Psychiatry Research: Neuroimaging 191 (2011) 68–75

Contents lists available at ScienceDirect

Psychiatry Research: Neuroimaging

journal homepage: www.elsevier.com/locate/psychresns

High contextual sensitivity of metaphorical expressions and gesture blending: A video event-related potential design

Agustín Ibáñez[a,b,c,d,*], Pablo Toro[e,g], Carlos Cornejo[g], Hugo Hurquina[a,b], Facundo Manes[a,b], Matthias Weisbrod[f], Johannes Schröder[e]

[a]Laboratory of Experimental Psychology and Neurosciences, Institute of Cognitive Neurology (INECO), Buenos Aires, Argentina
[b]Institute of Neuroscience, Favaloro University, Buenos Aires, Argentina
[c]Career of the National Scientific and Technical Research Council (CONICET), Buenos Aires, Argentina
[d]Neuroscience Laboratory, University Diego Portales, Santiago, Chile
[e]Section of Geriatric Psychiatry, University of Heidelberg, Heidelberg, Germany
[f]Section of Experimental Psychopathology, University of Heidelberg, Heidelberg, Germany
[g]P. Catholic University of Chile, Santiago, Chile

A R T I C L E I N F O

Article history:
Received 7 November 2008
Received in revised form 5 August 2010
Accepted 12 August 2010

Keywords:
N400
LPC
Contextual blending
Metaphor
Figurative language
Action sequences
Video clips

A B S T R A C T

Human communication in a natural context implies the dynamic coordination of contextual clues, paralinguistic information and literal as well as figurative language use. In the present study we constructed a paradigm with four types of video clips: literal and metaphorical expressions accompanied by congruent and incongruent gesture actions. Participants were instructed to classify the gesture accompanying the expression as congruent or incongruent by pressing two different keys while electrophysiological activity was being recorded. We compared behavioral measures and event related potential (ERP) differences triggered by the gesture stroke onset. Accuracy data showed that incongruent metaphorical expressions were more difficult to classify. Reaction times were modulated by incongruent gestures, by metaphorical expressions and by a gesture–expression interaction. No behavioral differences were found between the literal and metaphorical expressions when the gesture was congruent. N400-like and LPC-like (late positive complex) components from metaphorical expressions produced greater negativity. The N400-like modulation of metaphorical expressions showed a greater difference between congruent and incongruent categories over the left anterior region, compared with the literal expressions. More importantly, the literal congruent as well as the metaphorical congruent categories did not show any difference. Accuracy, reaction times and ERPs provide convergent support for a greater contextual sensitivity of the metaphorical expressions.

1. Introduction

Human interaction in a natural context implies a continuity of sense among actions (i.e., body language) and communication where gestures are combined with the spoken language and diverse contextual clues in a dynamic, temporal sequence (Cosmelli and Ibáñez, 2008; Ibáñez and Cosmelli, 2008). Verbal utterances are in fact tightly interwoven with gestural information and expressive context in ordinary conditions. This feature of language is particularly evident in figurative language, where a considerable amount of contextual information is required for its understanding (Gibbs, 1994; Giora, 2003). In the case of metaphorical expressions, evidence shows that context and information beyond the lexical content are critical for comprehension (Pynte et al., 1996; Coulson and Van Petten, 2002). Thus, in opposition to the literal-first hypothesis, which considers metaphorical meaning as activating only after the contextual failure of the necessarily first literal meaning, many language researchers consider that metaphorical meaning can be as available as the literal one, depending on the expressive context (Giora, 2003). Context is also increasingly conceived as having a much more central role in meaning construction than simply serving as a test of viability for the literal interpretation of an expression (Coulson, 2006). Central for the traditional studies of metaphor is the distinction between the tenor (or "topic") and the vehicle of a metaphor (Cornejo, 2004, 2007). Tenor or topic is what is described by the metaphor, while vehicle is the term used to describe the topic. So, in the metaphorical expression 'Physicians are gods', 'physicians' is the topic, which is described by means of the vehicle 'gods'. We consider for our study short metaphoric sentences (called novels or unfamiliar; Pynte et al., 1996) with final words comprising the metaphoric meaning (i.e., "Those fighters are LIONS").

* Corresponding author. Laboratory of Experimental Psychology and Neurosciences, Institute of Cognitive Neurology (INECO) & CONICET, Castex 3293 (CP 1425), Buenos Aires, Argentina. Tel./fax: +54 11 4807 4748.
E-mail address: aibanez@neurologiacognitiva.org (A. Ibáñez).
URL: http://www.neurologiacognitiva.org (A. Ibáñez).

0925-4927/$ – see front matter © 2010 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.pscychresns.2010.08.008

In recent years the technique of event related potentials (ERPs) has approached the study of multimodal contextual blending of stimuli (Hurtado et al., 2009), figurative language and action sequences (i.e. videos) offering a more ecological approximation to the study of language and meaning. A component that has been studied intensely in relation to the more ecological clues of language is the N400. The N400 is an ERP characterized by a negativity generated about 400 ms after the presentation of semantically anomalous information (Kutas and Hillyard, 1980). The N400 modulation has been shown to be simultaneously context-sensitive and automatic (Ibáñez et al., 2006, 2010). Another component, the so-called 'late positive complex' (LPC, sometimes elicited together with the N400), has been related to a process of re-analysis of the incongruent situation produced by the inconsistent meaning (Sitnikova et al., 2003). The N400 has shown more amplitude in so far as metaphorical phrases are concerned (Coulson and Van Petten, 2002), as well as a reduction of its negativity when the metaphors are inserted in a previous congruent context with metaphorical sense (Pynte et al., 1996). Studies done with videos in other areas have shown a modulation of the N400 in response to incongruent or unexpected action (Sitnikova et al., 2003; Reid and Striano, 2008). Other studies have shown that the gesture in itself modulates the N400 component, producing more amplitude when the gesture is incongruent with the linguistic context (Gunter and Bach, 2004; Holle and Gunter, 2007). The N400 and LPC were modulated by videos of incongruent gestures (Neville et al., 1997; Kelly et al., 2004; Wu and Coulson, 2005; Kelly et al., 2007; Özyüreck et al., 2007; Wu and Coulson, 2007). All these studies from different perspectives suggest that a continuity of multimodal sense arises from the context, the action and the language, advocating that both N400 and LPC are sensitive to this continuity.

The N400 and the LPC have typically been reported in response to linguistic and static stimuli. Nevertheless, processing of extralinguistic information has also been investigated in terms of integrating information from pictures to previous context (Barrett and Rugg, 1990; Ganis et al., 1996; McPherson and Holcomb, 1999; Federmeier and Kutas, 2001; West and Holcomb, 2002). Recent studies have been reported of N400 and LPC components elicited by meaningful but non-linguistic stimuli such as faces (Hannula et al., 2006), objects (Ganis and Kutas, 2003), music (Daltrozzo et al., 2010); pictures (West and Holcomb, 2002; Guerra et al., 2009); gestures (Proverbio and Riva, 2009) and hand actions (Aravena et al, 2010). In addition, N400 and LPC are triggered by dynamic events without a time-critical window (e.g., videos of real life situations: Sitnikova et al., 2003; or video gestures: Cornejo et al., 2009). The N400/LPC reported from dynamic events seems to have a more left and anterior topography and a greater latency compared with the classical N400/LPC elicited by static and linguistic stimuli. Since it is not possible to clearly identify both kinds of ERPs as the same component with the same neural generators (N400/LPC elicited by static and linguistic stimuli and the N400/LPC elicited by dynamic and non-linguistic stimuli), we prefer to call these effects N400-like and LPC-like components.

In the case of the metaphor, there is a possibility that the influence of contextual clues, particularly gestures, presented in a dynamic and synchronized sequence with linguistic expressions modulates the N400 and LPC component. To our knowledge, only a study with ERPs has come close to this question. An N400 and an LPC modulation were related to video clips showing incongruent gestures with metaphorical meaning (Cornejo et al., 2009). This study is the first to relate the ERP investigation about the coordination between action sequences (gestures) and figurative language (metaphors). Nonetheless, this study presents limitations because of the absence of contrasts (between literal and metaphorical stimuli) that reduce the conclusions about the processing of metaphors and related topics (contextual sensitivity, figurative language and gesture integration).

Our study evaluates the contextual coordination between gesture and literal/figurative language. Previous research has suggested that multimodal and co-occurring speech and gestures are integrated by the brain simultaneously into a preceding sentence context (Özyüreck et al., 2007). Understanding an utterance implies that the brain does not restrict itself to language information alone but also integrates semantic information conveyed through other modalities such as co-speech gestures (Ozyürek and Kelly, 2007; Willems et al., 2009). Previous research has shown that the stroke phase of gesture in particular conveys the meaning of a gesture (McNeill, 1992). Speakers produce the stroke simultaneously with the relevant speech segment (e.g. Levelt et al., 1985). Recently, this meaningful property of gesture stroke has been used in neuroscience. For example, functional magnetic resonance imaging (fMRI) studies have shown that unimodal (gesture only) and multimodal integration of gestures and speech increases activation in the classical left hemispheric language areas (Kircher et al., 2009). Although scarcer, recent ERP studies have triggered the N400 and LPC with the gesture stroke onset. For example, Özyüreck et al. (2007) used ERPs to assess the integration of speech and gesture simultaneously triggered by word onset and static gesture stroke onset. Despite the difference in modality and in the specificity of meaning conveyed by spoken words and gestures, the latency, amplitude, and topographical distribution of both word and gesture onset mismatches were found to be similar, indicating that the brain integrates both types of information similarly and simultaneously. Moreover, although the dynamic presentation of stimuli reduces early visual components such as P1/N1 due to a lack of discrete transient visual events separated by time (Sitnikova et al., 2003; Wu and Coulson, 2005), it seems to preserve similar effects of N400 and LPC (Sitnikova et al., 2003; Cornejo et al., 2009). Using the stroke of the gesture as a time-locking event allows the investigation of relevant semantic coordination processes drawn from a temporally dynamic event stroke gesture. In a previous report of co-gesture speech paradigms, N400 and LPC components elicited by gesture stroke onset resulted in strong modulation of congruent vs. incongruent stimuli compared with the N400/LPC component elicited by word onset (Cornejo et al., 2009). Since stroke occurs 200–400 ms after the relevant word onset, it constitutes a better temporal window compared with word onset because it implies the temporal integration between sentence meaning and gesture meaning. In brief, although N400 has been more frequently studied with word onset, recent neuroimaging and electrophysiological data suggest that speech and gesture convey related and similar information and that stroke expresses the meaning of the utterance and can be effective for studying the coordination between gesture and language in ERP paradigms.

With this in mind, the following questions arise: Does the gesture–expression interaction influence the metaphorical utterance differentially? (Thus, it might be in the case of the contextual sensitivity rather than in the literal expression). And therefore, the modulation of the N400/LPC-like components… would this be affected differentially by congruent/incongruent gestures of metaphorical expressions compared with literal ones? Our study proposes to respond to these questions to throw some light on the hypothesis of the linguistic nature of the metaphor. We therefore hypothesized that if the metaphorical expressions are highly dependent on the context of gestures/actions, the differences between the congruent/incongruent conditions would be larger in the metaphoric than the literal condition. This would occur due to the contextual blending of gesture and expressions that should act with more effect for figurative sentences, thereby incrementing the differences between congruent and incongruent actions for this condition. On the other hand, if the contextual blending of gesture and expressions had no effect of greater sensitivity in figurative expressions, then a differential effect would not be observed in the metaphorical expressions; the N400 modulation would be similar to the literal congruent condition. Otherwise, both categories (congruent and incongruent) should generate greater negativity than in the literal ones.

To test the aforementioned hypothesis, we constructed a paradigm with four types of video clips: literal and metaphorical expressions accompanied by congruent and incongruent gesture actions. ERPs were recorded while participants viewed short video clips of a person uttering metaphorical and literal expressions co-produced with bodily gestures that were either congruent or incongruent with the meaning of these expressions on the clips. In this way the effects of gestural congruency as well as expression type were observed.

2. Methods

2.1. Subjects

Twelve undergraduates from Heidelberg University (three of whom were excluded from the study due to a low number of remaining trials per category — fewer than 25 after artifact rejection) voluntarily participated in the experiment and signed a written consent after complete description of the study to the subjects. The study was approved by the ethics committee of the University of Heidelberg. All of the nine participants (five women) were right-handed, with an average age of 23 years (S.D. = 3.4). None of them had any personal history of neurological or psychiatric disease. All of them were native German speakers.

2.2. Procedure

2.2.1. Stimuli validation

Our paradigm was constructed following successive steps of validation through verbal stimuli: construction of sentences based on previous studies of co-gesture speech processing, selection of categories based on previous studies, frequency studies, cloze probability rating of target words and metaphorical ends, last word testing of the frequency of use, and the control of final words in content and length (see Supplementary data, Section 1, for more details). A rating study was carried out in which 55 students classified the linguistic expressions according to their degree of metaphoricity (with a minimum of 0% and a maximum of 100%). In order to validate the lists of metaphors that would be used later in the video study, each of the lists was separated into two sub-groups (Literal 1 and Literal 2; and Metaphoric 1 and Metaphoric 2). The mean (M, hereafter) score of the metaphorical sentences was ($M = 88.01$, S.D. = 20.32); these were clearly distinguished from the literal ones ($M = 9.24$, S.D. = 92.53; $F(41, 14) = 8.47$, $P < 0.01$). Post hoc comparisons (HSD test, Bonferroni corrected; MS = 362.51, d.f. = 54.00) indicate that congruent and incongruent sub-groups of metaphorical ($P = 0.567$) and literal expressions ($P = 0.840$) did not present significant differences in the rating averages.

The final sample of verbal expressions was used to record the video clips of an actress uttering them with congruent and incongruent gestures. Half of the video streams contained gestures that were congruent with the expressions and the other half contained incongruent gestures. All expressions had the basic structure: "Diese X sind Y" (Those X are Y). Several steps were performed in order to control the quality of the videos: actor training in timing, recording settings, clips' average duration, co-gesture speech timing, category differences on gesture starting time and stroke and word-gesture time differences (see Supplementary data, Section 1). For the EEG Experiment, we removed nine videos (five incongruent, four congruent) in which accuracy levels were low (below 70%). After the validation process, 176 (44 clips for each category) video streams of gestures were recorded following previous studies (Kelly et al., 2004; Ozyürek and Kelly, 2007; Cornejo et al., 2009). Before the EEG recording we performed a pilot study with reaction time measures (RTs hereafter). The same actress performed all video clips. The sentences were delivered via headphones, and the EEG measurements took place at the Section of Experimental Psychopathology at Heidelberg University.

2.2.2. Experimental design

Participants were instructed to classify the gesture accompanying the expression as congruent or incongruent pressing two different keys while electrophysiological activity was being recorded. The videos were recorded with a digital Mini-DV Panasonic camera and were later displayed with the software Presentation (Neurobehavioral Systems) on a PC. The participants were instructed to pay attention to several streams of short video clips, watching and listening to a person saying different expressions. The task consisted on keeping attention on the videos and indicating whether each one was congruent (if co-gesture and speech convey the same meaning) or incongruent (if the co-gesture and speech are not related semantically). The accuracy of the responses and RTs were measured. The production of the gesture began approximately in the middle of the first word of the expression (literal or metaphoric), having its stroke when pronouncing the final word. For instance, when the actor uttered *Diese Personen sind glänzend* ("Those people are enlightened"), he moved his right hand, opening his fingers from the head upwards, indicating open-mindedness. Regarding the incongruent gestures, they had the same timing as the congruent ones, but they suggested the opposite meaning of the sense of the expression (literal or metaphoric — see Supplementary data, Section 2, for more details and examples).

The final experimental stimulus presentation consisted of 176 videos (44 clips for each category). The task (Fig. 1) involved the classification of the video clips and began with a brief explanation of how each category should be assigned to each response key (congruent or incongruent). After a brief practice (five video clips), the trials were presented pseudo-randomly (the sequence was counterbalanced controlling for not more than two trials of the same category based on the type of gesture or type of word). Each trial began with a "+" at the center of the screen for ocular fixation (250 ms), followed by the video stream ($M = 5.4$, S.D. = 1.3) and subsequent classification task. In this task participants had to press key 1 or 2 (congruent or incongruent) with the right hand once they comprehended the utterance. After the response, an interval of 1500 ms preceded the next trial. There were no breaks during the task.

2.3. Electrophysiological recordings

Signals were recorded on-line using a 64-channel amplifier (Neuroscan Synamp; Neuroscan) and the Brain Vision software. Analog filters were 0.1 and 100 Hz. A band pass digital filter between 0.5 and 30 Hz was later applied to remove unwanted frequency components. Signals were sampled at 1000 Hz and then re-sampled to 250 Hz. During recordings the reference was set by default to vertex but then was re-referenced off-line to link mastoids. Two bipolar derivations were designed to monitor vertical and horizontal ocular movements (EOG). Continuous EEG data were segmented from 200 ms before the stimulus to 800 ms after it. All segments with eye movement contamination were removed from further analysis, using automatic (Gratton, Coles, and Donchin method for removing eye-blink artifacts) and visual procedures. Artifact free segments were averaged to obtain the ERPs. In order to assure an equivalent number of trials in each category, all responses were included in the ERP analysis. Waveforms were averaged separately for each experimental condition: (a) congruent gesture with literal expression; (b) congruent gesture with metaphorical expression; (c) incongruent gesture with literal expression; and (d) incongruent gesture with metaphorical expression. We triggered the onset of the stroke of the accompanying gesture (see Kita and Özyürek, 2003; Ozyürek and Kelly, 2007; Willems et al., 2007; Cornejo et al., 2009). The time frames when the stroke of the gesture occurred were further indicated

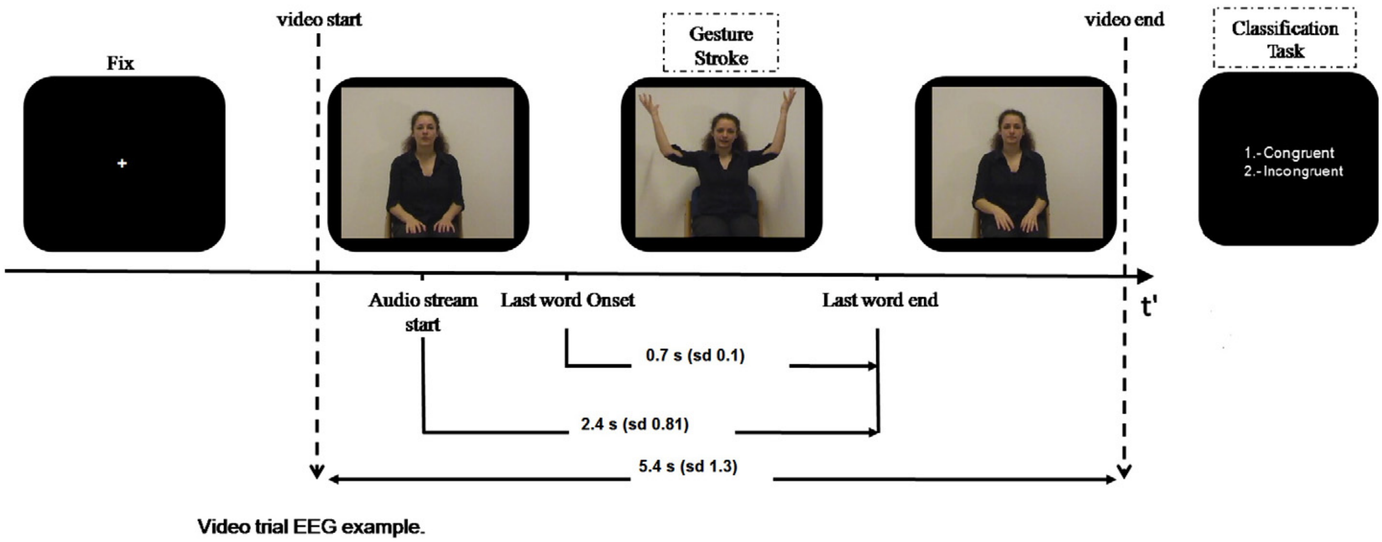

Fig. 1. Experimental procedure. In this example of a video with a congruent gesture and with a metaphorical expression, the actress says, “Diese Ärger sind Vulcane” (Those angers are volcanoes; like an exploding volcano). The gesture begins after the spoken expression has been initiated and before the last word is said. At the end of each video clip, the participants had to do a classification task indicating whether the gesture was or was not congruent with the expression. The values in the figure indicate averages and standard deviations in seconds.

in a frame by frame edition of the video stream. In the stimuli the gesture strokes are also located slightly after the word onset (see Supplementary data, Construction of video clips section). However, considering the starting time of gestures are considered as a whole, this relation reverses, i.e., the preparatory phase of the gestures anticipates the corresponding co-expressive speech. This anticipation of gesture has been previously demonstrated (McNeill, 1992). See Cornejo et al. (2009), and Supplementary data (Section 3), for more details of this procedure. The software Analyzer from Brain Vision was used for signal pre-processing, the EEGLab toolbox and T-BESP software (http://neuro.udp.cl/software) were used for EEG off-line processing. Matlab software was used for statistical analysis.

2.4. Statistical analysis

Behavioral measures (accuracy and RTs) were grouped into four categories (two × two design: congruent and incongruent gestures; and literal and metaphorical expressions). For ERP analysis five regions of interest (ROI) were used to represent and analyze the scalp topography of the ERP components as recommended for dense arrays. Groups of electrodes were collapsed into specific regions in order to avoid loss of statistical power (Oken and Chiappa, 1986) and the regions were based on the N400/LPC scalp topography reported in previous studies: left anterior (LA), right anterior (RA), central midline (Cz), left posterior (LP) and right posterior (RP; see Supplementary data, Section 4). Two ERP windows (N400 window: 300–500 ms; and LPC window: 500–700 ms) for all conditions were selected. For each comparison an analysis was effected through repeated measurements with the following three within-subject factors: Gesture (congruent vs. incongruent); Expression (literal vs. metaphoric); and ROI (LA, RA, CM, LP and RP). For both the behavioral and ERP analyses, a repeated measures analysis of variance was done (RM-ANOVA). For significant effects and interactions ($P=0.05$), averages and contrasts were calculated with a post hoc Tukey HSD test. This test is used for testing the significance of unplanned pairwise comparisons and is based on a studentized range distribution, q. The HSD test is performed in order to preserve the family-wise type I error as it applies simultaneously to the set of all pairwise comparisons. Univariate comparisons were done when necessary. Results were corrected with the Greenhouse–Geisser and Bonferroni methods to adjust the univariate output of repeated measurements of ANOVA for violations of the compound symmetry assumption.

3. Results

3.1. Behavioral measures

3.1.1. Accuracy

The behavioral measures revealed high levels of accuracy (literal congruency, 89.11%; literal incongruency, 83.66%; metaphorical congruency, 89.00%; and metaphorical incongruency, 76.43%). A significant main effect of *type of gesture* ($F(1, 8)=41.65$, $P<0.01$) but not *type of expression* ($F(1, 8)=4.15$, $P=$ n. sig.) was observed. Moreover, a significant *type of gesture* and *type of expression* interaction ($F(1, 8)=6.73$, $P<0.05$) evidenced a differential pattern. Post hoc comparisons (HSD test, MS = 16.90, d.f. = 80.00) showed significant accuracy differences only in the incongruent metaphorical category with respect to literal congruent ($P<0.01$), literal incongruent ($P=0.024$), and metaphorical congruent categories ($P<0.01$).

Summarizing, despite the main effects of gesture (better classification of congruent gestures), accuracy data showed a gesture–expression interaction, evidencing that incongruent metaphorical expressions were more difficult to classify. Compared with RTs and ERPs (see below), no accuracy differences were observed for type of expression.

3.1.2. Reaction times

The RTs were greater for metaphorical expressions ($M=1811$, S.D. = 223 ms) than literal ones ($M=1712$, S.D. = 202 ms); also, for the incongruent ones ($M=1780$, S.D. = 217 ms) with respect to the congruent ones ($M=1743$, S.D. = 207 ms). An ANOVA yielded significant main effects of *type of gesture* ($F(1, 8)=8.84$, $P<0.05$) and *type of expression* ($F(1, 8)=5.93$, $P<0.05$). More important, a significant interaction between *type of gesture* and *type of expression* ($F(1, 8)=8.02$, $P<0.05$) was observed. Post hoc comparisons (HSD test; MS = 1042.90, d.f. = 80.00) indicate that all the conditions were significantly distinguished except the literal congruent category with respect to the metaphor congruent category ($P=0.981$). Table 1 summarizes the behavioral results.

In brief, RTs showed delayed responses of incongruent gestures and incongruent expressions (in a similar vein than ERPs, see later).

Table 1
Reaction times (RTs) and accuracy.

	RTs (ms)		Accuracy	
	Mean	S.D.	%	S.D.
Type of gesture				
Congruent	1743	207	89.05	7.43
Incongruent	1780	217	80.04	7.98
Type of expression				
Literal	1712	202	86.38	7.72
Metaphorical	1811	223	80.45	8.32
Gesture and expression interactions				
Literal congruent	1708	218	89.11	7.54
Literal incongruent	1758	204	83.66	6.52
Metap. congruent	1719	243	89.00	8.43
Metap. incongruent	1834	276	76.43	8.18

Note. The table shows the RT averages in ms, the percentage of accurate responses, and the standard deviation (S.D.) for both, types of gesture (congruent and incongruent) and types of expression (literal and metaphorical). Metap: metaphorical.

Moreover, consistent with accuracy data, a gesture–expression interaction indicated a differential discrimination of incongruent metaphorical expressions (being more delayed than literal expressions when incongruent but as rapid as literal ones when congruent).

3.2. ERPs

3.2.1. N400-like effects

Fig. 2 shows the ERPs from five representative ROIs. An important effect of type of gesture and type of expression on the left anterior ROI was observed. In this ROI the incongruent metaphorical category produced greater negativity ($M=-2.63$; S.D. $=0.50$ μv), followed by the incongruent literal category ($M=-0.58$, S.D. $=0.32$ μv). In contrast, no differences arose under the congruent condition (congruent metaphorical: $M=1.42$, S.D. $=0.30$ μv; and congruent literal category: $M=1.64$, S.D. $=0.40$ μv). An ANOVA confirmed a significant main effect of *type of gesture* (congruent vs. incongruent; $F(1, 8)=7.97$, $P<0.01$); *type of expression* (literal vs. metaphoric; $F(1, 8)=6.76$, $P=0.031$); and *ROI* ($F(4, 32)=36.28$, $P<0.01$). In addition, an interaction between these three factors was obtained ($F(4, 32)=12.90$, $P<0.01$). Post hoc comparisons done on this last interaction (Tukey HSD test; Bonferroni corrected MS $=0.06$, d.f. $=32.00$) showed that significant differences were observed only in the left anterior region. All the categories are significantly distinguished except for the literal congruent categories in comparison to metaphorical congruent categories ($P=0.465$). Incongruent gestures from metaphorical expressions produced greater amplitude differences regarding the metaphor congruent category ($P<0.001$). In addition, significant differences were obtained between the literal congruent category and the literal incongruent category ($P<0.05$). No significant differences were observed in this temporal window for other ROIs. See all significant effects summarized in N400-like results of Table 2.

3.2.2. LPC-like effects

For the temporal window of the 500–700 ms a similar pattern was observed. An ANOVA yielded a significant main effect of *type of gesture* (congruent vs. incongruent; $F(1, 8)=13.94$, $P=0.005$); and *ROI* ($F(4, 32)=45.44$, $P<0.01$). In addition, an interaction between *type of*

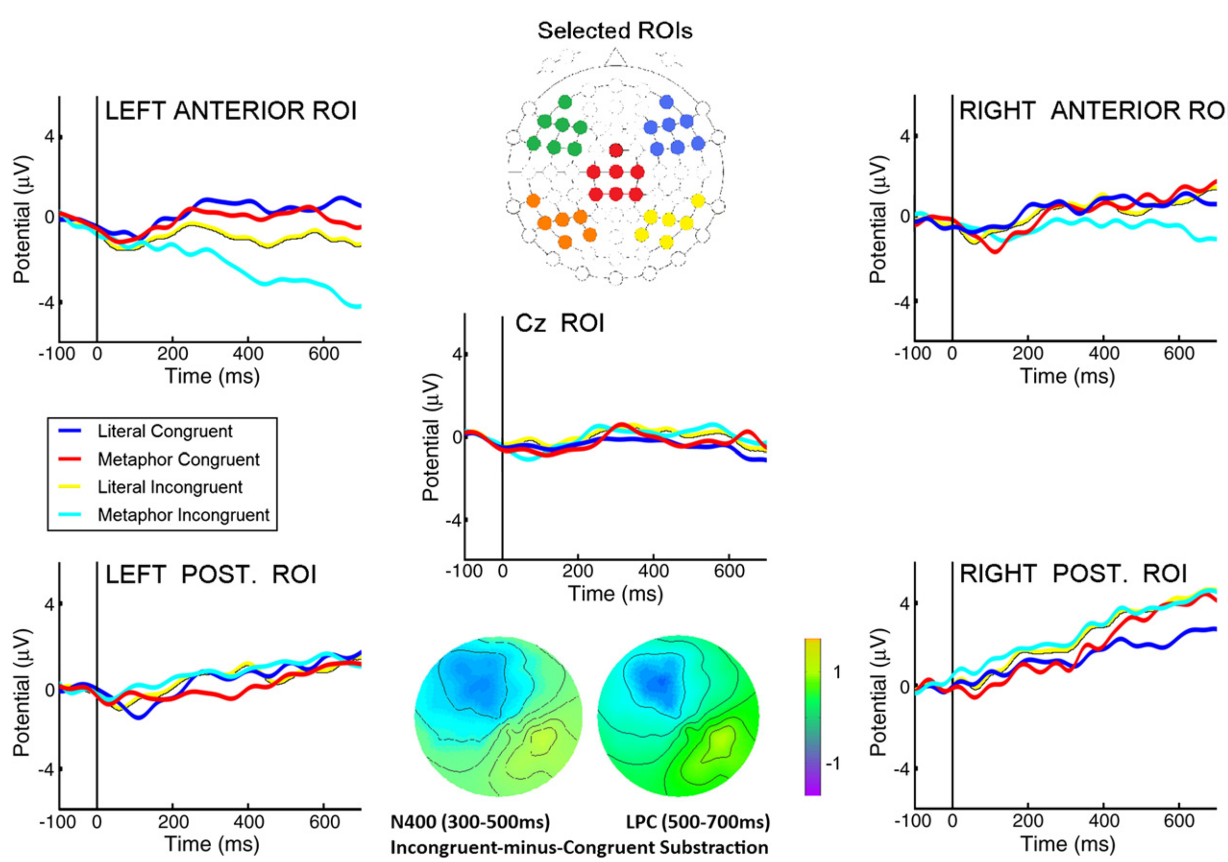

Fig. 2. ERPs of gesture congruency and expression. The figure shows the waveforms of four categories from five ROIs (left anterior, left posterior, Cz, right anterior and right posterior). In addition, the scalp localizations of each ROI and the topographic map (incongruent minus congruent subtraction) of N400 (300–500 ms) and LPC (500–750 ms) are shown. The bar indicates the color amplitude values of voltage maps in μv.

Table 2
Principal ERP effects for N400 and LPC.

Factor	N400 effects	LPC effects
Expression	Literal < metaphoric*	Literal < metaphoric**
Gesture	Congruent < incongruent**	
ROI	RP < LP < Cz < RA < LA**	LA < RP**
Expression × gesture × ROI	LA**:	LA*:
	LC < LI*	LC < LI*
	MC < MI***	MC < MI***
	LI < MI**	LI < MI*
		RA:
		MC < MI**
		LC < MI**
		LI < MI**
		RP:
		LI < LC*

Note. Principal ERP effects for N400 and LPC. Only statistically significant differences are presented. The "<" indicates the direction of increased microvolt amplitude. ROI abbreviations: LA (left anterior ROI); RA (right anterior ROI); CZ (central midline ROI); LP (left posterior ROI); and RP (right posterior ROI). Effect abbreviations: LC (literal congruent); LI (literal incongruent), MC (metaphor congruent), and MI (metaphor incongruent). Comparisons were statistically significant at *$P<0.05$, **$P<0.01$ and ***$P<0.001$.

gesture, *type of expression* and *ROI* was observed ($F(4, 32)=3.08$, $P=0.029$). Post hoc comparisons of this interaction (Tukey HSD test; Bonferroni corrected, MS = 0.29, d.f. = 32.00) showed the same effect as previously reported in the left anterior region. Again, all the categories were significantly distinguished except for literal congruent categories with respect to metaphorical congruent categories ($P=1.00$). In the right anterior ROI, only the incongruent metaphorical category induced a significantly greater negativity than the congruent metaphorical ($P<0.01$), literal congruent ($P<0.01$) and literal incongruent categories ($P<0.01$). In the right posterior ROI the literal congruent category generated the lowest amplitude, only significantly distinguishing this from the literal incongruent category ($P=0.03$). Significant differences were not observed in other ROIs. See all significant effects in LPC-like results of Table 2.

4. Discussion

4.1. General results

Accuracy levels were high in all conditions (above 76%). These results assured the quality of the stimuli for the electrophysiological study. Nevertheless, there were significant factorial effects in the accuracy measures supporting our hypothesis. At the same time, and consistent with the ERP results, a gesture–expression interaction was observed in the RT results. No behavioral differences were found between the literal and metaphorical expression when gesture was congruent. Reaction times and ERP results suggest the high contextual sensitivity of metaphorical expressions based on the gesture blending. Convergent results of accuracy, RTs and ERPs regarding the gesture–expression interaction suggest that, consistent with our hypothesis, metaphorical expressions were more sensitive to gestural information. When contextual information is not matching (e.g., incongruent gesture), metaphorical expression is more difficult to process. But when contextual information is matching, metaphorical expression is processed equivalently as neutral expressions.

4.2. Contextual effects on N400-like and LPC-like

The 300–500 ms epoch yielded the following results: (i) in the left anterior ROI the N400-like component of incongruent gestures from metaphorical expressions produced greater negativity. (ii) The N400-like modulation of metaphorical expressions showed a greater difference between congruent and incongruent categories in the left anterior region too. In addition, the literal expressions showed a smaller modulation on the basis of the gesture congruency in the same region. (iii) More importantly, the literal congruent as well as the metaphorical congruent categories did not show any difference in any region.

The 500–700 ms epoch (LPC-like) potentials showed a similar pattern obtained with the N400-like potentials. Additionally, on the right side of the scalp, the LPC-like potentials from incongruent metaphorical categories and the literal congruent categories were distinguished on the frontal and occipital regions, respectively. This suggests a differential pattern in the right hemisphere for the LPC-like component.

This study confirms previous results: (a) N400 component and LPC amplitude modulation obtained with incongruent videos (Sitnikova et al., 2003; Reid and Striano, 2008), (b) N400-like component and LPC-like amplitude modulation obtained with incongruent video gestures (Neville et al., 1997; Kelly et al., 2004; Wu and Coulson, 2005; Kelly et al., 2007; Özyüreck et al., 2007; Wu and Coulson, 2007) and moreover (c) N400 component and LPC amplitude modulation obtained with incongruent gestures of metaphorical expressions with a left anterior pattern (Cornejo et al., 2009). Nevertheless, our data are the first report suggesting a high contextual sensitivity of N400/LPC-like potentials to be presented that suggest to metaphorical language and gesture blending. This is the first report showing a differential modulation based on the integration of contextual information and metaphorical expressions. A major difference was observed on the N400/LPC-like amplitude in response to the gesture congruency in the metaphorical expressions giving the suggestion of a more intense effect of the context on this type of figurative language.

4.3. The relevance of this research

Our result adds additional evidence to previous results on the interaction between speech and gestures in the brain (Willems and Hagoort, 2007). In this sense, our results deploy an image of metaphor as a linguistic construction highly sensitive to the context of use (Cornejo, 2004, 2007). Our N400/LPC-like components strongly resemble the left anterior waveform and topography from a previous study of N400 video clips with figurative expression modulations. This particular scalp region, presented in a previous study by Cornejo et al. (2009), has been attributed to the complex and composite presentation of semantic stimuli (metaphorical and literal sentences and video gestures that generate a semantic expectancy) displayed in cross-modal stimulation (Cornejo et al., 2009).

In agreement also with the Cornejo et al. study we observed a latency of the obtained N400-like component greater than classical studies reported. This prolonged effect in the time course could be explained by the characteristics of the stimulus, considering that the incoming stimuli are continuous rather than discreet, and present differentiations in the precision for identification of global sense information for each gesture (see Sitnikova et al. (2003) for a similar interpretation as well).

The posterior-positivity effect found in this study strongly resembles the one reported for videos (Sitnikova et al., 2003) and co-gesture metaphorical expression modulation (Cornejo et al., 2009). Both previous studies, as well as our results, used video clips displaying a contextually inappropriate action in an otherwise coherent real-world situation. This fact suggests that the late positivity effect is expressing a process of re-analysis of the incongruent situation produced by the inconsistent meaning (Sitnikova et al., 2003).

Our results have two possible confounds to be addressed in further studies. First, we did not trigger the word onset. Although gesture stroke onset has been shown to be more effective than word onset for co-gesture speech interaction (Cornejo et al., 2009), further studies should compare both components, in addition to possible similar neural generators, using a source analysis technique. Additionally, our sample size was very small and was reduced by subject elimination due to the high number of artifacts from ocular movements induced by the video stimulation format. Further studies with larger samples should confirm these results.

We confirm that the use of the gesture stroke onset as time-locked is relevant because it implies the possibility of obtaining information from a *temporally* dynamic event (Cornejo et al., 2009). Since the stroke can be marked in a specific frame of the video, it is possible to analyze a dynamic event with a static point of reference (although early components of the visual ERP will likely be refractory due to the continuous stimulus presentation format; e.g., Sitnikova et al., 2003; Wu and Coulson, 2005; Cornejo et al., 2009).

The ERPs have been shown to be sensitive to age-related changes (Friedman, 2003). Moreover, the N400 component has been reported abnormal in dementia and mild cognitive impairment (MCI, Olichney et al., 2002). Moreover, dementia research suggested deficits in gesture comprehension and related praxias (i.e., Dumont et al., 2000). Potentially, the application of this paradigm to aging and participants with MCI may provide insights for aging research and markers of early electrophysiological deficits related to dementia vulnerability.

4.4. Conclusions

This result confirms previously obtained reports that allow understanding of the co-gesture figurative language blending as part of a highly contextualized comprehension. For our knowledge this is the first study of contextual sensitivity of gesture and figurative expressions using video clips and ERPs. Despite some disparities in the main effects of gesture and expression main effects (e.g., no effects of type of expression in accuracy data), both the behavioral and the ERP data evidenced a convergent interaction of gesture and expression a convergent interaction of gesture and expression. This interaction supports the hypothesis of a more contextual sensitivity of metaphorical expressions compared to literal ones. Our results suggest that metaphorical expression is clearly sensitive to the gesture information, suggesting that a supposedly derivative meaning, such as the metaphorical meaning, is active from early stages of processing. This evidence is coherent with contextual theories of meaning comprehension: metaphorical expression is processed in a similar way (or even more sensitively) to literal expressions when contextual or paralinguistic clues facilitate its meaning (Gibbs, 1994; Giora, 2003). In fact, when context information is coherent, electrophysiological patterns for literal and metaphorical expressions are undistinguishable. On the contrary, contextual inadequacy produces incongruence not only with figurative expressions (more accentuated), but also with literal ones. Context – gestures, in this case – is early and tightly interwoven with the lexical information (Coulson, 2006; Cornejo, 2007) supporting recent reports of contextual effects on cognition (Ibanez et al., 2006; Cornejo et al., 2009; Hurtado et al., 2009; Ibanez et al., 2009; Aravena et al., 2010; Barutta et al., 2010; Dufey et al., 2010; Ibanez et al., 2010; San Martín et al., 2010). This evidences a strong sensitivity of the human brain to the circumstances in which language is used.

Acknowledgments

This research was partially supported by two postdoctoral grants (CONICET and DAAD/CONICYT), and by a Neuroscience Grant CG90326010 (Diego Portales University) and a FINECO grant. P.T. and J.S. were supported by the Marsilius Kolleg, a center of advanced studies, University of Heidelberg. We wish to thank Stephan Walther, Esteban Hurtado and especially Kerstin Herwig for their technical support. We are indebted to Ela Kuzma and Christine Sattler for stimuli preparation.

Appendix A. Supplementary data

Supplementary data to this article can be found online at doi:10.1016/j.pscychresns.2010.08.008.

References

Aravena, P., Hurtado, E., Riveros, R., Cardona, F., Manes, F., Ibáñez, A., 2010. Applauding with closed hands: neural signature of action sentence compatibility effects. PLoS ONE 5 (7), e11751. doi:10.1371/journal.pone.0011751.

Barrett, S.E., Rugg, M.D., 1990. Event-related potentials and the semantic matching of pictures. Brain and Cognition 2, 201–212.

Barutta, J., Aravena, P., Ibáñez, A., 2010. The machine paradigm and alternative approaches in cognitive science. Integrative Psychological and Behavioral Science 44 (2), 176–183.

Cornejo, C., 2004. Who says what the words say?: the problem of linguistic meaning in psychology. Theory and Psychology 14, 5–28.

Cornejo, C., 2007. Review essay: conceptualizing metaphors versus embodying the language. Culture & Psychology 13, 474–487.

Cornejo, C., Simonetti, F., Ibáñez, A., Aldunate, N., Ceric, F., López, V., Núñez, R.E., 2009. Gesture and metaphor comprehension: electrophysiological evidence of cross-modal coordination by audiovisual stimulation. Brain and Cognition 1, 42–52.

Cosmelli, D., Ibáñez, A., 2008. Human cognition in context: on the biologic, cognitive and social reconsideration of meaning. Integrative Psychological and Behavioral Sciences 42, 233–244.

Coulson, S., 2006. Constructing meaning. Metaphor & Symbol 21, 245–266.

Coulson, S., Van Petten, C., 2002. Conceptual integration and metaphor: an event-related potential study. Memory and Cognition 30, 958–968.

Daltrozzo, J., Tillmann, B., Platel, H., Schön, D., 2010. Temporal aspects of the feeling of familiarity for music and the emergence of conceptual processing. Journal of Cognitive Neuroscience 22, 1754–1769.

Dufey, M., Hurtado, E., Fernández, A.M., Manes, F., Ibáñez, 2010. Exploring the relationship between vagal tone and event-related potentials in response to an affective picture task. Social Neuroscience 23, 1–15.

Dumont, C., Ska, B., Joanette, Y., 2000. Conceptual apraxia and semantic memory deficit in Alzheimer's disease: two sides of the same coin? Journal of International Neuropsychological Society 6, 693–703.

Federmeier, K.D., Kutas, M., 2001. Meaning and modality: influences of context, semantic memory organization, and perceptual predictability on picture processing. Journal of Experimental Psychology. Learning, Memory, and Cognition 1, 202–224.

Friedman, D., 2003. Cognition and aging: a highly selective overview of event-related potential (ERP) data. Journal of Clinical and Experimental Neuropsychology 25, 702–720.

Ganis, G., Kutas, M., Sereno, M.I., 1996. The search for common sense: an electrophysiological study of the comprehension of words and pictures in reading. Journal of Cognitive Neuroscience 8, 89–106.

Ganis, G., Kutas, M., 2003. An electrophysiological study of scene effects on object identification. Brain research. Cognitive Brain Research 2, 123–144.

Gibbs, R., 1994. The Poetics of Mind; Figurative Thought, Language, and Understanding. Cambridge University Press, Cambridge.

Giora, R., 2003. On our Mind: Salience, Context and Figurative Language. Oxford University Press, Oxford, England.

Guerra, S., Ibáñez, A., Martín, M., Bobes, M.A., Reyes, A., Mendoza, R., Bravo, T., Domínguez, M., Sosa, M.V., 2009. N400 deficits from semantic matching of pictures in probands and first-degree relatives from multiplex schizophrenia families. Brain and Cognition 2, 221–230.

Gunter, T., Bach, P., 2004. Communicating hands: ERPs elicited by meaningful symbolic hand postures. Neuroscience Letters 30, 52–56.

Hannula, D.E., Federmeier, K.D., Cohen, N.J., 2006. Event-related potential signatures of relational memory. Journal of Cognitive Neuroscience 11, 1863–1876.

Holle, H., Gunter, T., 2007. The role of iconic gestures in speech disambiguation: ERP evidence. Journal of Cognitive Neuroscience 19, 1175–1792.

Hurtado, E., Gonzalez, R., Haye, A., Manes, F., Ibanez, A., 2009. Contextual blending of ingroup/outgroup face stimuli and word valence: LPP modulation and convergence of measures. BMC Neuroscience 10, 69. doi:10.1186/1471-2202-10-69.

Ibanez, A., Riveros, R., Aravena, P., Vergara, V., Cardona, J.F., García, L., Hurtado, E., Martin Reyes, M., Barutta, J., Manes, F., 2010. When context is difficult to integrate: cortical measures of congruency in schizophrenics and healthy relatives from multiplex families. Schizophrenia Research. doi:10.1016/j.schres.2010.04.008.

Ibáñez, A., Cosmelli, D., 2008. Moving beyond computational cognitivism: understanding intentionality, intersubjectivity and ecology of mind. Integrative Psychological and Behavioral Sciences 42 (2), 129–136.

Ibáñez, A., Haye, A., González, R., Hurtado, E., Henríquez, R., 2009. Multi-level analysis of cultural phenomena: The role of ERP approach to prejudice. The Journal for Theory in Social Behavior 39, 81–110.

Ibáñez, A., López, V., Cornejo, C., 2006. ERPs and contextual semantic discrimination: evidence of degrees of congruency in wakefulness and sleep. Brain and Language 98, 264–275.

Kelly, S., Kravitz, C., Hopkins, M., 2004. Neural correlates of bimodal speech and gesture comprehension. Brain and Language 89, 243–260.

Kelly, S.D., Ward, S., Creigh, P., Bartolotti, J., 2007. An intentional stance modulates the integration of gesture and speech during comprehension. Brain and Language 101, 222–233.

Kircher, T., Straube, B., Leube, D., Weis, S., Sachs, O., Willmes, K., Konrad, K., Green, A., 2009. Neural interaction of speech and gesture: differential activations of metaphoric co-verbal gestures. Neuropsychologia 1, 169–179.
Kita, S., Özyürek, A., 2003. What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: evidence for an interface representation of spatial thinking and speaking. Journal of Memory and Language 48, 16–32.
Kutas, M., Hillyard, S., 1980. Reading senseless sentences: brain potentials reflect semantic incongruity. Science 207, 203–205.
Levelt, W.J.M., Richardson, G., La Heij., W., 1985. Pointing and voicing in deictic expressions. Journal of Memory and Language 24, 133–164.
McNeill, D., 1992. Language and Gesture. The University of Chicago Press, Chicago.
McPherson, W.B., Holcomb, P.J., 1999. An electrophysiological investigation of semantic priming with pictures of real objects. Psychophysiology 1, 53–65.
Neville, H., Coffey, S., Lawson, D., Fischer, A., Emmorey, K., Bellugi, U., 1997. Neural systems mediating American sign language: effects of sensory experience and age of acquisition. Brain and Language 57, 285–308.
Oken, B., Chiappa, K., 1986. Statistical issues concerning computerized analysis of brain wave topography. Annals of Neurology 19, 493–494.
Olichney, J., Morris, S., Ochoa, C., Salmon, D., Thal, L., Kutas, M., Iragui, V., 2002. Abnormal verbal event related potentials in mild cognitive impairment and incipient Alzheimer's disease. Journal of Neurology Neurosurgery and Psychiatry 73, 377–384.
Özyüreck, A., Willems, R., Kita, S., Hagoort, P., 2007. On-line integration of semantic information from speech and gesture: insights from event-related brain potentials. Journal of Cognitive Neuroscience 19, 605–616.
Ozyürek, A., Kelly, S.D., 2007. Gesture, brain, and language. Brain and Language 3, 181–184.
Proverbio, A.M., Riva, F., 2009. RP and N400 ERP components reflect semantic violations in visual processing of human actions. Neuroscience Letters 3, 142–146.
Pynte, J., Besson, M., Robichon, F., Poli, J., 1996. The time-course of metaphor comprehension: an event-related potential study. Brain and Language 55, 293–316.
Reid, M., Striano, T., 2008. N400 involvement in the processing of action sequences. Neuroscience Letters 12, 93–97.
San Martín, R., Manes, F., Hurtado, E., Isla, P., Ibáñez, A., 2010. Size and probability of rewards modulate the feedback error-related negativity associated with wins but not losses in a monetarily rewarded gambling task. NeuroImage 51 (2010), 1194–1204.
Sitnikova, T., Kuperberg, G., Holcomb, P., 2003. Semantic integration in videos of real-world events: an electrophysiological investigation. Psychophysiology 40, 160–164.
West, W.C., Holcomb, P.J., 2002. Event-related potentials during discourse-level semantic integration of complex pictures. Cognitive Brain Research 3, 363–375.
Willems, R., Hagoort, P., 2007. Neural evidence for the interplay between language, gesture and action: a review. Brain and Language 101, 278–289.
Willems, R., Özyürek, A., Hagoort, P., 2007. When language meets action: the neural integration of gesture and speech. Cerebral Cortex 17, 2322–2333.
Willems, R.M., Ozyürek, A., Hagoort, P., 2009. Differential roles for left inferior frontal and superior temporal cortex in multimodal integration of action and language. NeuroImage 4, 1992–2004.
Wu, Y., Coulson, S., 2007. How iconic gestures enhance communication: an ERP study. Brain and Language 101, 234–245.
Wu, Y.C., Coulson, S., 2005. Meaningful gestures: electrophysiological indices of iconic gesture comprehension. Psychophysiology 6, 654–667.