity associated with an updating of the situation model. Conflicting follow-ups trigger a late posterior positivity associated with a failure to update the situation model and possible second pass (Kuperberg et al., 2020). In the current experiment, we present a first stimulus that establishes a situation model, and a second stimulus that is either coherent or incoherent with respect to the first. The incoherent stimuli depict the same people in the first stimulus, but in a different situation. This implies an update of the situation model, but not an anomaly that would require a second pass.

Accordingly, in the results of our experiment, we thus observed late frontal positivities for the sentence and image stimuli. This allows the validation of the hypothesis that common neurophysiological processes are at work in the comprehension of sentence and visual narrative. Of course, while there is evidence for some common processing, we do not intend to argue that the processing for sentence and visual narrative is identical. Interestingly, only for the sentences, the positivity was also more widespread covering both frontal and posterior sites. Such a combined frontal and posterior positivity for sentences may reflect a form of compromise where the second stimulus can be integrated but in parallel triggers a more extensive revision of the situation model. This suggests that the image stimuli may create a more elaborate situation model than those created by sentence stimuli. Indeed, it is likely that our first sentences establish a situation model that is a compromise between

the high and low constraint contexts evoked in initial discourses in Kuperberg et al. (2020), thus producing the dual frontal-posterior positivity response.

4.3. Behavioral correlates of late semantic processing

This leads to the question of what it is that subjects are actually doing in our task. The constructivist theory of comprehension would hold that without any intention or will to do so, subjects naturally engage with the stimuli and attempt to make sense of them, to connect them (Graesser et al., 1994). In separate studies we have begun to behaviorally investigate this integrative process in the domain of sentence processing (Madden-Lombardi et al., 2015). There, we exposed subjects to successive sentences in pairs that were either sequentially coherent or not. Subjects found the second sentence in the sequentially incoherent pairs less easy to mentally imagine than those in the sequentially coherent condition, and they also required additional time to make the judgement (Madden-Lombardi et al., 2015). Good imagers responded at around 800 ms to coherent sentences and significantly slower at around 1200 ms for incoherent sentences. This suggests that in the sequentially coherent pairs, the representation of the second sentence is already (partially) included in the representation evoked by the first sentences, whereas in the sequentially incoherent pairs, additional processing is required to align the second in the context of the first. This additional processing may be reflected in the late positivities we observe in the present study, and may require a more extensive revision for incoherent sentences vs images. This processing may be part of attempting to make sense in the context of narrative integration (Graesser et al., 1994). According to the event indexing model of Zwaan et al. (1995) and Zwaan (1999) coherence can be measured along five dimensions: time, space, causation, motivation and protagonist. In the sequentially incoherent conditions, while relations in the dimensions of time and space may be broken, the protagonist relation remains, so that the sequentially incoherent conditions can be considered as a shift in time or storyline rather than a completely incoherent or unrelated event (Madden-Lombardi et al., 2015).

4.4. Underlying neurophysiology

The current results contribute to the ongoing effort to characterize the functional neurophysiology of comprehension. In our previous fMRI study of the comprehension of sentence and images, we identified a common semantic network made up of fronto-temporo-parietal network that included the middle and inferior frontal gyri, the parahippocampal-retrosplenial complex, the anterior and middle temporal gyri, the inferior parietal lobe in particular the temporo-parietal cortex (Jouen et al., 2015). Analysis of the functional and structural connectivity of this network revealed white matter fibers linking ventral neural structures including the parietal and temporal cortices through inferior and middle longitudinal fasciculi, the temporal and parahippocampal gyrus through the cingulate bundle, and the temporal and prefrontal structures through the uncinate fasciculus (Jouen et al., 2018). Together these results suggest that during analysis of the event evoked by our pictures or sentences (persons engaged in everyday actions), the participant will recognize and interpret the situation through several cognitive operations: self-projection in the scene, recall from episodic memory and application of Theory of Mind (Jouen et al., 2015, 2018).

Further advancing this research, Hu et al. (2019) performed a cross-linguistic fMRI study using coherent-incoherent sentence and image stimuli based on those in the current study. They identified brain regions that successfully discriminated between coherent and incoherent stimuli across languages and across modalities. These regions include left inferior parietal gyrus (IPG), which extends from the supramarginal gyrus (SMG; BA 40) to the AG (BA 39/7), and the left precental gyrus extending to the pars opercularis and pars triangularis of Broca's area (BA 44/45. Interestingly, these regions which play a role in the default mode network (Raichle, 2015) coincide with those which we have suggested play a role in comprehension by allowing the individual to situate themselves in an embodied representation of the narrated events.

In a study that examined relations between fMRI and ERP responses (Herzmann et al., 2012) observed a late (600–1000) frontal positivity for previously remembered vs. new pictures. This was correlated with activation in the left prefrontal gyrus, and the parietal cortex angular gyrus and supramarginal gyrus. This is consistent with our ERP results, and previous fMRI results, and provides an explanatory link between the fronto-temporo-parietal activation and the late positivities we observe in the current study.

5. Limitations and future work

Cross-modal comparisons of language and images provides new insights into the neural basis of meaning representations, but such comparisons are not without limitations. One limitation is related to the time course of modality specific processing. While we observe temporal correspondence of ERPs for sentence and image processing, it seems clear that there are modality specific processes at work, particularly in the earliest processing of the stimuli. This can be seen in Fig. 3, where in the 400–500 ms window the ERP in response to images precedes that for images. Future research can examine in more detail the time course of the modality specific processing and the convergence of these modality specific processing pathways toward the common processing.

In order to directly compare responses to images and sentences, we chose to present the sentences holistically. This can be contrasted with protocols that use serial word presentations. Ganis et al. (1996) examine the processing of words vs. simple line drawings of objects at the end of a sentence of serially presented words. These sentence final items were semantically congruous or incongruous. Incongruous stimuli produced similar N400-like effects. Since the objective in Ganis et al. (1996) was to compare single words to single object images, word by word presentation (particularly for the last word) was required. In our study, we were interested in brain activity induced by complex events described by complex images (i.e. images of people in the world, performing actions) and multiword sentences. Thus while Ganis et al. (1996) is pertinent to our study, because the objectives are different, we are constrained not to employ a serial word presentation of sentences. One of the reasons would be that it would not allow the direct comparison between sentences and image. Importantly, we recall that in the crucial second sentence, the verb that allows detection of the narrative coherence immediately follows a pronoun, and thus is immediately accessed. This is confirmed in the ERP responses, where we observe that the task related effects for sentences and images are temporally aligned, indicating that there is no or minimal impact of the holistic sentence presentation.

Experimental comparison of visual and language narrative has taken a significant step in research that examines brain activity when subject watch movies vs listen to narration of the same stories (Baldassano et al., 2017). Using fMRI this research demonstrated that multiple brain regions associated with high level semantic processing including angular gyrus, temporoparietal junction, posterior medial cortex and inferior frontal cortex displayed similar event segmentation properties for movie and narrative presentations of the same story. These areas overlap with the default mode network, and correspond well to the areas we identified in the semantic network for sentences and images (Jouen et al., 2015, 2018). Future research should use higher temporal resolution of EEG or MEG to further address common processing of meaning in this extended network.

Finally, we can ask what neurocomputational processes can underlie this common activation revealed in the late positivity. According to the retrieval-integration model of Brouwer et al. (2017), semantic retrieval load is indexed by the N400, and subsequent integration into the unfolding utterance is indexed by the P600. We could consider the extension of this model to include integration of multiple utterances into

a situation model. Such a model would be consistent with our current observations where the incoherent stimuli are not difficult to process in terms of semantic retrieval, but rather require a change of time and space in the situation model. Likewise, such a model should take into account the integrative processing of new inputs with an existing situation model as characterized in the generative model of (Brothers, Wlotko, Warnke, & Kuperberg, 2020; Kuperberg, Brothers, & Wlotko, 2020). We have taken initial steps in modeling this integrative processing (Uchida, Lair, Ishiguro, & Dominey, 2021).

6. Conclusion

The groundbreaking work of Vandenberghe et al. (1996) provided evidence for a common semantic system for words and pictures. This was further confirmed at the semantic category level by cross-modality decoding of semantic categories (Shinkareva et al., 2011). We (Jouen et al., 2015) extended this line of research, revealing a broadly extended semantic system common to the representation of the semantics of human event evoked by sentences and complex images, with extended anatomical and functional connectivity (Jouen et al., 2018). The current research reveals evidence for common neural dynamics for semantic processing of short narrative sequences made up of sentences or images. We observe common ERP patterns of brain activity when subjects encounter sequentially incoherent stimuli in the image or sentence modalities, in the form of a late frontal positivity. This is coherent with related research examining sentence and image processing separately. Importantly, in the stimuli that we use to asses this sequential narrative processing, there is nothing inherently wrong with the second stimulus in these sequentially incoherent pairs: it is only their relation with the first stimulus that is manipulated in order to produce a sequential incoherence. We interpret these results as a form of narrative integration, where the subject must make sense of an unexpected but possible follow-up (Burkhardt, 2007). This corresponds to the proposal of Brouwer et al. (2012, 2017) whereby this late central-frontal positivity reflects additional processing required to arrive at a coherent representation of the intended meaning. More specifically, the frontal positivity corresponds to an integrative process where the current situation model is updated to accommodate the incoming information (Brothers, Wlotko, Warnke, & Kuperberg, 2020; Kuperberg, Brothers, & Wlotko, 2020).

CRediT authorship contribution statement

Anne-Lise Jouen: Conceptualization, Formal analysis, Writing - original draft. **Nicolas Cazin:** Formal analysis, Methodology. **Sullivan Hidot:** Formal analysis, Methodology, Software, Visualization. **Carol Madden-Lombardi:** Conceptualization, Methodology, Writing - review & editing. **Jocelyne Ventre-Dominey:** Conceptualization, Supervision, Writing - original draft, Writing - review & editing. **Peter Ford Dominey:** Conceptualization, Formal analysis, Funding acquisition, Writing - original draft, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was supported by funding from the French ANR Comprendre, from the European Community through grants FP7-ICT 231267 Project Organic, FP7-ICT-270490 Project EFAA, and FP7-ICT-612139 Project WYSIWYD, and from the Région Bourgogne-Franche-Comté Projet ANER RobotSelf. An earlier version of this manuscript has been released as a pre-print at www.biorxiv.org, (Jouen et al., 2019).

References

Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., & Norman, K. A. (2017). Discovering event structure in continuous narrative perception and memory. *Neuron, 95*(3), 709-721.e705.

Biederman, I. (1981). On the semantics of a glance at a scene. In M. Kubovy, & J. Pomerantz (Eds.), *Perceptual Organization* (pp. 213–252). Hillsdale, NJ: Erlbaum.

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex, 19*(12), 2767–2796.

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2008). An alternative perspective on "semantic P600" effects in language comprehension. *Brain Research Reviews, 59*(1), 55–73.

Brothers, T., Wlotko, E. W., Warnke, L., & Kuperberg, G. R. (2020). Going the extra mile: Effects of discourse context on two late positivities during language comprehension. *Neurobiology of Language, 1*(1), 135–160.

Brouwer, H., Crocker, M. W., Venhuizen, N. J., & Hoeks, J. C. (2017). A neurocomputational model of the N400 and the P600 in language processing. *Cognitive Science, 41*, 1318–1352.

Brouwer, H., Fitz, H., & Hoeks, J. (2012). Getting real about semantic illusions: Rethinking the functional role of the P600 in language comprehension. *Brain Research, 1446*, 127–143.

Bruner, J. S. (1990). *Acts of meaning*. Harvard University Press.

Burkhardt, P. (2007). The P600 reflects cost of new information in discourse memory. *NeuroReport, 18*(17), 1851–1854.

Chen, J., Leong, Y. C., Honey, C. J., Yong, C. H., Norman, K. A., & Hasson, U. (2017). Shared memories reveal shared structure in neural activity across individuals. *Nature neuroscience, 20*(1), 115–125.

Cohn, N., & Foulsham, T. (2020). Zooming in on the cognitive neuroscience of visual narrative. *Brain and Cognition, 146*, Article 105634.

Cohn, N., Jackendoff, R., Holcomb, P. J., & Kuperberg, G. R. (2014). The grammar of visual narrative: Neural evidence for constituent structure in sequential image comprehension. *Neuropsychologia, 64*, 63–70.

Cohn, N., & Maher, S. (2015). The notion of the motion: The neurocognition of motion lines in visual narratives. *Brain Research, 1601*, 73–84.

DeLong, K. A., Quante, L., & Kutas, M. (2014). Predictability, plausibility, and two late ERP positivities during written sentence comprehension. *Neuropsychologia, 61*, 150–162.

Friedman, A. (1979). Framing pictures: The role of knowledge in automatized encoding and memory for gist. *Journal of experimental psychology: General, 108*(3), 316.

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*(2), 89–106.

Gernsbacher, M. A., & Faust, M. E. (1991). The mechanism of suppression: A component of general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(2), 245.

Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science, 17*(3), 311–347.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101*(3), 371.

Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes, 8*(4), 439–483.

Herzmann, G., Jin, M., Cordes, D., & Curran, T. (2012). A within-subject ERP and fMRI investigation of orientation-specific recognition memory for pictures. *Cognitive Neuroscience, 3*(3–4), 174–192.

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*(1), 59–73.

Hu, Z., Yang, H., Yang, Y., Nishida, S., Madden-Lombardi, C., Ventre-Dominey, J., Dominey, F., & Ogawa, K. (2019). Common neural system for sentence and picture comprehension across languages: A Chinese-Japanese bilingual study. *Frontiers in human neuroscience, 13*, 380.

Jouen, A.-L., Cazin, N., Hidot, S., Madden-Lombardi, C., Ventre-Dominey, J., & Dominey, P. F. (2019). Beyond the word and image: III. Neurodynamic properties of the semantic network, 767384 *BioRxiv*.

Jouen, A., Ellmore, T., Madden-Lombardi, C., Pallier, C., Dominey, P., & Ventre-Dominey, J. (2018). Beyond the word and image: II-Structural and functional connectivity of a common semantic system. *NeuroImage, 166*, 185–197.

Jouen, A., Ellmore, T., Madden, C., Pallier, C., Dominey, P., & Ventre-Dominey, J. (2015). Beyond the word and image: Characteristics of a common meaning system for language and vision revealed by functional and structural imaging. *NeuroImage, 106*, 72–85.

Kaan, E., & Swaab, T. Y. (2003). Repair, revision, and complexity in syntactic analysis: An electrophysiological differentiation. *Journal of Cognitive Neuroscience, 15*(1), 98–110.

Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language, 52*(2), 205–225.

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research, 1146*, 23–49.

Kuperberg, G. R., Brothers, T., & Wlotko, E. W. (2020). A tale of two positivities and the N400: Distinct neural signatures are evoked by confirmed and violated predictions at different levels of representation. *Journal of Cognitive Neuroscience, 32*(1), 12–35.

Kuperberg, G. R., Sitnikova, T., Caplan, D., & Holcomb, P. J. (2003). Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research, 17*(1), 117–129.
Kutas, M., & Federmeier, K. D. (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. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*(4427), 203–205.
Lallee, S., & Dominey, P. F. (2013). Multi-modal convergence maps: From body schema and self-representation to mental imagery. *Adaptive Behavior, 21*(4), 274–285.
Leckey, M., & Federmeier, K. D. (2020). The P3b and P600 (s): Positive contributions to language comprehension. *Psychophysiology, 57*(7), Article e13351.
Lio, G., & Boulinguez, P. (2013). Greater robustness of second order statistics than higher order statistics algorithms to distortions of the mixing matrix in blind source separation of human EEG: Implications for single-subject and group analyses. *NeuroImage, 67*, 137–152.
Madden-Lombardi, C., Jouen, A.-L., Dominey, P. F., & Ventre-Dominey, J. (2015). Sequential coherence in sentence pairs enhances imagery during comprehension: An individual differences study. *PLoS ONE, 10*(9), Article e0138269.
Mandler, J. M., & Johnson, N. S. (1976). Some of the thousand words a picture is worth. *Journal of Experimental Psychology: Human Learning and Memory, 2*(5), 529.
Meyer, K., & Damasio, A. (2009). Convergence and divergence in a neural architecture for recognition and memory. *Trends in Neurosciences, 32*(7), 376–382.
Neville, H., Nicol, J. L., Barss, A., Forster, K. I., & Garrett, M. F. (1991). Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of cognitive Neuroscience, 3*(2), 151–165.
New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du français contemporain sur internet: LEXIQUE™//A lexical database for contemporary french: LEXIQUE™. *L'année psychologique, 101*(3), 447–462.
Nolan, H., Whelan, R., & Reilly, R. B. (2010). FASTER: Fully automated statistical thresholding for EEG artifact rejection. *Journal of Neuroscience Methods, 192*(1), 152–162.
Osterhout, L., & Hagoort, P. (1999). A superficial resemblance does not necessarily mean you are part of the family: Counterarguments to Coulson, King and Kutas (1998) in the P600/SPS-P300 debate. *Language and Cognitive Processes, 14*(1), 1–14.
Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language, 31*(6), 785–806.
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*(4), 426–448.
Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology, 118*(10), 2128–2148.
Raichle, M. E. (2015). The brain's default mode network. *Annual Review of Neuroscience, 38*, 433–447.
Shinkareva, S. V., Malave, V. L., Mason, R. A., Mitchell, T. M., & Just, M. A. (2011). Commonality of neural representations of words and pictures. *Neuroimage, 54*(3), 2418–2425.
Sitnikova, T., Holcomb, P. J., Kiyonaga, K. A., & Kuperberg, G. R. (2008). Two neurocognitive mechanisms of semantic integration during the comprehension of visual real-world events. *Journal of cognitive neuroscience, 20*(11), 2037–2057.
Uchida, T., Lair, N., Ishiguro, H., & Dominey, P. F. (2021). A model of online temporal-spatial integration for immediacy and overrule in discourse comprehension. *Neurobiology of Language, 2*(1), 83–105.
Vandenberghe, R., Price, C., Wise, R., Josephs, O., & Frackowiak, R. S. J. (1996). Functional anatomy of a common semantic system for words and pictures. *Nature, 383*(6597), 254–256.
Verleger, R., Jaśkowski, P., & Wascher, E. (2005). Evidence for an integrative role of P3b in linking reaction to perception. *Journal of Psychophysiology, 19*(3), 165–181.
Weissman, B., & Tanner, D. (2018). A strong wink between verbal and emoji-based irony: How the brain processes ironic emojis during language comprehension. *PLoS ONE, 13*(8), Article e0201727.
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*(7), 1235–1249.
Xiang, M., & Kuperberg, G. (2015). Reversing expectations during discourse comprehension. *Language, Cognition and Neuroscience, 30*(6), 648–672.
Zadbood, A., Chen, J., Leong, Y. C., Norman, K. A., & Hasson, U. (2017). How we transmit memories to other brains: Constructing shared neural representations via communication. *Cerebral Cortex, 27*(10), 4988–5000.
Zwaan, R. A. (1999). Five dimensions of narrative comprehension: The event-indexing model. *Narrative comprehension, causality, and coherence: Essays in honor of Tom Trabasso*, 93–110.
Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science*, 292–297.