Brain Topography
https://doi.org/10.1007/s10548-018-0641-z

ORIGINAL PAPER

Action-Related Speech Modulates Beta Oscillations During Observation of Tool-Use Gestures

Yifei He[1] · **Miriam Steines**[1] · **Gebhard Sammer**[3] · **Arne Nagels**[2] · **Tilo Kircher**[1] · **Benjamin Straube**[1]

Received: 7 December 2017 / Accepted: 27 February 2018

Abstract

Language and action have been thought of as closely related. Comprehending words or phrases that are related to actions commonly activates motor and premotor areas, and this comprehension process interacts with action preparation and/or execution. However, it remains unclear whether comprehending action-related language interacts with action observation. In the current study, we examined whether the observation of tool-use gesture subjects to interaction with language. In an electroencephalography (EEG) study (n = 20), participants were presented with video clips of an actor performing tool-use (TU, e.g., hammering with a fist) and emblematic (EM, e.g., the thumb up sign for 'good job') gestures accompanied by either comprehensible German (G) or incomprehensible Russian sentences (R). Participants performed a semantic judging task, evaluating whether the co-speech gestures were object- or socially-related. Behavioral results from the semantic task showed faster response for the TU versus EM gestures only in the German condition. For EEG, we found that TU elicited beta power decrease (~ 20 Hz) when compared to EM gestures, however this effect was reduced when gestures were accompanied by German instead of Russian sentences. We concluded that the processing of action-related sentences might facilitate gesture observation, in the sense that motor simulation required for TU gestures, as indexed by reduced beta power, was modulated when accompanied by comprehensible German speech. Our results corroborate the functional role of the beta oscillations during perception of hand gestures, and provide novel evidence concerning language–motor interaction.

Keywords Action observation · Beta oscillations · Tool-use gesture · Language–motor interaction · Embodied cognition

Introduction

People use a variety of co-speech gestures during daily communication (McNeill 1992, 2008). Among those, tool-use gestures stand out in particular as they depict a specific physical form of an action by pantomiming the use of an action-related tool (Higuchi et al. 2007; Johnson-Frey 2004; Johnson-Frey et al. 2005). As an example, a hand action such as imitating "hitting a nail with a hammer" could be easily transferred to a corresponding tool-use gesture. Therefore, tool-use gestures are also commonly referred to as transitive gestures (Villarreal et al. 2008). On the contrary, there are types of gestures that are intransitive. For example, emblematic gestures convey meanings, which are readily understood and highly conventionalized, but they are usually arbitrary symbols not translated from any actions (e.g., the thumbs up sign for 'good job'). Clearly, although both tool-use and emblematic gestures are meaningful hand actions, they differ in the degree of transitivity and thus reflect differential level of motor simulation during observation (Agnew et al. 2012; Lindenberg et al. 2012; Villarreal et al. 2008).

Greater degree of motor simulation in the motor and premotor cortices has been closely associated with the observation of tool-use gestures when compared with other gesture types. A number of recent functional magnetic resonance imaging (fMRI) studies have directly investigated the observation of tool-use gestures (Agnew et al. 2012; Buccino et al.

Handling Editor: Micah M. Murray.

✉ Yifei He
yifei.he@staff.uni-marburg.de

[1] Department of Psychiatry and Psychotherapy, Philipps-University Marburg, Rudolf-Bultmann-Str. 8, 35039 Marburg, Germany

[2] Department of General Linguistics, Johannes-Gutenberg University Mainz, Jacob-Welder-Weg 18, 55122 Mainz, Germany

[3] Cognitive Neuroscience at Centre for Psychiatry, Justus Liebig University Giessen, Am Steg 28, 35385 Giessen, Germany

2001; Newman-Norlund et al. 2010; Villarreal et al. 2008). Those report that the observation of tool-use versus non-tool-use gestures engenders greater activation in pre/motor and parietal regions (Agnew et al. 2012; Villarreal et al. 2008), thus suggesting the greater degree of motor simulation specifically for tool-use gesture processing. In parallel, similar findings were reported by a number of M/EEG studies. It is well established that gesture observation elicits attenuation of oscillatory activity in the alpha (8–13 Hz) and/or beta (14–30 Hz) bands (Arnstein et al. 2011; Hari et al. 1998; Muthukumaraswamy et al. 2004; Perry and Bentin 2009), and that power decrease in the alpha/beta bands directly reflects increased activity in the premotor and the primary motor cortices (Arnstein et al. 2011). More importantly, with an MEG study, Järveläinen et al. (2004) specifically reported that observing tool-use gestures elicited a beta power decrease (20 Hz) in the primary motor cortex, thus suggesting that the beta band may directly reflect motor simulation, as required by observing tool-use gestures. In a similar vein, scalp EEG studies comparing the observation of more concrete and transitive gesture types commonly report a power decrease in either the alpha or the beta frequency bands when compared with gesture types that are less or non-transitive, e.g., deictic gestures (Avanzini et al. 2012; Quandt et al. 2012; Wriessnegger et al. 2013). Quandt et al. (2012) compared iconic versus deictic (pointing) gestures, and they reported power decrease in parietal–occipital electrodes in both the alpha (8–13 Hz) and beta (14–30 Hz) bands for iconic gestures; similarly, while looking at four different gesture types (grasping, pointing, supinating, and clenching), Avanzini et al. (2012) found that although power decrease in both the central parietal alpha (8–13 Hz) and the beta (18–25 Hz) bands were related to action observation, the alpha band was sensitive to the target versus non-target comparison, whereas the beta band was closely correlated to the velocity profile of the observed gesture types. Wriessnegger et al. (2013) compared the observation of grasping by manipulating the grasped object (realistic vs. abstract). What they observed was parietal power decrease in both the alpha (10–12 Hz) and the beta (16–20 Hz) bands for realistic objects when compared to abstract objects. A synthesis of these M/EEG studies suggests that, oscillations in both the alpha band (8–12 Hz) and the beta band (16–24 Hz, centering around 20 Hz) were sensitive to the types of gestures during observation, and the effects primarily had central/parietal scalp distribution, which may reflect differential degree of motor simulation within the pre/motor cortices: specifically, more concrete, transitive, and targeted-directed gestures commonly elicited the most pronounce power decrease in both frequency bands during gesture observation.

However, despite the increasing interests in action and gesture observation, previous research has focused more on gesture observation without any corresponding language context, even if gestures are rarely encountered in isolation in real life. With regard to the relationship between gesture and language, a consensus has been achieved jurying that gesture and speech are closely intertwined cognitive processes (Goldin-Meadow 2005; McNeill 2006), and that the two processes can be considered as two aspects of a single process (Kendon 1980). In fact, more recent evidence from neuroimaging suggests that gesture and speech share partially overlapping neural substrates (Andric et al. 2013; Straube et al. 2012; Xu et al. 2009). More importantly, increasing evidence suggests that gesture/action and language interact with each other (for a review, see Willems and Hagoort 2007). On the one hand, recent behavioral and neural evidence has demonstrated that gesture is able to facilitate speech production and comprehension (Biau and Soto-Faraco 2013; Holle et al. 2012; Kita and Özyürek 2003; Wang and Chu 2013). On the other, language has been shown to have substantial impact on motor behavior such as action planning and execution, even if this impact can be either facilitative or inhibitive. Studies have shown that presenting action-related language can facilitate action execution (Boulenger et al. 2006; Scorolli and Borghi 2007). Meanwhile, it is reported that action language can interfere with motor behavior (Buccino et al. 2005; Sato et al. 2008). For example, Buccino et al. (2005) performed an experiment asking participants to respond with either the hand or the foot if a presented verb was concrete, and to refrain from responding if the verb was abstract. Results showed that both reaction times and MEPs (motor evoked potentials) recorded from hand muscles were specifically modulated by listening to hand-action-related sentences, as were reaction times and MEPs recorded from foot muscle by listening to foot-action-related sentences. The findings are clear indication of interference effect to motor behavior from language contents that describe identical motor event. Similar effects were also reported by Klepp et al. (2015) showing modulated beta-power decrease for hand action preparation when participants were presented with hand-related action verbs, thus suggesting that the motor-related beta activity is sensitive to language–motor interaction. The line of studies form robust evidence supporting the embodied cognition theories, which suggest that language processing and motor behavior may share comparable neural basis (Barsalou 2008; Pulvermuller 2005), and that motor simulation, as required for action language perception (Hauk et al. 2004), may affect action observation in an interactive manner. Notably, in contrasts to action execution, no prior research has directly investigated the effect of action-related language on action and gesture observation. However, given the close resemblance between the role of motor regions during action execution and observation (Gallese et al. 1996; Hari et al. 1998; Rizzolatti et al. 2001), one might correspondingly speculate similar language–motor interaction during action/gesture observation. This interaction may be reflected by the differential level of activation in the pre/motor cortices, and potentially modulated/enhanced alpha/beta band activities under the influence of action-related language context.

The goal of the present study was twofold: firstly, we aimed at replicating previous M/EEG studies on tool-use gestures, by directly testing whether motor simulation, as reflected by the level of alpha and beta power change, is sensitive to the observation of tool-use gestures. Second, we precisely examined whether oscillatory activities related to tool-use observation subject to potential interaction from sentences describing identical action-related events. We used emblematic gestures (EM) to be compared with tool-use gestures (TU), because EM gestures are comparable to TU gestures in the sense that they are also meaningful without speech, however they can provide the maximum contrast to tool-use gestures in terms of transitivity (Lindenberg et al. 2012; McNeill 2008). Additionally, we manipulated the language contexts in which the gestures were observed. All gestures were accompanied by two different languages: comprehensible German (G) and incomprehensible Russian (R). In comparison to the German conditions, the Russian sentences do not provide native German participants with a genuine language context; however they serve as viable control in terms of auditory input. We tested two major hypotheses in the current study: (1) EEG responses elicited by TU and EM gestures differ. This difference will be observed as power decrease for the TU gestures in the alpha and/or beta bands with central/parietal scalp distribution. (2) Action-related sentences, as conveyed in the German condition, modulate/enhance the EEG effects obtained from the comparison between the two gesture types. We refrain from hypothesizing the direction of this interaction, as no comparable studies have directly investigated the role of language context on action observation, and additionally because of the heterogeneous reports on the direction of the effect of language on action execution.

Methods

Participants

Twenty subjects (14 female, mean age = 24.1, range 19–34 years) participated in this experiment. All participants were right-handed (Oldfield 1971). They were all native German speakers and had no Russian proficiency. None of the participants reported any hearing or visual deficits. Exclusion criteria were history of relevant medical or psychiatric illnesses of the participants. All subjects gave written informed consent prior to participation in the experiment. The study was approved by the local ethics committee.

Stimuli

The participants were shown video clips of co-speech gestures selected from a large pool of videos (Green et al. 2009; He et al. 2015; Kircher et al. 2009; Straube et al. 2011). Other material from this pool has been used in previous fMRI studies from our research group, focusing on different aspects of speech and gesture processing. All videos lasted 5 s and contained only one simple subject–verb–object sentence (e.g., 'The blacksmith hammers on the metal plate') spoken by an actor, who aimed at communicating a simple event with spontaneous performance with speech and/or gesture. In the current study, the video clips were selected based on a 2 × 2 factorial design. The actor was a highly proficient German–Russian early-bilingual speaker. He performed two types of gestures: these were tool-use gestures (TU) and emblematic gestures (EM). The two types of gestures were performed together with either comprehensible German sentences (G) or incomprehensible Russian sentences (R) correspondingly. Crucially, in the Russian condition, participants could understand the event only with visual-input from gesture, whereas in the German condition both speech and gesture were understandable. Overall, we used 208 videos (26 videos per condition × 4 conditions × 2 sets). Additionally, two control conditions with German sentences without gesture (26 videos per gesture type × 2 sets) and 52 filler videos containing Russian sentences with meaningless gestures were also presented. For a sample illustration of the videos, please refer to Fig. 1. For a complete list of the presented co-speech gestures, please refer to He et al. (2015).

Timing parameters (gesture onset, gesture offset, gesture apex, speech offset) for the four experimental conditions are listed in Table 1. Of note, all videos were trimmed to ensure that all speech sentences' onsets start at 0.5 s after video onsets. Analyses of variance with two factors (gesture type and language contexts) were performed for the all stimulus parameters. Of all stimulus parameters, we observed significant effect of *gesture-type* for gesture onset [$F_{(1,51)} = 15.56$, $p < 0.001$], gesture offset [$F_{(1,51)} = 102{,}81$, $p < 0.001$], as well as gesture apex [$F_{(1,51)} = 2.26$, $p = 0.002$]. No main effect of *language-context* and no interaction between the two factors were observed. Therefore, the stimulus-related differences were predominantly associated with gesture types, namely, TU gestures compared to EM gestures have earlier onsets (0.89 vs. 0.98 s), later offsets (3.93 vs. 3.32 s), and their apex are earlier (1.83 vs. 2.12 s). As our EEG analyses were based on gesture onsets and not video onsets, differences between stimulus duration are unlikely to have a relevant impact on the result pattern. More importantly, these parameters were not affected by the accompanying language context.

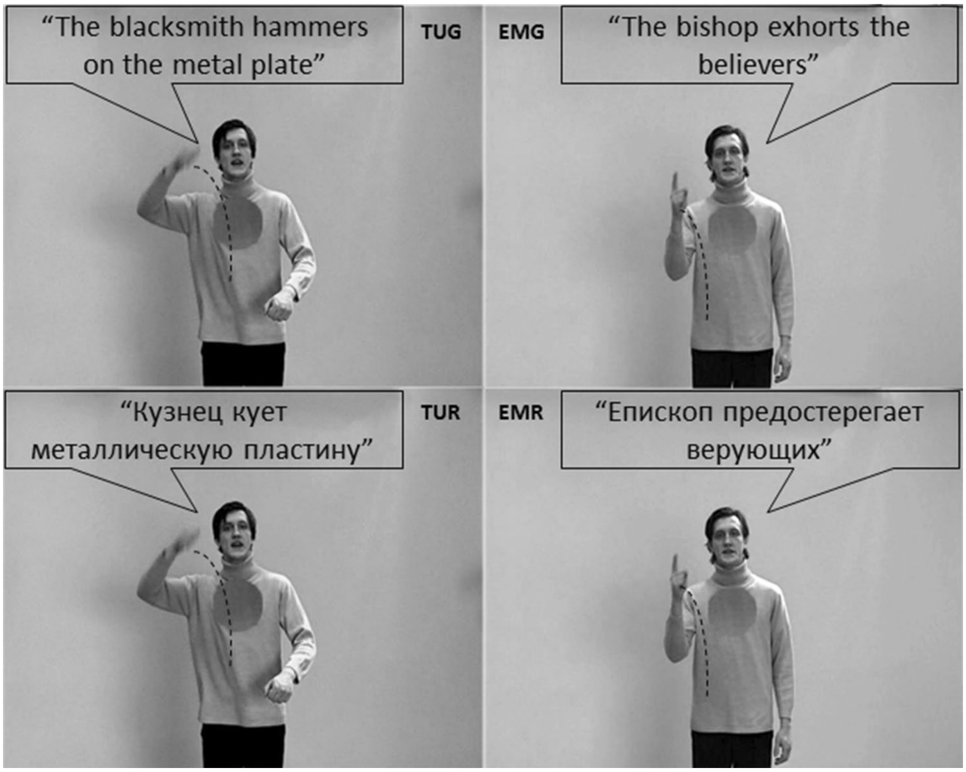

Fig. 1 Picture illustration for the tool-use gesture (TU) and emblematic gesture (EM) conditions in familiar (G, German, up) and foreign (R, Russian, bottom) language contexts. Additionally, German sentences without gestures had been presented. Note: For illustrative purposes the spoken German sentences were translated into English. For this figure all spoken sentences were written into speech bubbles

Table 1 Mean stimulus durations and rating parameters for videos in the four experimental conditions (*SD* standard deviation), all speech sentences onsets are 0.5 s after video onsets

	Gesture onset		Gesture offset		Gesture apex		Speech offset	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
TUG	0.87	0.20	3.93	0.43	1.83	0.29	2.71	0.42
EMG	1.09	0.37	3.40	0.44	1.99	0.48	2.79	0.42
TUR	0.91	0.18	3.92	0.41	1.82	0.31	2.73	0.42
EMR	0.97	0.31	3.24	0.45	2.24	0.6	2.65	0.34

Experimental Procedure

After participants gave written informed consent, an EEG-cap (Brain Products GmbH, Munich, Germany) was fastened to the participant's head and the electrodes were attached at their according sites. Abrasive electrode gel was administered and care was taken to keep all impedances below 5 kΩ. Participants were comfortably seated in front of a standard 19″ TFT monitor. An experimental session comprised 182 videos (26 for each condition and 26 additional filler videos) and consisted of two 14-min blocks. Each block contained 91 trials with a matched number of items from each condition (26) and 13 filler trials. The stimuli were presented in pseudo-randomized order and were counterbalanced across participants. Each video clip was followed by a gray background with a variable duration of 2154–5846 ms (jitter average: 4000 ms). Participants performed a content judgment task, indicating via button press whether the presented event was either object- or social-related (see Fig. 1). Crucially, all tool-use gestures are object-related whereas emblematic gestures are social/person-related. By performing this task, the participants focused directly on the semantic representation of the gestures and the corresponding speech. Participants were instructed to respond to the task as soon as they had decided on an answer. Of note, this task is independent of experimental conditions, namely, participants were asked to focus on the performance of the actor, and not specifically on either gesture or speech.

EEG Recording

EEG was collected from 29 sintered Ag/AgCl Electrodes attached to the BrainCap (Brain Products GmbH, Munich, Germany) according to the international 10-20-System. The reference electrode was located at the vertex between Fz and Cz and the ground electrode was located at the forehead in front of Fz. All input impedances were kept below 5 kΩ. Additionally, the vertical electrooculogram was recorded from one electrode located underneath the left eye. The 'Brain Amp' (Brain Products) amplifier was used to sample data at 500 Hz with a resolution of 0.1 µV. Data were analogously band-pass filtered between 0.016 and 250 Hz during recording.

EEG Data Analysis

All analyses were carried out using the Brain Vision Analyzer software (Brain Products GmbH, Munich, Germany) and the Fieldtrip toolbox for EEG/MEG analysis (Maris and Oostenveld 2007). Data were firstly high-pass filtered at 0.1 Hz and low-pass filtered at 125 Hz and then re-referenced to the average of the two mastoids (TP9 and TP 10). EOG and muscle artifacts were identified and rejected via an infomax independent component analysis. The continuous data were segmented around the onset of each gesture with a time interval between − 0.5 and 4 s. After the segmentation, additional jump artifacts were automatically detected and rejected based on the amplitude distribution across trials and channels (as implemented in Fieldtrip toolbox). Cutoffs for the jump artifacts were set at $z = 20$. The average numbers of trials were 23.06 (out of 26) per condition, with no significant difference between conditions. No participants were ruled-out because of excessive artifacts.

Based on previous M/EEG studies on gesture and action language processing (Järveläinen et al. 2004; Salmelin et al. 1995), we decomposed EEG time-series within the frequency bands between 2 and 30 Hz in order to capture the alpha and the beta band effects. For each gesture trial, during the − 0.5 to 4 s interval around the onset of the gesture, total power was computed using a constant Hanning taper of 400 ms in frequency steps of 1 Hz and time steps of 0.1 s. For statistical analyses, we used a cluster-based permutation approach so as to directly test the interaction between the two experimental factors in terms of time–frequency-representations. This approach naturally handles type-1 error resulting from multiple comparisons, and is therefore optimal solution for statistical analysis with time–frequency data which involve data points with multiple dimensions (Groppe et al. 2011; Maris and Oostenveld 2007). Based on M/EEG studies on the observation of TU gestures (Järveläinen et al. 2004; Wriessnegger et al. 2013), we focus our analysis directly on the alpha (8–13 Hz) and beta (16–24 Hz) bands on the baseline-corrected (− 0.5 to − 0.15 s) relative power changes. The detailed procedure is listed below:

For each frequency band, in each data point (time-electrode), firstly, we calculated the power difference between TU and EM gestures within each language contexts (TUG–EMG, and TUR–EMR). Secondly, a simple dependent-samples *t*-test is performed for each data point, which results in uncorrected *p* values. By testing the TU–EM difference in two language context conditions, uncorrected significant two-way interaction will be identified for each data point. Thirdly, all significant data points ($p < 0.05$) are grouped as clusters (here: clusters of at least three neighboring electrodes, with maximum distance of 2.5 cm). For each cluster the sum of the *t*-statistics is used in the cluster-level statistics. Finally, a Monte-Carlo non-parametrical permutation method with 1000 repetitions is implemented to estimate type I-error controlled cluster significance probabilities in the time-electrode space. Clusters reaching $p < 0.025$ are considered significant. We also conducted planned comparison for TUG versus EMG and TUR versus EMR separately: firstly, raw relative power changes were directly tested dependent-samples t-test for each data point; all following statistical procedures were identical to the interaction analysis.

Results

Behavioral Results

Participants were instructed to indicate via button press whether the actor in the video described and demonstrated a person- or object-related action/experience. Means and standard deviations of the correct responses and their reaction times in the EEG experiment are illustrated in Fig. 2. With respect to accuracy rates, descriptively, it is clear that the overall behavioral performance was generally high, with at least 80% correct responses. Repeated-measures

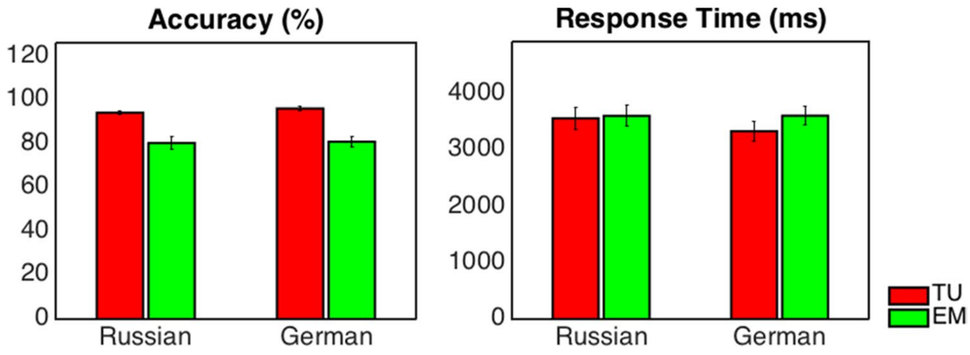

Fig. 2 Accuracy rates (%) and response times (ms) for the semantic task in all experimental conditions. The standard error within each condition is displayed with the error bar

ANOVA revealed significant main effect of gesture type ($F_{(1,19)} = 13.93$, $p < 0.01$): the accuracy rates for TU condition were generally higher than that of EM condition, irrespective of the language contexts. As for reaction times, there was a significant main effect of gesture type ($F_{(1,19)} = 37.16$, $p < 0.001$). Additionally, there was an interaction between gesture type and language contexts ($F_{(1,19)} = 7.68$, $p < 0.05$). Within the German condition, the difference between the TUG and EMG conditions was significant ($F_{(1,19)} = 28.87$, $p < 0.001$); however the difference between the TUR and EMR conditions was not significant. In general, in terms of accuracy rates, participants responded more accurately to the TU conditions in comparison to the EM conditions, irrespective of accompanying speech; as for reaction times, the faster processing of the TU gestures was only observable when they are accompanied by the German speech. The general higher accuracy rates for TU versus EM gesture may suggest that TU gestures are generally easier to categorize, potentially owing to their higher degree of transitivity (and thus more imagable) when compared to emblems.

EEG Results

We firstly directly tested the interaction between the two experimental factors [(TUG–EMG)–(TUR–EMR)] within both the alpha and the beta bands. We observed no significant clusters for the alpha band (8–13 Hz); however for the beta band (16–24 Hz), we observed a significant cluster ($p = 0.009$) over central–parietal electrodes, the effect was found between latency range of 1.0–2.2 s, as illustrated in Fig. 3b. Results from the planned pair-wise comparisons showed a significant negative cluster for TUR–EMR in the beta band; for TUG–EMG, no significant effects were observed. The TUR–EMR effect ($p = 0.001$) was observed largely over parietal electrodes, and the effect was found between 1.1 and 2.4 s, as in Fig. 3a. In sum, statistical results showed beta power decrease for the TU–EM comparison only when the accompanying language was foreign and not intelligible.

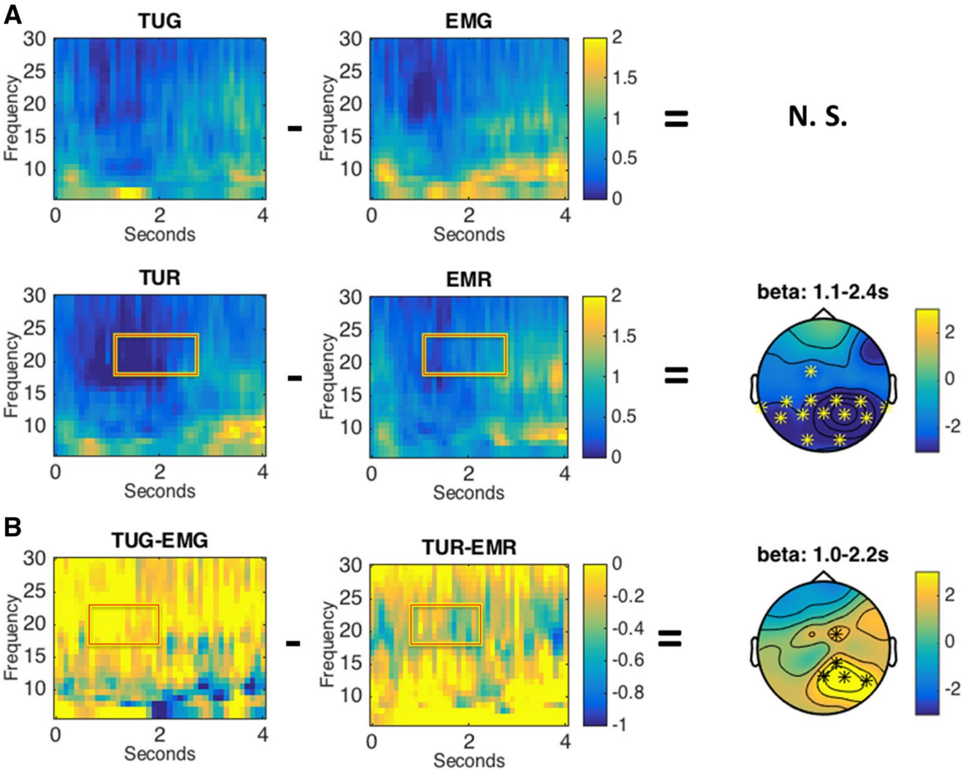

Fig. 3 Results of the time–frequency (TF) analysis for all experimental conditions compared at gesture onsets. **a** The panel shows the TF representations of all conditions at Pz electrode. The colorbar displays the relative power change after the gesture onsets when compared to the baseline activity (0.5–0.15 s before gesture onset). Significant difference in the beta band (16–24 Hz) was observed between the TUR and the EMR conditions, ranging between 1.1 and 2.4 s, and the effect has a parietal scalp distribution: the electrodes forming the significant cluster are marked with asterisks on the topographic plot. **b** The panel illustrates the significant interaction resulted from [(TUG–EMG)–(TUR–EMR)] in the beta band, the latency of this effect was between 1.0 and 2.2 s, and has a parietal scalp distribution: the electrodes forming the significant cluster are marked with asterisks on the topographic plot

Discussion

The aim of the current study was to investigate the oscillatory activities associated with tool-use gestures when comprehended with and without action-related sentences. Behaviorally, we found not only increased accuracy for all TU versus EM gestures, but also faster response times for TU gestures when they are accompanied by comprehensible German speech. With the EEG results from time–frequency analysis, we observed a significant two-way interaction between factors gesture-type and language-context, this effect occurred in the beta band with a central–parietal scalp distribution. Pairwise comparisons suggest that the interaction resulted from a significant beta band effect only for the TUR versus EMR comparison. Our findings corroborate the sensitivity of the beta oscillations to gesture types differing in the level of transitivity. More interestingly, they suggest possible interaction between gesture and speech during online comprehension.

Our findings concerning beta power decrease during tool-use gesture observation in the Russian condition are consistent with previous M/EEG studies reporting similar effects, which suggest the sensitivity of the beta oscillations to the degree of related transitivity in hand actions and gestures (Järveläinen et al. 2004; Moreno et al. 2013; Quandt et al. 2012; Schaller et al. 2017; Wriessnegger et al. 2013). For example, with an MEG experiment, Järveläinen et al. (2004) specifically reported that observing goal-directed versus non-goal-directed tool-use gestures elicited a beta power decrease (~ 20 Hz) in the primary motor cortex. This finding is direct evidence of the sensitivity of beta oscillations in the pre/motor regions to gestures differing in the degree of motor experience. Additionally, with an EEG experiment, Moreno et al. (2013) compared the observation of tool-use hand actions such as ‘cut the bread’ to the comprehension of abstract sentences, and found power decrease in the beta band. Similar beta band effects were also reported when the observation of tool-use gestures were compared against videos showing no movement (Schaller et al. 2017). In a similar vein, Quandt et al. (2012) compared iconic versus deictic gestures, in which the related sensorimotor experience of the gestures vary. They also observed a parietal power decrease in the beta band. Similar beta band effects originated from the comparisons between gesture types were also observed in a number of other studies (Avanzini et al. 2012; Wriessnegger et al. 2013). More importantly, despite slight topographic differences, all these EEG studies reported beta band effect with either central or central–parietal scalp distribution. These results, together with the present study, are in line with Järveläinen et al. (2004) that central–parietal beta band oscillations may be motor-related, and may be sensitive to the motor-related characteristic of the perceived gesture type such as transitivity or concreteness: namely, more transitive and concrete gestures, e.g., tool-use gestures, are more likely to elicit beta power decrease. They are also supportive to the assumption that action observation leads to the activation of motor representations, and possibly elicits action simulation processes, which underlie the premotor and primary motor cortices (Hari et al. 1998).

Interestingly, despite the reported relevance of the alpha band to action observation and execution (Avanzini et al. 2012; Moreno et al. 2015), we did not observe any oscillatory effects in the alpha band when comparing TU versus EM gestures. This might be an indication that the alpha and the beta bands contribute to differential aspects of motor action (Salmelin et al. 1995). It has been suggested that alpha band oscillations may be more related to sensory processing whereas the beta band is more specific for motor processing (Brinkman et al. 2014; Salmelin et al. 1995): crucially, as we only compared gesture types within modalities here, only beta band effects were observed. However, in another study focusing at a different aspect of speech and gesture processing, when we compared between modalities, clear alpha band effect were observed together with the beta band effect (He et al. 2015). This may be considered as putative evidence for the functional dissociation between the alpha and beta oscillations.

What is novel and most interesting in the current experiment is the observed interaction between gesture-type and language-context in the beta band: we observed beta power decrease for tool-use gestures in the Russian language contexts which was significantly reduced in the German language contexts. In this case, one might consider that comprehending sentences with comparable semantic content may interact with action observation. There can be several explanations regarding the underlying mechanisms which result in this interaction. Firstly, one might speculate that the participants predominantly resorted to the speech-input in the German condition to fulfill the requirement of the semantic task, and ignored the visual content delivered via gestures. If this is the case, it is then not surprising why we observed no oscillatory effects for the motor simulation difference between the observations of tool-use versus emblematic gestures. However, in a comparable study, while we looked at the modality difference between co-speech conditions with either German or Russian sentences versus the speech-only control condition, we observed a comparable alpha and beta power decrease for the both co-speech conditions (He et al. 2015). This may be an indication that while looking at co-speech gestures with comprehensible German speech, participants attended to both gesture and speech to fulfill the semantic task. This may also be the case in the current experiment.

Another viable explanation may be derived from the well-established action–language interaction literature embedded under the theoretical framework of embodied cognition (Barsalou 2008; Pulvermuller 2005): theoretical claims by embodied cognition theories suggest that higher-level cognitive functions such as language are grounded by sensorimotor activities, and that language and action are processed in an interactive manner (Buccino et al. 2005; Hauk et al. 2004). Empirically, it has been reported that action words and sentences can interfere with action execution (Buccino et al. 2005; Sato et al. 2008). For example, Buccino et al. (2005) used single-pulse TMS to simulate the hand/foot area when participants were listening to action-related sentences. The authors record reaction times and MEP from hand/foot muscles, and found that MEPs recorded from hand/foot muscles were modulated by listening to hand/foot-related actions sentences. More specifically, there is evidence showing that the beta oscillations revealed by action preparation/execution can be modulated by comprehending action languages: in a study from Klepp and colleagues (2015), participants were asked to perform a semantic decision task, by pressing the button, on single action verbs describing actions executed with the hands or the feet. Behaviorally, the reaction times were longer for the hand versus foot verbs; and MEG results showed that the beta power decrease, as elicited by the preparation of the button-press response, was observable only for foot verbs but diminished for the hand verbs. This finding clearly suggests a potential semantic interference from language processing to action preparation, which occurs when the verb's effector matches the response. Similarly, in the current study, our observation of lack of oscillatory effects in the German condition may be considered as motor-related beta oscillations subject to interference effect from language.

Of note, previous studies using language–motor interaction paradigms have shown decreased cortical excitability (and thus diminished oscillatory effects) in the motor cortex for action preparation/execution (Boulenger et al. 2008; Buccino et al. 2005; Klepp et al. 2015), and they seem to suggest a competition for shared resources in the motor areas between action language processing and action preparation/execution. However, it has to be noted that language–motor interaction is not always inhibitive. Cases were also commonly reported with regard to the facilitation effects of language on action preparation/perception (Boulenger et al. 2006; Pulvermüller et al. 2005; Willems et al. 2011). In the current experiment, even if we observed diminished beta band effects for the German condition, our behavioral results, on the contrary, clearly suggest that the comprehension of visual–auditory stimuli was in fact facilitated in the German condition when compared with the Russian condition: besides the general processing advantage for the TU gestures (as reflected in the higher accuracy rates), the response time advantage for TU was only observable in the German but not the Russian condition. Consequently, if we attribute the beta effects to the differential level of motor simulation during action observation for TU versus EM, it is then more likely that, having a complementary auditory input channel may facilitate action comprehension, so that motor simulation is less necessary. Crucially, this process is not a direct product of reduced access to the gestures: as we have discussed previously, clear EEG evidence suggesting gesture observation was shown for both the German and Russian conditions when compared against a speech-only control condition (He et al. 2015). Our results seem to indicate that, even if the processing of action-related sentences may recruit the motor regions encoding the corresponding actions, and thus may interact with action execution and observation, different tasks, stimulus sets, and timing may lead to either interference or facilitation effects (Chersi et al. 2010; Klepp et al. 2015). They may be also indicative of distinct language–motor interaction patterns between action execution versus observation, even if both action execution and observation may be tied with comparable brain networks within the pre/motor cortices. For action observation, unlike action execution, its interaction with language may be more facilitative in nature.

This argument also echoes recent theoretical framework, as well as neural evidence suggesting the facilitative interaction between hand-action/gesture and speech comprehension/production (Biau and Soto-Faraco 2013; Holle et al. 2012; Kita and Özyürek 2003; Skipper et al. 2009; Wang and Chu 2013). Essentially, as both gesture and speech can be considered as two sides of a unitary cognitive process (Kendon 1980), gesture thus packages spatial-motoric information which may or may not be present in speech (Kita and Özyürek 2003). This information can thus facilitate speech production and comprehension (Holle et al. 2012; Krahmer and Swerts 2007; Skipper et al. 2009). Vice versa, as speech entails semantic information that partially overlaps with hand actions and gestures, one might as well expect facilitation from speech to gesture observation. Our findings may thus be initial evidence for this facilitation. Clearly, further experiments are expected, to systematically examine which aspects of speech (related/unrelated/overlapping semantic content, matched/mismatched prosody) facilitates the observation of gestures, and what types of gestures can benefit from having an additional speech channel during observation. Nevertheless, our results, together with the line of research concerning the language–motor/gesture–speech interaction, not only corroborate that gesture and speech form the two aspects of a unitary communicative process, but also lend further support to the functional role of the beta band oscillations for both action and language within the framework of embodied cognition (Barsalou 2008; Pulvermüller et al. 2005). Based on this theoretical framework, more fine-grained mechanisms are expected from future studies with regard to the specified dynamics between language and action, especially in the case of action observation.

Conclusion

Taken together, our EEG results showed that the beta oscillations were sensitive to the observation of tool-use gestures. More importantly, the beta band effects were specific to the condition without semantic input from speech. This sensitivity of the beta band oscillations to the language context may be considered as evidence of facilitation from language to action observation, as indicated by the behavioral results. Our results corroborate the functional sensitivity of the beta oscillations to gesture types differing in the degree of transitivity, and provide novel evidence concerning language–motor and gesture–speech interaction during action/gesture observation.

Acknowledgements This research project is supported by a grant from the 'Von Behring-Röntgen-Stiftung' (Project No. 59-0002; 64-0001) and by the 'Deutsche Forschungsgemeinschaft' (Project No. DFG: STR 1146/9-1 and SFB/TRR135 project A3). M.S. is supported by the DFG (Project No. STR 1146/4-1). B.S. is supported by the DFG (Project No. STR 1146/8-1). Part of the work of this paper has been presented at the Tilburg Gesture Research meeting (TiGER 2013).

References

Agnew ZK, Wise RJ, Leech R (2012) Dissociating object directed and non-object directed action in the human mirror system; implications for theories of motor simulation. PLoS ONE 7:e32517

Andric M, Solodkin A, Buccino G, Goldin-Meadow S, Rizzolatti G, Small SL (2013) Brain function overlaps when people observe emblems, speech, and grasping. Neuropsychologia 51:1619–1629

Arnstein D, Cui F, Keysers C, Maurits NM, Gazzola V (2011) μ-suppression during action observation and execution correlates with BOLD in dorsal premotor, inferior parietal, and SI cortices. J Neurosci 31:14243–14249

Avanzini P, Fabbri-Destro M, Volta D, Daprati R, Rizzolatti E, Cantalupo G, G (2012) The dynamics of sensorimotor cortical oscillations during the observation of hand movements: an EEG study. PLoS ONE 7:e37534

Barsalou LW (2008) Grounded cognition. Annu Rev Psychol 59:617–645

Biau E, Soto-Faraco S (2013) Beat gestures modulate auditory integration in speech perception. Brain Lang 124:143–152

Boulenger V, Roy AC, Paulignan Y, Deprez V, Jeannerod M, Nazir TA (2006) Cross-talk between language processes and overt motor behavior in the first 200 msec of processing. J Cogn Neurosci 18:1607–1615

Boulenger V, Silber BY, Roy AC, Paulignan Y, Jeannerod M, Nazir TA (2008) Subliminal display of action words interferes with motor planning: a combined EEG and kinematic study. J Physiol (Paris) 102:130–136

Brinkman L, Stolk A, Dijkerman HC, de Lange FP, Toni I (2014) Distinct roles for alpha-and beta-band oscillations during mental simulation of goal-directed actions. J Neurosci 34:14783–14792

Buccino G, Binkofski F, Fink GR, Fadiga L, Fogassi L, Gallese V, Seitz RJ, Zilles K, Rizzolatti G, Freund HJ (2001) Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. Eur J Neurosci 13:400–404

Buccino G, Riggio L, Melli G, Binkofski F, Gallese V, Rizzolatti G (2005) Listening to action-related sentences modulates the activity of the motor system: a combined TMS and behavioral study. Cogn Brain Res 24:355–363

Chersi F, Thill S, Ziemke T, Borghi AM (2010) Sentence processing: linking language to motor chains. Frontiers in neurorobotics 4:4

Gallese V, Fadiga L, Fogassi L, Rizzolatti G (1996) Action recognition in the premotor cortex. Brain 119:593–609

Goldin-Meadow S (2005) Hearing gesture: how our hands help us think. Harvard University Press, Cambridge

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. Hum Brain Mapp 30:3309–3324

Groppe DM, Urbach TP, Kutas M (2011) Mass univariate analysis of event-related brain potentials/fields I: a critical tutorial review. Psychophysiology 48:1711–1725

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. Proc Natl Acad Sci 95:15061–15065

Hauk O, Johnsrude I, Pulvermüller F (2004) Somatotopic representation of action words in human motor and premotor cortex. Neuron 41:301–307

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

Higuchi S, Imamizu H, Kawato M (2007) Cerebellar activity evoked by common tool-use execution and imagery tasks: an fMRI study. Cortex 43:350–358

Holle H, Obermeier C, Schmidt-Kassow M, Friederici AD, Ward J, Gunter TC (2012) Gesture facilitates the syntactic analysis of speech. Front Psychol 3:74

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

Johnson-Frey SH (2004) The neural bases of complex tool use in humans. Trends Cogn Sci 8:71–78

Johnson-Frey SH, Newman-Norlund R, Grafton ST (2005) A distributed left hemisphere network active during planning of everyday tool use skills. Cereb Cortex 15:681–695

Kendon A (1980) Gesticulation and speech: Two aspects of the process of utterance. Relat Verbal Nonverbal Commun 25:207–227

Kircher T, Straube B, Leube D, Weis S, Sachs O, Willmes K, Konrad K, Green A (2009) Neural interaction of speech and gesture: differential activations of metaphoric co-verbal gestures. Neuropsychologia 47:169–179

Kita S, Özyürek A (2003) What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. J Mem Lang 48:16–32

Klepp A, Niccolai V, Buccino G, Schnitzler A, Biermann-Ruben K (2015) Language–motor interference reflected in MEG beta oscillations. Neuroimage 109:438–448

Krahmer E, Swerts M (2007) The effects of visual beats on prosodic prominence: acoustic analyses, auditory perception and visual perception. J Mem Lang 57:396–414

Lindenberg R, Uhlig M, Scherfeld D, Schlaug G, Seitz RJ (2012) Communication with emblematic gestures: shared and distinct neural correlates of expression and reception. Hum Brain Mapp 33:812–823

Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. J Neurosci Methods 164:177–190

McNeill D (1992) Hand and mind: What gestures reveal about thought. University of Chicago Press, Chicago

McNeill D (2006) Gesture: a psycholinguistic approach. In: Brown K (ed) The encyclopedia of language linguistics. Elsevier, Boston, pp 58–66

McNeill D (2008) Gesture and thought. University of Chicago Press, Chicago

Moreno I, De Vega M, León I (2013) Understanding action language modulates oscillatory mu and beta rhythms in the same way as observing actions. Brain Cogn 82:236–242

Moreno I, De Vega M, León I, Bastiaansen M, Lewis AG, Magyari L (2015) Brain dynamics in the comprehension of action-related language. A time-frequency analysis of mu rhythms. Neuroimage 109:50–62

Muthukumaraswamy SD, Johnson BW, McNair NA (2004) Mu rhythm modulation during observation of an object-directed grasp. Cogn Brain Res 19:195–201

Newman-Norlund R, van Schie HT, van Hoek ME, Cuijpers RH, Bekkering H (2010) The role of inferior frontal and parietal areas in differentiating meaningful and meaningless object-directed actions. Brain Res 1315:63–74

Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9:97–113

Perry A, Bentin S (2009) Mirror activity in the human brain while observing hand movements: a comparison between EEG desynchronization in the μ-range and previous fMRI results. Brain Res 1282:126–132

Pulvermuller F (2005) Brain mechanisms linking language and action. Nat Rev Neurosci 6:576–582

Pulvermüller F, Hauk O, Nikulin VV, Ilmoniemi RJ (2005) Functional links between motor and language systems. Eur J Neurosci 21:793–797

Quandt LC, Marshall PJ, Shipley TF, Beilock SL, 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 50:2745–2751

Rizzolatti G, Fogassi L, Gallese V (2001) Neurophysiological mechanisms underlying the understanding and imitation of action. Nat Rev Neurosci 2:661–670

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

Sato M, Mengarelli M, Riggio L, Gallese V, Buccino G (2008) Task related modulation of the motor system during language processing. Brain Lang 105:83–90

Schaller F, Weiss S, Müller HM (2017) EEG beta-power changes reflect motor involvement in abstract action language processing. Brain Lang 168:95–105

Scorolli C, Borghi AM (2007) Sentence comprehension and action: Effector specific modulation of the motor system. Brain Res 1130:119–124

Skipper JI, Goldin-Meadow S, Nusbaum HC, Small SL (2009) Gestures orchestrate brain networks for language understanding. Curr Biol 19:661–667

Straube B, Green A, Bromberger B, Kircher T (2011) The differentiation of iconic and metaphoric gestures: Common and unique integration processes. Hum Brain Mapp 32:520–533

Straube B, Green A, Weis S, Kircher T (2012) A supramodal neural network for speech and gesture semantics: an fMRI study. PLoS ONE 7:e51207

Villarreal M, Fridman EA, Amengual A, Falasco G, Gerscovich ER, Ulloa ER, Leiguarda RC (2008) The neural substrate of gesture recognition. Neuropsychologia 46:2371–2382

Wang L, Chu M (2013) The role of beat gesture and pitch accent in semantic processing: an ERP study. Neuropsychologia 51:2847–2855

Willems RM, Hagoort P (2007) Neural evidence for the interplay between language, gesture, and action: a review. Brain Lang 101:278–289

Willems RM, Labruna L, D'Esposito M, Ivry R, Casasanto D (2011) A Functional role for the motor system in language understanding evidence from theta-burst transcranial magnetic stimulation. Psychol Sci 22:849–854

Wriessnegger S, Leeb R, Kaiser V, Neuper C, Müller-Putz G (2013) Watching object related movements modulates mirror-like activity in parietal brain regions. Clin Neurophysiol 124:1596–1604

Xu J, Gannon PJ, Emmorey K, Smith JF, Braun AR (2009) Symbolic gestures and spoken language are processed by a common neural system. Proc Natl Acad Sci 106:20664–20669