RP and N400 ERP components reflect semantic violations in visual processing of human actions

Alice Mado Proverbio and Federica Riva

Since their discovery during the late decades of the last century, event-related brain potentials (ERPs) have contributed greatly to understanding the neural bases of cognitive processes and especially of language [14], [13] and [12]. In particular, the N400 component, a large negative deflection peaking at about 400 ms over centro-parietal scalp areas, has been related to semantic integration [1] and lexical access processes [11] and [15]. It is a rather well established notion that N400 is sensitive to word frequency, class, concreteness, orthographic neighbours, close probability, semantic relatedness, contextual constraint, prototypicity (see an exhaustive review in [10]), idiomaticity [2], age of language acquisition [26] and language proficiency [27], with higher amplitudes to less familiar or expected items. N400 neural generators seem to include the left temporal cortex (both posterior (VWFA) and anterior) [24] and [21], the left inferior frontal gyrus and the angular gyrus [15]. The study of N400 behaviour has helped us to understand how meanings are accessed, stored and integrated in the lexical semantic system. It has also been demonstrated that N400 is sensitive not only to word meanings but also to violations of world knowledge learned during everyday life [4] or to semantic violations in deaf native signers [23].

However, to our knowledge, the observation of linguistic components has not been applied so far to the study of gesture coding, for which it is known that there are permanent representational units in the inferior parietal and inferior frontal cortex. Indeed, apart from communicative gestures such as sign language (e.g., ASl or BSL), goal-directed gestures whose intent is not communicative are also recognized as unitary meaningful units by premotor and somatosensory mirror neurons. Indeed a left inferior parietal lesion (BA40) is associated with the inability to recognize or imitate a gesture (such as brushing teeth or flipping a coin) or to perform skilled actions (such as lighting a cigarette or making coffee), which is a deficit called apraxia (the linguistic counterpart of this might correspond to a posterior aphasia). Interestingly, an ERP study on ASL processing reported greater amplitude signals originating in the parietal cortices of native than of late signers [22].

The available neurometabolic literature [5], [8], [16], [25], [30], [33] and [34] provides evidence that a fronto-parietal mirror system, including the inferior frontal gyrus, left inferior parietal lobule and superior temporal sulcus, is involved in action coding and comprehension in humans. The evidence comes from the observation that goal-directed vs. non-goal-directed actions (e.g., picking up vs. just reaching), or more salient (e.g., grasping a glass to drink) vs. less salient (e.g., grasping a glass to clean up) actions, specifically activate the mirror neuron circuits. Human data are paralleled by neurophysiological recordings of macaque mirror neuron activity showing, e.g., a differential neural coding in area F5 for grasping to eat vs. grasping to throw away. However, no clear and direct evidence has ever been offered to show that reduced/absent activation of areas coding the intentions and outcomes of actions corresponds to a lack of goal/intention comprehension for human observers.

The aim of the present study was to determine whether N400 was sensitive to semantic violations in action representation using visual processing of comprehensible actions vs. incomprehensible actions. The incongruent actions used in this research were chosen to be not just infrequent or rare, but incomprehensible, purposeless, and in some cases socially inappropriate or impossible to understand in terms of the agent's intentions. In other words they violated so-called ‘world knowledge’ about typical human actions in ecological environments. For this reasons we expected the congruent/incongruent contrast to elicit a N400 similar to the one discovered by Hagoort et al. [4] for linguistically mediated pragmatic knowledge. As for earlier ERP components, we did not

have specific expectations since, to our knowledge; incongruent human actions vs. congruent human actions were never previously compared in ERP investigations.

Twenty-three healthy right-handed Italian University students (12 males and 11 females) were recruited for this experiment. Their ages ranged from 20 to 35 years (mean = 24.8 years). All had normal or corrected-to-normal vision and reported no history of neurological illness or drug abuse. The experiments were conducted with the understanding and the written consent of each participant. The experimental protocol was approved by the ethical committee of the National Research Council in Milan.

Two hundred and sixty ecological colour pictures representing persons differing in number, age and gender, engaged in goal-directed actions, were presented to participants (see some examples in Fig. 1).

Fig. 1. Examples of congruent and incongruent actions performed by young female agents. Pictures were balanced between classes for type of action, age, gender, number of persons and body parts depicted.

One hundred and thirty of the pictures displayed actions belonging to the typical human repertoire (e.g., young woman driving car; woman relaxing in bath, eyes closed; man tying shoelace; doctor writing prescription; mature man praying in pew; man cutting log in forest using chainsaw; smiling couple clinking glasses of champagne; girl eating spaghetti; 10-year-old girl practicing her flute; boy fishing on pier). Perception of these meaningful actions was contrasted with perception of as many pictures (130) showing humans engaged in actions lacking any understandable goal (e.g., businesswoman balancing on one foot in desert; young woman cutting jewellery on plate with fork and knife; woman cutting bread with saw in kitchen; young woman, eyes closed, sucking through straw placed in car engine; woman cutting man's hair with garden shears; young man playing cello with saw; man splashing face with pebbles; surgeon dissecting book; well dressed woman posing with a lamp on her head; woman in grey suit bending at waist, head in shopping bag). The action incongruity was preliminarily established by a group of ten judges. Congruent and incongruent actions were balanced for gender, age and number of persons, body parts depicted, picture size (11° 28 in. in length and 8° 36 in. in height) and average luminance (492.22 cd/m^2). Stimuli were presented randomly mixed on a PC screen for 1500 ms with an ISI of 1800–1900 ms on a grey background.

The participants, seated comfortably in a dimly lit, electrically and acoustically shielded room, faced a window behind which a PC monitor was positioned 114 cm from their eyes. A small bright dot (3 in. in size) located at the centre of the screen served as a fixation point to minimize eye movements. The subjects were instructed to fixate the centre of the screen and to avoid any eye or body movements during the recording session. The task consisted in signalling the rare presentation of a natural landscape without visible humans (44 in all) by pressing a button as accurately and rapidly as possible with the index finger of the left or right hand. The two hands were used alternately during the recording session, and the hand and sequence order were counterbalanced across subjects.

The EEG was continuously recorded from 128 scalp sites at a sampling rate of 512 Hz. Horizontal and vertical eye movements were also recorded. Linked ears served as the reference lead. The EEG and electro-oculogram (EOG) were amplified with a half-amplitude band pass of 0.016–100 Hz. Electrode impedance was kept below 5 kΩ. The artefact rejection criterion was peak-to-peak amplitude exceeding 50 μV, and the rejection rate was ~5%. ERPs were averaged off-line from −100 ms before to 1000 ms after stimulus onset. ERP components were identified and measured with reference to the average baseline voltage over the interval from −100 ms to 0 ms relative to stimulus onset. A preliminary inspection of the data showed a strong task-related modulation of posterior (visual) N270 response, and of an anterior late negative deflection (N420). The ERP component of interest were quantified and statistically treated.

The mean amplitude of the occipital N2 component (recognition potential, RP) was measured at posterior sites (OL1h, OL2h, POO3h, and POO4h) in the time window 250–350 ms. The mean amplitude of the N400 component was measured at anterior sites (F1, F2, FC1, and FC2) between 350 and 500 ms. The electrode sites considered are visible in maps of Fig. 2. ERP data were subjected to multifactorial repeated-measures ANOVA. The factors were “semantic congruence” (congruent and incongruent), “electrode” (dependent on ERP component of interest), and “hemisphere” (left and right). Multiple comparisons of means were done by post hoc Tukey tests.

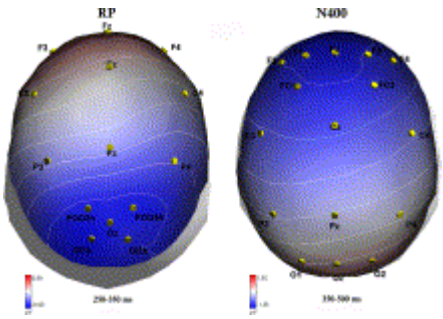

Fig. 2. Topographical difference maps performed by plotting difference voltages obtained by subtracting ERPs to incongruent from ERP to congruent actions in the Recognition Potential latency range, and ERPs to congruent from ERPs to incongruent actions in the N400 latency range. Note that ERPs were recorded from 128 sites which are not all represented here.

Fig. 3 shows the grand-average waveforms ($N = 23$) recorded in response to congruent and incongruent actions over anterior and posterior scalp sites. Strong effects of action meaningfulness

are evident over the occipito/parietal cortex as early as 170 ms (N1), reaching their maximal amplitude at about 250 ms (RP), and more anteriorly at about 400 ms.

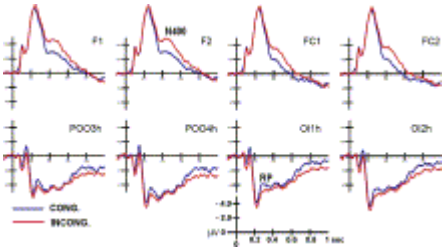

Fig. 3. Grand-average ERP waveforms ($N = 23$) recorded over frontocentral and occipito/parietal sites as a function of action meaningfulness.

ANOVA performed on the RP amplitude values recorded at the occipito/parietal electrode sites showed an effect of semantic congruence ($F(1,22) = 12.43$, $p < 0.002$), with much greater RP responses to congruent (3.69 μV) than incongruent (4.32 μV) actions. The significance of the hemisphere ($F(1,22) = 9.29$; $p < 0.006$) and hemisphere × electrode ($F(1,22) = 4.24$; $p = .05$) factors indicated larger RP potentials over the left (3.64 μV) than the right (4.37 μV) hemisphere, especially over the parieto/occipital area, as also indicated by post hoc comparisons.

N400 was strongly affected in amplitude by action significance ($F(1,22) = 52.5$; $p < 0.00001$), which was much larger to incongruent (−3.35 μV) than congruent (−1.84 μV) actions. Overall, N400 was greater at frontal (−3.04 μV) than frontocentral (−2.14 μV) sites, as indicated by the significance of electrode ($F(1,22) = 14.8$; $p < 0.0009$), and relative post hoc comparisons.

The aim of the study was to establish where semantic violations in action coding (obtained by presenting human actions difficult to understand/recognize or to integrate with previous world knowledge) elicited ERP responses similar to the linguistic components described in the literature, thus supporting the view that a common mode processes incoming conceptual information, whether linguistic, face-based, auditory or action-based.

The ERP results indicated the emergence of a posterior recognition potential (RP) to congruent actions and of a late N400 response to incongruent actions (see their scalp distribution in maps of Fig. 2). The finding of a negative N250 larger to meaningful recognizable actions agrees with the neurolinguistic and ERP literature describing a posterior electrical response of the brain peaking around 200–250 ms after stimulus onset and obtained when subjects view recognizable images such as words, pictures or faces [18]. It has been proposed that the RP component reflects early semantic processing [7], and is sensitive to semantic expectancy [3] and semantic category [17]. Furthermore, RP amplitude is larger for concrete than for abstract words [20], for open than for closed-class words [6], and for easy than for difficult words [31] and [32]. In this context, incongruent actions in our study might be conceptualized as legitimate pseudo-actions as opposed to real actions, thus mirroring the word/pseudo-word distinction. Alternately, one may hypothesize that incongruent actions were processed as inappropriate relative to the “semantic context” (for example, in the scene where a young woman, eyes closed, sucks through straw coming from a car

engine). This hypothesis is supported by an ERP study by Martín-Loeches et al. [19] in which words congruent with the previous semantic context elicited larger RP amplitudes as compared to incongruent words.

It follows then that perception of incongruent actions should trigger a later N400 deflection, which was present over all scalp sites in our study to a significant degree.

The N400 data represent, to our knowledge, the first electrophysiological evidence that non-linguistic gestures are coded as single meaningful units and automatically recognized by skilled adult brains, probably by visuomotor mirror neurons of the fronto-parietal system. This hypothesis is suggested by parallel source localization data indicating premotor, motor and inferior parietal cortices for RP and orbito-frontal cortex as possible neural generators for N400 effects [28].

It should be considered that our action-related N420 showed an anterior distribution, which is different from the typical right centro-parietal distribution of linguistic N400 (probably subtending a left temporal generator [24]). It is true that in the literature evidence has been offered for a "lexical processing negativity" (LPN) peaking at about 350 ms, and with a marked anterior distribution, which is more negative to pseudo-words than words [26] and [29], and later to less familiar or frequent words than to more familiar words [9], but its behaviour probably suggests lexical access rather than semantic integration processes. Since the incongruent actions of our study were not just infrequent or rare, but socially inappropriate, we hypothesized that action-related N400 reflected a difficulty integrating the action meaning with the semantic context and previous world knowledge, thus suggesting a functional similarity with linguistic N400 response [4].

In conclusion, our ERP findings indicate that comprehensible actions vs. incomprehensible actions elicit a posterior negativity at about 250 ms in the form of recognition potential, similar to the RPs described in the linguistics literature. The lack of action comprehension (relative to the agent's context or to the viewer's world knowledge) triggered a large and long-lasting N400 response, probably indicating difficulty in understanding the other's behaviour or integrating the incoming information with previous knowledge.

Acknowledgements

The authors wish to thank Patricia Klaas her insightful comments on an earlier version of this manuscript. They are also grateful to Alberto Zani for his invaluable help and to Roberta Adorni, Marzia Del Zotto and Nicola Crotti for their kind support. Granted from 2008 to 2009 FAR funds to AMP.

References

[1] C. Brown and P. Hagoort, The processing nature of the N400: evidence from masked priming, *J. Cogn. Neurosci.* **5** (1993), pp. 34–44.
[2] S. Coulson and C. Van Petten, A special role for the right hemisphere in metaphor comprehension? ERP evidence from hemifield presentation, *Brain Res.* **1146** (2007), pp. 128–145.
[3] J. Dien, G.A. Frishkoff, A. Cerbone and D.M. Tucker, Parametric analysis of event-related potentials in semantic comprehension: evidence for parallel brain mechanisms, *Cogn. Brain Res.* **15** (2003), pp. 137–153.
[4] P. Hagoort, L. Hald, M. Bastiaansen and K.M. Petersson, Integration of word meaning and world knowledge in language comprehension, *Science* **304** (2004), pp. 438–441.

[5] A.F. Hamilton and S.T. Grafton, Action outcomes are represented in human inferior frontoparietal cortex, *Cereb. Cortex* **18** (2008), pp. 160–1168.
[6] J.A. Hinojosa, M. Martín-Loeches, P. Casado, F. Muñoz, L. Carretié and M.A. Pozo, Semantic processing of open and closed-class words: an event-related potentials study, *Cogn. Brain Res.* **11** (2001), pp. 397–407.
[7] J.A. Hinojosa, M. Martín-Loeches, F. Muñoz, P. Casado and M.A. Pozo, Electrophysiological evidence of automatic early semantic processing, *Brain Lang.* **88** (2004), pp. 39–46.
[8] M. Iacoboni, Neural mechanisms of imitation, *Curr. Opin. Neurobiol.* **15** (2005), pp. 632–637.
[9] J.W. King and M. Kutas, Neural plasticity in the dynamics of human visual word recognition, *Neurosci. Lett.* **244** (1998), p. 61.
[10] M. Kutas, C. van Petten and R. Kluender, Psycholinguistics electrified II. In: M. Traxler and M. Gernsbacher, Editors, *Handbook of Psycholinguistics* (2008).
[11] M. Kutas and K.D. Federmeier, Electrophysiology reveals semantic memory use in language comprehension, *Trends Cogn. Sci.* **4** (2000), pp. 463–470.
[12] M. Kutas and S.A. Hillyard, Reading senseless sentences: brain potentials reflect semantic incongruity, *Science* **207** (1980), pp. 203–205.
[13] M. Kutas and S.A. Hillyard, Event-related brain potentials to grammatical errors and semantic anomalies, *Mem. Cogn.* **11** (5) (1983), pp. 539–550.
[14] M. Kutas and V. Iragui, The N400 in a semantic categorization task across 6 decades, *Electroencephalogr. Clin. Neurophysiol.* **108** (1998), pp. 456–471.
[15] E.F. Lau, C. Phillips and D. Poeppel, A cortical network for semantics: (de)constructing the N400, *Nat. Rev. Neurosci.* **9** (2008), pp. 920–933.
[16] V. Lestou and F.E. Pollick, Neural substrates for action understanding at different description levels in the human brain, *J. Cogn. Neurosci.* **20** (2008), pp. 324–341.
[17] B. Marí-Beffa, D.J. Valdés, D. Cullen, A. Catena and G. Houghton, Erp analyses of task effects on semantic processing from words, *Cogn. Brain Res.* **23** (2005), pp. 293–305.
[18] M. Martín-Loeches, The gate for reading: reflections on the recognition potential, *Brain Res. Rev.* **53** (2007), p. 89.
[19] M. Martín-Loeches, J.A. Hinojosa, P. Casado, F. Muñoz and C. Fernández-Frías, Electrophysiological evidence of an early effect of sentence context in reading, *Biol. Psychol.* **65** (2004), p. 265.
[20] M. Martín-Loeches, J.A. Hinojosa, C. Fernández-Frías and F.J. Rubia, Functional differences in the semantic processing of concrete and abstract words, *Neuropsychologia* **39** (2001), p. 1086.
[21] G. McCarthy, A.C. Nobre, S. Bentin and D.D. Spencer, Language-related field potentials in the anterior-medial temporal lobe. I. Intracranial distribution and neural generators, *J. Neurosci.* **15** (1995), pp. 1080–1089.
[22] H.J. Neville, S.A. Coffey, D.S. Lawson, A. Fischer, K. Emmorey and U. Bellugi, Neural systems mediating American Sign Language: effects of sensory experience and age of acquisition, *Brain Lang.* **57** (1997), pp. 285–308.
[23] H.J. Neville, D.L. Mills and D.S. Lawson, Fractionating language: different neural subsystems with different sensitive periods, *Cereb. Cortex* **2** (1992), pp. 244–258.
[24] A.C. Nobre and G. McCarthy, Language-related field potentials in the anterior-medial temporal lobe. Effects of word type and semantic priming, *J. Neurosci. Methods* **15** (1995), pp. 1090–1098.
[25] K.A. Pelphrey and J.P. Morris, Grasping the intentions of others: the perceived intentionality of an action influences activity in the superior temporal sulcus during social perception, *J. Cogn. Neurosci.* **16** (2004), pp. 1706–1716.
[26] A.M. Proverbio, R. Adorni and A. Zani, Inferring native language from early bio-electrical activity, *Biol. Psychol.* **80** (2009), pp. 52–63.
[27] A.M. Proverbio, G. Leoni and A. Zani, Language switching mechanisms in simultaneous interpreters: an ERP study, *Neuropsychologia* **42** (2004), p. 1636.

[28] Proverbio, A.M., Riva, F., Zani, A., When neurons do not mirror the agent's intentions: sex differences in neural coding of goal-directed actions, Biol. Psychol. (in revision).
[29] A.M. Proverbio, A. Zani and R. Adorni, The left fusiform area is affected by written frequency of words, *Neuropsychologia* **46** (2008), p. 2292.
[30] G. Rizzolatti, L. Fadiga, M. Matelli, V. Bettinardi, E. Paulesu, D. Perani and F. Fazio, Localization of grasp representations in humans by PET. 1. Observation versus execution, *Exp. Brain Res.* **111** (1996), pp. 246–252.
[31] A.P. Rudell, The recognition potential and the word frequency effect at a high rate of word presentation, *Cogn. Brain Res.* **8** (1999), pp. 173–175.
[32] A.P. Rudell and J. Hua, The recognition potential, word difficulty, and individual reading ability: on using event-related potentials to study perception, *J. Exp. Psychol. Hum. Percept. Perform.* **23** (1997), pp. 1170–1195.
[33] R. Saxe and D.K. Xiao *et al.*, A region of right posterior superior temporal sulcus responds to observed intentional actions, *Neuropsychologia* **42** (2004), pp. 1435–1446.
[34] M. Villarreal and E.A. Fridman *et al.*, The neural substrate of gesture recognition, *Neuropsychologia* **46** (2008), pp. 2371–2382.