Motor resonance during linguistic processing as shown by EEG in a naturalistic VR environment

Ana Zappa, Deirdre Bolger, Jean-Marie Pergandi, Pierre Mallet, Anne-Sophie Dubarry, Daniel Mestre, Cheryl Frenck-Mestre

HAL Id: hal-02309994

https://hal.science/hal-02309994

Submitted on 11 Oct 2019

Motor resonance during linguistic processing as shown by EEG in a naturalistic VR environment

Ana Zappa [a, e], Deirdre Bolger [a, e, f], Jean-Marie Pergandi [a, b, c], Pierre Mallet [a, b, c], Anne-Sophie Dubarry [a, d, e], Daniel Mestre [a, b, c, d, f], Cheryl Frenck-Mestre [a, d, e, f]

[a] *Aix-Marseille Univ*
[b] *Institute of Movement Sciences*
[c] *Mediterranean Virtual Reality Center, Institute of Language, Communication and the Brain, France*
[d] *Centre National de Recherche Scientifique*
[e] *Laboratoire Parole et Langage, Institute of Language, Communication and the Brain, France*
[f] *Brain & Language Research Institute, Institute of Language, Communication and the Brain, France*

ARTICLE INFO

ABSTRACT

Embodied cognition studies have shown motor resonance during action language processing, indicating that linguistic representations are at least partially multimodal. However, constraints of this activation linked to linguistic and extra-linguistic context, function and timing have not yet been fully explored. Importantly, embodied cognition binds social and physical contexts to cognition, suggesting that more ecologically valid contexts will yield more valid measures of cognitive processing. Herein, we measured cortical motor activation during language processing in a fully immersive Cave automatic virtual environment (CAVE). EEG was recorded while participants engaged in a Go/No-Go task. They heard action verbs and, for Go trials, performed a corresponding action on a virtual object. ERSP (event-related spectral perturbation) was calculated during verb processing, corresponding to the pattern of power suppression (event-related desynchronization – ERD) and enhancement (event-related synchronization – ERS) relative to the reference interval. Significant ERD emerged during verb processing in both the μ (8–13 Hz) and beta band (20–30 Hz) for both Go and No-Go trials. μ ERD emerged in the 400–500 msec time window, associated with lexical-semantic processing. Greater μ ERD emerged for Go compared to No-Go trials. The present results provide compelling evidence in a naturalistic setting of how motor and linguistic processes interact.

1. Introduction

A currently debated topic in cognitive psychology is the involvement of motor processes in language processing. Studies that approach language from an embodied cognition perspective have produced evidence that language comprehension involves perceptual and motor systems, indicating that linguistic representations are either partially or completely multimodal (Barsalou, 1999, 2008; Wilson & Golonka, 2013). Both neuroimaging and behavioral results have pointed to the recruitment of sensorimotor systems during semantic access and the overlap of these processes has often been interpreted as evidence that one performs mental simulations of situations to understand language (Barsalou, 1999; Glenberg & Kaschak, 2002; Pulvermüller, 2005). However, classical models of language comprehension posit that language representation is amodal and independent of perceptual and motor systems, suggesting that these motor activations are post-lexical and do not play a causal role in language processing (Fodor, 1980, 1987;

Mahon & Caramazza, 2008). Moreover, recent studies have shown that motor activation during language processing can be modified or neutralized by changes in linguistic context, task and timing (Aravena et al., 2012; Boulenger et al., 2006; Sato, Mengarelli, Riggio, Gallese, & Buccino, 2008). Hence, the implications of the involvement of motor processes in language comprehension are still not fully understood. Importantly, studies focusing on motor and linguistic interactions have yet to use set-ups that take into account the multimodality of language, to provide a closer-to-life experience under which to observe the bilateral influence between action and cognition (Peeters, 2019). The present study sought to fill this gap by placing participants in a three-dimensional virtual environment in which they manipulated virtual objects in response to linguistic cues. This allowed us to examine the hypothesis that motor representations are part and parcel of the linguistic representation of action verbs (Aravena et al., 2012; Pulvermüller, 2005) in a realistic environment, as opposed to the impoverished conditions that are most often used. The recording of EEG allowed us to clearly examine the neurological signature of motor planning (Aravena

et al., 2010; Funderud et al., 2012) as well as any evidence of motor activation during linguistic processing.

1.1. Behavioral evidence of motor involvement in linguistic processing

One way to observe the influence of motor processes on language comprehension is to manipulate motor activation through movement or motor planning and observe its influence on lexico-semantic processing. Behavioral studies using an Action-Sentence compatibility effect paradigm (ACE) have shown that compatibility between action language and the movement needed to produce a manual response can either hinder or facilitate response times depending upon timing. In semantic decision tasks, a negative ACE is found, i.e. the inhibition of response times for a compatible movement, within the first 400 msec during or after single action verb processing (Sato et al., 2008; Spadacenta, Gallese, Fragola, & Mirabella, 2014). When sentence comprehension is required, a negative ACE occurs within 700 msec post onset of the critical verb (García & Ibáñez, 2016, but see Repetto, Cipresso, & Riva, 2015). However, if response planning occurs at sentence onset *or* if movement onset is delayed to after the single verb or action sentence is fully processed, a positive ACE is obtained (Boulenger et al., 2006; de Vega, Moreno, & Castillo, 2013; Diefenbach, Rieger, Massen, & Prinz, 2013; Kaschak & Borreggine, 2008; for a review see García & Ibáñez, 2016). These results suggest that interference occurs when motor and linguistic processes overlap temporally due to competition for shared neural resources. Facilitation, on the other hand, seems to be due to a priming effect, upstream of competition but nonetheless indicating shared neural resources. Given that most ACE studies reveal effects late in sentence comprehension (700–2200 msec post stimulus), they do not rule out the possibility that the ACE effect is caused by post-lexical motor imagery (Toni, de Lange, Noordzij, & Hagoort, 2008).

1.2. Electrophysiological evidence

To overcome the limitations of behavioral measures as concerns the onset of motor influences on linguistic processing, several studies have adopted the recording of electrophysiological responses. Aravena and colleagues focused on the precise timing of motor-semantic effects using EEG to examine early motor-related ERPs (Readiness Potential (RP)) as well as later, linguistically related ERPs (N400) in an ACE paradigm (Aravena et al., 2010). Participants listened to action sentences while performing congruent and incongruent actions. Congruent actions led to an increase in the magnitude of the RP, revealing an effect of linguistic processing on motor preparation. In addition, in comparison to congruent trials incongruent actions elicited a greater N400-like response, thus revealing the interference of motor execution in semantic processing. This pattern of results was interpreted as illustrating a robust ACE which, crucially, could not be attributed to post-lexical effects but supports the hypothesis of early interactions between sensorimotor and semantic processing. In the current study, we built upon these results first, by using a virtual environment in which participants performed movements actually related to the action verbs as opposed to a rather impoverished environment or only imagined movement (cf. Peeters, 2019). Second, we examined motor activation during language processing as revealed by time-frequency analyses as opposed to ERP components alone. As outlined below, while several studies have used time-frequency analyses to quantify motor cortex activity proper, this approach has been applied to the study of embodied cognition less frequently (cf. Fargier et al., 2012, for a discussion).

Numerous studies have used EEG to quantify motor cortex activity by measuring oscillatory activity via time-frequency decomposition of the EEG signal to examine the pattern of cortical response to motor planning. More recently, post-stimulus spectral estimation methods

have been applied to study the pattern of event-related desynchronization (ERD) and synchronization (ERS), which corresponds to power suppression and enhancement respectively, in the time interval following stimulus onset compared to a pre-stimulus baseline. One method, event-related spectral perturbation (ERSP) (Grandchamp & Delorme, 2011; Makeig, 1993; Makeig, Debener, Onton, & Delorme, 2004), which we applied in the present study, groups ERD and ERS and calculates the power spectrum on a trial-by-trial basis. This approach can be applied to study the cortical response to motor events (for a review see Hobson & Bishop, 2016). Specifically, mu, or μ (8–13 Hz), and beta (13–30 Hz) rhythms are synchronized patterns of electrical activity recorded over the sensorimotor cortex whose suppression is associated with performing and observing movement (Caetano, Jousmäki, & Hari, 2007; Koelewijn, van Schie, Bekkering, Oostenveld, & Jensen, 2008; Pfurtscheller & Lopes da Silva, 1999; Pineda, 2005). A decrease in the μ rhythm has also been linked to motor imagery (Matsumoto et al., 2010).

In addition to movement proper, various recent studies have shown μ rhythm ERD as a function of action language processing. Language studies using time-frequency analysis have found μ and beta ERD for action-related sentences, directly showing motor resonance *during* the retrieval of lexical-semantic information as opposed to post-lexical retrieval of kinematic imagery (van Elk, van Schie, Zwaan, & Bekkering, 2010). Reading single verbs related to the body caused soma topical μ ERD (von Nicolai et al., 2014). Moreover, greater μ rhythm suppression has been observed for action language compared to abstract language (Alemanno et al., 2012; Moreno et al., 2015). To examine how the acquisition of lexical items might be directly influenced by the motor system, Fargier et al. (2012) conducted a learning paradigm using EEG. Participants learned novel words, either in association with motor actions or with abstract animated images. After two training sessions on a first day of learning, participants showed greater μ suppression while processing words learned in the motor action condition compared to the control condition. Nonetheless, based on the distribution of their effects across 2 days of training, Fargier et al. (2012) argued that the cortical regions conjointly activated by motor and linguistic processing were confined to convergence areas (i.e. more frontal regions as opposed to the central parietal areas assumed to subserve sensorimotor activity). Moreover, although these studies have all used time-locked μ suppression as a marker of motor neuron activity, it is important to keep in mind that the μ frequency band (8–13 Hz) overlaps with the alpha frequency band (8–12 Hz), which is reflective of attentional fluctuation, and it has been claimed that the two are often confounded (Hobson & Bishop, 2016). On the other hand, several researchers working with μ band oscillations argue that they reflect neural activity in the motor and premotor cortex and can be measured in fronto-central sites (but see Fargier et al., 2012) versus occipital sites for the alpha band (Moreno et al., 2015). Therefore, a consensus in terms of how to distinguish between these two has not been reached.

1.3. Virtual reality as a tool to study embodiment

A caveat of investigating motor activation during language processing, especially when working within an embodied cognition framework, is linked to the physical and environmental limitations imposed by neurolinguistic study protocols. Laboratory experimental tasks are generally performed in isolated and decontextualized environments, due to the need to control variables that could influence participants' responses (Peeters, 2019). These experiments often use single words or sentences presented in isolation, along with, if any, simplistic visual information on a computer screen. This very decontextualization could in turn affect how language is processed. Indeed, real-world language processing generally occurs in much richer environments and, importantly, interlocutors, social context and physical cues have a strong influence

on how language is understood (Knoeferle, 2015). In the last two decades, the discrepancy between real-life language processing and that which takes place in an experimental environment has been brought to light in the context of embodied cognition (Tromp, Peeters, Meyer, & Hagoort, 2018), according to which our bodily states and actions are heavily implicated in how we communicate and process information (Atkinson, 2010). When we communicate, speech and gesture systems interact to convey and comprehend meaning; the two systems have been suggested to comprise an integrated system (Goldin-Meadow, 2011; Graziano & Gullberg, 2018). Therefore, the more real-world and situated the language processing environment is, the more physically implicated and natural participants will feel and more applicable the results obtained will be to real-life processing (Peeters, 2019). Modern technological advances render possible the study of cognitive processes in their actual contexts (Ladouce, Donaldson, Dudchenko, & Ietswaart, 2017). Notably, in the present study we capitalized on a CAVE Automatic Virtual Environment (CAVE) system to examine the interaction of language and motor processes during the lexical access of auditory verbs, in the aim of testing the hypothesis that motor representations are part and parcel of the linguistic representations of action verbs (Pulvermüller, 1999). Several studies of language acquisition in adults have demonstrated clear benefits of a virtual environment (Legault et al., 2019; Repetto, Colombo, & Riva, 2015). As concerns the interaction of linguistic and motor processing, Repetto, Cipresso et al. (2015) found that simulating actions in a virtual environment facilitated the semantic processing of action verbs that involved the same effector (i.e. simulated running in a virtual park facilitated processing of verbs entailing movement of the foot). Simply viewing the motion did not influence semantic processing. Hence, virtual motion can elicit stimulation of the motor system, which can in turn affect linguistic processing of overlapping information.

Virtual reality (VR) experimental paradigms have indeed gained popularity as they offer a more ecological and yet controlled environment in which to test a wide variety of phenomena including language processing (Peeters, 2019; Repetto, Cipresso et al., 2015; Tromp et al., 2018). VR paradigms consist of 3-D environments that provide participants with visual and auditory stimuli, while allowing them to interact and receive real-time feedback from their actions via a graphic rendering system. Participants' movements are often tracked and recorded using input tools (trackers, gloves, a mouse or joystick) (Burdea & Coiffet, 2003). Three basic types of VR environments exist: a computer monitor, a head-mounted display (HMD) and the CAVE Automatic Virtual Environment (CAVE) system. Computer monitors are considered non-immersive due to the small percentage of the participant's visual field that they occupy (Repetto, 2014). HMDs provide an immersive experience. One interesting aspect of HMDs is that, contrary to a "classical" 3-sided CAVE, they visually isolate the participant from the real world. However, the downside is that participants no longer see their own bodies. This commonly leads to a sensation of self-floatation. Indeed, we carried out experiments comparing CAVE and HMD in a simple spatial task, i.e. walking through an aperture (Mestre, Louison, & Ferlay, 2016). We found that, with an HMD, participants were not correctly calibrated, in spatial terms, which resulted in many occurrences of collisions with the virtual environment. Adding an avatar of the self in the HMD view resulted in significantly fewer collisions. Collisions with the virtual environment did not occur in the CAVE, where the participants' own body was always present in the visual field (Lepecq, Bringoux, Pergandi, Coyle, & Mestre, 2009). This type of result suggests that participants need a colocalized representation of their own bodies to achieve precise spatial behavior, while wearing an HMD. There are many problems associated with this requirement, such as the need for precise biomechanical modeling and a realistic colocalized avatar.

The results from the above cited studies drove our choice to use a CAVE, in which participants naturally see their own bodies. The CAVE provides elements that are crucial to VR effectiveness, i.e. the sense of immersion and presence (Moore, Wiederhold, Wiederhold, & Riva, 2002). The sense of immersion, defined as "a sensorimotor coupling between a participant and a virtual environment" and presence, defined as "a psychological, attentional and cognitive state in which the participant, immersed within a virtual environment, behaves in accordance with the affordances provided by this environment […]", are contingent on ecological validity and can lead to real-life behavior (Mestre, 2015, p.1). The sense of presence in the CAVE stems from the fact that not only are participants immersed visually and auditorily, but they can see their own bodies. Visually, they are surrounded by virtual images projected onto 3 or 4 screens (the floor and surrounding walls), providing a sensory illusion that creates a credible environment (for a review see Bohil, Alicea, & Biocca, 2011). In the CAVE, participants experience the sense of "agency" that arises from being able to gesture and move their arms and hands freely (Johnson-Glenberg, 2018), allowing them to perform more naturalistic and interactive tasks. The sensorimotor system is therefore much more fully engaged than in traditional experiments and elicited responses are closer to what probably occurs in real life (Bohil et al., 2011). Finally, the engaging aspect of this rich environment can also act as a motivational tool (Bayliss & Ballard, 2000).

For all of the above stated reasons, virtual reality is an attractive methodology to pair with EEG to study the interaction of motor and linguistic processing. Researchers have full control over multimodal sensory stimulation, making it possible to directly observe brain activity that correlates with specific types of sensory input, whether visual or auditory, in a more ecologically valid environment where naturalistic actions can be planned and performed. The intrinsic multimodality of human communication makes virtual reality paradigms particularly beneficial when studying language processing. In a recent review article of studies that used virtual reality in psycholinguistic research, Peeters (2019) claimed that what is most promising about virtual reality as an experimental tool is that it will "shift theoretical focus towards the interplay between different modalities in dynamic and communicative real-world environments, moving beyond and complementing studies that focus on one modality in isolation" (Peeters, 2019, p.6). Combining EEG with CAVE simulation of movement during linguistic processing can provide a novel and compelling view into how motor and linguistic systems may interact, which was the aim of the present work.

1.4. The current study

In the present study, we measured participants' cortical activity while they listened to auditory action verbs and subsequently manipulated virtual objects or not in a CAVE. The main aim of this study was to provide evidence of μ ERSP during action language processing, in an ecological environment. In an effort to take into account theories of embodied cognition that argue that cognition is strongly constrained by one's surrounding environment and physical state (Atkinson, 2010), we chose to use an ecologically realistic environment that required real and varied action. We used a CAVE which, compared to a real-word setting, allows for the controlled and synchronized presentation of stimuli with EEG, to the same degree as computerized experiments. This paradigm has the benefit of being more ecologically valid than traditional computerized set-ups. However, as outlined in a previous case report, recording EEG in a VR environment presents a particular technical challenge, due to both possible crosstalk between systems and participants' movement (Török et al., 2014). Combining EEG recording and virtual stimulation also requires a precise synchronization process (Repetto, Cipresso et al., 2015). We were thus interested in providing a proof of concept, in addition to testing specific hypotheses about the role of motor activation during linguistic retrieval. To exam-

ine how motor processes may affect early linguistic processing, we explored the neural activity in the sensorimotor cortex during the auditory processing of verbs, prior to actual movement, via time-frequency analyses. We also examined ERP language related components. As concerns the overlap of linguistic and motor processes, we were specifically interested in ERD in the μ frequency bands (8–13 Hz) during the auditory processing of the verb. To address recent observations that μ ERD can be confounded with alpha ERD and that beta ERD provides a way to ensure that what is being observed is motor activation as opposed to alpha, we focused on both μ and beta ERD to show motor resonance (Hobson & Bishop, 2016). We hypothesized that we would observe μ ERD during verbal processing (Moreno et al., 2015); the distribution of the μ effect, whether central-parietal or more frontally located, as well as simultaneous beta ERD, should inform us of its nature (Hobson & Bishop, 2016).

Extant literature has revealed that activity in the 8–13 Hz frequency range is not a unitary measure (Klimesch, Doppelmayr, Pachinger, & Russegger, 1997) and can be divided into a lower range (8–10 Hz) and an upper range (11–13 Hz) and, in each range, ERD differs both in terms of its spatial distribution and in the processes thought to underlie it. ERD in the 8–10 Hz frequency band has been found to have a wide spatial distribution and to reflect processes related to attention and general task demands, which may thus be more reflective of alpha than of μ. In contrast, ERD in the 11–13 Hz frequency band has been revealed as being more topologically restricted and related to specific cognitive tasks, most notably processes related to semantic or long-term memory processes (Klimesch et al., 1997; Neuper & Pfurtscheller, 2001). We also examined the beta-band (13–30 Hz), which we subdivided into the following sub-bands: beta1 (13–18 Hz), beta2 (19–25 Hz) and beta3 (25–30 Hz). Previous research has revealed greater beta-band power suppression in response to action verbs compared to non-action verbs in the lower beta band, 13–25 Hz (Weiss, Berghoff, Rappelsberger, & Müller, 2001), such that sub-dividing the beta-band may allow us to disentangle beta activity related to language processing and that linked to motor activity.

We used a Go No-Go design in which participants either enacted the auditory verb upon subsequently presented virtual objects or not. This allowed us to examine sensorimotor activity during verbal processing (prior to movement) for both types of trials and to directly compare sensorimotor activity during verbal processing as a function of trial type. While we did not have a strong hypothesis concerning the effect of trial type, previous behavioral work has shown that manual responses are inhibited, on Go trials, if the go signal is presented simultaneously with a verbal stimulus denoting a hand movement (Sato et al., 2008). It is thus possible that greater μ ERD would be found in the present study for Go than No-Go trials if indeed motor preparation was inhibited by the semantic processing of the action verbs. The design also allowed us to determine whether variation in the μ ERD as a function of the type of trial (Go vs. No-Go) would be accompanied by a modification of ERP components, notably the contingent negative variation (CNV). Indeed there is debate concerning the direct coupling between these two responses (cf. Funderud et al., 2012; Filipović, Jahanshahi, & Rothwell, 2001; Zaepffel, Trachel, Kilavik, & Brochier, 2013).

2. Method

2.1. Participants

Twenty right-handed French native speakers (10 women, aged 20–26) participated in the study. Participants were volunteers from the student population of the Aix-Marseille Université, enrolled in the Science and Technique of Physical Sports Activities (STAPS) department. They had no history of neurological insult and received course credit in exchange for their participation. None had taken part in any prior VR experiment nor were they informed of the purpose of the experiment prior to the debriefing at the end of the session. All participants gave their written informed consent prior to the experiment, in keeping with the 1964 Helsinki Declaration, and the study was approved by the local ethics committee at Aix-Marseille Université.

2.2. Stimuli

Auditory stimuli consisted of 16 transitive French verbs (average number of phonemes = 5, ±1) denoting actions that can be performed using one's hand and arm ("attraper" [catch], "cacher" [hide], "coucher" [lay down], "déplacer" [move], "empiler" [stack], "frotter" [rub], "lâcher" [let go of], "lancer" [throw], "faire pivoter" [pivot or twist], "pousser" [push], "relever" [make stand up], "secouer" [shake], "soulever" [pick up], "tapoter" [tap], "tirer" [pull], "faire tomber" [drop]. The auditory stimuli were produced by a trained female speaker and digitally recorded at 48 kHz (32-bit float) in a professional sound booth in a single session. They were subsequently spliced into individual tracks (audacity software) and the duration of each auditory verb was determined. The verbs ranged in frequency per million from 1.16 to 415, with half being low frequency (average frequency = 17, ±16) and the other half high frequency (average frequency = 211, ±141). The choice of verbs was dictated both by their discriminability as concerns movement parameters and by the feasibility of tracking these movements with the finger-tracking glove. Visual stimuli consisted of 8 virtual geometric shapes (sphere, cube, cone, cylinder, rectangular prism, triangular prism, hexagonal prism, triangular pyramid). They were selected such that they did not provide affordances in relation to the set of verbs. The 16 auditory verbs were each presented 4 times (twice for each type of trial) in one of three pseudorandom orders. Each of the 8 objects was presented 16 times, equally often as a target and a distractor and across 10–12 verbs. The target object was color-coded green and the distractor was color-coded white.

2.3. Apparatus

2.3.1. The CAVE and the finger tracker

The Mediterranean Virtual Reality Center (CRVM) CAVE system consists of a cubic space measuring $3 \times 3 \times 3 \times 4$ m, with 3 vertical and 1 horizontal screen (floor). A graphic cluster of 4 video projectors delivers 4 stereoscopic projected images onto the 4 screens in real time (60 Hz). A movement capture system consisting of 8 infra red cameras provides the graphic cluster with the position of reflective targets on the user or on an entry peripheral device, allowing for interactions with the virtual environment. Participants wore 3D glasses and a 3-digit finger tracker (thumb, index and middle finger) on their right hand; the 3D glasses allowed participants to see their hand (and entire body) throughout the experiment and the finger tracker allowed for motion capture online as well as for participants to manipulate objects. The finger tracker was calibrated for each participant at the outset of the experiment to ensure the capture of acceptable movements, as predefined for each verb using UNITY. The apparatus used in this study is presented in Fig. 1.

Participants were visually surrounded by the virtual environment projected onto the four screens (the floor and 3 surrounding walls). The environment consisted of a virtual office containing a physical Plexiglas table on which the geometric objects to be manipulated were projected. Facing the participant was a large bookshelf containing objects typically found in offices such as books, filing boxes and framed pictures. To the left of the bookshelf was a virtual door and in the corner of the office was a virtual plant. Against the wall on the participant's left side was a chest of drawers bearing a vase and other office-type objects such as a filing folder. To the right of the participant was another chest of drawers.

Fig. 1. Participant in the CAVE, wearing a finger-tracker and 3D glasses while manipulating virtual objects.

2.3.2. Software

The UNITY software engine was used for stimulus presentation and a 64-channel Biosemi system (Actiview) was used for acquisition. The two systems were synchronized via a photodiode, which detected a change in luminosity (from black to white) of a square projected on the left bottom corner on the left vertical screen of the CAVE at the onset of each trial and again simultaneously to the onset of the virtual objects. The change in luminance was detected by the photodiode and the signal was sent to the acquisition system via one of the channels of the Biosemi AD system. The duration of the led signal was varied to distinguish the type of trial (Go vs. No-Go) and to indicate the side (left or right) of target object presentation.

2.3.3. EEG data acquisition

Electroencephalographic (EEG) activity was recorded continuously from 64 scalp electrodes located at left and right hemisphere positions over frontal, central, parietal, occipital, and temporal areas by means of a 64-channel electrode cap mounted with silver-chloride active electrodes (BioSemi Active Two system AD box). Individual electrodes were adjusted to a stable offset lower than 20 KΩ. Blinks and vertical eye movements (VEOG) were monitored via two external electrodes placed under each eye and horizontal eye movements. (HEOG) were monitored via two electrodes positioned at the outer canthus of both eyes. External electrodes were placed over both the left and right mastoids and the left mastoid served as reference during EEG acquisition. EEG was sampled online at a rate of 2048 Hz; a band-pass filter (0.16–100 Hz) was applied online for visualization purposes only.

2.4. Procedure

Participants sat comfortably behind a Plexiglas desk, wearing the finger-tracking glove and 3D glasses. The session began with a 10-minute training phase during which participants learned how to manipulate the virtual objects. For this, they learned to use different hand positions: pinch using the index finger and the thumb to manipulate smaller objects, C-shaped hand to manipulate larger objects, flat open hand to tap, pull or push objects.

During the experimental phase, stimuli were presented in two blocks of 32 trials. At the beginning of each trial, the participant sat with his/her right hand in resting position, on top of a small textured circle placed in the center of the Plexiglas desk. A trial was initiated only when the participant's hand was detected in this position. A trial sequence began with the presentation of a visual prompt, projected onto the Plexiglas table, signaling the type of trial (✓ = Go; ✗ = No-Go), 500 msec prior to and throughout the auditory presentation of the

verb (ex. "Tapote" [Tap]). Two and a half seconds after auditory verb onset, 2 different geometrical objects were projected onto the Plexiglas table: the target and distractor. On Go trials participants were told to perform the appropriate action on the target object. On No-Go trials they were instructed to simply listen to the verb. Post onset of the 2 objects, a constant 10-second period was allotted during which participants performed the action. A visual prompt was displayed for 2 s at the end of each trial during which participants were instructed to blink. The next trial was initiated when the participant's hand was detected in the resting position. If a movement was performed incorrectly on a Go trial or if the participant executed a Go during a No-Go trial, the trial was repeated at the end of the block. The experimental phase lasted roughly 20 min, with a short pause between the two blocks.

2.5. Data pre-processing

We used EEGLAB (Delorme & Makeig, 2004) to pre-process raw data. EEG data was downsampled to 512 Hz and bandpass filtered between 0.3 Hz and 80 Hz. The filtered data was re-referenced offline to the average of the two mastoids. Noisy electrodes were determined by calculating a robust noise adjusted z-score for each, as implemented in the ADJUST plugin for artifact detection (Mognon, Jovicich, Bruzzone, & Buiatti, 2011). This method calculates the ratio of the median absolute deviation of high frequency components (>50 Hz) to low frequency components (<50 Hz) for each electrode, expressed as a z-score relative to all other electrodes. Those electrodes with a robust z-score exceeding 5 were marked for possible rejection. This was complemented by visual examination of the power spectral density of each electrode to determine those with excessive low and high frequency activity or contaminated by line noise.

The continuous data was segmented into 3200 msec epochs, spanning 1200 msec before to 2000 msec after auditory verb onset. This trial length ensured sufficient data to resolve the low frequencies when carrying out time-frequency decomposition. However, for baseline correction a pre-stimulus interval of 200 msec (-250 msec to -50 msec) was applied and subsequent data analyses were limited to the 0 msec to 1000 msec post-stimulus interval.

Noisy electrodes marked for rejection were removed. Before carrying out independent components analysis (ICA), to correct for ocular movements, epochs that were highly contaminated with noise due, in particular, to movement were removed from the dataset. ICA was carried out on the segmented data of all 64 scalp electrodes for each participant using the infomax algorithm (Bell & Sejnowski, 1997) implemented in EEGLAB. Components corresponding to eye-blinks were determined automatically via the ADJUST Toolbox (Mognon et al., 2011) and, generally, only the first component was rejected. Epochs were then visually inspected again and those contaminated by noise were re-

moved. A minimum of 53 (average of 57, ±4) electrodes and 43 epochs (average of 52 ± 5) were retained per participant. At this point, rejected electrodes were interpolated using spherical spline interpolation. The data was then separated into Go and No-Go conditions. A total of 434 Go trials and 415 No-Go trails were retained overall, with an average of 25.5 ± 2.6 Go trials and 24.4 ± 4 No-Go trials per participant.

2.6. Event-Related Potential (ERP) analysis

For each participant, the mean over all trials was calculated to yield the subject-level ERP data. The ERP data of each subject was low-pass filtered, with a cutoff of 30 Hz. In line with previous electrophysiological studies of language processing (McLaughlin et al., 2010; Sneed-German, Herschensohn, & Frenck-Mestre, 2014) the 64 scalp electrodes were divided into 7 regions of interest (ROI): left frontal electrodes (AF3, F1, F3, F5), right frontal electrodes (AF4, F2, F4, F6), left frontal-central electrodes (FC1, FC3, FC5, C1, C3, C5), right frontal-central electrodes (FC2, FC4, FC6, C2, C4, C6), left central-parietal electrodes (CP1, CP3, CP5, P1, P3, P5), right central-parietal electrodes (CP2, CP4, CP6, P2, P4, P6), and midline electrodes (AFz, Fz, FCz, Cz, CPz, Pz). To determine the time-windows in which a significant difference between Go and No-Go trials emerged, a permutation test with false discovery rate (FDR) correction was carried out on all time points of the post-stimulus interval (0–1000 msec) for each electrode; to carry out the permutation, 1000 random partitions were performed. A significant difference was only taken into consideration ($q \leq 0.05$) if its duration exceeded 10 msec (8 consecutive time samples).

2.7. Time-frequency decomposition

The ERSP was computed for each participant both for Go and No-Go trials merged and for Go and No-Go trials separately using the Matlab toolbox, FieldTrip (Oostenveld, Fries, Maris, & Schoffelen, 2011). Time-frequency decomposition was carried out on a single trial basis as the squared norm of the convolution of the complex Morlet wavelet for each of the 64 scalp electrodes and for the 3–40 Hz frequency band. Within this frequency range, the wavelet width varied linearly from 3 to 10 cycles as a function of increasing frequency. This yielded a spectral bandwidth of 2 Hz and a temporal bandwidth of 318 msec at the lowest frequency of interest (3 Hz) and a spectral bandwidth of 8 Hz and a temporal bandwidth of 80 msec at the highest frequency of interest (40 Hz). At the single trial level, post-stimulus time-frequency data was z-score normalized relative to the pre-stimulus baseline interval (−250 to −50 msec). Then, for each participant, the grand average ERSP was calculated by averaging the trial-level ERSP and this data was entered into the subsequent statistical analyses.

2.8. Time frequency statistical analyses

2.8.1. Post stimulus activity versus baseline

To assess the statistical significance of ERSP in relation to the pre-stimulus baseline for Go and No-Go trials merged and for Go and No-Go trials separately, we applied a baseline permutation method (Delorme et al., 2004) for each of 9 frontal central electrodes (*FC3, FC4, C3, C4, CP3, CP4, FCz, Cz and CPz*). These electrodes were chosen based on the hypothesis and findings of previous studies showing motor activation at these sites (Fargier et al., 2012; van Elk et al., 2010). This method involves permuting the pre-stimulus baseline values across both time and trials for each frequency to generate a surrogate distribution for each frequency value; we carried out 2000 permutations at each frequency. For each time-frequency point, values that fell within the 97.5% tail of the surrogate distribution were considered significant at $q \leq 0.05$. The comparison of Go and No-Go merged to baseline was

computed across the 3–40 Hz range. In contrast, the independent comparison of Go to baseline and No-Go to baseline focused on the lower and upper μ bands, 8–10 Hz and 11–13 Hz, respectively.

2.8.2. Cluster based permutation analyses: Go versus No-Go trials

To directly compare Go and No-Go trials, the participant-level grand-average ERSP for each type of trial was entered into a cluster-based permutation analyses (Maris & Oostenveld, 2007). This non-parametric test simplifies the resolution of the multiple comparisons problem by correcting at the level of clusters that are determined based on an adjacency criterion and its calculation involved a multi-level statistical approach. At the first level, a dependent-samples *t*-test was performed for every data sample across conditions; data points corresponded either to time × frequency (for a given electrode) or electrode × time (for a given frequency band of interest) samples. A pre-set threshold of 5% (two-tailed) was used to group neighboring electrodes into clusters; neighbors were determined based on an adjacency criterion of a minimum of 2 electrodes calculated using the Delaunay triangulation function implemented in FieldTrip. To calculate cluster-level statistics, t-statistics were summed in each cluster and the maximum of the cluster-level statistic was determined. On the second level, we created a Monte-Carlo permutation distribution to calculate the significance probability. Participants' grand averages were randomly assigned to one of two conditions 1000 times and, for each random partition, the largest cluster-level statistic was determined. The Monte-Carlo permutation distribution was then constructed. The cluster-level test statistics were then compared to this permutation distribution and clusters in the highest or lowest 2.5th percentile of the distribution were considered significant (Fonteneau, Bozic, & Marslen-Wilson, 2015; Mazaheri et al., 2018).

Two comparisons were carried out. Cluster-based permutation analyses were carried out for two specific frequency bands of interest, the lower and upper μ bands (8–10 Hz and 11–13 Hz, respectively), for all 64 electrodes and the entire post-stimulus time interval (0–1000 msec), in which case each data sample constituted a spatial-temporal sample. The same analysis was carried out for individual electrodes of interest over the entire post-stimulus time interval and for the entire 3–40 Hz frequency band, in which case each data sample constituted a time × frequency sample.

3. Results

3.1. Event-related potentials (ERPs)

The mean ERPs over 6 ROIs as well as the 95% confidence intervals for both Go and No-Go conditions are presented in Fig. 2. A clear N1-P2 complex followed by an N400 can be seen, revealing clean recording of ERPs during the linguistic processing of action verbs in the CAVE. No statistically significant differences were revealed between the Go and No-Go conditions at any ROI or any time point.

3.2. Event-related spectral perturbation (ERSP)

3.2.1. Go and No-Go merged

Fig. 3 presents the ERSP of both trial types merged (Go + No-Go) for the 3–40 Hz frequency band and over the post-stimulus interval (0–1000 msec). We examined the mean oscillatory activity over groups of electrodes analyzed in previous studies that examined motor and semantic interactions (Fargier et al., 2012): Left frontal-central (FC1, FC3, FC5), Right frontal-central (FC2, FC4, FC6), Left Central (C1, C3, C5), Right Central (C2, C4, C6) and Left central-parietal (CP1, CP3, CP5) and Right central-parietal (CP2, CP4, CP6) and Midline electrodes (FCz, Cz, CPz).
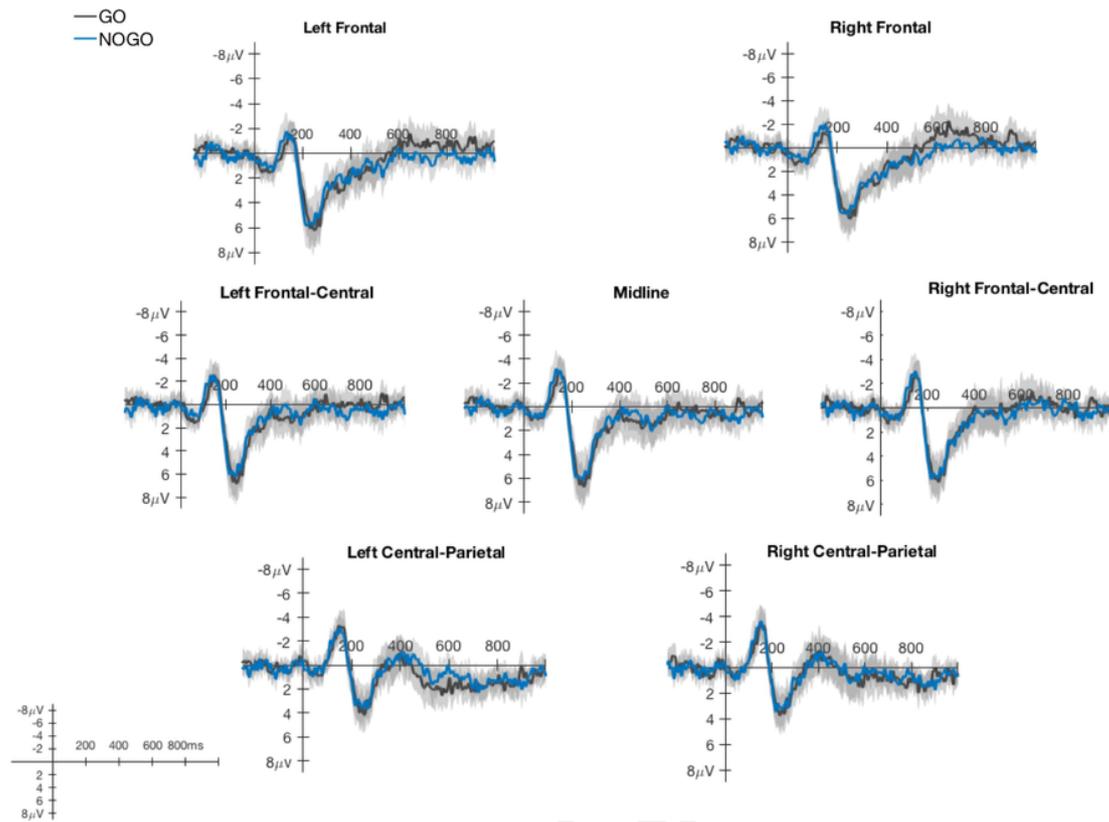
Fig. 2. Grand average ERPs and 95% confidence intervals for frontal electrodes (top), frontal-central electrodes (middle), and central-parietal electrodes (bottom), for Go and No-Go conditions.

The visual examination of the time-frequency maps presented in Fig. 3 revealed two patterns of activity. First, we see an early strong ERS emerging within the first 100 msec of the post-stimulus interval, which corresponds to the N1-P2 complex that we observed in our ERP results. In relation to this, it is important to underline that the time-frequency decomposition was carried out on a single-trial basis such that grand-average time-frequency activity includes both trial varying or induced activity (non-phase locked) and evoked activity (time locked but not phase locked in relation to stimulus onset) (Roach & Mathalon, 2008). Secondly, we observed an ERD in the μ band (8–13 Hz) that emerged in the 400–500 msec time window. Visual examination of this ERD revealed that it was stronger over the left hemisphere than the right, which is undoubtedly related to the fact that all participants were right-handed and executed actions with their right hand.

These patterns were substantiated by the statistical analysis of the post-stimulus interval (0–1000 msec) for all trials (Go + No-Go). Fig. 4 presents the results of this analysis; it reveals post-stimulus activity that was statistically significant ($q \leq 0.01$) compared to the baseline period. Only significant activity is shown. The μ-band ERD was statistically significant from the 400–500 msec time window up to 1000 msec after stimulus onset. Results also showed significant ERD in the beta band, in particular in the 20–30 Hz frequency band, spanning beta 2 and 3.

3.2.2. Go vs. No-Go comparison: permutation analyses

To test the hypothesis of a difference in ERSP between Go and No-Go conditions, a two-tailed spatio-temporal cluster-based permutation test was carried out for each of the 9 electrodes of interest for the 1-second post-stimulus time window across the 3–40 Hz frequency band. Fig. 5 presents those 5 electrodes out of the 9 that revealed statistically significant differences. The time-frequency maps present the raw effect (Go/No-Go difference) and only statistically significant activity ($q \leq 0.025$) is shown. We found significantly greater ERD for Go vs. No-

Go trials in the μ band from 700 to 1000 msec, for all electrodes except FC3 where it emerged earlier. In the 3 left hemisphere electrodes (FC3, C3 and CP3), the significant difference between conditions extended into the beta band (20–30 Hz, spanning beta 2 and beta 3) and, for FC3 and C3, emerged in an early time window (200–300 msec).

To test the hypothesis of a difference in ERD between conditions specifically in the μ band, a one-tailed spatio-temporal cluster-based permutation test was carried out for the low μ (8–10 Hz) and high μ (11–13 Hz) bands, over the entire post-stimulus time window (0–1000 msec) and over all 64 electrodes. The topographies in Fig. 6 present the log normalized Monte-Carlo significance probability or p-value ($-\log10(p)$) of the first significant cluster over consecutive 100 msec time windows for both μ bands. It is important to point out that, while the statistical analysis was carried out for each data sample ($\Delta t = 15.6$ msec), the results are presented in consecutive 100 msec time steps for visualization purposes. This reveals the spatial-temporal points for which the Go vs. No-Go difference was statistically significant ($q \leq 0.05$).

Greater ERD for Go trials compared to No-Go trials was revealed in both the 8–10 Hz ($p = .048$) and 11–13 Hz ($p = .05003$) frequency bands. For the 11–13 Hz frequency band, a significant difference emerged from 400 to 600 msec over a small number of central electrodes; the difference was also reliable from 800 msec to the end of the 1 sec time period. For the 8–10 Hz frequency band, the difference in μ ERD reached significance later, in the 600–700 msec time window over posterior electrodes. However, from the 700–800 msec time window to the end of the trial, this difference emerged over frontal-central electrodes.

3.2.3. Significant ERSP for Go and No-Go independently

In a final comparison, we calculated the significant ERSP for each condition independently, but focusing on significant ERD in the upper
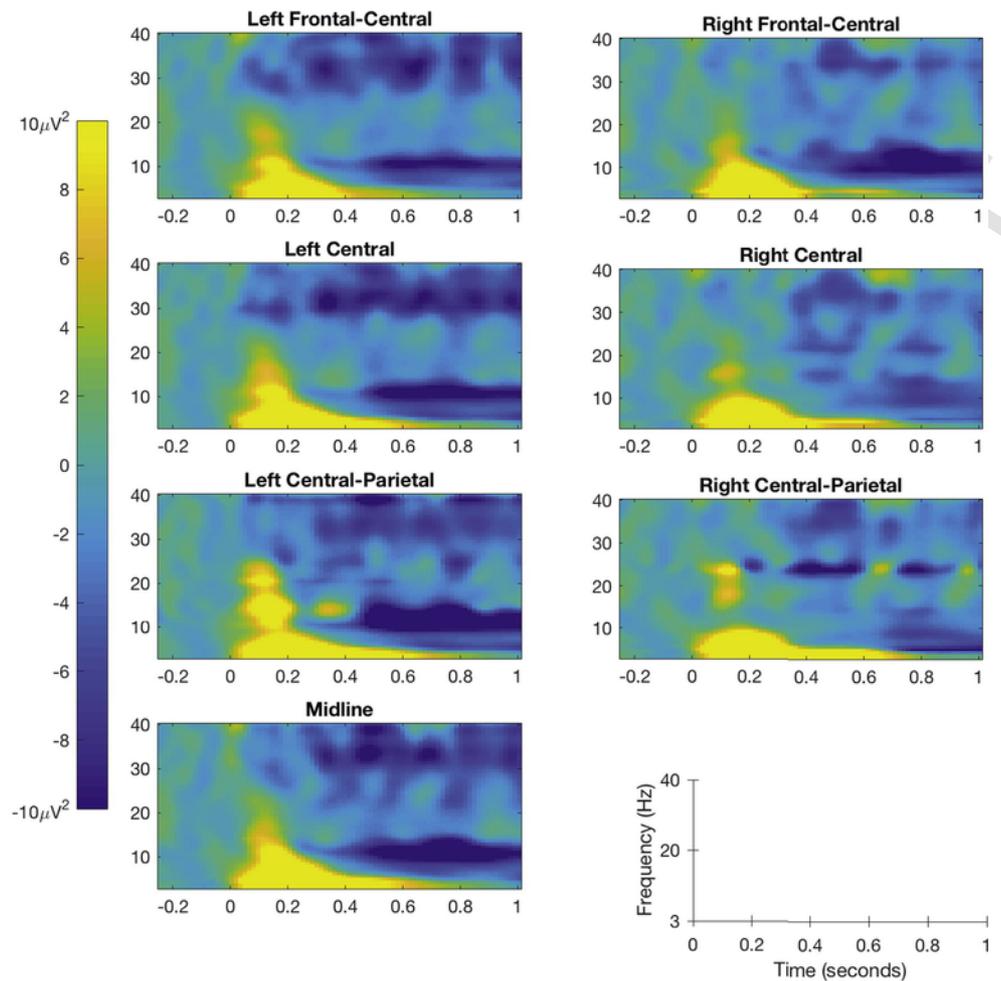
Fig. 3. Time-Frequency maps of all trial types (Go + No-Go).

(11–13 Hz) and lower (8–10 Hz) μ bands. Results showed significant ERD in both Go and No-Go conditions and for both μ bands. Figs. 7a and 7b show significant ERSP post-stimulus activity for the low μ-band (8–10 Hz) and the upper μ band (11–13 Hz), respectively. Only significant ($q \leq 0.05$) post-stimulus activity is shown, all non-significant activity relative to the baseline is masked.

4. Discussion

4.1. EEG-VR combination

The present study examined the synergy between motor and semantic processes during language processing, in a novel protocol using EEG in an ecologically valid environment. Participants performed a Go/No-Go task in an interactive CAVE environment. They heard action verbs and subsequently saw virtual objects which they either manipulated or not. Our results showed, first, a clear pattern of language related ERPs during verb processing for all trials, i.e. an N1/P2 complex followed by an N400. As discussed in greater length below, we found no variation in the ERP response as a function of trial type (Go/No-Go). Second, as detailed below, we found clear evidence of interactions between motor and linguistic processing, as shown by event-related desynchronization in language processing time windows. Significant ERD emerged during verb processing in both the μ (8–13 Hz) and beta band (20–30 Hz) for both Go and No-Go trials. While μ ERD emerged in the 400–500 msec time window, beta ERD emerged earlier (starting at 200 ms) and per-

sisted. The direct comparison of Go to No-Go trials revealed greater μ ERD for Go trials. We did not find concomitant variation of the CNV and ERD. These results highlight the advantage of combining EEG and CAVE for the study of motor-language processes under well-controlled conditions and offer a novel, ecologically valid methodology to study these processes.

In comparison to the present study, it could be argued that previous EEG experiments that have focused on the neural signatures of motor-semantic interactions used relatively impoverished environments (Alemanno et al., 2012; Fargier et al., 2012; Moreno et al., 2015; von Nicolai et al., 2014; van Elk et al., 2010) and do not provide a clear reflection of how language is processed in real life (Knoeferle, 2015; Tromp et al., 2018). The present study overcame this limitation by presenting participants with a realistic albeit virtual environment in which they not only saw but actually manipulated virtual objects. In general, VR uses digital images and sound to create a credible and immersive sensory experience. Manual control tools such as finger-trackers allow participants to interact with objects and receive real-time feedback (Burdea & Coiffet, 2003) and head movements are tracked such that the visual environment responds to the participant's movement in a similar way to the real world. Compared to computer screen or head-mounted display systems, CAVE environments are highly effective in providing participants with a sense of presence and immersion (Juan & Pérez, 2009). In the current experiment, participants benefitted from an embodied experience as they were immersed in a virtual environment depicting an office, complemented by a physical Plexiglas table.
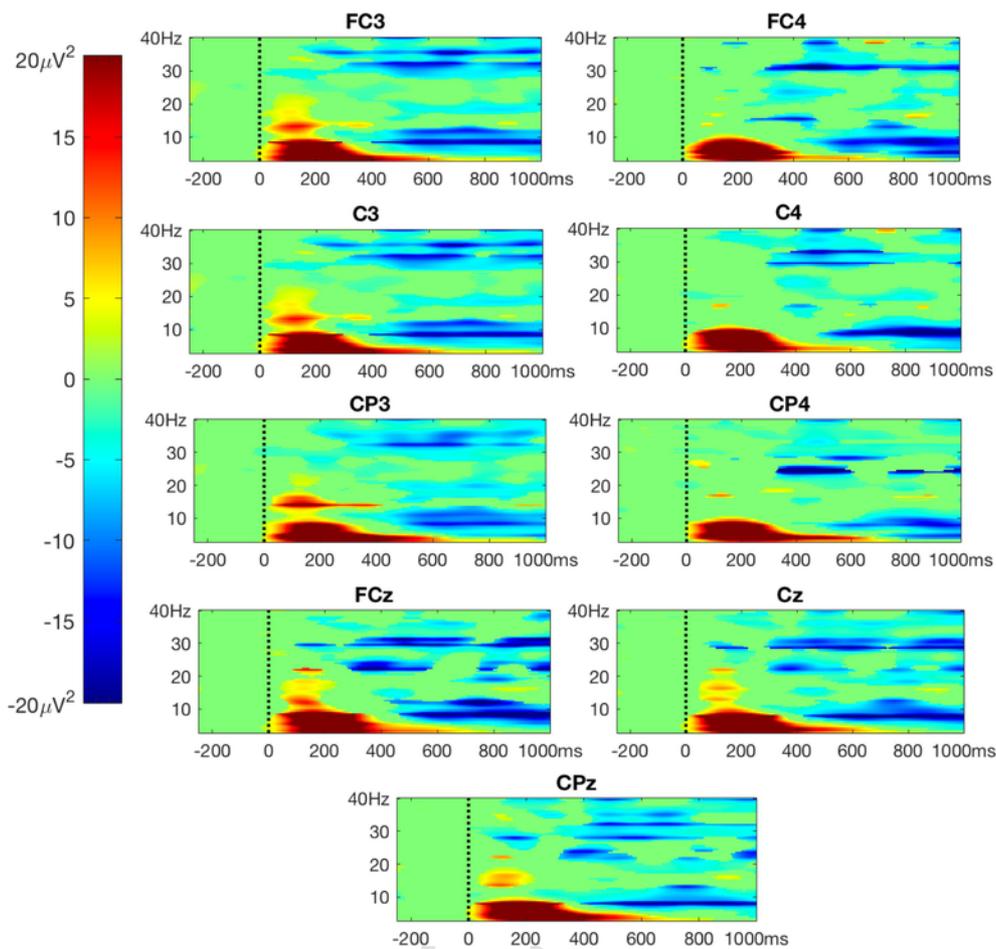
Fig. 4. Time-frequency maps showing statistically significant (q ≤ .05, FDR corrected) post-stimulus power relative to the baseline period for 9 electrodes over both hemispheres.
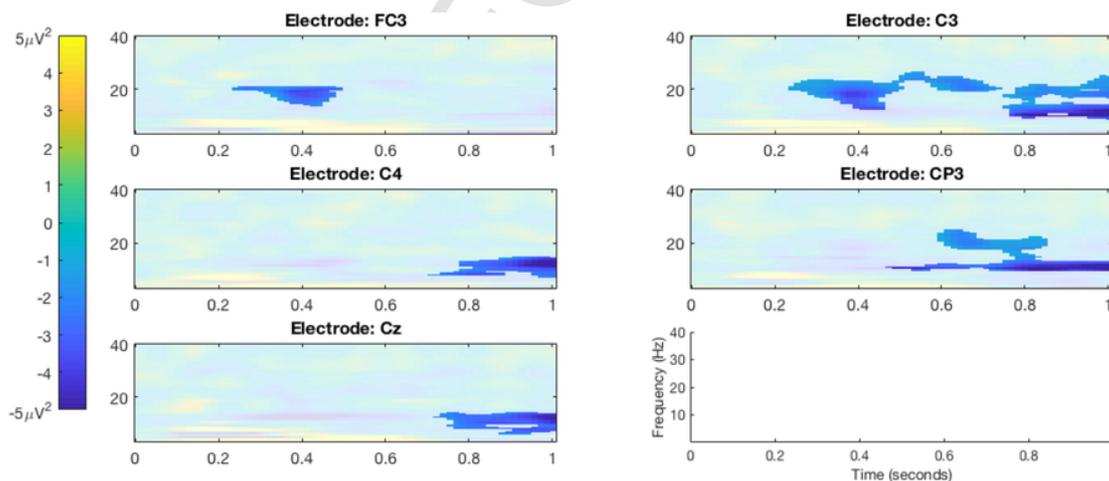

Fig. 5. Result of cluster-based permutation test comparing Go and No-Go trials across 3–40 Hz, showing the 5 electrodes that revealed statistically significant differences.

They were equipped with a finger tracker, increasing the experience of object manipulation, although sensory feedback was not provided to avoid any confounds with the recording of motor activation. They were free to move their arms and hands and performed 16 naturalistic actions (ex. *throw, drop, push*) on virtual objects that obeyed natural physical constraints. Both the objects and the visual environment responded in real-time to participants' movements, providing them with a sense of agency (Johnson-Glenberg, 2018) and leading to real-life behavior (Mestre, 2015). Presence, immersion and agency all contribute to a greater involvement of the sensorimotor system, such that the responses elicited in these conditions are closer to what probably occurs in real life (Bohil et al., 2011). Combining the CAVE and EEG therefore enabled us to control multimodal sensory stimulation while observing the brain correlates of motor and linguistic interaction in an ecologically valid environment, where participants could plan and perform naturalistic movements. We did not, however, measure presence and thus have no concrete evidence that the present design afforded a
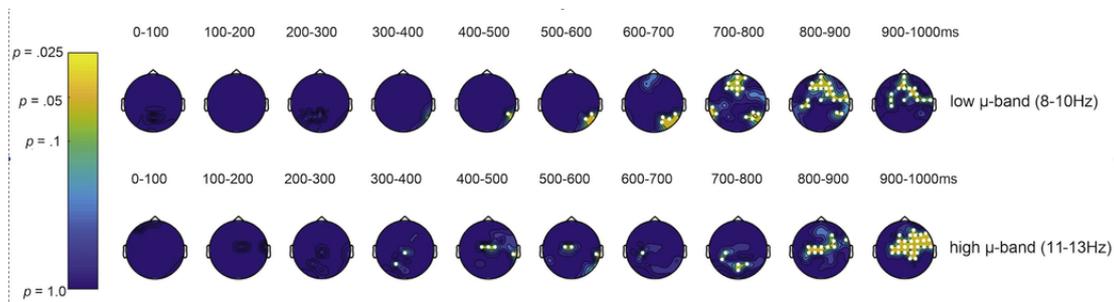
Fig. 6. Result of the spatio-temporal cluster-based permutation test comparing Go to No-trials over the entire post-stimulus time window and all 64 electrodes for the low (8–10 Hz) and high (11–13 Hz) μ band.
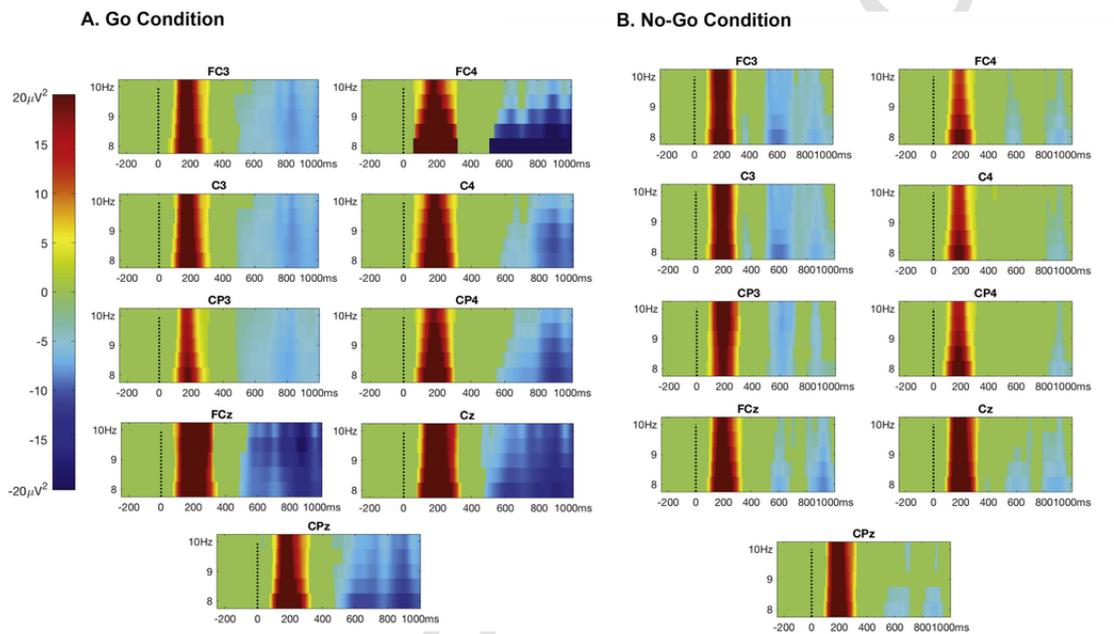

Fig. 7a. Post-stimulus oscillatory activity over the 8–10 Hz frequency range for (A) Go trials and (B) No-Go trials. Only statistically significant ($q \leq 0.05$, FDR corrected) ERSP values are shown.
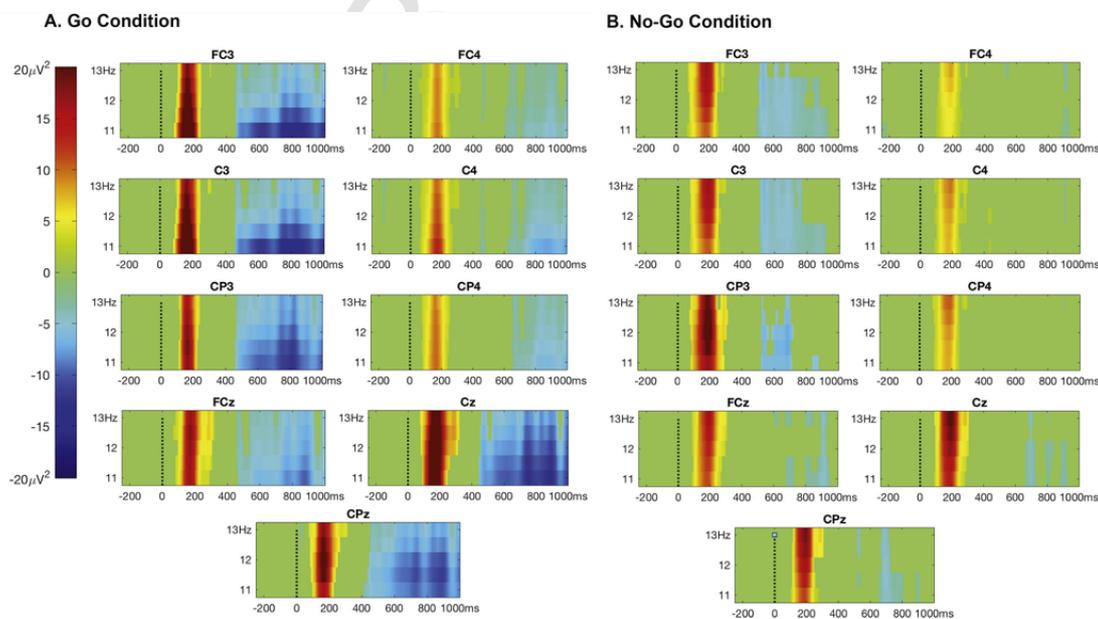

Fig. 7b. Post-stimulus oscillatory activity over the 11–13 Hz frequency range for (A) Go trials and (B) No-Go trials. Only statistically significant ($q \leq 0.05$, FDR corrected) ERSP values are shown.

greater sense of involvement than, say a flat screen presentation. Future studies would benefit from this added measure.

Combined EEG and VR can provide for a richer and more representative illustration of what occurs when motor and language processes overlap, and hence a better understanding of how language is embodied. In relation to the technical constraints of the present work, it is noteworthy that we implemented a design in which participants remained stationary during the period of interest of EEG recording. This indeed afforded the extremely high quality of the ERP traces that were obtained in the present study. This choice in fact mirrors previous VR studies on language processing (Repetto, 2014; Repetto, Cipresso et al., 2015; Tromp et al., 2018) and is linked to the constraints of EEG as concerns the need to eliminate spurious noise. Nonetheless, in contrast to previous work, participants performed a variety of naturalistic movements. This thus afforded a far more interactive and ecological situation than previously used. One might nonetheless argue against the ecological validity of our experimental setup, because participants did not receive haptic feedback during the manipulation of virtual objects. This might create some cognitive dissonance for the participants and thereby influence the data. To assess this, a direct comparison with the manipulation of real objects would be required (cf. Repetto, Cipresso et al., 2015). However, our results refer to EEG recordings acquired *prior to* the onset of virtual objects and hence manipulation, which tends to counter the argument of cognitive dissonance. Nevertheless, Invitto, Faggiano, Sammarco, De Luca, and De Paolis (2016) report both behavioral and EEG results from a study in which they compared mental imagery, virtual manipulation and real grasping of objects. They found significant differences in ERP components (primarily N1) between real and virtual manipulation such that virtual action failed to facilitate perceptual processes compared to real action. These findings suggest a need for further investigation of the role of multimodal stimulation in deciphering the links between embodiment and linguistic processing. However, two important limitations of the comparison between our study and Invitto et al. (2016) are that they measured ERP in a "Go/No-Go" recognition task (respond only if the stimulus had been presented) and *after* the manipulation (training) phase. Hence, their "Go/No-Go" task was not similar to ours and undoubtedly measured different aspects of processing.

4.2. Interaction between motor and semantic processes

The current study explored motor activation in response to auditory action verbs as well as the effect of motor planning on linguistic processing. This was investigated in part by modulations in the ERP signature. Previous behavioral (Buccino et al., 2005; Glenberg & Kaschak, 2002; Sato et al., 2008) and EEG (Aravena et al., 2010) studies have shown that processing action verbs is facilitated when congruent actions are planned prior to verb presentation but hindered when congruent actions are planned simultaneously. We therefore expected No-Go trials to elicit a greater N400 compared to Go trials, possibly due to interference effects linked to inhibiting action (García & Ibáñez, 2016). In line with Aravena et al. (2010), we posited that Go trials might also increase early, motor-related ERPs, showing enhanced motor preparation. Contrary to these predictions, the analyses of ERPs revealed no differences in motor preparation (RP) or semantic processing (N400) across conditions. One possible explanation for the absence of differences in motor preparation is that, whereas in Aravena et al.'s experiment (2010) participants planned a specific and repeated action (closed or open-handed manual response throughout the experiment), in our experiment participants were requested to produce one of numerous hand movements, which they had to plan during a 2.5 s period before object presentation. It is possible that planning to perform a pre-specified repetitive movement allowed participants to pre-program movements in a way that was not possible under the current conditions, in

which participants performed specific but variable actions on Go trials. The lack of an N400 effect is open to speculation; however, in the current design all actions were coherent, the only variation being the specific hand movement executed for a given verb. This differs from previous work in which participants knew a priori which movement to perform across all trials and had to semantically integrate the specific verb with a specific action on a trial to trial basis (Aravena et al., 2010). In the same line of argumentation, it is possible that the facilitation of semantic integration by coherent motor preparation only emerges under optimal conditions in which participants know a priori the movement to execute. It is noteworthy that the finding of facilitation in ACE studies is indeed not systematic but subject to numerous constraints, including the timing of motor preparation (Sato et al., 2008), the linguistic context (Aravena et al., 2012; Boulenger et al., 2006) and even the syntactic formulation (Aravena et al., 2014).

4.3. Modulation of ERSP

ERPs do not tell the whole story when it comes to motor activation as they use phase-locked linear averaging methods and hence do not capture the large portion of motor cortex signal that is not phase-locked to the stimulus (Pfurtscheller & Lopes da Silva, 1999; Vukovic & Shtyrov, 2014). Event-related ERS/ERD, however, considers both phase-locked and unlocked cortical activity. In the present study we capitalized on ERSP to explore neural activity in the sensorimotor cortex, notably as reflected by ERD in the μ frequency bands (8–13 Hz) in response to action verb processing and prior to physical movement. Our results revealed greater motor-related cortical activity (μ ERD) during verb processing for both Go and No-Go trials, starting in the 400–500 msec time window.

Previous studies have provided evidence of early activation of motor areas during linguistic processing. Pulvermüller, Härle, and Hummel (2001) used high-density EEG to measure cortical activity while participants read action verbs performed using different body parts. Somato-specific activation was found along the motor strip starting at 250 msec after verb presentation. In line with these findings, one could hypothesize that ERD in the μ frequency bands during verb processing should have occurred in our experiment in this window. Unlike the Pulvermüller et al. (2001) study, which used written single verbs as stimuli, we used auditory stimuli. The timing of lexical access is notoriously more difficult to determine for auditory language compared to written language (Hauk, Shtyrov, & Pulvermüller, 2008); this could account for the differences across studies.

The comparison of Go and No-Go trials to baseline revealed significant μ ERD starting in the 400-500 msec time window, which is generally considered to be associated with lexical access (Indefrey & Levelt, 2004; Kutas & Federmeier, 2011). This result is in line with that reported by Fargier et al. (2012) in a word learning paradigm, where no ERD was observed in the μ band for novel words prior to training but emerged following two training sessions in which the novel words were associated with hand movements, and was observed for centro-parietal electrode starting at 450 msec post stimulus onset. It is important to note that Fargier et al. (2012) included a second day of training and test sessions for which the results were less conclusive. Our study looked at motor activation during the processing of well-known words (verbs) in the native language, such that we would not expect the pattern of results to vary over time. We also found significant ERD in the beta band, in particular in the 20–30 Hz band (spanning beta 2 and 3) starting from 200 msec for frontal central electrodes. As discussed below, the co-occurrence of ERD for the μ and beta bands provides an argument against the hypothesis that the μ ERD we report was in fact alpha. The fact that motor resonance was found in both conditions when they were tested independently also suggests that motor activation was not (only) caused by action planning but, in line with previous studies,

was involved in semantic processing (Aravena et al., 2010; Boulenger et al., 2006; Pulvermüller et al., 2001).

We also investigated differences in motor activation during verb processing as a function of Trial type (Go vs. No-Go) for both the lower (8–10 Hz) and higher μ band (11–13 Hz). The lower μ band has been associated with action observation and the higher μ band with action execution (Aridan, Ossmy, Buaron, Reznik, & Mukamel, 2018). Analyses revealed greater μ ERD for Go compared to No-Go trials across the two frequency bands but for different time windows and with different distributions. In the higher μ band (11–13 Hz), greater ERD for Go trials emerged in the 400–500 msec time window for central-parietal electrodes. Although actual movement was delayed to 2500 msec after verb onset during Go trials, the go signal was presented only 200 msec prior to verb presentation. Previous Go No-Go studies have shown that when the go signal was presented simultaneously with a verb describing a motion with a specific effector, responses using that same effector were slower (Buccino et al., 2005; Sato et al., 2008). In the present experiment, it is possible that Go trials produced greater μ ERD because processing action verbs inhibited motor preparation. In the lower μ band (8–10 Hz), a greater ERD for Go trials emerged later, starting at 700 msec, and was concentrated primarily over central and frontal electrodes. This activity is likely not a direct reflection of motor activation but indicative of activation in "convergence zones" of language and motor structures (Damasio, 1989; Fargier et al., 2012).

Finally, it is of interest to note that the significant variations we observed in ERSP were not accompanied by significant modulations in the contingent negative variation (CNV). Various studies have examined the relationship between the reduction in spectral power in the alpha and beta bands (ERD) and the increase CNV in Go/No-Go paradigms (Filipović et al., 2001; Funderud et al., 2012; Mento, 2013; Zaepffel et al., 2013). The CNV is typically seen in paradigms where participants receive a warning signal (S1) followed by a target stimulus (S2) and is thought to reflect a series of intentional motor, preparatory and decisional processes (Funderud et al., 2012; Mento, 2013). It is often separated into an earlier and a later component. The late (or terminal) CNV, beginning up to 1.5 s before S2, causes activity over frontal and prefrontal cortices and has been associated with sustained alpha (8–13 Hz) and beta (14–30 Hz) ERD/ERS (Morash, Bai, Furlani, Lin, & Hallett, 2008). The association of increased CNV amplitudes and reduced event-related spectral power in the alpha, beta, theta and low gamma bands for Go versus the No-Go trials has been interpreted as possibly illustrating a coordinated, dynamic change in neural networks involved in motor preparation (Funderud et al., 2012). However, a number of studies have provided evidence against a direct coupling between these two phenomena by showing discrepancies between them and arguing that they reflect different cognitive and motor processes (Filipović et al., 2001; Zaepffel et al., 2013). In our experiment, greater μ synchronization was found for Go versus No-go trials but no significant differences between conditions emerged for the CNV. This could be taken as evidence that the ERD and CNV reflect different cognitive and motor processes. However, our study did show some variation in the CNV albeit very small, which may have been too slight to detect with the sample size of the current study (a G-power test based on our sample size and Cohen's d suggested a population of over 400 would be needed to produce a significant result). As such, our study cannot adjudicate this question.

4.4. Caveats

One important limitation of the present study is the lack of a control condition using abstract verbs. Such a condition would have allowed us to see whether motor resonance emanated from processing action verbs specifically. Based on previous studies, we posit that we would have found greater μ rhythm ERD for action verbs compared to

abstract verbs (Alemanno et al., 2012; Moreno et al., 2015). Importantly, a non-action verb control condition would also have allowed us to observe motor activation during Go vs. No-Go trials for action versus non-action verbs, once again illustrating whether the greater motor activation we found for Go compared to No-Go trials originated from interference from the action semantic content of the verbs as opposed to motor preparation alone. Note that Sato et al. (2008) based their conclusions of semantic interference on motor processing on the comparison of verbs related to a specific effector (hand vs. foot), not on action verbs compared to abstract verbs, despite having included abstract controls (but see Buccino et al., 2005). This could also provide an avenue to explore, i.e. the inclusion of movements for different effectors (cf. Buccino et al., 2005).

Another caveat to bear in mind is the current debate on what μ ERD reveals (Hobson & Bishop, 2016). Despite the growing use of time-locked μ ERD as a marker of motor neuron activity (Moreno et al., 2015), μ and alpha frequency bands (8–13 Hz) overlap and it is therefore important to dissociate the two (Hobson & Bishop, 2016). One way to accomplish this is to focus on distribution. Mu band activity is thought to originate in the motor and premotor cortex, measured in centro-parietal sites. Alpha band activity, on the other hand, is reportedly found over occipital and frontal sites (Fargier et al., 2012; Moreno et al., 2015). In addition, as beta band activity is thought to reflect motor activation directly (Pfurtscheller et al., 1999), ERD in the beta band generally accompanies μ ERD and has often been considered as an indicator that what is being detected is indeed μ and not alpha ERD (van Elk et al., 2010). We found greater μ ERD compared to baseline for both Go and No-Go trials as well as both beta 2 (19–25 Hz) and beta 3 (25–30 Hz) bands compared to baseline. This, along with the finding that ERD in the μ band was greater for central compared to posterior electrodes, indicates that the effects are indeed linked to μ and were not confounded with posterior alpha.

5. Conclusion

The present study used a naturalistic setting to investigate motor activation during language processing within an embodied framework. Results showing motor activation in time windows associated with semantic processing are in line with the language studies that found μ-band ERD compared to baseline during lexical-semantic retrieval of action language (Alemanno et al., 2012; Fargier et al., 2012; Moreno et al., 2015; von Nicolai et al., 2014; van Elk et al., 2010). The greater action-related μ ERD during verb processing for all trials, but prior to movement proper, also bolsters the claim that sensorimotor processing is involved in the conceptual representation of linguistic information. Our results indicating greater μ ERD related to single verb processing for Go compared to No-Go trials are in line with ACE studies showing that movement preparation interacts with semantic processing (Aravena et al., 2010; Buccino et al., 2005; Sato et al., 2008). The present results, which used auditory verbs in conjunction with the execution of the specified action to measure the overlap of motor and linguistic processing, bear strong similarity to those reported by Fargier et al. (2012) who also used a varied set of linguistic materials and associated actions to examine this question. However, in the present study participants engaged with virtual objects rather than observed movements. Our results further validate the use of time-frequency analysis to measure motor activation in this novel EEG-CAVE experimental paradigm.

Uncited references

Delorme, Westerfield, and Makeig (2007), Efron and Tibshirani (1991), Medendorp et al. (2007), Pulvermüller, Hauk, Nikulin, and Ilmoniemi (2005), Pulvermüller, Shtyrov, and Hauk (2009).

References

Alemanno, F., Houdayer, E., Cursi, M., Velikova, S., Tettamanti, M., Comi, G., Leocani, L., 2012. Action-related semantic content and negation polarity modulate motor areas during sentence reading: An event-related desynchronization study. Brain Research 1484, 39–49.

Aravena, P., Courson, M., Frak, V., Cheylus, A., Paulignan, Y., Deprez, V., Nazir, T., 2014. Action relevance in linguistic context drives word-induced motor activity. Frontiers in Human Neuroscience 8, 163.

Aravena, P., Delevoye-Turrell, Y., Deprez, V., Cheylus, A., Paulignan, Y., Frak, V., Nazir, T., 2012. Grip force reveals the context sensitivity of language-induced motor activity during "action words" processing: Evidence from sentential negation. PLoS One 7 (12), e50287.

Aravena, P., Hurtado, E., Riveros, R., Cardona, J.F., Manes, F., Ibáñez, A., 2010. Applauding with closed hands: Neural signature of action-sentence compatibility effects. PloS One 5 (7), e11751.

Aridan, N., Ossmy, O., Buaron, B., Reznik, D., Mukamel, R., 2018. Suppression of EEG mu rhythm during action observation corresponds with subsequent changes in behavior. Brain Research 1691, 55–63.

Atkinson, D., 2010. Extended, embodied cognition and second language acquisition. Applied Linguistics 31 (5), 599–622.

Barsalou, L.W., 1999. Perceptions of perceptual symbols. Behavioral and Brain Sciences 22 (4), 637–660.

Barsalou, L.W., 2008. Grounded cognition. Annual Review of Psychology 59 (1), 617–645.

Bayliss, J.D., Ballard, D.H., 2000. A virtual reality testbed for brain-computer interface research. IEEE Transactions on Rehabilitation Engineering 8 (2), 188–190.

Bell, A.J., Sejnowski, T.J., 1997. The "independent components" of natural scenes are edge filters. Vision Research 37 (23), 3327–3338.

Bohil, C.J., Alicea, B., Biocca, F.A., 2011. Virtual reality in neuroscience research and therapy. Nature Reviews. Neuroscience 12 (12), 752–762.

Boulenger, V., Roy, A.C., Paulignan, Y., Deprez, V., Jeannerod, M., Nazir, T.A., 2006. Cross-talk between language processes and overt motor behavior in the first 200 msec of processing. Journal of Cognitive Neuroscience 18 (10), 1607–1615.

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. Brain Research. Cognitive Brain Research 24 (3), 355–363.

Burdea, G.C., Coiffet, P., 2003. Virtual reality technology. John Wiley & Sons.

Caetano, G., Jousmäki, V., Hari, R., 2007. Actor's and observer's primary motor cortices stabilize similarly after seen or heard motor actions. Proceedings of the National Academy of Sciences of the United States of America 104 (21), 9058–9062.

Damasio, A.R., 1989. Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. Cognition 33 (1–2), 25–62.

de Vega, M., Moreno, V., Castillo, D., 2013. The comprehension of action-related sentences may cause interference rather than facilitation on matching actions. Psychological Research 77 (1), 20–30.

Delorme, A., Makeig, S., 2004. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of Neuroscience Methods 134 (1), 9–21.

, , , , , , , . 0, –.

Diefenbach, C., Rieger, M., Massen, C., Prinz, W., 2013. Action-sentence compatibility: The role of action effects and timing. Frontiers in Psychology 4, 272.

, , , , . 0, –.

Fargier, R., Paulignan, Y., Boulenger, V., Monaghan, P., Reboul, A., Nazir, T.A., 2012. Learning to associate novel words with motor actions: Language-induced motor activity following short training. Cortex; a Journal Devoted to the Study of the Nervous System and Behavior 48 (7), 888–899.

Filipović, S.R., Jahanshahi, M., Rothwell, J.C., 2001. Uncoupling of contingent negative variation and alpha band event-related desynchronization in a go/no-go task. Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology 112 (7), 1307–1315.

Fodor, J.A., 1980. Methodological solipsism considered as a research strategy in cognitive psychology1. Behavioral and Brain Sciences 3 (1), 63–73.

Fodor, J.A., 1987. Psychosemantics: The Problem of Meaning in the Philosophy of Mind. MIT Press.

Fonteneau, E., Bozic, M., Marslen-Wilson, W.D., 2015. Brain network connectivity during language comprehension: Interacting linguistic and perceptual subsystems. Cerebral Cortex (New York, N.Y.: 1991) 25 (10), 3962–3976.

Funderud, I., Lindgren, M., Løvstad, M., Endestad, T., Voytek, B., Knight, R.T., Solbakk, A.-K., 2012. Differential Go/NoGo activity in both contingent negative variation and spectral power. PLoS ONE 7 (10).

García, A.M., Ibáñez, A., 2016. A touch with words: Dynamic synergies between manual actions and language. Neuroscience and Biobehavioral Reviews 68, 59–95.

Glenberg, A.M., Kaschak, M.P., 2002. Grounding language in action. Psychonomic Bulletin & Review 9 (3), 558–565.

Goldin-Meadow, S., 2011. Learning through gesture. Wiley Interdisciplinary Reviews: Cognitive Science 2 (6), 595–607.

Grandchamp, R., Delorme, A., 2011. Single-trial normalization for event-related spectral decomposition reduces sensitivity to noisy trials. Frontiers in Psychology 2, 236.

Graziano, M., Gullberg, M., 2018. When speech stops, gesture stops: Evidence from developmental and crosslinguistic comparisons. Frontiers in Psychology 9.

Hauk, O., Shtyrov, Y., Pulvermüller, F., 2008. The time course of action and action-word comprehension in the human brain as revealed by neurophysiology. Journal of Physiology-Paris 102 (1), 50–58.

Hobson, H.M., Bishop, D.V.M., 2016. Mu suppression – A good measure of the human mirror neuron system?. Cortex 82, 290–310.

Invitto, S., Faggiano, C., Sammarco, S., De Luca, V., De Paolis, L.T., 2016. Haptic, virtual interaction and motor imagery: Entertainment tools and psychophysiological testing. Sensors (Basel, Switzerland) 16 (3).

Johnson-Glenberg, M.C., 2018. Immersive VR and education: Embodied design principles that include gesture and hand controls. Frontiers in Robotics and AI 5.

Juan, M.C., Pérez, D., 2009. Comparison of the levels of presence and anxiety in an acrophobic environment viewed via HMD or CAVE. Presence: Teleoperators and Virtual Environments 18 (3), 232–248.

Kaschak, M.P., Borreggine, K.L., 2008. Temporal dynamics of the action–sentence compatibility effect. Quarterly Journal of Experimental Psychology 61 (6), 883–895 https://doi.org/1.

Klimesch, W., Doppelmayr, M., Pachinger, T., Russegger, H., 1997. Event-related desynchronization in the alpha band and the processing of semantic information. Brain Research. Cognitive Brain Research 6 (2), 83–94.

Knoeferle, P., 2015. Language comprehension in rich non-linguistic contexts: Combining eye-tracking and event-related brain potentials. In: Willems, R.M. (Ed.), Cognitive neuroscience of natural language use. Cambridge University Press, Cambridge, pp. 77–100 https://doi.org/.

Koelewijn, T., van Schie, H.T., Bekkering, H., Oostenveld, R., Jensen, O., 2008. Motor-cortical beta oscillations are modulated by correctness of observed action. NeuroImage 40 (2), 767–775.

Ladouce, S., Donaldson, D.I., Dudchenko, P.A., Ietswaart, M., 2017. Understanding minds in real-world environments: Toward a mobile cognition approach. Frontiers in Human Neuroscience 10.

Legault, J., Zhao, J., Chi, Y.-A., Chen, W., Klippel, A., Li, P., 2019. Immersive virtual reality as an effective tool for second language vocabulary learning. Languages 4 (1), 13.

Lepecq, J.-C., Bringoux, L., Pergandi, J.-M., Coyle, T., Mestre, D., 2009. Afforded actions as a behavioral assessment of physical presence in virtual environments. Virtual Reality 13 (3), 141–151.

Mahon, B.Z., Caramazza, A., 2008. A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. Journal of Physiology, Paris 102 (1–3), 59–70.

Makeig, S., 1993. Auditory event-related dynamics of the EEG spectrum and effects of exposure to tones. Electroencephalography and Clinical Neurophysiology 86 (4), 283–293.

Makeig, S., Debener, S., Onton, J., Delorme, A., 2004. Mining event-related brain dynamics. Trends in Cognitive Sciences 8 (5), 204–210.

Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG- and MEG-data. Journal of Neuroscience Methods 164 (1), 177–190.

Matsumoto, J., Fujiwara, T., Takahashi, O., Liu, M., Kimura, A., Ushiba, J., 2010. Modulation of mu rhythm desynchronization during motor imagery by transcranial direct current stimulation. Journal of Neuroengineering and Rehabilitation 7, 27.

Mazaheri, A., Segaert, K., Olichney, J., Yang, J.-C., Niu, Y.-Q., Shapiro, K., Bowman, H., 2018. EEG oscillations during word processing predict MCI conversion to Alzheimer's disease. NeuroImage: Clinical 17, 188–197.

McLaughlin, J., Tanner, D., Pitkänen, I., Frenck-Mestre, C., Inoue, K., Valentine, G., Osterhout, L., 2010. Brain potentials reveal discrete stages of L2 grammatical learning. Language Learning 60 (Suppl. 2), 123–150.

, , , , , , , , , , , , , . 0, –.

Mento, G., 2013. The passive CNV: Carving out the contribution of task-related processes to expectancy. Frontiers in Human Neuroscience 7.

Mestre, D.R., 2015. On the usefulness of the concept of presence in virtual reality applications. In: The engineering reality of virtual reality 2015. International Society for Optics and Photonics, p. 93920J.

Mestre, D.R., Louison, C., Ferlay, F., 2016. The contribution of a virtual self and vibrotactile feedback to walking through virtual apertures. Proceedings of HCII, Toronto, 17–22 July.

Mognon, A., Jovicich, J., Bruzzone, L., Buiatti, M., 2011. ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. Psychophysiology 48 (2), 229–240.

Moore, K., Wiederhold, B.K., Wiederhold, M.D., Riva, G., 2002. Panic and agoraphobia in a virtual world. Cyberpsychology & Behavior: The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society 5 (3), 197–202.

Morash, V., Bai, O., Furlani, S., Lin, P., Hallett, M., 2008. Classifying EEG signals preceding right hand, left hand, tongue, and right foot movements and motor imageries. Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology 119 (11), 2570–2578.

Moreno, I., de Vega, M., León, I., Bastiaansen, M., Glen Lewis, A., Magyari, L., 2015. Brain dynamics in the comprehension of action-related language. A time-frequency analysis of mu rhythms. NeuroImage 109, 50–62.

Neuper, C., Pfurtscheller, G., 2001. Event-related dynamics of cortical rhythms: Frequency-specific features and functional correlates. International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology 43 (1), 41–58.

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, 156869.

Peeters, D., 2019. Virtual reality: A game-changing method for the language sciences. Psychonomic Bulletin & Review 26 (1), 1–7.

Pfurtscheller, G., Lopes da Silva, F.H., 1999. Event-related EEG/MEG synchronization and desynchronization: Basic principles. Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology 110 (11), 1842–1857.

Pineda, J.A., 2005. The functional significance of mu rhythms: Translating "seeing" and "hearing" into "doing". Brain Research. Brain Research Reviews 50 (1), 57–68.

Pulvermüller, F., 2005. Brain mechanisms linking language and action. Nature Reviews. Neuroscience 6 (7), 576–582.

Pulvermüller, F., Härle, M., Hummel, F., 2001. Walking or talking? Behavioral and neurophysiological correlates of action verb processing. Brain and Language 78 (2), 143–168.

, , , , , , II, , . . (3), –.

, , , , , , , . . (2), –.

Repetto, C., 2014. The use of virtual reality for language investigation and learning. Frontiers in Psychology 5.

Repetto, C., Cipresso, P., Riva, G., 2015. Virtual action and real action have different impacts on comprehension of concrete verbs. Frontiers in Psychology 6.

Repetto, C., Colombo, B., Riva, G., 2015. Is motor simulation involved during foreign language learning? A virtual reality experiment. SAGE Open.

Roach, B.J., Mathalon, D.H., 2008. Event-related EEG time-frequency analysis: An overview of measures and an analysis of early gamma band phase locking in schizophrenia. Schizophrenia Bulletin 34 (5), 907–926.

Sato, M., Mengarelli, M., Riggio, L., Gallese, V., Buccino, G., 2008. Task related modulation of the motor system during language processing. Brain and Language 105 (2), 83–90.

Sneed-German, E., Herschensohn, J., Frenck-Mestre, C., 2014. Pronoun processing in Anglophone late L2 learners of French: Behavioral and ERP evidence. EuroSLA 20. Presented at the York, United Kingdom.

Spadacenta, S., Gallese, V., Fragola, M., Mirabella, G., 2014. Modulation of arm reaching movements during processing of arm/hand-related action verbs with and without emotional connotation. PLOS ONE 9 (8), e104349.

Toni, I., de Lange, F.P., Noordzij, M.L., Hagoort, P., 2008. Language beyond action. Journal of Physiology-Paris 102 (1), 71–79.

Török, , Sulykos, I., Kecskés-Kovács, K., Persa, G., Galambos, P., Kóbor, A., Honbolygó, F., 2014. Comparison between wireless and wired EEG recordings in a virtual reality lab: Case report. In: 2014 5th IEEE conference on cognitive infocommunications (CogInfoCom). pp. 599–603.

Tromp, J., Peeters, D., Meyer, A.S., Hagoort, P., 2018. The combined use of virtual reality and EEG to study language processing in naturalistic environments. Behavior Research Methods 50 (2), 862–869.

van Elk, M., van Schie, H.T., Zwaan, R.A., Bekkering, H., 2010. The functional role of motor activation in language processing: Motor cortical oscillations support lexical-semantic retrieval. NeuroImage 50 (2), 665–677.

von Nicolai, C., Engler, G., Sharott, A., Engel, A.K., Moll, C.K., Siegel, M., 2014. Corticostriatal coordination through coherent phase-amplitude coupling. Journal of Neuroscience 34 (17), 5938–5948.

Vukovic, N., Shtyrov, Y., 2014. Cortical motor systems are involved in second-language comprehension: Evidence from rapid mu-rhythm desynchronisation. NeuroImage 102 (Pt 2), 695–703.

Weiss, S., Berghoff, C., Rappelsberger, P., Müller, H.M., 2001. Elektrophysiologische Hinweise zur Kategorisierung von Verben. In: Tagungsband der 1. Jahrestagung der Gesellschaft für Aphasieforschung und -behandlung GAB. p. S.64.

Wilson, A.D., Golonka, S., 2013. Embodied cognition is not what you think it is. Frontiers in Psychology 4, 58.

Zaepffel, M., Trachel, R., Kilavik, B., Brochier, T., 2013. Modulations of EEG beta power during planning and execution of grasping movements. PLoS ONE 27, 11877–11877.