WILEY

Gesture's body orientation modulates the N400 for visual sentences primed by gestures

Yifei He[1] | Svenja Luell[2] | R. Muralikrishnan[3] | Benjamin Straube[1] | Arne Nagels[2]

[1]Department of Psychiatry and Psychotherapy, Philipps-University Marburg, Marburg, Germany

[2]Department of General Linguistics, Johannes-Gutenberg University Mainz, Mainz, Germany

[3]Department of Neuroscience, Max Planck Institute for Empirical Aesthetics, Frankfurt, Germany

Correspondence
Yifei He, Department of Psychiatry and Psychotherapy, Philipps University Marburg, Rudolf-Bultmann-Str. 8, 35039 Marburg, Germany.
Email: yifei.he@staff.uni-marburg.de

Funding information
Deutsche Forschungsgemeinschaft, Grant/Award Numbers: HE8029/2-1, STR1146/11-2, STR1146/15-1; Von-Behring-Röntgen-Stiftung, Grant/Award Number: 64-0001

Abstract

Body orientation of gesture entails social-communicative intention, and may thus influence how gestures are perceived and comprehended together with auditory speech during face-to-face communication. To date, despite the emergence of neuroscientific literature on the role of body orientation on hand action perception, limited studies have directly investigated the role of body orientation in the interaction between gesture and language. To address this research question, we carried out an electroencephalography (EEG) experiment presenting to participants (n = 21) videos of frontal and lateral communicative hand gestures of 5 s (e.g., raising a hand), followed by visually presented sentences that are either congruent or incongruent with the gesture (e.g., "the mountain is *high/low...*"). Participants underwent a semantic probe task, judging whether a target word is related or unrelated to the gesture-sentence event. EEG results suggest that, during the perception phase of handgestures, while both frontal and lateral gestures elicited a power decrease in both the alpha (8–12 Hz) and the beta (16–24 Hz) bands, lateral versus frontal gestures elicited reduced power decrease in the beta band, source-located to the medial prefrontal cortex. For sentence comprehension, at the critical word whose meaning is congruent/incongruent with the gesture prime, frontal gestures elicited an N400 effect for gesture-sentence incongruency. More importantly, this incongruency effect was significantly reduced for lateral gestures. These findings suggest that body orientation plays an important role in gesture perception, and that its inferred social-communicative intention may influence gesture-language interaction at semantic level.

KEYWORDS
beta oscillations, body orientation, gesture, N400, semantics, social perception

1 | INTRODUCTION

In every day face-to-face communication, individuals deploy multiple ways to express their communicative intentions. These encompass linguistic expressions in the form of verbal speech, as well as a wide range of non-linguistic cues, for example, gesture, facial expression, eye-gaze, or even body orientation and postures. For example, if an individual wants to keep someone else away, except for verbally expressing *go away*, s/he may additionally produce an accompanying gesture to emphasize this communicative intention. This intention may be even enhanced by a negative facial expression, together with a change of body orientation by turning away from the addressee.

Notably, despite the wealth and heterogeneity of these communicative cues, the human brain seems to be remarkably good not only in perceiving, but also in an immediate and efficient integration of all of them. Therefore, an interesting question arises concerning the brain mechanism underlying this integrative process.

Of all these communicative cues, gesture and language are among the most intensively investigated with neuroscientific methods. Literature suggests that human brain is able to effortlessly integrate semantic information from these two input channels. With functional magnetic resonance imaging (fMRI), it is suggested that two regions in the left hemisphere, the left posterior superior temporal sulcus and the left inferior frontal gyrus are crucially involved in the semantic combination and semantic integration between gesture and verbal speech (Dick, Mok, Beharelle, Goldin-Meadow, & Small, 2012; Green et al., 2009; He et al., 2015, He, Steines, Sommer, et al., 2018; Holle, Obleser, Rueschemeyer, & Gunter, 2010; Kircher et al., 2009; Straube, Green, Bromberger, & Kircher, 2011; Willems, Özyürek, & Hagoort, 2009). With electroencephalography (EEG) methods such as event-related potentials (ERPs), it has been suggested that the human brain is able to rapidly integrate semantic representations from gesture and language, as reflected by the N400 component. The N400 has been consistently observed for a variety of experimental manipulations, for example, semantic mismatch or semantic disambiguity (Holle & Gunter, 2007; Özyürek, Willems, Kita, & Hagoort, 2007), both for language stimuli that were presented either in auditory or visual form (Fabbri-Destro et al., 2015; Özyürek et al., 2007), and for gesture and speech stimuli that are presented simultaneously or consecutively, across adults and children (Fabbri-Destro et al., 2015; Habets, Kita, Shao, Özyurek, & Hagoort, 2011; Kelly, Kravitz, & Hopkins, 2004; Sekine et al., 2020; Wu & Coulson, 2005). Of note, despite consistent reports (see Özyürek, 2014 for review), it remains an open question whether these N400 effects reflect the cost of semantically integrating gesture and language, or the differential level of semantic prediction from gesture to language and vice versa, or a combination of both processes. In fact, in language processing, although the N400 has been classically considered as a semantic integration index (Kutas & Hillyard, 1984), recent functional accounts of the component suggest either the prediction alternative (Bornkessel-Schlesewsky & Schlesewsky, 2019; Kuperberg, Brothers, & Wlotko, 2020; Kutas & Federmeier, 2011; Lau, Phillips, & Poeppel, 2008), or a hybrid of both processes (Baggio & Hagoort, 2011; Nieuwland et al., 2020).

In addition, recent studies looking at the oscillatory domain of EEG also showed that both the alpha (8–13 Hz) and beta (14–26 Hz) bands may be functionally related to the integration between speech and communicative gesture (Biau, Torralba, Fuentemilla, de Diego Balaguer, & Soto-Faraco, 2015; Drijvers, Özyürek, & Jensen, 2018a, 2018b; He et al., 2015; He, Steines, Sommer, et al., 2018). Interestingly, these frequency bands were also highly relevant to the perception of action and hand gestures in particular (Avanzini et al., 2012; He, Steines, Sammer, et al., 2018; Järveläinen, Schuermann, & Hari, 2004; Quandt, Marshall, Shipley, Beilock, & Goldin-Meadow, 2012). It has been reported that gestures differing in social-communicative intention (goal-directed vs. non-goal-directed) (Hari et al., 1998; Järveläinen et al., 2004)—as well as the level of motor simulation (Quandt et al., 2012)—elicited differential level of alpha or beta band power decrease. Additionally, these effects may be even modulated by accompanying auditory speech, as in the case of co-speech gesture (He, Steines, Sammer, et al., 2018).

However, despite the emergence of neuroscientific inquiries into semantic integration between gesture and language, the role of body orientation during this process remains largely unclear. In fact, body orientation (e.g., frontal vs. lateral view) not only presents as physical difference between view-points, but may also indicate differential level of social-communicative intention. The social role of body orientation has been investigated intensively with fMRI studies, the results of which show that it differentially affects either facial emotion or communicative intention (Ciaramidaro, Becchio, Colle, Bara, & Walter, 2013; Schilbach et al., 2006) by consistently activating the medial prefrontal cortex (mPFC). Notably, the mPFC is a crucial region within the mentalizing network, which supports the perception of social and emotional features and mentalizing others' social-communicative intentions (Frith & Frith, 2006; Van Overwalle, 2009). More relevantly, the mPFC is even found to be activated while observing hand gestures that differ in the degree of communicative intentions (Trujillo, Simanova, Ozyurek, & Bekkering, 2020).

Nevertheless, body orientation may have a direct impact, physically and socially, upon (a) the perception of hand gestures and (b) the semantic processing during gesture-language interaction. However, both research questions remain elusive with current literature. Research in the domain of action observation has shown that, non-communicative hand actions (reaching, grasping, and etc.) observed from egocentric vs. allocentric view-point elicited greater sensorimotor-related alpha or beta band power decrease (Angelini et al., 2018; Drew, Quandt, & Marshall, 2015). This line of research indicates that body orientation (or view-point) has direct impact on hand action perception, which could be related to differential level of power decrease in the alpha and the beta bands. Of note, although alpha and beta oscillations as elicited during motor execution and imagery tasks are suggested to be functionally dissociable (Brinkman, Stolk, Dijkerman, de Lange, & Toni, 2014; Quandt et al., 2012; Stolk et al., 2019), that is, alpha power decrease might be related to sensory/motor perception in general, whereas the beta band is more sensitive to fine-grained motor-related features (Brinkman et al., 2014; He, Steines, Sammer, et al., 2018; Salmelin, Hämäaläinen, Kajola, & Hari, 1995). Importantly, prior studies investigated mostly non-communicative and thus semantic-free gestures, and participants from these studies usually only see a fraction of the actor/actress (typically one hand), but not the entire body. By contrast, daily communicative gestures are clearly richer than reaching or grasping hand actions in terms of semantic representations and social-communicative intentions. Therefore, it remains an open question whether the perception of communicative gestures will be affected by body orientation (or view-point) in a similar way when compared to non-communicative hand actions. These caveats have been addressed by previous studies using fMRI. For example, in Straube, Green, Jansen, Chatterjee, and Kircher (2010), a direct comparison between frontal and lateral co-speech gestures activated the mPFC and other regions in

the mentalizing network, thus indicating the potential social-cues conveyed by body orientation (see also Nagels, Kircher, Steines, & Straube, 2015; Redcay, Velnoskey, & Rowe, 2016; Saggar, Shelly, Lepage, Hoeft, & Reiss, 2014). However, for these communicative gestures, in the EEG domain, whether body orientation affects sensorimotor alpha or beta band power remains unclear.

A more intriguing question concerns the role of body orientation on the semantic interaction between gesture and language in terms of semantic prediction and integration. It is fairly established that gesture and language both share comparable neural underpinnings (Xu, Gannon, Emmorey, Smith, & Braun, 2009), and that the two communicative channels interact with each other (Holler & Levinson, 2019). Neural evidence from both M/EEG and fMRI also shows that gesture modulates language processing, most commonly in a facilitative manner (Biau & Soto-Faraco, 2013; Cuevas et al., 2019; Drijvers et al., 2018b; Zhang, Frassinelli, Tuomainen, Skipper, & Vigliocco, 2020). Therefore, given the social nature of communicative gestures, it might be hypothesized that gestures differing in social aspects, as in the case of facing versus not facing the addressee, may differentially affect the semantic processing of gesture as well as the semantic prediction/integration between gesture and language. This hypothesis is being supported by emerging behavioral and neuroscientific evidence. For example, it has been shown that the semantic processing of gestures—during a gesture recognition task for pantomimes—is facilitated when they are instructed as from a more communicative context (Trujillo, Simanova, Bekkering, & Özyürek, 2019). Moreover, the latter hypothesis is supported by a line of research comparing different gesture styles (Holle & Gunter, 2007; Obermeier, Kelly, & Gunter, 2015): with a disambiguation paradigm, the authors reported that the N400 effect on auditory speech for disambiguating gestures was modulated by adding non-straightforward grooming gestures (Holle & Gunter, 2007). Additionally, the N400 can also be modulated when participants watched an actor who more often produces meaningless grooming gestures, when compared to an actor who always produce straightforward non-grooming gestures. Despite providing initial evidence for social influence on semantic processing of gesture and language, however, these previous studies did not directly compare *identical* gestures differing in body orientation or viewpoints. In addition, it is suggested that language processing may be influenced by other socially relevant factors, for example, having versus not having a co-comprehender (Jouravlev et al., 2019; Rueschemeyer, Gardner, & Stoner, 2015), and being directed or not by eye-gaze (Holler et al., 2015; Holler, Kelly, Hagoort, & Ozyurek, 2012; McGettigan et al., 2017). Analogously, one might hypothesize similar interference effect to semantic processing by the body orientation of gestures.

The purpose of the current study was to therefore (a) find out the electrophysiological markers that differentiate gestures with frontal versus lateral body orientations during gesture perception, and (b) investigate how body orientation impacts upon semantic interaction between gesture and language. To these ends, we employed a cross-modal priming paradigm: participants were presented with visual sentences in an RSVP (rapid serial visual presentation) manner, primed by videos of frontal and lateral communicative gestures, and

their EEG data were recorded for both gesture perception and sentence processing. For gesture perception, although we expected to observe both alpha and beta power decrease for both types of gestures, we hypothesized that frontal versus lateral gesture observation will result in differential level of power decrease only in the beta band. In the visual sentence, we manipulated the semantic congruency between a target word and their gesture prime. We focused our analyses on the target words: for those primed with frontal gestures, we expected an N400 effect resulted from word-gesture incongruency. In the lateral gesture condition, however, we hypothesized that the implied lower degree of social-communicative intention could interfere with the level of semantic prediction from lateral gestures, thus leading to a modulated N400 congruency effect. Of note, this cross-modal priming paradigm, although appears to be more artificial and to depart from the majority of prior literature using more natural co-speech stimuli (see Section 4), provides (a) for gesture perception, EEG signals that are not affected by language/speech related processes (as in He, Steines, Sammer, et al., 2018), and (b) a testing case for dissociating the predictive aspects of the N400 during gesture-speech interaction: The target words (bottom-up information) are identical between the frontal and lateral conditions. Therefore, if an N400 difference is observed between frontal versus lateral gestures, it may most likely originate from a differential level of top-down semantic prediction.

2 | METHODS

2.1 | Participants

Twenty-four participants (15 female, mean age = 24.95, range 19–35) participated in this experiment and were monetarily compensated for participation. The sample size, although appears to be moderate, is comparable to the extensive body of N400 and brain oscillations literature on multimodal language processing (Habets et al., 2011; He, Nagels, Schlesewsky, & Straube, 2018; He, Steines, Sammer, et al., 2018; Wu & Coulson, 2005). All participants were right-handed as assessed by a questionnaire on handedness (Oldfield, 1971). They were all native German speakers. None of the participants reported any hearing or visual deficits. Exclusion criteria were history of relevant medical or psychiatric illnesses. All participants gave written informed consent prior to taking part in the experiment. The data from three participants were discarded because of excessive movement artifacts during recording.

2.2 | Materials and procedure

A sample of an experimental trial is illustrated in Figure 1c. A trial started with a fixation mark of 1,000 ms, followed by a video of an actor performing a hand gesture with two different body orientations (frontal vs. lateral, Figure 1a). The frontal and lateral videos were adopted from a line of fMRI studies investigating social aspects of
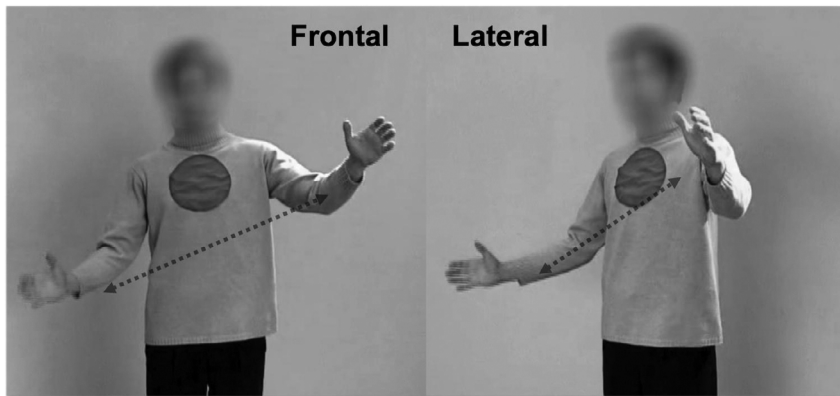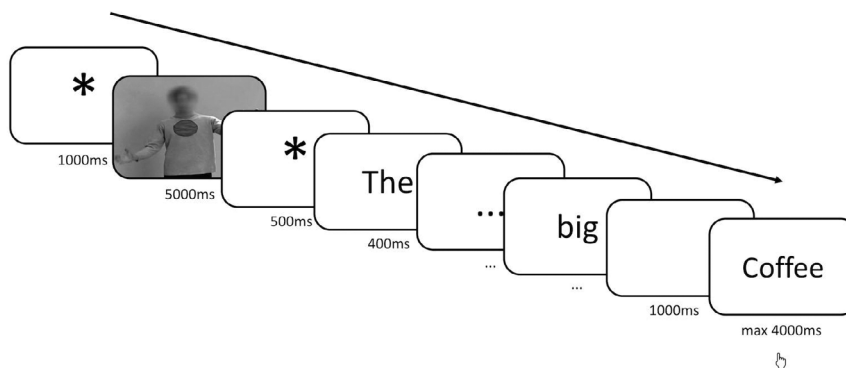
(a) Gesture video

Frontal   Lateral

(b) Target RSVP sentence

Congruent:	The/fish/is/<u>big</u>/thinks/Lara
Incongruent:	The/fish/is/<u>small</u>/thinks/Lara

(c) Sample trial

* 1000ms
5000ms
* 500ms
The 400ms
...
big
...
Coffee 1000ms
max 4000ms

FIGURE 1 Illustration of the experimental paradigm. (a): Each communicative gesture was recorded by two cameras and videos were presented with frontal (F) and lateral (L) body orientations. (b): Congruent (C) and incongruent (I) sentences of an experimental item, the critical word is underlined. (c): A sample of an exemplary trial. Each word was presented for 300 ms + 100 ms ISI. ISI, inter stimulus interval

gesture processing (Nagels et al., 2015; Straube et al., 2010; Straube, Green, Chatterjee, & Kircher, 2011). They were recorded by two video cameras simultaneously when the actor was gesturing. As a result, frontal and gesture videos are only different in terms of view-point, and all semantic and temporal aspects of the gestures remain comparable between the two conditions. All videos lasted 5,000 ms and did not contain any auditory information. For all videos, different from previous fMRI studies, we blurred the face area of the actor so that potential differences in eye gaze between frontal and lateral videos were not available to the participants, so as to avoid potential confounds of communicative intention as resulted from eye-gaze (Holler et al., 2015). We identified gesture onsets and offsets for all videos: the mean gesture onset time was about 1.10 s (SD = 0.29 s), and mean offset was about 3.38 s (SD = 0.46 s). The videos were followed by sentences presented in an RSVP (rapid-serial-visual-presentation) manner, with 300 ms per word and 100 ms inter stimulus interval (ISI). Sentences were formulated in two experimental conditions, such that they were either clearly congruent or incongruent with the prime video. All sentences were of the same structure as in Figure 1b: *Dass der Fisch groß/klein ist, findet*

Lara (The fish is <u>big/small</u>, thinks Lara). Therefore, the congruency between the sentence and the prime video becomes only apparent at the critical adjective, and that the critical words do not occur at the sentence-final position (Stowe, Kaan, Sabourin, & Taylor, 2018). We controlled word length and frequency (Wortschatz Leipzig) at the position of the critical words between both congruency conditions. The length of the critical words in the congruent condition was marginally shorter than that of the incongruent condition (6.87 vs. 7.81, $t = -1.86$, $p = .065$). The frequency of critical words was not different between the two conditions (12.44 vs. 12.53, $t = -0.19$, $p = .85$). One second after the presentation of the RSVP sentence, a semantic probe word (related or unrelated, 50% each) appeared for a maximum of 4,000 ms, prompting the participants to judge, as fast as possible, whether the word is semantically related or unrelated to the previous event (irrespective of gesture or sentence). Altogether, we used 32 videos (17 iconic gestures and 15 emblematic gestures) differing in body orientation. For each video, two sets of (congruent/incongruent) sentences were created. Therefore, the 256 (32-videos × 2-orientation × 2-congruency × 2-sets) experimental items were split into two lists of 128 items each. Each

participant was presented with only one pseudo-randomized list consisting of 128 experimental items together with an additional 128 filler items of no interests. The fillers items were formed by a combination of 32 different frontal/lateral videos followed by congruent/incongruent RSVP sentence with varying sentence structure (e.g., a combination of a raising hands and shredding shoulder gesture with sentences *The young man does not know the answer/The student presented a suitable answer*). For each participant, an experimental session lasted for around an hour, distributed in eight blocks.

2.3 | EEG acquisition

EEG was collected from 64 Ag/AgCl passive electrodes attached to an elastic cap (EasyCap GmbH, Herrsching, Germany) according to the international 10–10 System. The reference electrode was located at the FCz and the ground electrode was located at the forehead in AFz. All input impedances were kept below 10 kΩ. Additionally, the vertical electrooculogram was recorded from one electrode located underneath the left eye. Two "Brain Amp" (Brain Products, Garching, Germany) amplifiers were used to sample data at 500 Hz with a resolution of 0.1 μV. Trigger signals from stimulus and participants responses were presented with Presentation Software (Neurobehavioral Systems, Berkeley, CA), and acquired together with the EEG using Brain Vision Recorder (Brain Products GmbH).

2.4 | EEG preprocessing and analyses

All EEG preprocessing and analyses were carried out using the Brain Vision Analyzer 2.1 (Brain Products GmbH) and the Fieldtrip toolbox for EEG/MEG analysis (Oostenveld, Fries, Maris, & Schoffelen, 2011). We applied two processing pipelines at the video onset (comparing time-frequency representation [TFR] between frontal and lateral videos) and at the onset of the critical words in the target sentence (comparing ERPs for the four experimental conditions).

2.4.1 | EEG preprocessing

For the TFR analyses at video onset, raw continuous EEG were firstly high-pass filtered at 0.1 Hz and low-pass filtered at 125 Hz, and then re-referenced to the average of all electrodes. EOG and muscle artifacts were manually identified and rejected via an infomax independent component analysis. The raw EEG was then segmented to segments around −1 to 5.5 s around the onset of video. After segmentation, we firstly applied automatic muscle artifact rejection based on the amplitude distribution across trials and channels (as implemented in Fieldtrip toolbox). Cutoffs for these artifacts were set at $z = 6$. Afterwards, additional trials were rejected manually. The average rejection rate for both frontal and lateral conditions was 13.61% ($SD = 9.48\%$), with no significant difference between the two conditions.

For ERP comparison time-locked to the critical word onset, we applied a band pass filter between 0.1 and 40 Hz to the raw EEG. The re-referencing and artifacts rejection procedure (for both continuous and segmented data) was analogous to the TFR analyses for video onset, except for the fact that we segmented a shorter time window between −0.2 and 1 s based on critical word onset. Data preprocessing resulted in an average rejection rate of 9.63% ($SD = 7.62\%$) for all four conditions, with no significant main effects or interaction between the two experimental manipulations.

2.4.2 | TFR and source analyses

As we are primarily interested in the alpha and beta frequency bands for a direct comparison between frontal and lateral videos, for TFR computation, we applied a sliding window Hanning taper approach (5 cycles per window), in frequency steps of 1 Hz in the range of 2–30 Hz, and time steps of 0.20 s. All TFRs were interpreted based on baseline corrected (−0.5 to −0.25 s) decibel change (dB). After time-frequency decomposition, we applied a cluster-based random permutation test (Maris & Oostenveld, 2007) with 1,000 permutations to test the difference in both the alpha (8–12 Hz) and beta bands (16–24 Hz) between frontal and lateral conditions. The test was conducted in the time-electrode space, with each electrode having on average 5.8 neighboring electrodes according to a template layout. A cluster in the permutation test contained at least three neighboring electrodes.

For source localization of the beta band effect (see Section 3), we conducted a frequency-domain adaptive spatial filtering imaging of coherent sources algorithm (Gross et al., 2001), as implemented in the Fieldtrip toolbox. Source analysis was performed for the time-frequency windows in the beta band (20 Hz with 4 Hz padding) in which significant results were obtained on the scalp level. Participants' electrode positions (defined with a templated electrode layout) were warped to the cortical mesh of a standard boundary element head model (BEM). The forward model was computed on an 8 mm grid of source positions covering the whole brain compartment of the BEM. For source analysis, common space filters were constructed using the leadfield of each grid point and the cross-spectral density matrix (CSD). The CSD matrix was computed based on an additional time-frequency analyses for data segments of interests plus their respective baseline. The source activity volumes of interests were firstly corrected against the baseline period, and were then compared by means of paired t tests between both experimental conditions. This procedure resulted in a source-level t statistics of alpha power change for each voxel in the volume grid, in which we thresholded at a level of $p < .05$, uncorrected. For identification of anatomical labels, we used the AAL atlas (Tzourio-Mazoyer et al., 2002).

2.4.3 | ERP analyses

In order to reveal the N400 difference at the critical word, for each segment, we applied a baseline-correction based on the −0.2 to 0 s time window. Segments within a condition were then averaged within

each participant across trials in the first step. Secondly, grand averaged ERPs were averaged across all participants. For ERPs, we used cluster-based permutation test (1,000 permutations) for statistical comparison in the electrode space for amplitudes averaged within the N400 (300–500 ms) time window. Neighborhood parameters were set comparable to the TFR analysis. As we were interested in the interaction between body orientation and sentence-gesture congruency, we firstly compared congruency effects within each body orientation conditions, and then compared the incongruent > congruent difference between frontal and lateral conditions to test for statistical significance of interaction.

3 | RESULTS

3.1 | Behavioral results

The reaction times and accuracy for the semantic probe task is reported in Figure 2. For reaction times, repeated-measures analysis of variance showed a significant main effect of body orientation ($F_{[1,20]}$ = 11.32, p = .003). For accuracy, we observed no main effects ($F_{[1,20]max}$ = 2.71, p = .12) but a significant interaction between body orientation and congruency ($F_{[1,20]}$ = 6.15, p = .03). However, pairwise

t tests showed that congruency effects in both frontal and lateral gesture conditions were not significant ($|t|_{max}$ = 1.70, p = .10).

3.2 | Video onset: Beta band oscillations

We directly compared time-frequency representations between frontal and lateral gestures (Figure 3). For both conditions, although we observed alpha and beta power decrease relative to the baseline, only the beta band showed significant difference between the frontal versus lateral conditions ($p_{cluster}$ = .005). This effect has a right scalp distribution. The source of this beta band effect was source localized to the right middle frontal cortex (MNI coordinates: x = 34, y = 53, z = 12, $|t|_{max}$ = 2.78).

3.3 | Critical target word onset: N400

We compared the N400 effect firstly within each body orientation conditions (Figure 4). For frontal gestures, congruent versus incongruent words elicited a significant N400 effect with a central-parietal scalp distribution ($p_{cluster}$ = .006, as in Figure 4a). For lateral gestures, however, no N400 effect for congruency was observed ($p_{cluster-min}$ = .46, Figure 4c). Secondly, we calculated the amplitude difference for congruency effects, and
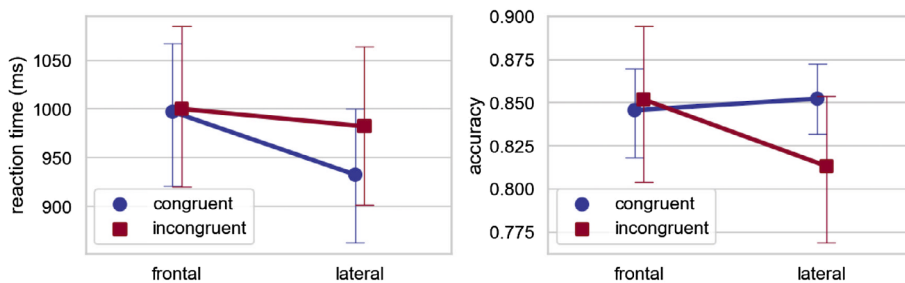
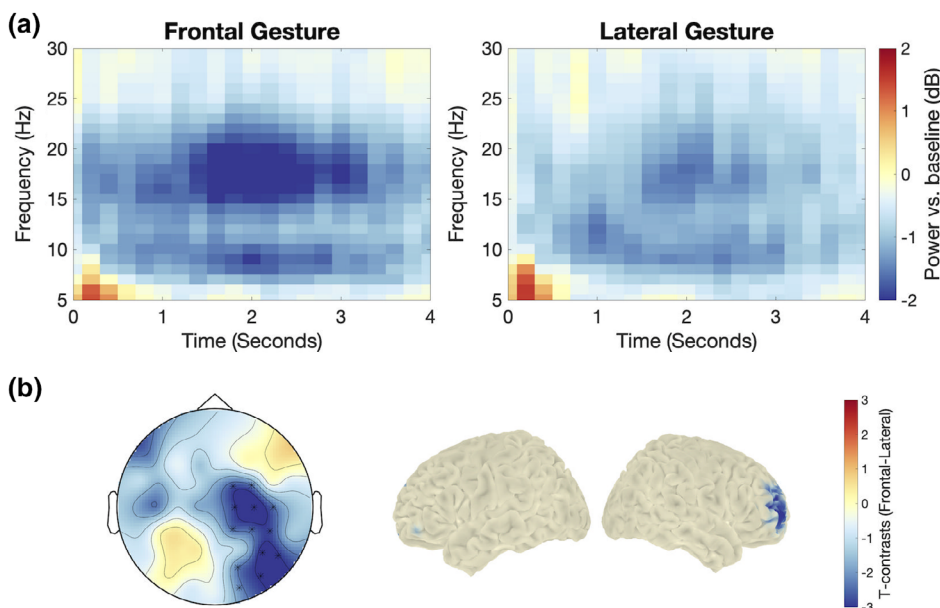**FIGURE 2** Reaction times (ms) and accuracy for the semantic probe task

FIGURE 3 Time-frequency (TF) results at video onset for frontal and lateral gestures. (a): Averaged TF representations across all significant electrodes showing significant difference (as in (b)—left). (b) (left): Scalp-level t-maps for frontal versus lateral conditions in the beta band (16–24 Hz) between 0.6 and 2.4 s. Electrodes showing significant difference between frontal and lateral conditions (p < .05, corrected) are marked with asterisks. (b) (right): Source-level t-maps in the beta band (16–24 Hz) between 0.6 and 2.4 s showing significant difference (p < .05, uncorrected) between frontal and lateral conditions
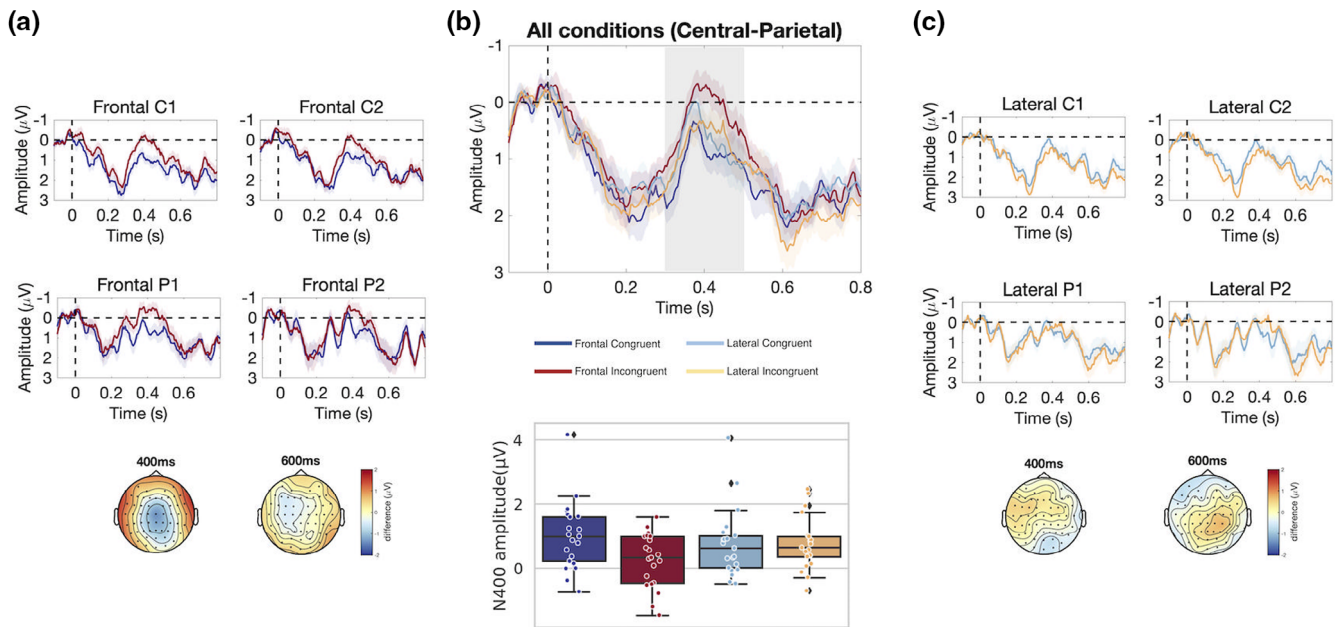
FIGURE 4 ERPs at the onset of the critical word. (a): ERPs for congruent versus incongruent words primed by frontal gestures at electrodes C1, C2, P1, and P2. Scalp maps display amplitude difference at 400 and 600 ms. (b): ERPs averaged from nine central-parietal electrodes (C1, Cz, C2, CP1, CP1, CPz, CP2, P1, Pz, P2), and box- and swarm-plots for individual subjects' N400 (300–500 ms) amplitudes from these electrodes, for all experimental conditions. (c): ERPs for congruent versus incongruent words primed by lateral gestures at electrodes C1, C2, P1, and P2. Scalp maps display amplitude difference at 400 and 600 ms. For all waveform maps, shaded areas indicate by-subject *SE*s for respective conditions. ERP, event-related potential

compared the congruency differences between frontal and lateral gestures to test for significance of interaction. This comparison led to a significant cluster also with central-parietal scalp distribution ($p_{cluster}$ = .002). We additionally analyzed the ERP effects at the noun-phrase preceding the critical word (e.g., *fish* as in *The fish is big*), and reported the results in the Figure S1. No difference has been observed between conditions at this position.

3.4 | Post-hoc correlation analyses between the beta power and the N400

As the modulation of the N400 effect for lateral gestures may result from perception of lateral gestures and its implied social-communicative intention (see Section 4), we set out a number of post-hoc, by-subject correlation analyses (Spearman's) between the trial-averaged beta power modulation (difference between frontal and lateral gestures) and the N400 amplitude difference between frontal and lateral gestures for the incongruent condition. However, we did not observe any significant correlation between these two measures.

4 | DISCUSSION

We conducted the current study using EEG, so as to investigate the role of body orientation on gesture observation and on the effect of semantic priming of gesture on visual language processing. During gesture observation, time-frequency analyses showed that frontal gestures elicited a

more pronounced power decrease in the beta band, source localized to the mPFC. Moreover, for critical words in the visual sentences, we observed a clear interaction between body orientation and semantic congruency: at the critical word, the semantic congruency N400 effect observed in the frontal condition was modulated in the lateral condition. Below, we argue that these results may originate from the social-communicative nature of gesture's body orientation.

4.1 | Social perception of gesture's body orientation

We compared gesture perception for identical gestures that only differ in body orientation. Our results showed that, although both the alpha and the beta power were decreased during gesture perception, frontal gestures elicited a more pronounced beta band power decrease than lateral gestures. This finding may not only supports the role of the beta band oscillations for perception of hand actions in general (Angelini et al., 2018; Avanzini et al., 2012; Järveläinen et al., 2004), but also corroborates the sensitivity of the beta band power to the observation of different types of gestures (He, Steines, Sammer, et al., 2018; Quandt et al., 2012). Previous studies that directly tested perceptual differences of body orientation or view point of non-communicative hand actions (e.g., reaching) also showed a modulated decrease of beta power (and alpha power) for more allocentric view point (Angelini et al., 2018; Drew et al., 2015). Our study compliments this line of research by showing that this pattern

could be extended to more communicative hand actions, namely gestures. However importantly, in the current study, although we observed an alpha band power decrease for both frontal and lateral gestures, unlike those studies examining non-communicative hand actions (Angelini et al., 2018; Drew et al., 2015), the alpha power did not differ between frontal and lateral orientation in our study. This data pattern attests to the hypotheses that alpha and beta power are dissociable during action or gesture observation, in the sense that alpha power decrease might be more related to sensory/motor perception in general, whereas the beta band is more sensitive to fine-grained motor-related features (Brinkman et al., 2014; He, Steines, Sammer, et al., 2018; Salmelin et al., 1995).

Despite commonalities in the beta band power modulation to non-communicative hand actions, the gestures that we investigated are apparently more social. Importantly, as the beta band difference between frontal and lateral gestures were source-located to the mPFC, our finding supports the hypothesis that the difference of gesture's body orientation is both physical and social. The mPFC has been considered as one of the most important regions for social perception and interaction (Frith & Frith, 2006), and supports the perception of social-communicative intention (Ciaramidaro et al., 2013; Schilbach et al., 2006). Previous gesture studies using fMRI also showed that the mPFC is differentially activated for frontal versus lateral gestures (Nagels et al., 2015; Saggar et al., 2014; Straube et al., 2010), as well as gestures differing in their level of social-communicative intentions (Trujillo et al., 2020). In line with these fMRI findings, our results provide new evidence from EEG, suggesting that frontal gestures—as opposed to lateral gestures—direct to the comprehender, and may thus convey a higher degree of social-communicative intention.

4.2 | Social perception interacts with gesture-language semantic integration

At the critical word, we observed that, at least for frontal gestures, the processing of words primed by incongruent gestures elicited a classic N400 effect in terms of both latency and scalp distribution (Holcomb, 1988; Kutas & Hillyard, 1984). In addition, this N400 effect was modulated when participants integrate the critical word to laterally presented gesture primes. This finding, to our knowledge, is a first report showing that body orientation, or potentially social-communicative aspects of gestures, may interfere with the semantic interaction between gesture and language. Gestures, like word primes (Holcomb, 1988), semantically predict an upcoming event as described in the target sentences. At the critical word when this semantic prediction is verified/violated, one should have observed an N400 effect upon semantic incongruency, as observed in the frontal gesture condition. However, we failed to observe an N400 effect for the lateral gestures, even though bottom-up features between frontal and lateral gesture conditions are identical at the critical word. Here we offer two plausible explanations: Firstly, one might argue that semantic incongruency—as resulting from semantically validating the feature of the critical word to gesture's prediction—is only observable when

gestures are frontally presented and are more addressed to the comprehender. Alternatively, it is also reasonable to argue that when gestures are presented laterally, they *do not* semantically predict as much as frontal gestures. Importantly, under both arguments, the presence and absence of the N400 effect is only observable for semantic incongruency: for critical words that are semantically congruent, as well as non-critical words that are less relevant to gestures, body orientation of gesture does not seem to play an important role. This may reflect a processing hierarchy, that is, the N400 is only affected by lesser degree of semantic prediction from lateral gestures (or higher degree of semantic prediction from frontal gestures), when an *incongruency* between gesture and language is detected. However, further studies are necessary to test this hypothesis, especially in naturalistic co-speech scenarios (Holler & Levinson, 2019): In a co-speech context, from a predictive coding perspective (Lewis & Bastiaansen, 2015; Rao & Ballard, 1999), humans may be able to *tolerate* the decreased level of semantic prediction from lateral gestures, as semantic predictions are being updated by bottom-up checks when both gesture strokes and auditory streams unfold. However, the view-point (social) influence could still be so strong that it is immune from being updated, which would still lead to a modulated N400 for lateral gestures during online gesture-speech integration.

Nevertheless, being initial evidence, our results suggest a potential role of gesture's body orientation during semantic interaction between gesture and visual language processing. This finding echoes the studies by Holle and Gunter (2007) and Obermeier et al. (2015), who observed that sentence disambiguation N400 may be modulated by (a) additional grooming gestures that are non-communicative, and by (b) watching non-grooming gestures from an actor who produces more grooming gestures. Similar behavioral findings were also reported by Trujillo et al. (2019), showing that gestures that are rated with differential degree of communicative intentions have a direct impact on their semantic processing. In a broader sense, our study, together with others, imply that the established interaction between gesture and language may not only originates from their semantic, but also social-communicative nature (Holler & Levinson, 2019). Of note, the social impact of language processing was also observed by other "social" manipulations, as shown by two recent studies (Jouravlev et al., 2019; Rueschemeyer et al., 2015): both studies showed that sentence processing (the N400 effect caused by contextual facilitation) could be affected by having a co-comprehender. These so-called "social-N400" effects, together with the modulation effects on the N400 from our current and previous gesture studies, call for an extension of the current functional interpretation of the N400 component during sentence processing. Despite the prediction versus integration debate (Bornkessel-Schlesewsky & Schlesewsky, 2019; Kutas & Federmeier, 2011; Lau et al., 2008; Nieuwland et al., 2020), it becomes clearer that the role of social information, especially how it interacts with semantic prediction during sentence processing, will need to be addressed by the current language processing models: Across the reported N400 studies with social manipulations, it is apparent that top-down social aspects are the only resources that has led to the reported N400 modulations, and are therefore an important

component during linguistic prediction. The imperatives for further investigations are to find out: when, how, and at what levels of the linguistic hierarchy social influences are observable, and whether social influences can be interactively modulated by multimodal language inputs in a naturalistic setting. Such an extended interactive model will not only provide a more comprehensive framework for empirical testing, but is also beneficial for a better understanding and treatment of social-communicative dysfunctions across several types of mental disorders (Dawson, Meltzoff, Osterling, Rinaldi, & Brown, 1998; Green, Horan, & Lee, 2015; Suffel, Nagels, Steines, Kircher, & Straube, 2020; Yang et al., 2020).

CONFLICT OF INTEREST

The authors declare no potential conflict of interest.

AUTHOR CONTRIBUTIONS

Y.H., B.S., and A.N. conceptualized the study. Y.H., S.L., R.M. designed the experiments. S.L. acquired the data. Y.H. and S.L. analyzed the data. Y.H., B.S., A.N. interpreted the data. Y.H. wrote the manuscript. Y.H., S.L., R. M., B.S., and A.N. revised the manuscript.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Yifei He https://orcid.org/0000-0002-2826-4230

REFERENCES

Angelini, M., Fabbri-Destro, M., Lopomo, N. F., Gobbo, M., Rizzolatti, G., & Avanzini, P. (2018). Perspective-dependent reactivity of sensorimotor mu rhythm in alpha and beta ranges during action observation: An EEG study. *Scientific Reports*, 8(1), 12429. https://doi.org/10.1038/s41598-018-30912-w

Avanzini, P., Fabbri-Destro, M., Dalla Volta, R., Daprati, E., Rizzolatti, G., & Cantalupo, G. (2012). The dynamics of sensorimotor cortical oscillations during the observation of hand movements: An EEG study. *PLoS One*, 7(5), e37534. https://doi.org/10.1371/journal.pone.0037534

Baggio, G., & Hagoort, P. (2011). The balance between memory and unification in semantics: A dynamic account of the N400. *Language and Cognitive Processes*, 26(9), 1338–1367.

Biau, E., & Soto-Faraco, S. (2013). Beat gestures modulate auditory integration in speech perception. *Brain and Language*, 124(2), 143–152.

Biau, E., Torralba, M., Fuentemilla, L., de Diego Balaguer, R., & Soto-Faraco, S. (2015). Speaker's hand gestures modulate speech perception through phase resetting of ongoing neural oscillations. *Cortex*, 68, 76–85.

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2019). Toward a neurobiologically plausible model of language-related, negative event-related potentials. *Frontiers in Psychology*, 10(298), 10. https://doi.org/10.3389/fpsyg.2019.00298

Brinkman, L., Stolk, A., Dijkerman, H. C., de Lange, F. P., & Toni, I. (2014). Distinct roles for alpha-and beta-band oscillations during mental simulation of goal-directed actions. *The Journal of Neuroscience*, 34(44), 14783–14792.

Ciaramidaro, A., Becchio, C., Colle, L., Bara, B. G., & Walter, H. (2013). Do you mean me? Communicative intentions recruit the mirror and the mentalizing system. *Social Cognitive and Affective Neuroscience*, 9(7), 909–916.

Cuevas, P., Steines, M., He, Y., Nagels, A., Culham, J., & Straube, B. (2019). The facilitative effect of gestures on the neural processing of semantic complexity in a continuous narrative. *NeuroImage*, 195, 38–47.

Dawson, G., Meltzoff, A. N., Osterling, J., Rinaldi, J., & Brown, E. (1998). Children with autism fail to orient to naturally occurring social stimuli. *Journal of Autism and Developmental Disorders*, 28(6), 479–485.

Dick, A. S., Mok, E. H., Beharelle, A. R., Goldin-Meadow, S., & Small, S. L. (2012). Frontal and temporal contributions to understanding the iconic co-speech gestures that accompany speech. *Human Brain Mapping*, 35, 900–917.

Drew, A. R., Quandt, L. C., & Marshall, P. J. (2015). Visual influences on sensorimotor EEG responses during observation of hand actions. *Brain Research*, 1597, 119–128.

Drijvers, L., Özyürek, A., & Jensen, O. (2018a). Alpha and beta oscillations index semantic congruency between speech and gestures in clear and degraded speech. *Journal of Cognitive Neuroscience*, 30(8), 1086–1097. https://doi.org/10.1162/jocn_a_01301

Drijvers, L., Özyürek, A., & Jensen, O. (2018b). Hearing and seeing meaning in noise: Alpha, beta, and gamma oscillations predict gestural enhancement of degraded speech comprehension. *Human Brain Mapping*, 59(5), 617–613. https://doi.org/10.1002/hbm.23987

Fabbri-Destro, M., Avanzini, P., De Stefani, E., Innocenti, A., Campi, C., & Gentilucci, M. (2015). Interaction between words and symbolic gestures as revealed by N400. *Brain Topography*, 28(4), 591–605.

Frith, C. D., & Frith, U. (2006). The neural basis of mentalizing. *Neuron*, 50(4), 531–534.

Green, A., Straube, B., Weis, S., Jansen, A., Willmes, K., Konrad, K., & Kircher, T. (2009). Neural integration of iconic and unrelated coverbal gestures: A functional MRI study. *Human Brain Mapping*, 30(10), 3309–3324. https://doi.org/10.1002/hbm.20753

Green, M. F., Horan, W. P., & Lee, J. (2015). Social cognition in schizophrenia. *Nature Reviews Neuroscience*, 16(10), 620–631.

Gross, J., Kujala, J., Hämäläinen, M., Timmermann, L., Schnitzler, A., & Salmelin, R. (2001). Dynamic imaging of coherent sources: Studying neural interactions in the human brain. *Proceedings of the National Academy of Sciences of the United States of America*, 98(2), 694–699.

Habets, B., Kita, S., Shao, Z., Özyurek, A., & Hagoort, P. (2011). The role of synchrony and ambiguity in speech-gesture integration during comprehension. *Journal of Cognitive Neuroscience*, 23(8), 1845–1854.

Hari, R., Forss, N., Avikainen, S., Kirveskari, E., Salenius, S., & Rizzolatti, G. (1998). Activation of human primary motor cortex during action observation: A neuromagnetic study. *Proceedings of the National Academy of Sciences of the United States of America*, 95(25), 15061–15065.

He, Y., Gebhardt, H., Steines, M., Sammer, G., Kircher, T., Nagels, A., & Straube, B. (2015). The EEG and fMRI signatures of neural integration: An investigation of meaningful gestures and corresponding speech. *Neuropsychologia*, 72, 27–42. https://doi.org/10.1016/j.neuropsychologia.2015.04.018

He, Y., Nagels, A., Schlesewsky, M., & Straube, B. (2018). The role of gamma oscillations during integration of metaphoric gestures and abstract speech. *Frontiers in Psychology*, 9, 1348. https://doi.org/10.3389/fpsyg.2018.01348

He, Y., Steines, M., Sammer, G., Nagels, A., Kircher, T., & Straube, B. (2018). Action-related speech modulates beta oscillations during observation of tool-use gestures. *Brain Topography*, 31(5), 838–847. https://doi.org/10.1007/s10548-018-0641-z

He, Y., Steines, M., Sommer, J., Gebhardt, H., Nagels, A., Sammer, G., ... Straube, B. (2018). Spatial-temporal dynamics of gesture-speech integration: A simultaneous EEG-fMRI study. *Brain Structure & Function*, 223(7), 3073–3089. https://doi.org/10.1007/s00429-018-1674-5

Holcomb, P. J. (1988). Automatic and attentional processing: An event-related brain potential analysis of semantic priming. *Brain and Language*, 35(1), 66–85.

Holle, H., & Gunter, T. C. (2007). The role of iconic gestures in speech disambiguation: ERP evidence. *Journal of Cognitive Neuroscience*, 19(7), 1175–1192.

Holle, H., Obleser, J., Rueschemeyer, S.-A., & Gunter, T. C. (2010). Integration of iconic gestures and speech in left superior temporal areas boosts speech comprehension under adverse listening conditions. *NeuroImage*, 49(1), 875–884.

Holler, J., Kelly, S., Hagoort, P., & Ozyurek, A. (2012). *When gestures catch the eye: The influence of gaze direction on co-speech gesture comprehension in triadic communication*. Paper presented at the 34th Annual Meeting of the Cognitive Science Society (CogSci 2012).

Holler, J., Kokal, I., Toni, I., Hagoort, P., Kelly, S. D., & Özyürek, A. (2015). Eye'm talking to you: Speakers' gaze direction modulates co-speech gesture processing in the right MTG. *Social Cognitive and Affective Neuroscience*, 10(2), 255–261.

Holler, J., & Levinson, S. C. (2019). Multimodal language processing in human communication. *Trends in Cognitive Sciences*, 23, 639–652.

Järveläinen, J., Schuermann, M., & Hari, R. (2004). Activation of the human primary motor cortex during observation of tool use. *NeuroImage*, 23(1), 187–192.

Jouravlev, O., Schwartz, R., Ayyash, D., Mineroff, Z., Gibson, E., & Fedorenko, E. (2019). Tracking colisteners' knowledge states during language comprehension. *Psychological Science*, 30(1), 3–19.

Kelly, S. D., Kravitz, C., & Hopkins, M. (2004). Neural correlates of bimodal speech and gesture comprehension. *Brain and Language*, 89(1), 253–260.

Kircher, T., Straube, B., Leube, D., Weis, S., Sachs, O., Willmes, K., ... Green, A. (2009). Neural interaction of speech and gesture: Differential activations of metaphoric co-verbal gestures. *Neuropsychologia*, 47(1), 169–179. https://doi.org/10.1016/j.neuropsychologia.2008.08.009

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.

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. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, 307(5947), 161–163. https://doi.org/10.1038/307161a0

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. *Nature Reviews Neuroscience*, 9(12), 920–933. https://doi.org/10.1038/nrn2532

Lewis, A. G., & Bastiaansen, M. (2015). A predictive coding framework for rapid neural dynamics during sentence-level language comprehension. *Cortex*, 68, 155–168.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, 164(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

McGettigan, C., Jasmin, K., Eisner, F., Agnew, Z. K., Josephs, O. J., Calder, A. J., ... Scott, S. K. (2017). You talkin' to me? Communicative talker gaze activates left-lateralized superior temporal cortex during perception of degraded speech. *Neuropsychologia*, 100, 51–63. https://doi.org/10.1016/j.neuropsychologia.2017.04.013

Nagels, A., Kircher, T., Steines, M., & Straube, B. (2015). Feeling addressed! The role of body orientation and co-speech gesture in social communication. *Human Brain Mapping*, 36(5), 1925–1936. https://doi.org/10.1002/hbm.22746

Nieuwland, M. S., Barr, D. J., Bartolozzi, F., Busch-Moreno, S., Darley, E., Donaldson, D. I., ... Huettig, F. (2020). Dissociable effects of prediction and integration during language comprehension: Evidence from a large-scale study using brain potentials. *Philosophical Transactions of the Royal Society B*, 375(1791), 20180522.

Obermeier, C., Kelly, S. D., & Gunter, T. C. (2015). A speaker's gesture style can affect language comprehension: ERP evidence from gesture-speech integration. *Social Cognitive and Affective Neuroscience*, 10(9), 1236–1243.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9(1), 97–113.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, 2011, 1), 1–1), 9. https://doi.org/10.1155/2011/156869

Özyürek, A. (2014). Hearing and seeing meaning in speech and gesture: Insights from brain and behaviour. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1651), 20130296.

Özyürek, A., Willems, R. M., Kita, S., & Hagoort, P. (2007). On-line integration of semantic information from speech and gesture: Insights from event-related brain potentials. *Journal of Cognitive Neuroscience*, 19(4), 605–616.

Quandt, L. C., Marshall, P. J., Shipley, T. F., Beilock, S. L., & Goldin-Meadow, S. (2012). Sensitivity of alpha and beta oscillations to sensorimotor characteristics of action: An EEG study of action production and gesture observation. *Neuropsychologia*, 50(12), 2745–2751. https://doi.org/10.1016/j.neuropsychologia.2012.08.005

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2(1), 79–87.

Redcay, E., Velnoskey, K. R., & Rowe, M. L. (2016). Perceived communicative intent in gesture and language modulates the superior temporal sulcus. *Human Brain Mapping*, 37(10), 3444–3461.

Rueschemeyer, S.-A., Gardner, T., & Stoner, C. (2015). The social N400 effect: How the presence of other listeners affects language comprehension. *Psychonomic Bulletin & Review*, 22(1), 128–134.

Saggar, M., Shelly, E. W., Lepage, J.-F., Hoeft, F., & Reiss, A. L. (2014). Revealing the neural networks associated with processing of natural social interaction and the related effects of actor-orientation and face-visibility. *NeuroImage*, 84, 648–656.

Salmelin, R., Hámáaláinen, M., Kajola, M., & Hari, R. (1995). Functional segregation of movement-related rhythmic activity in the human brain. *NeuroImage*, 2(4), 237–243.

Schilbach, L., Wohlschlaeger, A. M., Kraemer, N. C., Newen, A., Shah, N. J., Fink, G. R., & Vogeley, K. (2006). Being with virtual others: Neural correlates of social interaction. *Neuropsychologia*, 44(5), 718–730.

Sekine, K., Schoechl, C., Mulder, K., Holler, J., Kelly, S., Furman, R., & Ozyurek, A. (2020). Evidence for online integration of simultaneous information from speech and iconic gestures: An ERP study. *Language, Cognition and Neuroscience*.

Stolk, A., Brinkman, L., Vansteensel, M. J., Aarnoutse, E., Leijten, F. S., Dijkerman, C. H., ... Toni, I. (2019). Electrocorticographic dissociation of alpha and beta rhythmic activity in the human sensorimotor system. *eLife*, 8. https://doi.org/10.7554/eLife.48065

Stowe, L. A., Kaan, E., Sabourin, L., & Taylor, R. C. (2018). The sentence wrap-up dogma. *Cognition*, 176, 232–247. https://doi.org/10.1016/j.cognition.2018.03.011

Straube, B., Green, A., Bromberger, B., & Kircher, T. (2011). The differentiation of iconic and metaphoric gestures: Common and unique

integration processes. *Human Brain Mapping, 32*(4), 520–533. https://doi.org/10.1002/hbm.21041

Straube, B., Green, A., Chatterjee, A., & Kircher, T. (2011). Encoding social interactions: The neural correlates of true and false memories. *Journal of Cognitive Neuroscience, 23*(2), 306–324. https://doi.org/10.1162/jocn.2010.21505

Straube, B., Green, A., Jansen, A., Chatterjee, A., & Kircher, T. (2010). Social cues, mentalizing and the neural processing of speech accompanied by gestures. *Neuropsychologia, 48*(2), 382–393. https://doi.org/10.1016/j.neuropsychologia.2009.09.025

Suffel, A., Nagels, A., Steines, M., Kircher, T., & Straube, B. (2020). Feeling addressed! The neural processing of social communicative cues in patients with major depression. *Human Brain Mapping*. 1–14.

Trujillo, J. P., Simanova, I., Bekkering, H., & Özyürek, A. (2019). The communicative advantage: How kinematic signaling supports semantic comprehension. *Psychological Research*, 1–15.

Trujillo, J. P., Simanova, I., Ozyurek, A., & Bekkering, H. (2020). Seeing the unexpected: How brains read communicative intent through kinematics. *Cerebral Cortex, 30*(3), 1056–1067.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., ... Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage, 15*(1), 273–289.

Van Overwalle, F. (2009). Social cognition and the brain: A meta-analysis. *Human Brain Mapping, 30*(3), 829–858.

Willems, R. M., Özyürek, A., & Hagoort, P. (2009). Differential roles for left inferior frontal and superior temporal cortex in multimodal integration of action and language. *NeuroImage, 47*(4), 1992–2004.

Wu, Y. C., & Coulson, S. (2005). Meaningful gestures: Electrophysiological indices of iconic gesture comprehension. *Psychophysiology, 42*(6), 654–667.

Xu, J., Gannon, P. J., Emmorey, K., Smith, J. F., & Braun, A. R. (2009). Symbolic gestures and spoken language are processed by a common neural system. *Proceedings of the National Academy of Sciences of the United States of America, 106*(49), 20664–20669. https://doi.org/10.1073/pnas.0909197106

Yang, Y., Lueken, U., Richter, J., Hamm, A., Wittmann, A., Konrad, C., ... Lang, T. (2020). Effect of CBT on biased semantic network in panic disorder: A multicenter fMRI study using semantic priming. *American Journal of Psychiatry, 177*(3), 254–264.

Zhang, Y., Frassinelli, D., Tuomainen, J., Skipper, J. I., & Vigliocco, G. (2020). More than words: The online orchestration of word predictability, prosody, gesture, and mouth movements during natural language comprehension. *bioRxiv*.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.