Time-course and neurophysiological underpinnings of metaphor in literary context

Valentina Bambini[a], Paolo Canal[b*], Donatella Resta[c*], Mirko Grimaldi[c]

[a] *Center for Neurocognition and theoretical Syntax (NEtS), School of Advanced Studies IUSS Pavia, Piazza della Vittoria 15, 27100 Pavia, Italy*

[b] *Laboratorio di Linguistica "G. Nencioni", Scuola Normale Superiore, Piazza dei Cavalieri 7, 56100 Pisa, Italy*

[c] *Centro di Ricerca Interdisciplinare sul Linguaggio (CRIL), University of Salento, Piazza Angelo Rizzo 1, 73100 Lecce, Italy*

** equally contributed*

Corresponding Author:

Valentina Bambini, Center for Neurocognition, Epistemology and theoretical Syntax (NEtS), School of Advanced Studies IUSS Pavia, Piazza della Vittoria 15, 27100 Pavia, Italy; Fax +39-0382375899; Phone +39-0382375841; Email: valentina.bambini@iusspavia.it

Other authors' email addresses:

Paolo Canal: paolo.canal@sns.it

Donatella Resta: donatellaresta@gmail.com

Mirko Grimaldi: mirko.grimaldi@unisalento.it

Abstract

Several theoretical proposals tried to account for the meaning open-endedness of metaphors in literature and for the effortful process that they trigger in the readers. Yet very few experiments have tackled the neurophysiological underpinnings of literary metaphor. Here we used Event-Related brain Potentials (ERPs) to explore the temporal dynamics of comprehending metaphors from Italian poems and novels (e.g., *grass of velvet*), presented in their original context, as compared with literal expressions (e.g., *throne of velvet*). Results evidence a more negative ERP response for metaphors, unfolding in an N400 followed by a sustained negativity over frontal sites, suggesting a long-lasting effort in elaborating figurative meanings. While the N400 might be indicative of lexical/semantic processes typical of metaphors and amplified by the literary context, the sustained negativity might reflect the manipulation of multiple meanings in working memory, possibly responsible for the poetic effect. Interestingly, the late negativity effect is driven by familiarity, with a more negative response for those metaphors that are less familiar. These findings offer material to discuss ideas put forward in pragmatics, literary studies, and cognitive neuroscience of literature, like the condensation of weak implicatures, the foregrounding, and the relation between a metaphor and its context.

Keywords: literary metaphor; ERP; neuropragmatics; experimental pragmatics; neurocognitive poetics; sustained negativity; context

Introduction

For a long time metaphor was seen as a rhetorical device mastered maximally by poets, a classic topic for literary studies and research in rhetoric, poetics, and stylistics. A radical change of perspective can be traced back to the publication of *Metaphors We Live By*, the seminal book by Lakoff and Johnson, which emphasized the pervasiveness of metaphors outside literature, in communication and cognition (Lakoff & Johnson, 1980). In the last decades research has contributed knowledge on the phenomenon of metaphor, described in pragmatics as a loose use of language commonly occurring in everyday communication (Sperber & Wilson, 2008) and in Cognitive Linguistics as an important conceptual mapping operation underlying our mental activity (Gibbs, 2011). It is in this new perspective that metaphor – seen as a communicative and cognitive tool in everyday life – has attracted the interest of experimental approaches in psychology and neuroscience, mainly in relation to fields such as experimental pragmatics, neuropragmatics, and embodied cognition (Bambini, 2010; Bambini & Resta, 2012; Glucksberg, 2003; Jamrozik, Mcquire, Cardillo, & Chatterjee, 2015).

While acknowledging the role of metaphor in everyday life communication, one should not neglect that metaphors remain a distinctive feature of literature, where it generates evocative images that hit the readers' minds. From the distributive point of view, the percentage of metaphorical expressions in literature is striking, and much higher than in everyday language. Attempts to count metaphors in language usage reported an estimation of 5% in TV debates and news programs (Graesser, Mio, & Millis, 1989), while it was calculated that 11% of the lexical units in fiction are metaphor-related (Steen, Dorst, Herrmann, Kaal, & Krennmayr, 2010), and similar or higher percentage are likely in lyrics or other literary genres (Goatly, 1997).

Metaphor in literature is commonly described as less familiar, richer and open-ended in meaning with respect to metaphor elsewhere, and several scholars embarked into the description of the linguistic and cognitive mechanisms underlying the phenomenon of literary metaphor (Semino & Steen, 2008). However, experimental approaches to literary metaphor are extremely scant, especially when it comes to its brain correlates, and "the lion's share of the empirical literature studied context-isolated, 'nonpoetic' sentential metaphors" (Jacobs & Kinder, 2017). In this study we aimed at providing initial evidence on the neurophysiological underpinnings and temporal dynamics of literary metaphor as presented in their original contexts. In constructing the experiment, we followed up on a previous work devoted to the construction of a dataset of literary metaphors in Italian, balanced for the main psycholinguistic variables (Bambini, Resta, & Grimaldi, 2014). The rationale of the study is based on cognitive accounts on literary metaphor, the role of context, and empirical evidence on conventional and novel metaphors.

Cognitive approaches to literary metaphor and the role of context

Several scholars tried to tackle the cognitive processes involved in the resolution of literary figures. While all approaches seem to agree that metaphor in literature is more novel and more creative than metaphor in ordinary language, scholars disagree on the source of these properties (Semino & Steen, 2008). On the one hand, some authors emphasized a discontinuity between literary and non-literary metaphors, claiming that metaphor in literature is radically different from metaphor in ordinary language (Leech, 1969; Tsur, 1992). On the other hand, some scholars argued for a continuity between the two, as based on the same mechanisms (Gibbs, 1994; Turner, 1996). Within those who emphasized continuity, two approaches stand out as the most elaborated accounts of poetic metaphors, one within the frame of Cognitive Linguistics and specifically Conceptual Metaphor Theory, and the other in the frame of pragmatics and specifically Relevance Theory. The Conceptual Metaphor approach posits that in the course of metaphor comprehension two conceptual domains are connected (the target and the source domain): poets use the same tools working for everyday language but in novel and talented ways (Lakoff & Turner, 1989). Although some metaphors reflect completely new source-to-target domains, new metaphors are often creative instantiations of conventional conceptual metaphors (Gibbs, 2006). Poets may take conceptual metaphors common in everyday language and manipulate them, challenging their features and boundaries, for instance by extending and elaborating them in unusual ways. For example, when Horace referred to death as the "eternal exile of the raft" he added further content to the conceptual metaphor DEATH IS DEPARTURE through the concepts of "exile" and "raft". The concept of "exile" is not simply "departure", but introduces the ideas of unwanted banishment, and the desire to come back, while the concept of "raft" evokes the ideas of uncertainty and danger. Thus, the poet elaborated an enduring conceptual metaphor in such a way that the reader thinks of death in a very different perspective (Lakoff & Turner, 1989).

Studies inspired by the Conceptual Metaphor Theory suggest that embodiment mechanisms play an important role in the way we understand metaphors (Gibbs, 2006; Jamrozik et al., 2015). Typically, bodily experience provides the source domains for metaphorically structuring aspects of abstract target domains, and metaphor understanding is based on representing abstract knowledge that arises from bodily experience. In the strong embodiment account, this representation is not amodal, but rather relies on the simulation of sensory-motor processes associated with the source domains (Gibbs, 2006). People construct an embodied simulation of what it must be like for speakers or writers to have certain thoughts, and they imaginatively engage in similar actions. When reading a poem about love that capitalizes on the metaphor RELATIONSHIPS ARE JOURNEY, the reader

imaginatively simulates the journey and the travelling, and this facilitates the metaphorical understanding of the abstract idea that romantic relationships move along a path.

A different perspective is taken by research in pragmatics, which sees metaphor as understood through the formation of ad hoc concepts and the derivation of implicatures. In this frame, poetic metaphors have stronger communicative power than conventional metaphors, and much of this power comes from the indefinable range of weak implicatures generated by the figurative expressions, i.e., meanings that, differently from strong implicatures, do not fall under the speaker's communicative intentions but are largely dependent on the addressee (Sperber & Wilson, 2008). While conventional metaphors convey a straightforward implicature, creative metaphors communicate a condensation of acceptable implicatures, and the addressee is engaged in the exploration of at least some of these. Understanding this range of weakly communicated implicatures requires more effort, but produces more cognitive effects, in the form of aesthetic experience, i.e., "poetic effects" (Pilkington, 2000). Take the example "His ink is pale", a remark produced by Flaubert in relation to the poet Leconte de Lisle. The assumptions behind are explored using the encyclopedic entries "ink" and "pale" as a starting point. "Ink" is a poetic metonymy standing for "writing" or "style". Then the reader is prompted to explore the encyclopedic entries attached to "pale". The reader derives an implicature such as "Leconte de Lisle's writing is weak", and from that a wide range of weak implicatures is derived around the idea that the style of writing is boring and without creativity: "Leconte de Lisle's writing lacks contrast", "Leconte de Lisle's writing will not last", "Leconte de Lisle's writing may fade", and so on (Pilkington, 2000). Derivable implicatures are indefinite, there is no cut-off point that limits them, and the poetic power of literary metaphors derives from the indeterminacy of these weak implicatures. The addressee is asked to apply additional cognitive effort in understanding poetic metaphors, but to this effort greater cognitive effects correspond that cannot be achieved by simply saying, "Leconte de Lisle's writing is boring". The indefiniteness and indeterminacy of the range of implicatures also explains why interpretation might differ across individuals.

Crucially, context plays a fundamental role in preparing interpretation, in guiding the reader from the encyclopedic entries to the construction of the weak implicatures through the activation of many assumptions, or better networks of assumptions. Pilkington uses a nice example from Heaney's poem "Digging": "Between my finger and my thumb / The squat pen rests. I'll dig with it" (Pilkington, 2000). This metaphor may appear quite obscure out of context, yet its interpretation was prepared by the context set up earlier in the poem, where digging is described as the way in which the poet's forefathers earned their living, and as an activity with a long tradition in the community, hard and honest work that inspires awe. Based on these contextual assumptions the reader is able to access the properties of digging that might be transferred to writing poetry. A good

poem gives directions to the interpretation of the metaphorical expressions, by creating a context that enables the reader to explore the many concepts involved and interpret the metaphorical expression in a rich and creative way.

More generally, the importance of all aspects of context in shaping the meaning of expressions has been largely emphasized in research in pragmatics (Bezuidenhout, 2001; Carston, 2010; Clark, 1992) and recently also by approaches stemming from Conceptual Metaphor Theory (Kövecses, 2015). Moreover, studies have shown that metaphor recognition is largely influenced by the type of discourse context, defined as the use of linguistic objects in socially constructed settings (e.g., literary vs journalistic) (Steen, 1992). When readers adopt a literary mode, they recognize more metaphors compared with a journalistic reading mode (Steen, 1994). Metaphorical statements are considered meaningful even when only presumed to be written by poets (Gibbs, Kushner, & Mills, 1991). Also, several properties of metaphors interact with the type of discourse: unclear metaphors receive more attention than clear metaphors, and positive metaphors are recognized more often than negative metaphors in literary reading than in journalistic reading (Steen, 1994). In a rating study aiming at constructing a dataset of literary metaphors, we showed that the comprehension of literary expressions relies on a cluster of variables (mainly familiarity, difficulty, and meaningfulness) that are strongly influenced by the context (Bambini, Resta, & Grimaldi, 2014). Specifically, a literary metaphor presented in its original context is judged as more difficult yet more meaningful than the same metaphor presented out of context. In short, comprehending a literary metaphor is largely interconnected with the processing of its literary context, and this should be taken into account when planning neurophysiological investigations.

ERP studies on every-day metaphor and the effect of familiarity

Despite several works tackling literary metaphors in its cognitive aspects, and although some authors have advocated an empirical turn in poetics and literary studies (Semino & Steen, 2008; Willems & Jacobs, 2016), little research has been devoted to the experimental level, also due to the difficulty of testing literary materials while controlling for the relevant psycholinguistic variables. This gap is especially evident in the literature using the recording of Event-Related brain Potentials (ERPs), a methodology that offers great insights into processing mechanisms with millisecond precision (Luck & Kappenman, 2012), which by contrast has been largely used to test the processing of non-literary metaphors. From this literature, and especially from those studies employing novel and creative metaphors, one can nevertheless gain insights for the investigation of the neurophysiological underpinning of literary metaphor.

Since the work of Pynte and colleagues (Pynte, Besson, Robichon, & Poli, 1996), a large body of studies compared the processing of sentences ending either with a literal or with a metaphorical expression (Rataj, 2014). The most common paradigm employs nominal metaphors, contrasted with paired literal expressions, e.g., "Those fighters are lions" vs "Those animals are lions". Throughout studies, the most consistent result is a biphasic pattern of brain activity, with an earlier negativity (N400) followed by a later positivity (P600/LPC) (Bambini, Bertini, Schaeken, Stella, & Di Russo, 2016; Coulson & Van Petten, 2002; De Grauwe, Swain, Holcomb, Ditman, & Kuperberg, 2010; Schmidt-Snoek, Drew, Barile, & Agauas, 2015; Weiland, Bambini, & Schumacher, 2014).

In general, the N400 is a negative component linked to efforts in terms of lexical access and semantic representation. While some authors posit more emphasis on the lexical access aspects, others highlight the link between N400 and semantic memory and conceptual representations (Kutas & Federmeier, 2011a). A well-known aspect of the N400 is its context-sensitivity, as its amplitude is reduced depending on the degree of contextual expectancy, either within the sentence or the discourse (Kutas & Federmeier, 2011b; Van Berkum, 2012). In the case of metaphor comprehension, the N400 has been generally associated with the effort in meaning retrieval and operations at the conceptual level (Coulson & Van Petten, 2002; De Grauwe et al., 2010; Lai, Curran, & Menn, 2009).

The P600/LPC is a positive component originally observed for syntactic operations and later connected also to a spectrum of semantic and pragmatic phenomena (Bornkessel-Schlesewsky & Schlesewsky, 2008; Brouwer, Fitz, & Hoeks, 2012). It is observed, for instance, for the comprehension of ironic statement, where P600-related pragmatic effects are not accompanied by N400 effects (Regel, Gunter, & Friederici, 2011; Spotorno, Cheylus, Van Der Henst, & Noveck, 2013). For metaphor, the interpretation of the P600 varies, being associated either with further meaning elaboration (De Grauwe et al., 2010) or with the successful retrieval of the conceptual metaphor (Coulson & Van Petten, 2002). Recently, Bambini and colleagues compared metaphors and literal counterparts in different contextual conditions (Bambini et al., 2016). They replicated the biphasic N400-P600 pattern when metaphors are presented in a minimal context, and also showed that – when context is more supportive – the N400 for metaphor is reduced, while the P600 effect remains. In light of these findings, the authors argued that the N400 is linked to lexical/semantic operations as shaped by context, i.e., to meaning operations that become effortless when context is supportive enough, while the P600 is indicative of as an integrative, interpretative phase.

Within the experimental literature on metaphor, a number of studies have emphasized the role of familiarity in shaping metaphor processing (Schmidt & Seger, 2009). By familiarity the literature usually refers to the speaker's frequency of experience with a certain expression and its meaning, as judged on a scale ranging from very conventional to very novel (Bambini, Ghio, Moro, &

Schumacher, 2013; Cardillo, Schmidt, Kranjec, & Chatterjee, 2010). In terms of ERP patterns, familiarity seems to affect especially the negativity. Lai and colleagues reported that, compared with literal statements (e.g., *There was too much food to digest*), both conventional (e.g., *There was too much info to digest*) and novel metaphors (e.g., *There was too much love to digest*) elicited an N400 effect, but for conventional metaphors the negativity ended earlier (Lai et al., 2009). Specifically, the ERP response to conventional metaphors converged with that of literal expressions in the 440-560 time window, while for novel metaphors the negativity was prolonged. In another study, Lai and Curran showed that priming the underlying conceptual mapping structure affected the N400 effect differently for conventional and novel metaphors (Lai & Curran, 2013).

Furthermore – and importantly for our investigation on literary metaphors –, a study using novel metaphors extracted from literary texts showed a prominent effect of familiarity (Arzouan, Goldstein, & Faust, 2007). The authors compared the ERP response to novel metaphors drawn from poetry texts in Hebrew (e.g., *ripe dream*), conventional metaphoric expressions (e.g., *lucid mind*), literal semantically related (e.g., *burning fire*) and unrelated word pairs (e.g., *indirect blanket*) (Arzouan et al., 2007). The N400 response was sensitive to the type of expression, increasing from literal to conventional, novel, and unrelated conditions. Moreover, for novel metaphors, results also evidenced a right-sided late negativity in the 550-880 time window. The authors took the N400 followed by the late negativity pattern as evidence of sequential processing for novel metaphors, and specifically they interpreted the late negativity as suggesting further attempts to integrate meaning in a non-literal fashion. Although the authors did not discuss their findings in the perspective of literary but rather of novel metaphors, the result of the late negativity is especially intriguing, as it suggests that the ERP response to creative expressions such as literary metaphors might be extended, and possibly qualitatively different from the pattern observed for everyday metaphors. There is also another study on metaphor comprehension that, despite not including literary materials, employed a passive conceptual expansion task that enhanced creativity processes (Rutter et al., 2012). The authors compared literal, novel metaphorical and anomalous sentences, asking subjects to judge appropriateness and unusualness. There too results evidenced an N400 followed by a sustained negativity in the 500-900 time-window for novel metaphors compared with literal stimuli, not limited to right sites but broadly distributed, interpreted as ongoing effort in integrating concepts.

After this short detour across ERP studies on metaphor, a few elements appear as striking: (a) no ERP studies focused specifically on literary metaphor, considered not merely as a type of novel metaphor but also in light of its specific features as highlighted in the theoretical approaches; (b) the empirical literature on literary metaphor did not take into account the dimension of context, although context is a fundamental feature in determining the processing of literary metaphors.

Rationale of the study

Given the scant information on literary metaphor in the brain on the one hand and the increasing interest towards psychological testing of theoretical accounts (Steen, 2001) and cognitive neuroscience of literature (Jacobs, 2015; Willems & Jacobs, 2016) on the other hand, we designed a study to tackle the neural mechanisms underlying the comprehension of metaphors in literature. We aimed at increasing the empirical evidence on literary metaphor, in a more specific perspective compared with previous studies, i.e., by taking the processing of literary metaphor as the focus of our attention. More specifically, we aimed at identifying the neurophysiological processes involved in literary metaphor comprehension, presented in their context of use, and discussing the results in the light of available theoretical accounts. To this aim, we employed a set of materials extracted from Italian literary texts and we devised an ERP experiment where these expressions where compared with literal counterparts. We first adopted a traditional approach and run an analysis of variance (ANOVA) on the ERP data in the typical N400 time window and in a later time window (600-900 ms). We were interested in replicating the late negativity findings also for literary metaphors embedded in their original contexts. As an additional analysis, we dug further into the role of familiarity in shaping the ERP response. Considering that familiarity accounts for a great deal of variation in processing non-literary metaphors (conventional vs novel) and that literary metaphors are usually described as creative and rated as less familiar than non-literary metaphors (Katz, Paivio, Marschark, & Clark, 1988; Miall & Kuiken, 1994; Semino & Steen, 2008), investigating whether familiarity plays a role also within literary metaphors might lead to a better understanding of this phenomenon. To disentangle whether potential ERP differences between the literary metaphor and the literal conditions were moderated by the familiarity of the expressions, taking advantage from the familiarity rating values collected while constructing the database in Bambini et al. (2014), we adopted the framework of linear mixed models (Pinheiro & Bates, 2000) on single ERP trials. With this approach by-item variability is accounted for and the condition for generalizing the results beyond the sample of materials is more safely met compared with traditional ANOVA (Baayen, Davidson, & Bates, 2008; Clark, 1973). Results of both analyses will be mainly discussed in terms of temporal development and neurophysiological components, expanding to considerations in relation to cognitive and pragmatic accounts of metaphors in literature.

Material and methods

Participants

Twenty-two Italian students (14 females) participated in the experiment. Mean age was 22.32 (SD 2.44). All were right-handed (scores >85) according to the Edinburgh Handedness Inventory (Oldfield, 1971), with normal or corrected-to-normal vision and no history of neurological or psychiatric disorders. The protocol was approved by the ethical committee of the local health authority of Lecce and all subjects signed a written consent.

Materials

The metaphorical stimuli were extracted from a dataset of metaphors from Italian literary texts scored for the main psycholinguistic variable (Bambini et al., 2014). They consisted of 62 genitive construction (Goatly, 1997), i.e., metaphorical phrases in the "A of B" form (e.g., *grass of velvet*, *sky of pearl*, *eyes of steel*; see Table 1), embedded in their original context, i.e., excerpts from either poems or novels from the 19th and 20th century. To help anchor the responses, the literary metaphors in Bambini et al. 2014 were rated along with matched literal expressions, although these were not included in the published dataset. For this study, the 62 literal expressions paired to the 62 literary metaphors were retrieved from the experimental materials used in Bambini et al. 2014, along with their rating values. The literal stimuli (e.g., *throne of velvet*, *ring of pearl*, *door of steel*) had the same "A of B" structure as literary metaphors, with the B term unaltered and the A term varied with respect to the metaphorical counterparts, but matched for lexical frequency based on the EsploraCoLFIS resource (Bambini & Trevisan, 2012) (Table 2). Literal phrases were embedded in texts extracted from Italian newspaper and magazines. Context size was on average 45.08 (SD 13.32) words, in line with the literature on context effects in metaphor comprehension (Inhoff, Lima, & Carroll, 1984). Specifically, context excerpts included 46.16 (SD 13.40) words in the metaphorical condition and 44 words (SD 13.26) in the literal condition ($p>0.05$). The words preceding the target expressions were 35.15 (SD 13.33) in the metaphorical condition and 34.89 (SD 12.95) in the literal condition, while the post target region included 8.02 (SD 4.52) words in the case of metaphors and 6.11 (SD 2.96) in the case of literal expressions. In addition, passages in both conditions were preceded by an introductory sentence to supply a global topic and facilitate a coherent interpretation.

INSERT TABLE 1

To gain a better description of the properties of the materials, the main psycholinguistic measures used in metaphor research (Cardillo et al., 2010) were considered for each of the 62 metaphorical expressions, and compared with the values obtained for the literal counterparts, both presented in their original context (Table 2). On a scale from 1 to 5, literary metaphors were rated as significantly less familiar, less concrete, less meaningful and more difficult than literal counterparts (all ps<0.001). Similar results were reported in previous rating studies comparing novel metaphors and literal expressions (Bambini et al., 2013; Lai & Curran, 2013; Schmidt-Snoek et al., 2015), since figurativeness is intrinsically intertwined with increased effort in the dimensions captured by the psycholinguistic measures in questions, as distinct from literal usage.

A closer look at familiarity value showed that literary metaphors obtained low values of familiarity, indicating that on average they were perceived as novel by the subjects. Moreover, although globally judged as unfamiliar, literary metaphors obtained intermediate values on the 5-point scale for meaningfulness and difficulty, suggesting that subjects were still able to interpret the phrases and that the expressions were not processed as anomalous. Note that, when rated also out-of-context in the study of Bambini et al. (2014), the same metaphors received lower meaningfulness scores (2.62 vs 3.22 for the in-context presentation used here), indicating that the contexts helped subjects to interpret the expressions.

As for cloze probability, values were lower for literary metaphors than for literal phrases (p<0.001). Yet both sets were in the range of values classified as low contextual constraint (Kutas & Hillyard, 1984), suggesting that ERP differences cannot be solely explained in terms of different closure probabilities.

The mean text readability was 62 for literary contexts and 54 for literal contexts, based on the Gulpease index (Lucisano & Piemontese, 1988), both in the range of easy texts for Italian undergraduates.

INSERT TABLE 2

Procedure and task

Participants were instructed to read the stimuli for comprehension and to perform a word-matching task at the end of each trial. They were informed that some stimuli consisted of excerpts from literary texts, but were not alerted about the presence of metaphors. A short training of five trials to familiarize with the procedure was conducted.

For each trial, the introductory sentence and the context were presented in full in two successive screens and read in a self-paced modality. At 200 ms after the button press, a fixation cross appeared in the center of the screen for 400 ms. At 200 ms after the offset of the cross, a series of words presented one at a time and with appropriate punctuation started, with a duration of 400 ms

and an inter-word interval of 200 ms. The word-by-word series included a few words belonging to the context (not presented in the previous screen), the "A of B" expression, and the end of the passage. This solution served to avoid confounding effects due to sudden serial visual presentation of the critical expressions. ERP were time-locked to the B noun in the "A of B" expressions, either metaphorical or literal. To avoid word position effects (Van Petten & Kutas, 1990), the number of words preceding the target noun was controlled as much as possible. Within the word-by-word sequence, the position of the target was 8.24 (SD 1.45) for literary metaphors and always 8 for literal expressions.

Following each passage, two words were presented on the screen, one related and the other unrelated with respect to the "A of B" expression (e.g., for *throne of velvet*, the words were *king* and *folder*; for *grass of velvet*, the words were *flowers* and *balance*). The task was to select the word that better matched the previous text.

The 124 passages were organized in two lists in a Latin square fashion. 13 fillers were added to each list, with the same structure as metaphorical and literal passages. Each list was presented in 5 blocks of 15 trials each, randomized across participants.

Electrophysiological recording

EEG was recorded using 64 Ag/AgCl scalp electrodes according to the 10/20 system of electrode placement. Three EOG electrodes placed above and below the participant's left eye and at the outer cantus of the right eye, were used for detecting eye movements. All electrodes were referred to linked mastoids. Individual sensors were adjusted until impedances were less than 10 kΩ for all electrodes. All channels were sampled at 250 Hz and amplified using a BrainVision Brain-Amp amplifier with 1 s time constant high-pass filter. Eye movements were corrected with an ocular artifact correction algorithm (Gratton, Coles, & Donchin, 1983) implemented in the *Brain Vision Analyzer*® software.

ERPs were baseline-corrected with a 150-ms pre-stimulus interval and averaged off-line per experimental condition, participant, and electrode site in epochs spanning from -150 to 1000 ms relative to the onset of the target word. Epochs containing EEG values outside the +/-75μV threshold were not included in the individual averages (average rejection rate 13.77%, with no differences between conditions).

Statistical analysis

We ran statistical analyses for the behavioral data over accuracy rates and reaction times, over subjects and items. For the ERP data, we first performed an Analysis of Variance, analyzing electrodes along the Midline separately from Lateral electrodes. In both Midline and Lateralized

ANOVAs, electrodes were arranged along a Longitudinal factor with three levels: Frontal (Lateralized: AF3, AF4, F1 F2, F3, F4, F5, F6; Midline: Fpz, Fz), Central (Lateralized: C1, Cz, C2, C3, C4, CP1, CP2, CP3, CP4; Midline: Cz, CPz) and Parietal (Lateralized: P1, P2, P3, P4, PO3, PO4, O1, O2; Midline: Pz, POz). Time-windows of interest were identified based on the previous literature predicting an N400 effect followed by a later effect, either P600/LPC or sustained negativity. The N400 effect was measured in the canonical time window (300-500 ms) and the later time window ranged from 600 to 900 ms. Mean amplitudes in each time-window were analyzed using a 3 Longitude (Frontal, Central, and Parietal sites) x 2 Phrase Type (Metaphorical, Literal) repeated measures ANOVA. Significant interactions with Longitude were followed-up with one-way ANOVAs on each level of Longitude. When appropriate, results were corrected with the Greenhouse-Geisser method to adjust for sphericity violations.

To better investigate the ERP effects triggered by literary metaphors, we took into account the familiarity of the figurative expressions as resulting from the rating scores, and performed an additional analysis by using linear mixed models, which can simultaneously capture by-item and by-participant random variation. The analysis was carried out using lme4 package (Bates, Mächler, Bolker, & Walker, 2015). The data consisted in the average voltage amplitude recorded during the two time windows tested, for each epoch after artifact rejection, from each participant on the set of ten electrodes that constituted the Frontal level of Longitude (Lateralized and Midline). For technical problems (corrupted logfiles) data from 3 subjects could not be included in the analysis. Using likelihood ratio tests (χ^2 statistic is reported), we evaluated whether the inclusion of the Phrase Type by Familiarity interaction brought to an increase in the models' fit compared with a simpler (and nested) model including Phrase Type and Familiarity as (not interacting) main effects. For likelihood ratio tests, the random structure was the same for the two models and allowed for by-subject and by-item intercept adjustments, and also for adjustments of the effect of Phrase Type (random slopes) for both subjects and items. The model with the best fit is then described, after including in the by-subject random slope the highest order interaction tested in the fixed effect structure (e.g., Condition by Familiarity), thus using the maximal random structure (Barr, Levy, Scheepers, & Tily, 2013) in order to make the analysis conservative (Baayen & Milin, 2010).

Results

Behavioral results

Participants accurately performed the word-matching task both with literary metaphors (91%, range 81-100) and with literal phrases (93%, range 81-100), indicating that they paid attention to all stimuli, with no statistical differences between conditions [$t_1(21)=+1.64$, *ns*; $t_2(30)=+1.56$, *ns*].

Reaction times were slightly longer for literary metaphors (1266 ms, SD 229) than for literal phrases (1192 ms, SD 377), possibly indicating higher effort in metaphor processing that spill over the completion of the task [$t_1(21)=-4.57$, $p\leq0.001$; $t_2(30)=-2.25$, $p\leq0.05$].

ERP results: Analysis of Variance

The visual inspection of the ERP waveforms suggests that literary metaphors triggered more negative ERPs compared with literal expressions. Differences are evident during the N400 component and are sustained in time as affecting the ERPs up to the end of the ERP epoch (Figure 1).

INSERT FIGURE 1

During the N400 time-window (300-500 ms), ANOVAs confirmed that literary metaphors showed more negative voltages than literal expressions, as attested by a main effect of Phrase Type in the Lateralized analysis [$F(1,21)=6.09$, $p<0.05$, $\eta_G^2=0.046$] and in the Midline analysis [$F(1,21)=4.31$, $p<0.05$, $\eta_G^2=0.032$]. Moreover, a significant interaction between Phrase Type and Longitude emerged in the Lateralized analysis [$F(2,42)=7.38$, $p<0.01$, $\varepsilon=0.73$, $\eta_G^2=0.008$] and resulted only marginally significant in Midline electrodes [$F(2,42)=3.71$, $p=0.065$, $\varepsilon=0.64$, $\eta_G^2=0.006$], showing that the larger differences in the ERPs occurred over Frontal [Lateralized: -1.65μV, $F(1,21)=12.99$, $p<0.01$, $\eta_G^2=0.083$; Midline: -1.52μV, $F(1,21)=11.47$, $p<0.01$, $\eta_G^2=0.076$] and Central [Lateralized: -1.14μV, $F(1,21)=4.65$, $p<0.05$, $\eta_G^2=0.042$; Midline: -1.18μV, $F(1,21)=3.09$, $p=0.09$, $\eta_G^2=0.032$], but not Parietal [Lateralized: -1.21μV, $F(1,21)=1.46$, ns; Midline: -0.41μV, $F<1$] electrodes.

During the later time-window (600-900 ms), ANOVA showed similar effects attesting that literary metaphors were still associated with more negative voltages compared with literal expressions. The significant interaction between Longitude and Phrase Type [Lateralized: $F(2,42)=8.60$, $p<0.01$, $\varepsilon=0.61$, $\eta_G^2=0.023$; Midline: $F(2,42)=8.00$, $p<0.01$, $\varepsilon=0.74$, $\eta_G^2=0.027$] confirms that the difference between conditions were pronounced over Frontal [Lateralized: -1.33μV, $F(1,21)=4.75$, $p<0.05$, $\eta_G^2=0.071$; Midline: -1.43μV, $F(1,21)=5.26$, $p<0.05$, $\eta_G^2=0.071$] rather than Central [Lateralized: -0.47μV, $F<1$; Midline: -0.54μV, $F<1$] or Parietal [Lateralized: +0.11μV, $F<1$; Midline: +0.39μV, $F<1$] electrodes.

ERP results: Item-based analysis of familiarity

Likelihood ratio tests performed on the N400 time window did not reveal any significant interaction between Phrase Type and Familiarity [$\chi^2(1)<1$, ns]. This means that the N400 effect triggered by

literary metaphors compared with literal phrases is not influenced by the familiarity of the figurative expressions.

The same test performed on the time-window ranging from 600 to 900 ms revealed instead that the inclusion of the Phrase Type by Familiarity interaction brought to an increase of the model's likelihood, which was close to reach significance [$\chi^2(1)=2.83$, $p<0.1$; $\Delta\beta=+3.06$, $t=1.79$, $p<0.1$]. From the model we can see that the effect of Phrase Type has the same direction and magnitude as observed in the ANOVAs, being metaphorical expressions more negative (-1.47μV) than literal expressions, although the t test evaluating the differences between the two conditions did not reach the level of statistical significance ($t<1$). This is only apparently in contrast with the results reported in the ANOVAs, since, as mentioned in the Statistical Analysis section, mixed models take into account an additional source of variance that is simply neglected by ANOVAs, i.e., by-item variability. Such variability adds to the error term and this is too large to allow for detecting the effect of Phrase Type. The Phrase Type by Familiarity interaction might explain the pattern, suggesting that the variability in the response is actually captured by modulation in Familiarity (Figure 2). In particular, an increase in familiarity determined less positive ERPs in the 600-900 time window in the literal condition and more positive ERPs in the metaphor condition, indicating that the effect of Phrase Type is mainly driven by the larger negativity associated with those literary metaphors that were rated as less familiar.

INSERT FIGURE 2

Discussion

With this work we tackled the issue of the neural correlates of metaphor comprehension by focusing on literary metaphor. Specifically, we aimed at exploring the temporal dynamics of literary metaphor as embedded in its original context, as a fundamental feature in defining its neurophysiological underpinnings. Our findings showed that literary metaphors, compared with literal expressions, elicited a more negative ERP response during a long time interval, unfolding in an N400 followed by a sustained negativity over frontal sites, indicative of a long-lasting effort in elaborating figurative meanings in literary text. Moreover, the sustained negativity effect was affected by familiarity, with unfamiliar metaphors eliciting a more negative response than more familiar metaphors.

These results are in line with those few studies that approached literary metaphors from the perspective of novel metaphor (Arzouan et al., 2007) and creativity (Rutter et al., 2012). These studies reported an N400 followed by a late negativity, and the present results replicated the pattern

by using a set of metaphors that were presented not in isolation but rather embedded in their original literary context. Our study captured the late negativity over frontal electrodes, in a topographic configuration that is more in line with the sustained negativity reported in the literature outside the phenomenon of metaphor. Finally, through an additional item-based analysis we also reported a novel finding, i.e., that the late negativity seems to driven by the familiarity of the figurative expressions. In what follows we will discuss the functional characterization of the N400 and sustain negativity components in light of previous ERP literature, considering also the compatibility with the available theoretical accounts.

The N400 is normally assumed to index lexical/semantic processes as influenced by contextual expectations (Kutas & Federmeier, 2011a). Consistently, the N400 for every-day metaphor is taken as an index of operations at the lexical and semantic level (Bambini et al., 2016; Schmidt-Snoek et al., 2015; Weiland et al., 2014). It seems likely that also the N400 observed for literary metaphor might reflect meaning operations, i.e., efforts in lexical access and semantic representation. When presented with a literary metaphor, readers are prompted to retrieve and process words that, besides being unexpected, carry a conundrum of difficulty, novelty, and meaningful potential, resulting in more effort in meaning processing compared with literal expressions. This interpretation of the N400 is highly compatible with the theoretical accounts of literary metaphors, which all emphasize the richness of conceptual operations triggered by literary metaphors (Lakoff & Turner, 1989; Pilkington, 2000; Prandi, 2017). While Cognitive Linguistics focuses on the manipulation of conceptual domains (e.g., through extension), Relevance Theory focuses on the exploration of the encyclopedic properties of the concepts involved. In both cases, literary metaphor is said to posit a great load in terms of concept representation and semantic knowledge. As for the scalp distribution of the N400 effect, we observed a fronto-central rather than the canonical centro-parietal maximum. This is possibly driven by the negativity that follows the N400: the temporal overlap between the two different effects, one more central (the N400) and one more frontal (the sustained negativity) might result in the fronto-central topography of the former.

The N400 was followed by a sustained negativity, a negative component emerging 300–400 ms after word onset and lasting for several hundreds of milliseconds, differing from the typical N400 in duration (sustained) and scalp distribution (anterior) (Baggio, van Lambalgen, & Hagoort, 2008). A late negativity was also reported by Arzouan et al. (2007) for novel metaphorical word pairs extracted from literary texts and by Rutter et al. (2012) for creative metaphors, although these studies did not observe the more typical frontal distribution of the effect, possibly due to different choices in data processing (average reference versus mastoid reference). The authors interpreted the sustained negativity as an ongoing effort in searching for meaning, which extends the lexical/semantic processes indexed in the N400. Interestingly, a recent study reported an N400

followed by a late negativity (600-900 ms) also for scientific metaphors (Tang, Qi, Jia, Wang, & Ren, 2017). Again, the negativity was interpreted as a prolonged attempt to integrate meaning through knowledge-based inference. Indeed, scientific metaphors too are described as open-ended, although in an inductive rather than in a conceptual fashion as for literary metaphors (Boyd, 1993).

Considering that we observed a more typical sustained negativity, with the classic frontal distribution, we can dig further in the functional role of this component by considering also studies outside the metaphor field. Although the literature is not abundant, the sustained negativity is normally reported for high-level language processing, such as joke comprehension (Coulson & Kutas, 2001), solving referential ambiguity (Van Berkum, Brown, Hagoort, & Zwitserlood, 2003), modifying assumptions (Baggio et al., 2008) and elaborating temporal connectives in discourse (Münte, Schiltz, & Kutas, 1998). In those studies, its presence was taken as indicative of working memory involvement, linked to the representation of discourse-level information that must be kept in the focus of attention to solve ambiguity, connectives, as well as jokes. In light of this evidence, we can put forward a finer interpretation of the sustained negativity observed for literary metaphors, specifying further what kind of process might be going in the prolonged search for meaning. Given the link between the late negativity and working memory, a plausible interpretation is that the lexical/semantic operations indexed in the N400 are followed by a phase where working memory resources are called upon for the manipulation of the multitude of meanings evoked by the metaphors in their literary contexts, generating the typical effect of open-endedness in senses. In other words, the late negativity might arise when a multitude of meanings is at stake.

This interpretation is highly consistent with the Relevance Theoretic account, which assumes that literary metaphors carry a condensation of weakly manifest implicatures (Pilkington, 2000; Sperber & Wilson, 2008). Going through the array of weak implicatures is assumed to be an effortful process where the reader is engaged into an insightful interpretative activity, largely dependent on his/her side, taking into account different potential senses related to the literary expression, with the benefit of surprise and beauty effects. Such a process is likely to be taxing in terms of working memory resources and might indeed be captured by the sustained negativity. In this light, the N400-sustained negativity pattern would be linked to the lexical/semantic efforts followed by the processing of the halo of weakly implicated meanings.

Another aspect of our study was the item-based analysis of familiarity. Results revealed an interaction between Phrase Type (metaphorical vs literal) and Familiarity in the 600-900 time-window, with less familiar literary metaphors eliciting a larger negative effect in the late time-window. This additional analysis allows us to further dig into the nature of the ERP response observed for literary metaphor. First – from the processing point of view –, the familiarity effect emphasized that literary metaphors are not all alike: despite being on average unfamiliar, they vary

in the degree of (un)familiarity, which seem to affect processing efforts especially in the late time window. Specifically, the prolonged search for meaning indexed in the sustained negativity seems to be driven by the familiarity characteristics of the literary metaphors. In understanding a literary metaphor, thus, the reader is engaged in an extended search for interpretation and the effort of this search depends on the familiarity on the expression: the less familiar the metaphor the more effortful the later processing stage, possibly with a different involvement of working memory resources. This finding can be easily accommodated in the Relevance Theory account, as the size of the multiple meanings halo is likely to depend on experience with the verbal expressions, i.e., familiarity, and to increase for the most novel metaphors. Results are also in line with the recently proposed neurocognitive poetics model, which postulates a dual route system devoted to literary reading, where a shallow parsing handling backgrounded elements alternates with a more effortful and slower processing of foregrounded elements (Jacobs, 2015). In this perspective, metaphors are seen as foregrounded elements of language created by the author to convey a dense meaning potential and engage the reader in a search for senses. The model predicts, for foregrounding elements, slower reading tempo, depending also on specific psycholinguistic features of the metaphor and on individual characteristics. The N400-sustained negativity pattern is in line with this model, as the sustained negativity reflects the long-lasting elaboration process, as affected by the familiarity characteristics of the expressions.

Second, the familiarity effect observed here offers indirect evidence on the difference between literary and conventional metaphors. Previous studies reported extended negativities only for novel metaphors compared with conventional metaphors (Arzouan et al., 2007; Lai et al., 2009). In our study, the ANOVA analysis showed a larger negativity for literary metaphors compared with literal expressions, suggesting that the brain response to literary metaphors resembles the brain response to novel rather than to conventional metaphors. Moreover, by adopting an item-based perspective, we showed that familiarity not only differentiates between novel and conventional metaphors as attested in the literature, but also plays a role within the domain of literary metaphors, which are at the cutting edge of unfamiliarity. We thus take our finding of a larger negativity for less familiar literary metaphors as an indirect evidence of the differences between literary metaphors and conventional metaphors, as literary metaphors seem to behave more similarly to novel than conventional metaphors, and this difference seems to be guided precisely by the degree of familiarity. Of course, such an interpretation would be strengthened by a single study including both literary and non-literary metaphors with different degrees of familiarity. Such an enterprise is not easy, especially in the perspective of a tightly controlled design, as the same target word might not be applicable to different conditions (e.g., it might be difficult to use the same word to construct a

non-literary conventional, a non-literary novel and a literary metaphor). We will go back to the difficulty and of comparing different types of metaphors in the "Limitations and outlook" section.

Elaborating further on the issue of familiarity, a large literature beyond EEG showed that it is indeed a crucial parameter in metaphor comprehension (Schmidt & Seger, 2009). For instance, Mashal and Faust run a series of neuroimaging studies that highlighted differences in the neural network supporting novel vs familiar metaphors (Mashal, Faust, Hendler, & Jung-Beeman, 2009), and a special role of the right cerebral hemisphere in processing novel metaphors (Mashal, Faust, Hendler, & Jung-Beeman, 2007). Another fMRI study showed that unfamiliar metaphors elicited greater activity in bilateral temporal regions compared with familiar metaphors and literal sentences (Bambini, Gentili, Ricciardi, Bertinetto, & Pietrini, 2011). Based on these studies, one might hypothesize that also literary metaphors, being at the cutting edge of familiarity, might capitalize on right-sided activations. Familiarity might also be related to embodiment mechanisms. A series of studies on action verbs (Cacciari et al., 2011; Romero Lauro, Mattavelli, Papagno, & Tettamanti, 2013) showed that the involvement of sensory motor cortex is higher for non-conventionalized figurative uses (e.g., *Matilde throws her sadness far away*), but decreases when the figurative meanings are conventionalized (e.g., *The detective throws away the sponge immediately*, which corresponds to an idiomatic expression in Italian, the language of the original stimuli). The strength of sensorimotor engagement might thus reflect the degree of experience that speakers have with figurative senses: while embodiment mechanisms might be irrelevant to fully routinized figurative uses, they might be important for the construction of novel senses via sensorimotor simulation (Jamrozik et al., 2015). Given that literary metaphors are generally highly unfamiliar, their comprehension might involve embodiment mechanisms. The sensory motor grounding is also likely to depend on other aspects of the literary figurative expressions (e.g., semantic features like the association with the domain of seeing, hearing, smelling, etc.), and might be related to aesthetic appreciations. For instance, a study in neurocognitive poetics explored the effects of embodiment and affective features at the single word level, and their potential relation with poetry reception (Jacobs et al., 2015).

Although we did not compare literary and every-day metaphor, further considerations based on the literature might be of some interest. Two differences stand out when we consider the present findings on literary metaphor in light of previous ERP studies on non-literary metaphor: the absence of a late positivity (P600/LPC) and the large N400 effect despite the contextualized presentation of the materials. A P600/LPC effect is often reported for everyday metaphor (Bambini et al., 2016; De Grauwe et al., 2010; Weiland et al., 2014), as well as for other figurative expressions such as metonymy (Schumacher, 2011) and irony (Spotorno et al., 2013). The P600/LPC has been associated with the interpretation of the non-literal expression, i.e., the point where the speaker's

meaning is inferred. Its absence in this study might indicate that in the comprehension of literary metaphor the lexical effort indexed in the N400 does not evolve in a straightforward interpretation phase. Instead, a prolonged effort in elaborating the multitude of meanings is like to occur, possibly indexed by the sustained negativity as observed in our study.

Concerning the N400, its presence should be considered together with the issue of context. One important new aspect introduced in our experimental paradigm was the contextualized presentation of the stimuli, both literary and non-literary. Discourse induced modulations of the N400 have been widely reported in the literature, using normally multisentence passages compared with single sentences (Van Berkum, 2012). For instance, the simple addition of an informative title to locally coherent but globally opaque text induced a reduction of the N400 response recorded for all the words in the passage (St. George, Mannes, & Hoffman, 1994). Focusing on non-literal language, previous studies on metaphors (Bambini et al., 2016) and metonymies (Schumacher, 2011) showed that, when figurative expressions are presented in a supportive context, no N400 effects are observed, indicating that context reduces the efforts in lexical/semantic processing. Here we observed a larger N400 for literary metaphors, although they were provided with their original context. This finding suggests that literary context seems to function differently from non-literary discourse, possibly amplifying instead of reducing the effort in lexical access. This might be due to the fact that literary context, rather than simplifying lexical retrieval as in everyday language (Peleg, Giora, & Fein, 2001), promotes a global interpretative activity of the literary metaphor, which enhances also lexical processes. Consistently with the idea that literary context amplifying interpretation and processing effort, the comparison between out-of-context and in-context presentation of the set of metaphor used here showed higher costs in the latter case (Bambini et al., 2014). The N400 effect reported in this study might thus represent the neurophysiological underpinnings of lexical processing as enhanced by the literary context. Also, this finding might have implications at the theoretical level, as it points in the direction of the specificity of literary context, largely discussed by literary theorists (Miall & Kuiken, 1999), and it strengthens the need of considering literary metaphors in relation to the context in which they occur, as emphasized in distributional-linguistic accounts (Steen, 2001) and narratology approaches (Fludernik, 2011).

Limitations and outlook

The present study represents a first attempt towards the identification of the brain signature of literary metaphor, using a highly controlled paired design. Results are encouraging in showing an N400 followed by a sustained negativity, in line with previous literature on similar topics. Results are also interesting in pointing to a number of similarities and differences with respect to the ERP

patterns reported for non-literary metaphor, namely the effect of familiarity, the absence of a positivity. However, these considerations would find confirmation only in a combined experiment comparing literary and non-literary metaphors, with different values of familiarity, taking into account also the effect of context in shaping interpretation processes. This is not an easy enterprise, since metaphor, and especially literary metaphor, is a multidimensional phenomenon. As describe by Jacobs and Kinder (2017), there is on the one hand the need for more natural and ecologically experimental materials, and on the other hand the need of quantitative descriptors that account for the many features of metaphor (familiarity, difficulty, as well as imagery and beauty). A solution to these conflicting needs is to adopt item-based approach, as we tried to do in this study, although limited to literary metaphors and familiarity. A wider experiment, including different regressors to account for the variability in the materials, would be of paramount importance for a number of issues. First, a similar experiment could help in tackling theoretical issues, such as first of all discriminating between the continuity and the discontinuity views. The continuity view claims that literary metaphors simply extend and exploit the same mechanisms responsible for everyday metaphor, while the discontinuity view argues for radically different processes. When considering the findings of this study, the familiarity result seems in favor of the continuity account, since the same dimension accounts for differences between literary metaphors in the late negativity, as well as for differences between novel and conventional non-literary metaphors. Yet the qualitative differences between the ERP response to literary metaphors observed here and the ERP response to non-literary metaphors reported in the literature (with a late positivity) is more in line with the discontinuity view. In sum, our findings per se cannot be considered conclusive relative to the continuity vs discontinuity debate. To this respect, Semino and Steen's position seems very reasonable in assuming that both continuity and discontinuity views might capture some aspects of literary metaphors, and in calling for more experimental research (Semino & Steen, 2008). Other theoretical models that could be put to test are the neurocognitive poetics model, as well as the metarepresentational mode of metaphor processing. As already mentioned, the neurocognitive poetics model describes poetic reading as based on the alternation of a fast upper route triggered by backgrounded text elements, and a slower lower route responding to foregrounded elements, including creative and novel metaphors (Jacobs, 2015). Under the idea of the metarepresentational mode, Carston discusses the possibility that, especially for creative and extended metaphors, a special processing mode might be triggered, where literal meaning is metarepresented, i.e., held for a further process, and used as material for slower and more reflective interpretative inferences (Carston, 2010; Carston & Wearing, 2011). Both proposals address metaphor comprehension in the context of the original text, targeting immersed metaphorical thinking rather than sporadic

metaphorical comprehension. Similar proposals could be tested only by overcoming the size of the contexts used here (45 words on average), and considering longer texts and metaphorical chains.

Second, a wider experiment taking into account multiple metaphor types and multiple factors, from familiarity to beauty, could also offer empirical evidence useful for metaphor identification. Metaphor identification is a long-standing issue in the literature, due to the difficulty of finding a comprehensive definition and clear-cut criteria (Steen, 2016). Currently, two main approaches exist, one based on human annotation (Pragglejaz-Group, 2007), the other consisting in an automated classification based on a list of metaphorical domains (Dodge, Hong, & Stickles, 2015). We believe that identifying how the brain signature for metaphors varies based on their source (i.e., type of text, e.g., literary vs non-literary) and properties (e.g., familiarity) might be of some help in fine-tuning and validating a taxonomy of metaphorical cases to be used for annotation. Ideally, in order to go ecologically and account for the experience of reading, experimenters should use a continuous narrative as stimulus, and record the EEG along the discourse, time-locking the analysis of changes in microvolt amplitude or frequency power to the time points where the annotation system previously identified a metaphor. Similar approaches have been used in fMRI study to detect brain fingerprints of specific linguistic processes during naturalistic language comprehension, such as for instance changes in brain activity related to syntactic complexity as measured through an automated parser during the listening of *Alice in Wonderland* (Brennan et al., 2012). Furthermore, data driven approaches may be used to highlight peculiar processing correlates discriminating metaphors from literal language. First steps in this direction have been made applying machine learning to the frequency domain of the EEG: for instance, Murphy and colleagues were able to detect differences between semantic categories (mammals vs tools) with high degree of accuracy (Murphy et al., 2011). Clearly, moving from the word level and broad semantic categories to the discourse level to decode subtle differences in the metaphor taxonomy is a huge challenge.

Finally, future experiments should take into account not only item-based variation but also subject-based variation. Evidence showed that the acquisition of poetic literacy is a long, slow process of formal literary education (Peskin, 2010), and that metaphor comprehension in general is largely influenced by individual characteristics, such executive control capacities (Columbus et al., 2014). Individual differences are likely to impact also on the brain response, as shown for instance for IQ-related modulations of the N400 amplitude (Kazmerski, Blasko, & Dessalegn, 2003). Here we tested a sample of 22 participants, which is in line with the current ERP literature on metaphor processing. However, data from a larger sample of participants, tested also for individual characteristics, could shed light on how different individuals experience metaphors in reading.

Conclusions

In sum, when considered in terms of temporal dynamics and neurophysiological components, literary metaphor seems to unfold prolonged and costly over two time-windows, one typical of lexical phenomena, the other typical of processes taxing on working memory and linked to the familiarity of the expression. These two effects might be interpreted as enhanced lexical activity promoted by the literary context and manipulations of the halo of meanings typical of poetic experience. Although this investigation was exploratory and our findings cannot be taken as supportive of any cognitive model as a whole, we showed that they offer a basis to model future studies at the crossroads between literary studies and empirical approaches such as neuropragmatics and cognitive neuroscience of literature.

Authors contribution

Study concept and design: MG, DR, VB. Data collection: DR. Data analysis: PC, VB. Data interpretation: VB, PC, MG. Manuscript writing: VB, PC, MG.

Acknowledgments

The work was supported by the MIUR PRIN (Progetti di Ricerca di Rilevante Interesse Nazionale) 2017, project "The Interpretative Brain: Understanding and Promoting Pragmatic Abilities across Lifespan and in Mental Illness," project code 201577HA9M, awarded to VB.

References

Arzouan, Y., Goldstein, A., & Faust, M. (2007). Brainwaves are stethoscopes: ERP correlates of novel metaphor comprehension. *Brain Research*, *60*(1), 69–81. http://doi.org/10.1016/j.neuroimage.2007.02.015

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. http://doi.org/10.1016/J.JML.2007.12.005

Baayen, R. H., & Milin, P. (2010). Analyzing Reaction Times. *International Journal of Psychological Research*, *3*(2), 12–28. http://doi.org/10.21500/20112084.807

Baggio, G., van Lambalgen, M., & Hagoort, P. (2008). Computing and recomputing discourse models: An ERP study. *Journal of Memory and Language*, *52*(1), 36–53. http://doi.org/10.1016/j.jml.2008.02.005

Bambini, V. (2010). Neuropragmatics: A foreword. *Italian Journal of Linguistics*, *22*(1), 1–20.

Bambini, V., Bertini, C., Schaeken, W., Stella, A., & Di Russo, F. (2016). Disentangling metaphor from context: An ERP study. *Frontiers in Psychology*, *7*(559). http://doi.org/10.3389/fpsyg.2016.00559

Bambini, V., Gentili, C., Ricciardi, E., Bertinetto, P. M., & Pietrini, P. (2011). Decomposing metaphor processing at the cognitive and neural level through functional magnetic resonance imaging. *Brain Research Bulletin*, *86*(3–4), 203–216. http://doi.org/10.1016/j.brainresbull.2011.07.015

Bambini, V., Ghio, M., Moro, A., & Schumacher, P. B. (2013). Differentiating among pragmatic uses of words through timed sensicality judgments. *Frontiers in Psychology*, *4*(December), 938. http://doi.org/10.3389/fpsyg.2013.00938

Bambini, V., & Resta, D. (2012). Metaphor and experimental pragmatics: when theory meets empirical investigation. *Humana.Mente Journal of Philosophical Studies*, *23*, 37–60.

Bambini, V., Resta, D., & Grimaldi, M. (2014). A dataset of metaphors from the italian literature: Exploring psycholinguistic variables and the role of context. *PLoS ONE*, *9*(9). http://doi.org/10.1371/journal.pone.0105634

Bambini, V., & Trevisan, M. (2012). EsploraCoLFIS: Un'interfaccia Web per ricerche sul Corpus e Lessico di Frequenza dell'Italiano Scritto. *Quaderni Del Laboratorio Di Linguistica Della Scuola Normale Superiore*, (11), 1–16.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3). http://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using **lme4**. *Journal of Statistical Software*, *67*(1), 1–48. http://doi.org/10.18637/jss.v067.i01

Bezuidenhout, A. (2001). Metaphor and what is said: A defense of a direct expression view of metaphor. *Midwest Studies in Philosophy*, *25*(1), 156–186. http://doi.org/10.1111/1475-4975.00044

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2008). An alternative perspective on "semantic P600" effects in language comprehension. *Brain Research Reviews*, *59*(1), 55–73. http://doi.org/10.1016/j.brainresrev.2008.05.003

Boyd, R. (1993). Metaphor and theory change: What is "metaphor" a metaphor for? In A. Ortony (Ed.), *Metaphor and Thought* (pp. 481–532). Cambridge: Cambridge University Press. http://doi.org/10.1017/CBO9781139173865.023

Brennan, J., Nir, Y., Hasson, U., Malach, R., Heeger, D. J., & Pylkkänen, L. (2012). Syntactic structure building in the anterior temporal lobe during natural story listening. *Brain and Language*, *120*(2), 163–173. http://doi.org/10.1016/j.bandl.2010.04.002

Brouwer, H., Fitz, H., & Hoeks, J. (2012). Getting real about Semantic Illusions: Rethinking the functional role of the P600 in language comprehension. *Brain Research*, *1446*, 127–143.

http://doi.org/10.1016/j.brainres.2012.01.055
Cacciari, C., Bolognini, N., Senna, I., Pellicciari, M. C., Miniussi, C., & Papagno, C. (2011). Literal, fictive and metaphorical motion sentences preserve the motion component of the verb: a TMS study. *Brain and Language*, *119*(3), 149–57. http://doi.org/10.1016/j.bandl.2011.05.004
Cardillo, E. R., Schmidt, G. L., Kranjec, A., & Chatterjee, A. (2010). Stimulus design is an obstacle course: 560 matched literal and metaphorical sentences for testing neural hypotheses about metaphor. *Behavior Research Methods*, *42*(3), 651–64. http://doi.org/10.3758/BRM.42.3.651
Carston, R. (2010). XIII-Metaphor: Ad Hoc Concepts, Literal Meaning and Mental Images. *Proceedings of the Aristotelian Society (Hardback)*, *110*(3pt3), 295–321. http://doi.org/10.1111/j.1467-9264.2010.00288.x
Carston, R., & Wearing, C. (2011). Metaphor, hyperbole and simile: A pragmatic approach. *Language and Cognition*, *3*(2), 283–312. http://doi.org/10.1515/langcog.2011.010
Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, *12*(4), 335–359. http://doi.org/10.1016/S0022-5371(73)80014-3
Clark, H. H. (1992). Making sense of nonce sence. *Arenas of Language Use, Chicago and London: Chicago University Press and CSLI, Pp. 305–40.*, 305–40. http://doi.org/10.1016/S0277-5395(00)00081-9
Columbus, G., Sheikh, N. A., Côté-Lecaldare, M., Häuser, K., Baum, S. R., & Titone, D. (2014). Individual differences in executive control relate to metaphor processing: an eye movement study of sentence reading. *Frontiers in Human Neuroscience*, *8*, 1057. http://doi.org/10.3389/fnhum.2014.01057
Coulson, S., & Kutas, M. (2001). Getting it: Human event-related brain response to jokes in good and poor comprehenders. *Neuroscience Letters*, *316*(2), 71–74. http://doi.org/10.1016/S0304-3940(01)02387-4
Coulson, S., & Van Petten, C. (2002). Conceptual integration and metaphor: an event-related potential study. *Memory & Cognition*, *30*(6), 958–968.
De Grauwe, S., Swain, A., Holcomb, P. J., Ditman, T., & Kuperberg, G. R. (2010). Electrophysiological insights into the processing of nominal metaphors. *Neuropsychologia*, *48*(7), 1965–84. http://doi.org/10.1016/j.neuropsychologia.2010.03.017
Dodge, E., Hong, J., & Stickles, E. (2015). MetaNet: Deep semantic automatic metaphor analysis. In E. Shutova, B. B. Klebanov, & P. Lichtenstein (Eds.), *Proceedings of the 3rd Workshop on Metaphor in NLP* (pp. 40–49). NAACL HLT.
Fludernik, M. (Ed.). (2011). *Beyond Cognitive Metaphor Theory: Perspectives on Literary Metaphor*. London: Routledge.
Gibbs, R. W. (1994). *The Poetics of Mind: Figurative Thought, Language, and Understanding*. Cambridge University Press.
Gibbs, R. W. (2006). Metaphor interpretation as embodied simulation. *Mind and Language*, *21*(3), 434–458. http://doi.org/10.1111/j.1468-0017.2006.00285.x
Gibbs, R. W. (2011). Evaluating Conceptual Metaphor Theory. *Discourse Processes*, *48*(8), 529–562. http://doi.org/10.1080/0163853X.2011.606103
Gibbs, R. W., Kushner, J. M., & Mills, W. R. (1991). Authorial intentions and metaphor comprehension. *Journal of Psycholinguistic Research*, *20*(1), 11–30.
Glucksberg, S. (2003). The psycholinguistics of metaphor. *Trends in Cognitive Sciences*, *7*(2), 92–96. http://doi.org/10.1016/S1364-6613(02)00040-2
Goatly, A. (1997). *The Language of Metaphors*. London: Routledge.
Graesser, A., Mio, J., & Millis, K. (1989). Metaphors in persuasive communication. In D. Meutsch & R. Viehoff (Eds.), *Comprehension and Literary Discourse: Results and Problems of Interdisciplinary Approaches* (pp. 131–154). Berlin: De Gruyter.
Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, *55*(4), 468–84.

http://doi.org/10.1016/0013-4694(83)90135-9
Inhoff, A. W., Lima, S. D., & Carroll, P. J. (1984). Contextual effects on metaphor comprehension in reading. *Memory & Cognition*, *12*(6), 558–567. http://doi.org/10.3758/BF03213344
Jacobs, A. M. (2015). Neurocognitive poetics: methods and models for investigating the neuronal and cognitive-affective bases of literature reception. *Frontiers in Human Neuroscience*, *9*(April 2015), 186. http://doi.org/10.3389/fnhum.2015.00186
Jacobs, A. M., & Kinder, A. (2017). "The Brain Is the Prisoner of Thought": A Machine-Learning Assisted Quantitative Narrative Analysis of Literary Metaphors for Use in Neurocognitive Poetics. *Metaphor and Symbol*, *32*(3), 139–160. http://doi.org/10.1080/10926488.2017.1338015
Jacobs, A. M., Võ, M. L.-H., Briesemeister, B. B., Conrad, M., Hofmann, M. J., Kuchinke, L., … Braun, M. (2015). 10 years of BAWLing into affective and aesthetic processes in reading: what are the echoes? *Frontiers in Psychology*, *6*, 714. http://doi.org/10.3389/fpsyg.2015.00714
Jamrozik, A., Mcquire, M., Cardillo, E. R., & Chatterjee, A. (2015). Metaphor : Bridging embodiment to abstraction. *Psychonomic Bulletin & Review*, *23*(4), 1080–1089. http://doi.org/10.3758/s13423-015-0861-0
Katz, A. N. A., Paivio, A., Marschark, M., & Clark, J. J. M. (1988). Norms for 204 literary and 260 nonliterary metaphors on 10 psychological dimensions. *Metaphor and Symbol*, *3*(4), 191–214.
Kazmerski, V. A., Blasko, D. G., & Dessalegn, B. G. (2003). ERP and behavioral evidence of individual differences in metaphor comprehension. *Memory & Cognition*, *31*(5), 673–89.
Kövecses, Z. (2015). *Where Metaphors Come From: Reconsidering Context in Metaphor*. New York: Oxford University Press.
Kutas, M., & Federmeier, K. D. (2011a). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, *62*(August), 621–47. http://doi.org/10.1146/annurev.psych.093008.131123
Kutas, M., & Federmeier, K. D. (2011b). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, *62*, 621–47. http://doi.org/10.1146/annurev.psych.093008.131123
Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, *307*(5947), 161–3. http://doi.org/10.1038/307161a0
Lai, V. T., & Curran, T. (2013). ERP evidence for conceptual mappings and comparison processes during the comprehension of conventional and novel metaphors. *Brain and Language*, *127*(3), 484–496. http://doi.org/10.1016/j.bandl.2013.09.010
Lai, V. T., Curran, T., & Menn, L. (2009). Comprehending conventional and novel metaphors : An ERP study. *Brain Research*, *1284*, 145–155. http://doi.org/10.1016/j.brainres.2009.05.088
Lakoff, G., & Johnson, M. (1980). *Metaphors We Live By*. Chicago: University of Chicago.
Lakoff, G., & Turner, M. (1989). *Lakoff, G. and Turner, M. (1989). More than cool reason: A field guide to poetic metaphor. Chicago: University of Chicago Press.* Chicago: Chicago University Press.
Leech, G. N. (1969). *A linguistic guide to English poetry*. London: Longman.
Lucisano, P., & Piemontese, M. E. (1988). GULPEASE: Una formula per la predizione della difficolta` dei testi in lingua italiana. *Scuola E Città*, *3*(31).
Luck, S. J., & Kappenman, E. S. (2012). *The Oxford handbook of event-related potential components*. Oxford: Oxford University Press.
Mashal, N., Faust, M., Hendler, T., & Jung-Beeman, M. (2007). An fMRI investigation of the neural correlates underlying the processing of novel metaphoric expressions. *Brain and Language*, *100*(2), 115–126. http://doi.org/10.1016/j.bandl.2005.10.005
Mashal, N., Faust, M., Hendler, T., & Jung-Beeman, M. (2009). An fMRI study of processing novel metaphoric sentences. *Laterality: Asymmetries of Body, Brain and Cognition*, *14*(1), 30–54. http://doi.org/10.1080/13576500802049433
Miall, D. S., & Kuiken, D. (1994). Beyond Text Theory: Understanding Literary Response. *Discourse Processes*, *17*(3), 337–352. http://doi.org/10.1080/01638539409544873

Miall, D. S., & Kuiken, D. (1999). What is literariness? Three components of literary reading. *Discourse Processes*, *28*(2), 121–138. http://doi.org/10.1080/01638539909545076
Münte, T. F., Schiltz, K., & Kutas, M. (1998). When temporal terms belie conceptual order. *Nature*, *395*(6697), 71–73. http://doi.org/10.1038/25731
Murphy, B., Poesio, M., Bovolo, F., Bruzzone, L., Dalponte, M., & Lakany, H. (2011). EEG decoding of semantic category reveals distributed representations for single concepts. *Brain and Language*, *117*(1), 12–22. http://doi.org/10.1016/j.bandl.2010.09.013
Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113.
Peleg, O., Giora, R., & Fein, O. (2001). Salience and context effects: Two are better than one. *Metaphor and Symbol*, *16*, 173–192.
Peskin, J. (2010). The Development of Poetic Literacy During the School Years. *Discourse Processes*, *47*(2), 77–103. http://doi.org/10.1080/01638530902959653
Pilkington, A. (2000). *Poetic Effects: a Relevance Theory Perspective*. Amsterdam: John Benjamins.
Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-Effects Models in S and S-PLUS*. New York: Springer-Verlag. http://doi.org/10.1007/b98882
Pragglejaz-Group. (2007). MIP: A method for identifying metaphorically used words in discourse. *Metaphor and Symbol*, *22*, 1–39.
Prandi, M. (2017). *Conceptual conflicts in metaphors and figurative language*. New York: Routledge.
Pynte, J., Besson, M., Robichon, F.-H., & Poli, J. (1996). The Time-Course of Metaphor Comprehension : An Event-Related Potential Study. *Brain and Language*, *316*(55), 293–316.
Rataj, K. (2014). Surfing the brainwaves of metaphor comprehension. *Poznan Studies in Contemporary Linguistics*, *50*(1), 55–73. http://doi.org/10.1515/psicl-2014-0004
Regel, S., Gunter, T. C., & Friederici, A. D. (2011). Isn't it ironic? An electrophysiological exploration of figurative language processing. *Journal of Cognitive Neuroscience*, *23*(2), 277–93. http://doi.org/10.1162/jocn.2010.21411
Romero Lauro, L. J., Mattavelli, G., Papagno, C., & Tettamanti, M. (2013). She runs, the road runs, my mind runs, bad blood runs between us: Literal and figurative motion verbs: An fMRI study. *NeuroImage*, *83*, 361–371. http://doi.org/10.1016/j.neuroimage.2013.06.050
Rutter, B., Kröger, S., Hill, H., Windmann, S., Hermann, C., & Abraham, A. (2012). Can clouds dance? Part 2: An ERP investigation of passive conceptual expansion. *Brain and Cognition*, *80*(3), 301–310. http://doi.org/10.1016/j.bandc.2012.08.003
Schmidt-Snoek, G. L., Drew, A. R., Barile, E. C., & Agauas, S. J. (2015). Auditory and motion metaphors have different scalp distributions: an ERP study. *Frontiers in Human Neuroscience*, *9*(March), 1–9. http://doi.org/10.3389/fnhum.2015.00126
Schmidt, G. L., & Seger, C. A. (2009). Neural correlates of metaphor processing: The roles of figurativeness, familiarity and difficulty. *Brain and Cognition*, 101016/jbandc200906001. http://doi.org/10.1016/j.bandc.2009.06.001
Schumacher, P. B. (2011). The hepatitis called ...: Electrophysiological evidence for enriched composition. In J. Meibauer & M. Steinbach (Eds.), *Experimental Pragmatics/Semantics* (pp. 199–219). Amsterdam: John Benjamins.
Semino, E., & Steen, G. (2008). Metaphor in literature. In R. W. Gibbs (Ed.), *The Cambridge Handbook of Metaphor and Thought* (Vol. 6, pp. 57–70). Cambridge University Press.
Sperber, D., & Wilson, D. (2008). A deflationary account of metaphors. In R. W. Gibbs (Ed.), *The Cambridge Handbook of Metaphor and Thought* (pp. 84–105). New York: Cambridge University Press. http://doi.org/10.1017/CBO9781139028370.007
Spotorno, N., Cheylus, A., Van Der Henst, J.-B., & Noveck, I. A. (2013). What's behind a P600? Integration Operations during Irony Processing. *PLoS ONE*, *8*(6), e66839. http://doi.org/10.1371/journal.pone.0066839
St. George, M., Mannes, S., & Hoffman, J. E. (1994). Global Semantic Expectancy and Language

Comprehension. *Journal of Cognitive Neuroscience*, *6*(1), 70–83. http://doi.org/10.1162/jocn.1994.6.1.70
Steen, G. (1992). Literary and Nonliterary Aspects of Metaphor. *Poetics Today*, *13*(4), 687–704. http://doi.org/10.2307/1773294
Steen, G. (1994). *Understanding Metaphor in Literature: An Empirical Approach*. New York: Longman Publishing.
Steen, G. (2001). A rhetoric of metaphor: Conceptual and linguistic metaphor and the psychology of literature. In D. Schram & G. Steen (Eds.), *The Psychology and Sociology of Literature: In honor of Elrud Ibsch* (pp. 145–163). Amsterdam: John Benjamins. http://doi.org/10.1075/upal.35.10ste
Steen, G. (2016). Identifying metaphors in language. In E. Semino & Z. Demjén (Eds.), *The Routledge Handbook of Metaphor and Language*. London: Routledge.
Steen, G., Dorst, A. G., Herrmann, J. B., Kaal, A. A., & Krennmayr, T. (2010). Metaphor in usage. *Cognitive Linguistics*, *21*(4), 765–796. http://doi.org/10.1515/COGL.2010.024
Tang, X., Qi, S., Jia, X., Wang, B., & Ren, W. (2017). Comprehension of scientific metaphors: Complementary processes revealed by ERP. *Journal of Neurolinguistics*, *42*, 12–22. http://doi.org/10.1016/j.jneuroling.2016.11.003
Tsur, R. (1992). *Toward a theory of cognitive poetics*. Amsterdam: Elsevier (North Holland) Science Publishers.
Turner, M. (1996). *Turner, M. (1996). The Literary Mind: The Origins of Thought and Language*. New York: Oxford University Press.
Van Berkum, J. J. A., Brown, C. M., Hagoort, P., & Zwitserlood, P. (2003). Event-related brain potentials reflect discourse-referential ambiguity in spoken language comprehension. *Psychophysiology*, *40*(2), 235–248. http://doi.org/10.1111/1469-8986.00025
Van Berkum, J. J. A. Van. (2012). The electrophysiology of discourse and conversation. In M. J. Spivey, K. McRae, & M. F. Joanisse (Eds.), *The Cambridge Handbook of Psycholinguistics* (pp. 589–614). New York: Cambridge University Press.
Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory & Cognition*, *18*(4), 380–393. http://doi.org/10.3758/BF03197127
Weiland, H., Bambini, V., & Schumacher, P. (2014). The role of literal meaning in figurative language comprehension: Evidence from masked priming ERP. *Frontiers in Human Neuroscience*, *9*, 11–12. http://doi.org/10.3389/fnhum.2014.00583
Willems, R. M., & Jacobs, A. M. (2016). Caring About Dostoyevsky: The Untapped Potential of Studying Literature. *Trends in Cognitive Sciences*, *20*(4), 243–245. http://doi.org/10.1016/j.tics.2015.12.009

Table 1. Examples of literary metaphors and literal counterparts in their original contexts (original Italian and literal English translation). Target words in bold. Introductory sentences in italics.

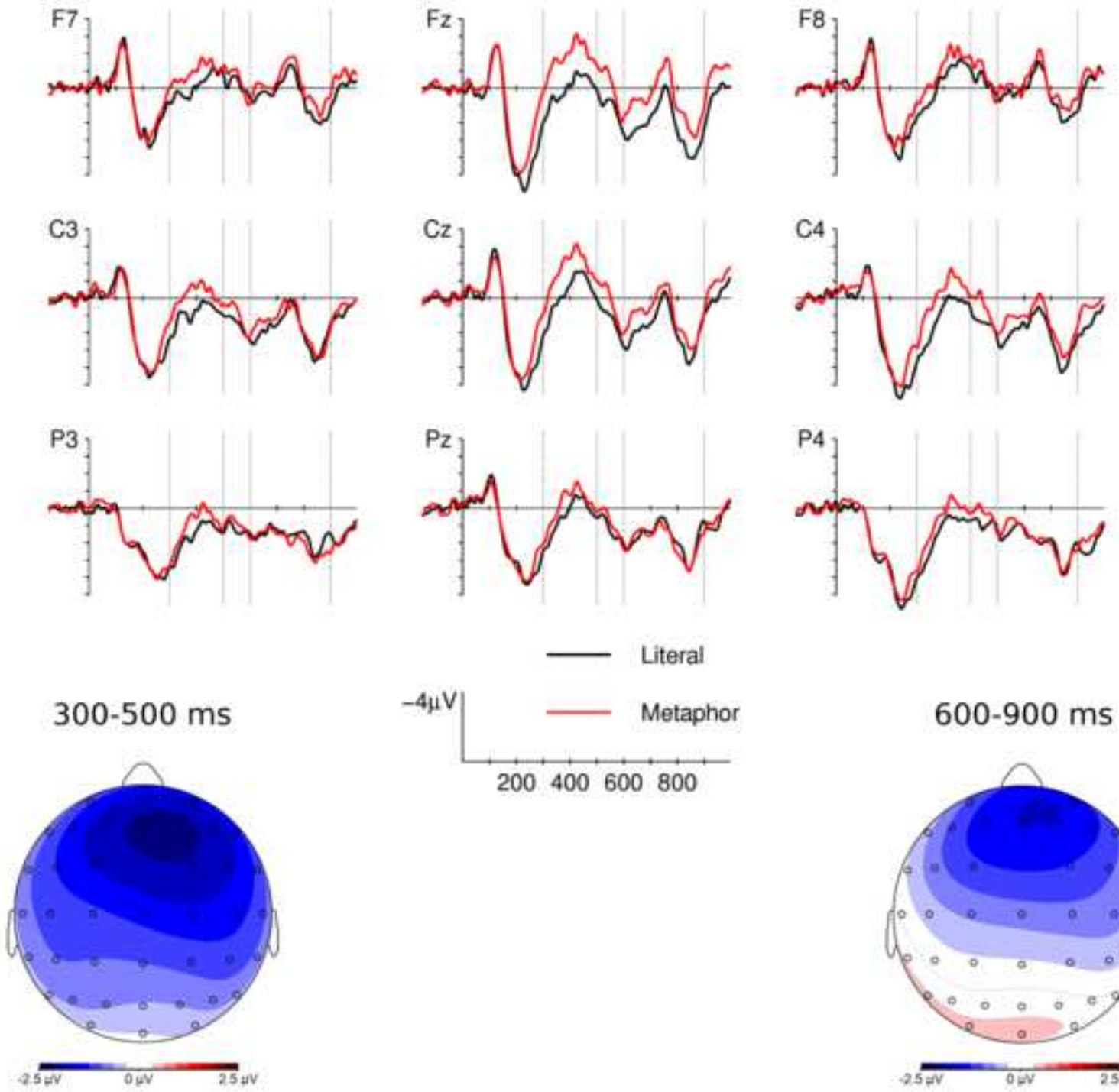
Literary metaphor Source: G. Gozzano, Il Giovenile Errore (The Youthful Error)	*In questo brano un uomo si interroga sulla sua esperienza.* Non so se veramente fu vissuto Quel giorno della prima primavera. Ricordo o sogno? Un prato di **velluto**, Ricordo o sogno? Un cielo che s'annera. English translation: *In this passage a man is wandering about his experience.* I don't know if really was lived that day of the early spring Do I remember or dream? A grass of **velvet**, Do I remember or dream? A sky that grows dark.
Literal expression Source: Ecowebnews.it	*Questo brano parla di una donna dello spettacolo.* Venti minuti appena la sua ospitata: seduta su un trono di **velluto** rosso e dorato ha sorseggiato una coppa di spumante. English translation: *This passage is about a show woman.* Her visit lasted only twenty minutes: sitting on a throne of (red and golden) **velvet**, she sipped a bowl of sparkling wine.
Literary metaphor Source: G. D'Annunzio, Orsola	*Questo brano descrive un uomo, Don Vincenzo, e una donna, Teodora.* Don Vincenzo Bucci, il medico, seduto, guardava il pomo d' argento della bella mazza, le belle corniole incise ch' egli aveva negli anelli delle dita, aspettando. Teodora La Iece, una tessitrice vicina, stava ritta, in silenzio, tutta intenta nell' atteggiare a dolore la faccia bianca e lentigginosa, gli occhi d'**acciaio**, la bocca crudele. English translation: *This passage describes a man, Don Vincenzo, and a woman, Teodora.* Don Vincenzo Bucci, the doctor, sitting, was looking at the silver knob of the nice stick, the beautiful carnelians engraved that he had in the finger rings, waiting. Teodora La Iece, a neighboring weaver, was standing, silent, intent in striking a pose of pain with the wait and freckled face, the eyes of **steel**, the cruel mouth.
Literal expression Source: LaRepubblica.it	*Questo brano parla dell' arresto di un malavitoso di nome Giuseppe Bastone.* È stato arrestato questa mattina nel quartiere napoletano di Scampia. Bastone, latitante da circa un anno, è stato trovato nel suo bunker di via Labriola. Al rifugio, delle dimensioni di tre metri per tre, si accedeva attraverso una botola posta al di sotto delle scale e oltrepassando una pesante porta d'**acciaio**, un cunicolo di oltre 200 metri. English translation: *This passage is about the arrest of a gangster named Giuseppe Bastone.* He was arrested this morning in the Scampia neighborhood of Naples. Bastone, fugitive for about one year, was found in his bunker in via Labriola. In the refuge, three squared meters, one could enter through a manhole below the stairs and overstepping a heavy door of **steel**, a tunnel of about 200 meters.

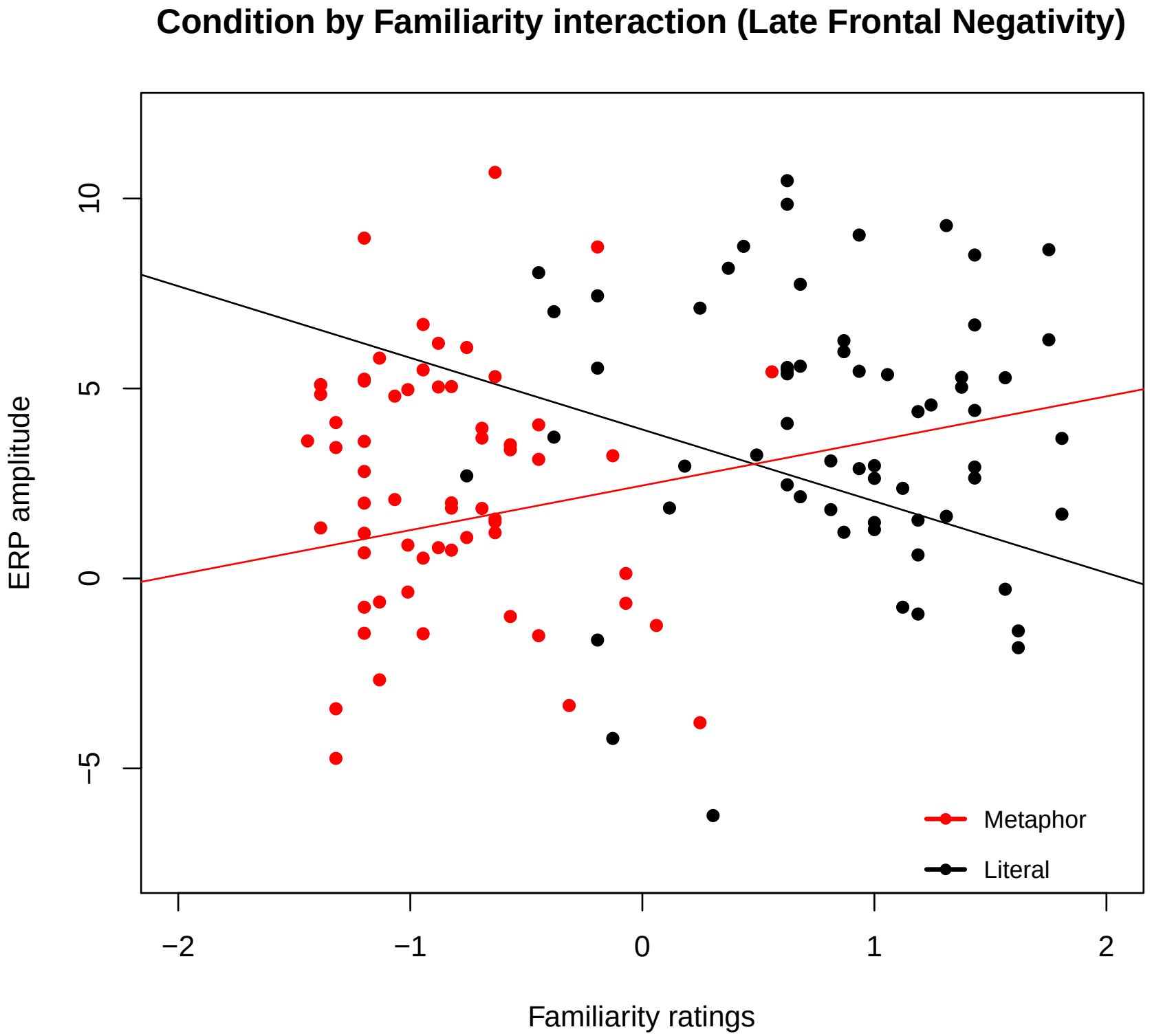
Table 2. Psycholinguistic properties of literary metaphors and literal counterparts, presented in their contexts: log-frequency values (mean and standard deviation) of both terms in the "A of B" expressions; rating values of familiarity, concreteness, difficulty, meaningfulness (mean and standard deviation on a 5-point scale), and cloze probability (mean and standard deviation) for the complete "A of B" expressions.

	Literary metaphors (from Bambini et al. 2014)	Literal expressions	Statistics
Log-frequency A terms	1.78 (0.52)	1.82 (0.52)	t(61)=-1.04, p=0.30
Log-frequency B terms	1.95 (0.59)		-
Familiarity	1.91 (0.46)	3.70 (0.66)	t(61)=-16.39, p<0.001
Concreteness	1.79 (0.38)	3.52 (0.90)	t(61)=-15.88, p<0.001
Difficulty	2.63 (0.66)	2.05 (0.57)	t(61)=+5.41, p<0.001
Meaningfulness	3.22 (0.68)	3.74 (0.62)	t(61)=-4.73, p<0.001
Cloze probability	0.07 (0.13)	0.26 (0.31)	t(61)=-4.56, p<0.001

Figure 1. Upper part: ERPs from nine representative electrodes time-locked to the presentation of the target noun for literary metaphors (red) and literal counterparts (black). Lower part: Topographical scalp distributions of the effects for the N400 (left) and the sustained negativity (right).

Figure 2. Effect of familiarity (based on the rating scores) on the ERP amplitude in the 600-900 ms time window for each condition. Individual dots represent the estimate voltage for each item. Familiarity ratings in the x axis are expressed in z scores around the mean.


F7
Fz
F8
C3
Cz
C4
P3
Pz
P4
Literal
Metaphor
−4μV
200 400 600 800
300-500 ms
600-900 ms
-2.5 μV
0 μV
2.5 μV
-2.5 μV
0 μV
2.5 μV


Condition by Familiarity interaction (Late Frontal Negativity)
ERP amplitude
10
5
0
−5
−2
−1
0
1
2
Familiarity ratings
Metaphor
Literal