The Neural Blending of Words and Movement: Event-Related Potential Signatures of Semantic and Action Processes during Motor–Language Coupling

Sabrina Cervetto, Mariano Diaz-Rivera, Agustín Petroni, Agustina Birba, Miguel Martorell Caro, Lucas Sedeño, Agustín Ibáñez, and Adolfo M. García

Abstract

Behavioral embodied research shows that words evoking limb-specific meanings can affect responses performed with the corresponding body part. However, no study has explored this phenomenon’s neural dynamics under implicit processing conditions, let alone by disentangling its conceptual and motoric stages. Here, we examined whether the blending of hand actions and manual action verbs, relative to nonmanual action verbs and non-action verbs, modulates electrophysiological markers of semantic integration (N400) and motor-related cortical potentials during a lexical decision task. Relative to both other categories, manual action verbs involved reduced posterior N400 amplitude and greater modulations of frontal motor-related cortical potentials. Such effects overlapped in a window of ~380–440 msec after word presentation and ~180 msec before response execution, revealing the possible time span in which both semantic and action-related stages reach maximal convergence. These results allow refining current models of motor–language coupling while affording new insights on embodied dynamics at large.

INTRODUCTION

One leading contribution of the embodied cognition framework (Moguilner et al., 2021; Pulvermüller, 2013a, 2018; de Vega, Moreno, & Castillo, 2013; Glenberg & Gallese, 2012; Barsalou, 2008) is the demonstration that language can variously affect overt behavior (Kogan, Muñoz, Ibáñez, & García, 2020; García & Ibáñez, 2016a; de Vega et al., 2013; Glenberg & Gallese, 2012). In particular, motor–language coupling paradigms show that action verbs can modulate movement-related mechanisms in an effector-specific fashion (Kogan et al., 2020; García & Ibáñez, 2016b). For example, in tasks requiring explicit semantic access, processing of manual action verbs (MaVs), compared to non-MaVs (nMaVs) and non-action verbs (nAVs), can facilitate or delay diverse hand actions, such as keyboard typing (García & Ibáñez, 2016b) and object reaching (Dalla Volta, Gianelli, Campione, & Gentilucci, 2009). Yet, little is known about the neural dynamics underlying such limb-selective interactions, and no study has explored their emergence under implicit processing conditions—let alone by considering time-locked signatures of their conceptual and motoric subprocesses. To bridge this gap, we assessed whether the blending of hand actions and MaVs, relative to nMaVs and nAVs, can modulate electrophysiological markers of semantic integration and motor preparation in a shallow processing task, while examining the temporal overlap of both subprocesses.

Motor–language coupling is a bidirectional phenomenon whereby action–language processes affect motor dynamics, and vice versa (Afonso et al., 2019; Melloni et al., 2015; Ibáñez et al., 2013; Aravena et al., 2010). Behavioral experiments show that MaVs can increase grip force in both unimanual and bimanual tasks (Da Silva, Labrecque, Caromano, Higgins, & Frak, 2018; Nazir et al., 2017; Frak, Nazir, Goyette, Cohen, & Jeannerod, 2010),
whereas neuroimaging studies show that these verbs differentially increase activation along partly somatotopic (hand-specific) cortical motor regions (Pulvermüller, 2013a, 2018). Yet, only limited research has examined the neurofunctional unfolding of motor–language coupling, mainly using overtly semantic tasks. In particular, manual responses to MaVs have been linked to early engagement of motor and premotor regions (Mollo, Pulvermüller, & Hauk, 2016; Klepp, Niccolai, Buccino, Schnitzler, & Biermann-Ruben, 2015; Pulvermüller, Härle, & Hummel, 2001), fast modulation of motor–language evoked-potentials (Gianelli & Dalla Volta, 2015; Papeo, Vallesi, Isaja, & Rumiati, 2009; Buccino et al., 2005), and reduced synchronization of movement-sensitive beta oscillations (Klep et al., 2015). Moreover, during explicit semantic access, this functional synergy is typified by canonical neural signatures of both semantic integration (Ibáñez et al., 2013; Aravena et al., 2010) and action initiation (Melloni et al., 2015; Ibáñez et al., 2013; Aravena et al., 2010), irrespective of task accomplishment (Dalla Volta, Avanzini, De Marco, Gentiliucci, & Fabbri-Destro, 2018). Thus, the dynamic neural signatures of motor–language coupling are fairly well established for tasks that overtly direct attention to word meanings.

However, this empirical corpus faces two critical shortcomings. First, none of these studies examined whether such effects also emerge under shallow-processing conditions, thus overlooking a key requisite to reveal primary embodied effects (García et al., 2019; Mollo et al., 2016; Kiefer, Sim, Herrmberger, Grothe, & Hoenig, 2008; Hauk, Johnsrude, & Pulvermüller, 2004). Second, relevant EEG research (Dalla Volta et al., 2018; Melloni et al., 2015; Ibáñez et al., 2013; Aravena et al., 2010) has failed to assess whether canonical markers of semantic integration and action planning during effector-specific (e.g., manual) movements are differentially sensitive to effector-congruent meanings (e.g., by contrasting MaVs with nMaVs). Promisingly, these shortcomings can be overcome by comparing how implicit semantic processing of different action-verb types during motor–language coupling modulates functionally critical event-related potentials (ERPs), such as the N400 and motor-related cortical potentials (MRCPs).

The N400 is a negative deflection peaking at approximately 400 msec after word presentation, with a centro-parietal scalp distribution (Lau, Phillips, & Poeppel, 2008). N400 modulations are a gold-standard signature of conceptual integration demands during processing of sentences, paired words, or even single words (Kutas & Federmeier, 2000, 2011; Lau et al., 2008). For their part, MRCPs comprise two primary (premovement) components (Castro, Díaz, & van Boxtel, 2005; Kutas & Donchin, 1980; Deecke, Scheid, & Kornhuber, 1969), which are typically maximal in central/fronto-central sites (Nguyen, Breakspear, & Cunnington, 2014; Smith & Staines, 2006) with sources in sensorimotor regions (Shibasaki & Hallett, 2006; Toma et al., 2002; Yazawa et al., 2000), namely, the readiness potential (RP), a slow-rising negativity that precedes movement onset by 100–400 msec in forced-choice tasks (Travers, Khalighinejad, Schurger, & Haggard, 2020; Haggard, 2008; Kutus & Donchin, 1980), and the motor potential (MP), a subsequent sharply negative peak occurring immediately before (~100 to −10 msec) the response (Shibasaki & Hallett, 2006; Deecke et al., 1969). Both these components are canonical indexes of increased sensorimotor cortical excitability during motor preparation (Shibasaki & Hallett, 2006; Toma et al., 2002; Yazawa et al., 2000; Deecke et al., 1969). Importantly, reduced N400 amplitude (Ibáñez et al., 2013; Aravena et al., 2010) and increased MRCP modulations (Melloni et al., 2015; Ibáñez et al., 2013; Aravena et al., 2010) have been shown to index compatibility between MaVs and (compatible) manual responses during explicit semantic tasks. Therefore, these ERPs emerge as suitable candidates to track movement-related and conceptual dimensions of implicit motor–language coupling.

Against this background, we employed high-density EEG (HD-EEG) to identify core neural signatures of semantic and motoric dynamics during motor–language coupling under implicit semantic processing conditions. We designed a lexical decision task requiring manual responses to MaVs, nMaVs, and nAVs. Building on previous findings (Ibáñez et al., 2013; Aravena et al., 2010), we predicted that MaVs would involve reduced N400 modulations relative to response-incongruent verb types (nMaVs and nAVs). Moreover, considering relevant evidence (Melloni et al., 2015; Ibáñez et al., 2013; Aravena et al., 2010), we predicted that MaVs would enhance the amplitude of MRCPs during manual response preparation. In addition, to ascertain the period of maximal semantic–motoric integration, we aimed to identify the temporal overlap between effector-specific modulations in both ERPs. Furthermore, on the basis of emergent results in the motor–language coupling literature (Klep et al., 2015), we performed an exploratory analysis of stimulus- and response-locked time–frequency modulations over the beta band. Briefly, with this novel approach, the present work seeks to illuminate the functional underpinnings of semantic and action dimensions in implicit motor–language coupling.

**METHODS**

**Participants**

The study was composed of 22 participants (13 women), a sample size that reaches a power of 0.93. All participants were adult native Spanish speakers who were enrolled in or had completed higher education programs. The sample had a mean age of 25.9 (SD = 4.7) years and an average of 17.1 years of education (SD = 3.5 years). All but two participants were exclusively right-handed. None of the participants reported a history of neuropsychiatric diseases or substance abuse, and all had normal or corrected-to-normal vision. Each participant read and signed an informed consent form before entering the study. All experiments and
procedures were performed in accordance with relevant guidelines and regulations of the Declaration of Helsinki. The study’s protocols were approved by the institutional ethics committee.

**Stimuli**

This study used 75 verbal stimuli belonging to four categories. Sixty items were real Spanish words, including 20 MaVs, denoting actions performed with the hands (e.g., *cut* [cortar]); 20 nMaVs, denoting actions performed with body parts other than the hands (e.g., *walk* [caminar]); and 20 nAVs, denoting cognitive or affective processes that do not involve bodily motion (e.g., *improve* [mejorar]). All these items appeared in infinitive form (most of them ending in -ar), which forced their interpretation as verbs. Finally, the task included 15 pseudowords (PsVs). These were created by choosing five real words from each list and replacing only one letter such that the resulting letter string, although phonotactically and graphotactically legal, did not represent a Spanish word (e.g., *collar*, *capinar*, *meborar*).

Psycholinguistic data for all stimuli were extracted from B-Pal (Davis & Perea, 2005), except for age-of-acquisition data, which were taken from a survey used in previous motor–language coupling research (Kogan et al., 2020; García & Ibáñez, 2016b). One-way ANOVA tests showed that all categories were similar in log frequency, $F(2, 57) = 0.39$, $p = .68$, familiarity, $F(2, 57) = 0.21$, $p = .82$, orthographic length, $F(2, 57) = 0.58$, $p = .56$, imageability, $F(2, 57) = 1.2823$, $p = .28$, and age of acquisition, $F(2, 57) = 1.01$, $p = .37$. As in previous motor–language coupling experiments, all items had four to eight letters; this guaranteed that all of them could be visualized in a single fixation, so that the time needed for their recognition would remain constant across categories (Lavidor & Ellis, 2002; Weekes, 1997). As expected, an additional test revealed significant differences in concreteness, $F(2, 57) = 18.974$, $p < .001$, among categories. A post hoc analysis (Tukey’s honestly significant difference test, $MSE = 0.30156$, $df = 57$) corroborated that nAVs were less concrete than both MaVs ($p < .001$) and nMaVs ($p < .001$), as expected; crucially, however, no significant differences were observed between the latter two categories ($p = .36$).

**Procedure**

Participants were tested individually in a dimly illuminated room. They sat comfortably at a desk, facing a laptop equipped with a 15.6-in. 16:9 HD (1366 × 768) LED backlight. As the HD-EEG electrodes were being placed on their scalps, they received oral instructions on the lexical decision task (instructions were then recapped on-screen before the start of the experiment).

The task was composed of three blocks of 75 trials. Each trial consisted of the presentation of a single stimulus belonging to one of the four categories (MaVs, nMaVs, nAVs, or PsVs). As in previous studies (García et al., 2019), PsVs were irrelevant for brain-signal analyses, but they served to ensure task compliance and attentional engagement by forcing linguistic decisional processes item after item. Each block included the same stimuli but in different pseudorandomized sequences (no more than three stimuli from the same category appeared in succession, and real words linked by meaning or form were separated by at least three trials). As in previous research (García et al., 2020), this allowed maximizing signal-to-noise ratio while retaining a strict control of psycholinguistic variables across conditions. Participants had to decide, as quickly and accurately as possible, whether the string constituted a real Spanish word or not by pressing the left arrow with the index finger or the right arrow with the middle finger of their right hand, respectively.

Each trial began with a fixation cross shown for a random period between 700 and 1000 msec at the center of the screen. Immediately afterward, the target item (verb or PsV) was shown until response and for a maximum of 2300 msec, with responses being allowed after the first 300 msec. A new trial was triggered upon the participant’s button press or if the trial’s overall duration elapsed without a response (Figure 1A). The use of a random period for the fixation cross minimized the chances of responses being driven by rhythmic motor habituation, with no biases for any lexical category. The fixation cross and the targets (font: Microsoft Sans Serif; color: white; size: 48; style: regular) were presented at the center of the screen against a black background. A custom-made script written in MATLAB’s (The MathWorks) Psychtoolbox was used to run the task and record behavioral responses (see below for details about simultaneous HD-EEG data acquisition). Before the actual task, 10 practice trials were presented with stimuli not included in the experimental blocks. The complete session for each participant lasted roughly 30 min.

**Behavioral Data Analysis**

Accuracy and RT data were analyzed via a linear mixed model, with Category as a fixed factor and Participant as a random factor. As in previous EEG studies on action-verb processing (Mollo et al., 2016; Pulvermüller et al., 2001), trials with RTs falling 2 SDs away from each participant’s mean were considered outliers and removed from analysis. All analyses were performed on MATLAB (R2015a) software.

**HD-EEG Data Acquisition and Processing**

HD-EEG data were acquired with a Biosemi ActiveTwo 128-channel system at 2048 Hz, resampled offline at 256 Hz. The EEGLAB (13.4.4b) toolbox (Delorme & Makeig, 2004) was used for offline processing and analysis. Three participants were excluded from analysis because of data acquisition problems. For all remaining participants, data were band-pass filtered during recording (0.1–100 Hz) and offline (0.5–40 Hz) to remove undesired frequency components. The latter
cutoff (40 Hz) was based on gold-standard recommendations to maximize data quality and temporal precision (Cohen, 2014) and to favor comparability with ERP studies on motor responses (Gentili et al., 2018; Fabi & Leuthold, 2017; Kadosh et al., 2007), action-verb processing (Sokoliuk, Calzolari, & Cruse, 2019; Casado et al., 2018), and other embodied language categories (García et al., 2020; Daly et al., 2019; Gentsch, Sel, Marshall, & Schütz-Bosbach, 2019). References for N400 and MRCP analyses were selected based on gold-standard recommendations (Luck, 2014) to maximize comparability with key studies in the field. N400 analyses were referenced to link mastoids and then re-referenced offline to the average of all electrodes, as done in previous studies on motor–language coupling (Aravena et al., 2010) and semantic effects at large (Kutas & Donchin, 1980; Lim et al., 2009). In addition, given the posterior topography of canonical N400 modulations (Kutas & Federmeier, 2011) and the location of our ROIs, especially in a dense-array setting like ours, this procedure guarantees a reference-independent estimation of scalp voltage for this ERP (Nunez & Srinivasan, 2006). Instead, signals for MRCP analyses were referenced to link mastoids, a common choice across embodied cognition studies targeting MRCPs and other response-locked ERPs (Wang et al., 2019; Melloni et al., 2015; Guan, Meng, Yao, & Glenberg, 2013; Santana & de Vega, 2013; Senderecka, Grabowska, Szewczyk, Gerc, & Chmłak, 2012; Aravena et al., 2010; Smith & Staines, 2006; Falkenstein, Hoormann, & Hohnsbein, 1999), including those which also tap on N400 effects by common average reference.
Importantly, mastoid reference also meets other gold-standard requisites (Luck, 2014) for MRCP analyses, as they offer a low electrical activity site away from the frontal ROIs capturing maximal effects (Nguyen et al., 2014; Smith & Staines, 2006) while avoiding hemispheric bias and reducing noise (e.g., Nguyen et al., 2014; Kutas & Donchin, 1980). All subsequent processing steps were identical for both components.

In line with reported procedures (Dottori et al., 2020; Fittipaldi et al., 2020; García et al., 2020; Vilas et al., 2019; Salamone et al., 2018), eye movements or blink artifacts were corrected with independent component analysis, and remaining artifacts were rejected offline from trials that contained voltage fluctuations exceeding ±200 μV. Bad channels were identified via visual inspection by two experts (S. C. and A. P.) following the exact same approach used by our team in previous EEG studies on semantic and motoric processes (Dottori et al., 2017, 2020; Vilas et al., 2019; Yoris et al., 2017; García-Cordero et al., 2016; Melloni et al., 2015; Ibáñez et al., 2010, 2013; Aravena et al., 2010). Once identified, such channels were replaced with statistically weighted spherical interpolation (based on all sensors), and then the variance of the signal across trials was calculated to guarantee stability of the averaged waveform (Courellis, Iversen, Poizner, & Cauwenberghs, 2016). HD-EEG data were then segmented offline into 1.5-sec epochs extending from 500 msec prestimulus to 1000 msec poststimulus for stimulus-locked segments in N400 analyses (where Time 0 corresponds to stimulus onset) and for hand response-locked segments in MRCP analyses (where Time 0 corresponds to motor response execution). The epochs were baseline-corrected, using a baseline of −300 to 0 msec for the N400 component and −600 to −300 msec for motor responses. Noisy epochs were rejected from the analysis using a visual inspection procedure as in previous studies (García-Cordero et al., 2016).

**Figure 2.** Electrophysiological markers of semantic integration (N400) effects during motor–language coupling. The topoplots show average modulations, locked to stimulus onset, within a canonical 350- to 450-msec window over the left (A), medial (B), and right (C) centro-posterior ROIs for MaVs (black line), nMaVs (red line), and nAVs (blue line). The baseline ranged from −300 to 0 msec before stimulus onset. Shadowed bars around the lines indicate standard errors. Vertical pink bars indicate significant differences at $p < .05$ (with a threshold of at least 10 contiguous points).
ERP Data Analysis

Analyses of N400 and MRCP modulations were performed considering only correct trials. As regards N400 analysis, upon removal of incorrect or artifactual trials, the number of remaining epochs averaged across participants was 47.4 (78.9%) for MaVs, 47 (78.3%) for nMaVs, and 49.6 (82.6%) for nAVs, $F(2, 54) = 0.74$, $p = .48$. In addition, after removal of incorrect trials and faulty segments, the number of MRCP-related epochs, averaged across participants, was 51.6 (86%) for MaVs, 52.3 (87.2%) for nMaVs, 53.5 (89.1%) for nAVs, and 35.5 (78.9%) for PsVs. The number of epochs analyzed did not differ significantly between categories, $F(2, 54) = 1.06$, $p = .35$.

N400 analyses were conducted over three canonical six-electrode ROIs, as reported in previous studies (Manfredi, Cohn, & Kutas, 2017; Schmidt-Snoek, Drew, Barile, & Agauas, 2015; Lim et al., 2009; Wu & Coulson, 2005), namely, a centro-posterior ROI (Channels A2, A3, A4, A19, D16, and B2), a left posterior ROI (Channels A5, A6, A7, D16, D17, and D28), and a right posterior ROI (Channels A32, B3, B4, B16, B18, and B19; Figure 2). Of note, these regions have proven sensitive to various semantic manipulations in action–language research, for example, expectancy or congruency between words, gestures, symbols, and images (Ibáñez et al., 2010; Lau et al., 2008). As regards MRCPs, given their typical markedly frontal distribution (Nguyen et al., 2014), these were examined in two symmetrical frontal ROIs, composed of six electrodes each: a right frontal ROI (Channels C4, C9, C10, C12, C13, and C14) and a left frontal ROI (Channels D4, C25, C26, C27, C31, and C32; Figure 3).

Point-by-point comparisons along the whole ERP signal were made via Monte Carlo permutation tests combined with bootstrapping, as done in previous works (Salamone et al., 2018; Yoris et al., 2017, 2018). This method circumvents the multiple-comparison problem and does not depend on multiple comparison corrections or Gaussian distribution assumptions (Nichols & Holmes, 2002). In addition, it avoids the selection of narrow a-priori windows for analysis, preventing circularity biases. The number of randomly simulations partitioning the data (permutations) was set to 5000. A $p$ value was thus obtained for each distance, but only those below .05 were considered significant. Then, a minimum extension of 10 consecutive points was used as the criterion to capture reliably significant effects (Mueller, Swainson, & Jackson, 2009). In line with standard practice for permutation test results in ERP research, results of significant windows are presented by reporting the limits of the corresponding time interval, the mean difference between condition, the associated $p$ value, and the effect size.

Exploratory Time–Frequency Analysis

In light of emerging research (Klepp et al., 2015), we explored stimulus- and response-locked oscillatory...
modulations in the beta band (12−30 Hz) within the same ROIs used for N400 and MRCP analyses, respectively. In line with this antecedent, we computed the time−frequency representation (TFR) on the targeted spectral power by convolving four-cycle complex Morlet wavelets with steps of 2 Hz for each single-trial epoch. Each stimulus-locked epoch ranged from −500 to 950 msec, whereas response-locked epochs ranged from −600 to 950 msec. A baseline subtraction was performed in every time point, and the ensuing difference was divided by the mean power of the baseline (set to −500 to −200 msec for stimulus-locked analyses and −600 to −300 msec for response-locked analyses). The single-trial TFRs were averaged for each condition (MaVs, nMaVs, and nAVs). The averaged TFRs were compared between each condition pair through a random pairwise permutation test across time over the same ROIs used for ERP analyses. The statistical significance threshold was set to \( p = .05 \).

RESULTS

Behavioral Results

Accuracy outcomes were high across all categories (MaVs = 96.9%, nMaVs = 97.1%, nAVs = 97.5%, PsVs = 87.3%). Statistical analysis revealed a main effect of Category. PsVs (\( \beta = .87, SE = 1.13, df = 66.39 \)) elicited more errors than all real-word categories (MaVs: \( \beta = 9.52, SE = 1.34, df = 63 \); nMaVs: \( \beta = 9.74, SE = 1.34, df = 63 \); nAVs: \( \beta = 10.12, SE = 1.34, df = 63 \); all \( ps < .001 \)), showing a canonical lexicality effect. However, there was no significant difference among the three categories of real verbs (all \( ps > .90 \); Figure 1B).

After removing outliers for RT analyses, the mean of remaining trials across participants was 19.4 (96%) for MaVs, 19 (94.6%) for nMaVs, 19 (95%) for nAVs, and 14.1 (96.1%) for PsVs. No trials were lost because of recording issues. RT results also revealed a main effect of category. PsVs (mean = 760 msec, \( \beta = 0.75, SE = 0.02, df = 66.39 \)) yielded slower responses than all real-word categories (MaVs: mean = 632 msec, \( \beta = -0.13, SE = 0.008, df = 63 \); nMaVs: mean = 617 msec, \( \beta = -0.14, SE = 0.008, df = 63 \); nAVs: mean = 626 msec, \( \beta = -0.13, SE = 0.008, df = 63 \); all \( ps < .001 \)), again showing a canonical lexicality effect. However, mean RTs did not differ between the real-verb categories (all \( ps > .30 \); Figure 1C).

N400 Results

Permutation analyses showed that N400 amplitudes were smaller for MaVs than nAVs in the right posterior ROI (from 328.1 to 410.2 msec, \( M_{\text{diff}} = 0.34, p < .05, d = 2.08 \)) and in the centro-posterior ROI (from 347.6 to 437.5 msec, \( M_{\text{diff}} = 0.39, p < .05, d = 0.79 \)). MaVs also elicited significantly smaller N400 amplitudes than nMaVs in a left posterior ROI (from 383.7 to 437.9 msec, \( M_{\text{diff}} = 0.34, p < .05, d = 0.99 \)). Significant differences between nMaVs and nAVs were found only in the right posterior ROI (from 347.7 to 418 msec, \( M_{\text{diff}} = 0.35, p < .05, d = 1.68 \)). For details, see Figure 2.

MRCP Results

As revealed by permutation analyses, MaVs elicited significantly more negative amplitudes than nAVs over the left frontal ROI (from −252.8 to 17.2 msec, \( M_{\text{diff}} = 1.18, p < .05, d = 1.10 \)) and over the right frontal ROI (from −248.4 to 71.9 msec, \( M_{\text{diff}} = 1.23, p < .05, d = 0.86 \)). Furthermore, MaVs elicited greater negative modulations than nMaVs before the onset of motor action, in both the left (from −252.3 to −139.1 msec, \( M_{\text{diff}} = 1.16, p < .05, d = 3.59 \)) and from −123.4 to −84.4 msec, \( M_{\text{diff}} = 0.83, p < .05, d = 1.82 \) and right (from −236.7 to −135.2 msec, \( M_{\text{diff}} = 1.35, p < .05, d = 5.52 \); and from −123.4 to −84.4 msec, \( M_{\text{diff}} = 0.84, p < .05, d = 1.71 \)) frontal ROIs. Differences between nMaVs and nAVs were solely observed over the right frontal ROI (from −84.4 to 5.5 msec, \( M_{\text{diff}} = 0.95, p < .05, d = 5.21 \); and from 25.0 to 75.8 msec, \( M_{\text{diff}} = 0.97, p < .05, d = 1.17 \)). For details, see Figure 3.

Exploratory Time−Frequency Results

Exploratory time−frequency analyses over the beta band revealed significant effects (\( p < .05 \)) over specific time segments. For stimulus-locked analyses, significant differences between MaVs and nAVs were observed only in the left posterior ROI, from 50 to 90 msec. The contrast between MaVs and nMaVs yielded significant differences only in the centro-posterior ROI (from 680 to 770 msec). Finally, differential modulations between nMaVs and nAVs were observed in both the centro-posterior (80−750 msec) and right posterior (120−760 msec) ROIs. For response-locked analyses, MaVs and nAVs yielded significantly different modulations over the left frontal ROI (in segments between −260 and 110 msec) and in the right frontal ROI (between −270 and 100 msec). Differences between MaVs and nMaVs reached significance over the left frontal ROI (in time segments comprised between −270 and 0 msec) and over the right frontal ROI (between −270 and 120 msec). Finally, differences between nMaVs and nAVs reached significance in the left frontal ROI (across successive time segments between −190 and 150 msec) and in the right frontal ROI (between −200 and 120 msec).

DISCUSSION

This study examined neural signatures of semantic and motoric dynamics during motor−language coupling in a shallow processing task. Relative to both nAVs and nMaVs, MaVs involved (a) reduced posterior N400 effects locked to stimulus onset and (b) greater modulations of frontal MRCPs locked to response execution. Moreover, only the former ERP discriminated between nMaVs and nAVs. Suggestively, effector-specific modulations for both components overlapped in a window of ∼380−440 msec after
word presentation and \( \sim 180 \text{ msec} \) before response execution. Such results provide new insights to understand the neural co-determinations of lexical and motoric processes, as discussed below.

**N400 Modulations: Semantic Dynamics during Motor–Language Coupling**

N400 results revealed two informative patterns. First, both MaVs and nMaVs elicited smaller N400 amplitude than nAVs. This resembles previous research (Dalla Volta, Fabbr-Destro, Gentilucci, & Avanzini, 2014) showing that ERP modulations for MaVs and nMaVs differed from those of nAVs in windows from roughly 220 to 400 msec. Given that N400 amplitude indexes semantic integration of linguistic material with active contextual signals (Lau et al., 2008), even during lexical access (Penolazzi, Hauk, & Pulvermüller, 2007), increased modulations for nAVs than for both action-verb categories probably signal the former’s greater incongruity with an ongoing physical response. In fact, as amply shown by speech/co-speech gesture paradigms, N400 amplitude is greater for incongruent than congruent pairings of word meaning and manual movement (Ibáñez et al., 2010, 2012, 2013; Kelly, Creigh, & Bartolotti, 2010; Cornejo et al., 2009; Holle & Gunter, 2007; Kelly, Ward, Creigh, & Bartolotti, 2007; Özyürek, Willems, Kita, & Hagoort, 2007; Wu & Coulson, 2005; Kelly, Kravitz, & Hopkins, 2004; for a review, see Amoroso et al., 2013).

More particularly, the N400 modulations yielded by MaVs were smaller than those of nMaVs. Comparable effects were reported in manual-response tasks showing enhanced beta suppression for MaVs relative to nMaVs (foot verbs) around 350 and 750 msec (Klepp et al., 2015), although the effect may also occur in later time windows (Klepp, Van Dijk, Niccolai, Schnitzler, & Biermann-Ruben, 2019). In addition, mismatch negativity effects around \( \sim 300 \text{ msec} \) can be modulated over fronto-central electrodes when a preceding action sound (e.g., footsteps) is incompatible with an effector-specific action word (e.g., kiss; Grisoni, Dreyer, & Pulvermüller, 2016). This suggests that, during motor–language coupling, N400 modulations are sensitive not only to general motor resonance but also to effector-specific dynamics.

This interpretation fits with evidence that N400 modulations capture compatibility between the hand positions evoked by words and those used for ongoing actions (Ibáñez et al., 2013; Aravena et al., 2010) and with hemodynamic (Kemmerer, Castillo, Talavage, Patterson, & Wiley, 2008; Assmus, Giessing, Weiss, & Fink, 2007; Buxbaum, Kyle, Tang, & Detre, 2006; Hamzei et al., 2003) and electromagnetic (Mollo et al., 2016; Pulvermüller, Shysterov, & Ilmoniemi, 2005) studies showing that motor–language coupling is indexed by modulation changes in different hubs of core semantic networks (e.g., posterior superior temporal cortex). In line with these antecedents, our results suggest that semantic components of motor–language coupling are sensitive to fine-grained (limb-specific) features and not only to coarse (limb-neutral) integrations of movement and meaning (Pulvermüller, 2013a, 2013b).

**MRCP Modulations: Action Dynamics during Motor–Language Coupling**

MRCP results also revealed two key patterns. First, compared with nAVs, MaVs yielded greater frontal modulations since 250 msec before until 75 msec after response execution, with nMaVs yielding similar effects in the latter part of that segment (from \( -85 \) until 75 msec). This extended pattern encompasses two critical ERPs, namely, the RP and the MP (Haggard, 2008; Shibasaki, Barrett, Halliday, & Halliday, 1980), even extending onto postmovement MRCPs (Shibasaki & Hallett, 2006; Deecke et al., 1969). Crucially, modulations of these components indexes heightened motor-system recruitment not only during action preparation and execution (Siemionow, Yue, Ranganathan, Liu, & Sahgal, 2000), respectively, but also during relevant higher-order processes, such as action imagery (Moran, Campbell, Holmes, & MacIntyre, 2012; Niazi et al., 2011; Sharma, Pomeroy, & Baron, 2006) and observation (Bozzacchi, Spinelli, Pitzalis, Giusti, & Di Russo, 2015). More particularly, RP amplitude discriminates between nAVs and diverse (mouth, leg, and hand related) action verbs, with additional modulations for each category around the MP window (Dalla Volta et al., 2018). In line with this evidence, our results suggest that verbs that evoke bodily motion, relative to those that do not, prime successive action-related cortical mechanisms.

Yet, when compared with nMaVs, MaVs elicited a greater negative deflection only between \( -250 \) and \( -85 \text{ msec} \), thus being mainly restricted to the canonical window of the RP (Haggard, 2008; Libet, Wright, & Gleason, 1993). Interestingly, previous motor–language coupling studies assessing response-locked activity (Melloni et al., 2015; Ibáñez et al., 2013; Aravena et al., 2010) reported MP effects without RP modulations. However, because these studies compared different subtypes of MaVs (denoting open and closed manual actions) following congruent hand shapes (i.e., open or closed), they were blind to the integration of same- and different-effector information. Accordingly, our findings suggest that RP modulations during motor–language coupling are particularly sensitive to effector-specific effects. In other words, this ERP seems to index the match between the limb implied by a verb and the one set in motion—a synergy that is also captured in hemodynamic (Kemmerer et al., 2008; Assmus et al., 2007; Buxbaum et al., 2006; Hamzei et al., 2003) and electromagnetic (Mollo et al., 2016; Pulvermüller et al., 2005) dimensions.

It is worth noting that such effector specificity was traced by the RP, as opposed to the MP. Unfortunately, the only previous EEG study on response-locked modulations during processing of mouth-, leg-, and hand-related action verbs did not compare among these categories, as it...
which motor-processing stages are engaged in motor
preparation until motor execution proper. Conversely,
results indicate that MRCPs would have a graded sensitivity
in the unfolding of motor–language coupling, with the RP
capturing both coarse-grained (general motor resonance)
and fine-grained (effector-specific) embodied dynamics
and the MP tracking only the former effects. Note, incidentally,
that this pattern was corroborated even when considering
a much longer window.

Theoretical Implications

The results above have a number of theoretical implications. First, they indicate that both the semantic and action
dimensions of motor–language coupling are typified by
effector-sensitive neural modulations. In other words,
both the conceptual integration and movement-related
stages of the phenomenon are influenced by fine-grained
(limb-specific) features of the word’s meaning and the
to-be-executed action. Note that previous studies have
reached similar conclusions by comparing different categories of action verbs relative to nAVs (but not with each other)
in manual-response tasks (Dalla Volta et al., 2018)
or by examining compatibility effects between subsets of
MaVs and different hand responses (Melloni et al., 2015;
Ibáñez et al., 2013; Aravena et al., 2010). Therefore, all this
previous evidence proves blind to limb-specific effects.
Conversely, in line with recent works (Klepp et al., 2015,
2019), our study indicates that motor–language coupling
is characterized by not only gross motor resonance and
effector-position effects but also the overlap between
the body part denoted by the verbal material and the
one used to perform an action.

In addition, to our knowledge, our study is the first to
dissociate specific MRCPs underlying these two embodied
phenomena. General motor resonance effects seem to cut
across the RP (Travers et al., 2020; Haggard, 2008; Kutas
& Donchin, 1980) and the MP (Shibasaki & Hallett, 2006;
Denecke et al., 1969), suggesting that coarse-grained
embodied dynamics extend from distal stages of motor
preparation until motor execution proper. Conversely,
effector-specific effects were indexed only by the RP, indicating
that fine-grained embodied dynamics are short-lived
and vanish as motor execution becomes imminent.
This suggests that semantic specificity might influence
which motor-processing stages are engaged in motor–
language coupling, opening new avenues for research.

Such findings allow extending the Hand-Action-Network
Dynamic Language Embodiment model (García & Ibáñez,
2016a), the most systematic account of motor–language
coupling phenomena to date. Succinctly, this framework
postulates that ongoing manual action processes can be modu-
lated in predictable ways if accompanied by verbal materials
evoking hand-specific meanings. In particular, the model
postulates that the temporal lag between word onset and action
execution can determine whether the latter will be facilitat-
ed (Afonso et al., 2019; Kelly et al., 2010; Springer &
Prinz, 2010; Masson, Bub, & Warren, 2008; Lindemann,
Stenneken, Van Schie, & Bekkering, 2006; Tucker & Ellis,
2004) or delayed (García & Ibáñez, 2016b; Spadacenta,
Gallese, Fragola, & Mirabella, 2014; Mirabella, Iaconelli,
Spadacenta, Federico, & Gallese, 2012; Bergen, Lau,
Narayan, Stojanovic, & Wheeler, 2010; Sato, Mengarelli,
Riggio, Gallese, & Buccino, 2008). However, the model is
exclusively rooted in behavioral evidence, and it fails to dis-
entangle the time course of motoric and semantic dynamics
underlying its target phenomenon. In this sense, the Hand-
Action-Network Dynamic Language Embodiment model
can be refined by incorporating core findings of our study:
Effectorspecific motor–language coupling modulates electrophysiological signatures of both action-related and
semantic integration processes, and these feature their
own, partially overlapping, temporal and functional dy-
namics (Figure 4).

On the basis of present results, the diagram in Figure 4
maps the temporal span in which semantic and motoric
processes coalesce during motor–language coupling.
First, effector-specific N400modulations, signaling se-
matic integration, were observed ~380–440 m sec after
stimulus onset. Second, considering that response laten-
cies fell in the order of ~620 m sec and that response-
locked MRCPs yielded effector-specific effects starting
~250 m sec before action execution, motoric stages of
the phenomenon began almost in synchrony with
semantic ones and extended for ~200 m sec thereon.
Accordingly, both stages seem to co-occur over a window
of ~380–440 m sec after stimulus presentation. This time
span, we propose, might signal the moment of conver-
gence between the two contributing streams shaping
effector-specific motor–language coupling—a conclusion
that simply cannot be derived from previous ERP studies
comparing MaVs and nMaVs with nAVs, but not with each other (e.g., Dalla Volta et al., 2018). Together with research
on the time course of MaV-induced muscular activity (Da
Silva et al., 2018; Frak et al., 2010), such a finding further
illuminates the complex embodied dynamics linking
meanings and actions.

Moreover, previous motor–language coupling studies
relied on explicit processing tasks, as they required particip-
ants to semantically categorize stimuli (Dalla Volta et al.,
2018) or indicate when they were understood (Melloni
et al., 2015; Ibáñez et al., 2013; Aravena et al., 2010).
Therefore, they overlooked a key requisite to detect pri-
mary embodied effects: their emergence under implicit
semantic conditions (García et al., 2019; Mollo et al.,
2016; Hauk, Shtyrov, & Pulvermüller, 2008; Kiefer et al.,
2008; Pulvermüller et al., 2001). Unlike those studies,
our experiment was based on a lexical decision task, so
that no explicit attention to meaning was required. As
proposed for other embodied phenomena (García et al.,
2016a), we consider that the Hand-Action-Network
Dynamic Language Embodiment model can be refined by incorporating core findings of our study.
In this study, motor variables failed to show a significant effect on the independent variable at large. Indeed, as shown in Figure 4, this reminds us that a null effect on one dimension typically yields null motor effects that broadly echoed the time windows capturing our main ERP results. Reduced beta oscillations were observed for all three contrasts starting at roughly the same time as MRCP effects and covering the entirety of their duration. However, they also extended beyond the window capturing such ERP effects. This might speak to the different temporal dynamics of motor–language coupling in each analytical dimension, a possibility that should be explored in new studies designed to compare both approaches.

Limitations and Avenues for Further Research

Notwithstanding its contributions, our study presents some limitations. First, our sample size was modest. Although the present number of participants provided good statistical power and although most previous EEG studies on motor–language coupling have employed similar or even smaller Ns (Dalla Volta et al., 2018; Mollo et al., 2016; Mellon et al., 2015), replications with a larger group would be desirable. In addition, although recent works with comparable experimental designs (Dalla Volta et al., 2018) have used similar numbers of stimuli, it would be ideal to extend this study with more items per category. Moreover, it would be useful to assess whether the modulations observed during our present shallow-level task manifest similarly when explicit semantic processing is required—as observed in behavioral motor–language coupling experiments (Afonso et al., 2019). Finally, although the question addressed in this study pertained to the temporal dimension, future studies should aim to complement our approach with high-spatial-resolution methods, as done in recent magnetoencephalography (García et al., 2019; Klepp et al., 2014, 2015, 2019) and intracranial (García et al., 2020; Ibáñez et al., 2013) EEG experiments.
Conclusion

This is the first ERP study assessing the semantic and action stages of motor–language coupling during a shallow-processing task. Our core finding is that the execution of a manual response after processing of MaVs (as compared with nMaVs) involves reduced N400 modulations and increased MRCP effects, signaling semantic and motoric dynamics, respectively. Such modulations overlapped in a window of ~380–440 msec after word presentation and ~180 msec before response execution, a pattern that motivates new insights on when both stages reach maximal convergence during effectorspecific motor–language coupling. Moreover, we showed that the RP is sensitive to both coarse-grained and effector-specific dynamics, whereas the MP only captures the former modulations, suggesting different functional roles for each ERP. Further work in this direction can hone our understanding of how language and movement coalesce in neural time.

Acknowledgments

The content is solely the responsibility of the authors and does not represent the official views of the National Institutes of Health, Alzheimer’s Association, Rainwater Charitable Foundation, or Global Brain Health Institute.

Reprint requests should be sent to Adolfo M. García, Universidad de San Andrés & CONICET, Vito Dumas 284, B1644BID Victoria, Buenos Aires, Argentina, or via e-mail: adolfomartingarcia@gmail.com.

Author Contributions

Sabrina Cervetto: Data curation; Formal analysis; Investigation; Software; Writing original draft. Mariano Díaz-Rivera: Investigation; Writing original draft. Agustin Petroni: Data curation; Formal analysis. Agustina Birba: Formal analysis; Writing review & editing. Miguel Martorell Caro: Data curation. Lucas Sedeño: Writing review & editing. Adolfo M. García: Conceptualization; Investigation; Writing–original draft. Agustín Ibáñez: Funding acquisition; Writing review & editing. Adolfo M. García: Conceptualization; Investigation; Methodology; Project administration; Resources; Validation; Writing–original draft; Writing–review & editing.

Funding Information

This work was supported by CONICET, FONCYT-PICT (https://dx.doi.org/10.13039/50110002663; grant numbers: 50110006668); grant number R01AG057254; an Alzheimer’s Association (https://dx.doi.org/10.13039/100000957) grant (SG-20-725707-ReDLat); the Rainwater Foundation; and the Global Brain Health Institute.

Diversity in Citation Practices

A retrospective analysis of the citations in every article published in this journal from 2010 to 2020 has revealed a persistent pattern of gender imbalance: Although the proportions of authorship teams (categorized by estimated gender identification of first author/last author) publishing in the Journal of Cognitive Neuroscience (JoCN) during this period were M(an)/M = .408, W(oman)/M = .335, M/W = .108, and W/W = .149, the comparable proportions for the articles that these authorship teams cited were M/M = .579, W/M = .243, M/W = .102, and W/W = .076 (Fulvio et al., JoCN, 33:1, pp. 3–7). Consequently, JoCN encourages all authors to consider gender balance explicitly when selecting which articles to cite and gives them the opportunity to report their article’s gender citation balance.

Data and Code Availability Statement

All experimental data and the scripts used for their collection and analysis are available online (Garcia, 2020).

REFERENCES


Cervetto et al. 1427