



Context matters: Novel metaphors in supportive and non-supportive contexts

Franziska Hartung^{a,*}, Yoed N. Kenett^{b,1}, Eileen R. Cardillo^a, Stacey Humphries^a, Nathaniel Klooster^a, Anjan Chatterjee^a

^a Department of Neurology, University of Pennsylvania, Philadelphia, PA, 19104, USA

^b Department of Psychology, University of Pennsylvania, Philadelphia, PA, 19104, USA

ARTICLE INFO

Keywords:

Narrative
Metaphor
Novelty
Creative language
Figurativeness
fMRI

ABSTRACT

Creative language is defined as linguistic output that is both novel and appropriate. Metaphors are one such example of creative language in which one concept is used to express another by highlighting relevant semantic features. While novelty is an inherent property of unfamiliar metaphors, appropriateness depends on the context. The current study tests the hypothesis that the context in which metaphors are encountered affects their processing. We examined the neural effects of comprehending metaphors in context by comparing neural activations in response to novel metaphors and literal sentences that were either embedded in a meaningful narrative or in matched jabberwocky contexts. We found that the neural correlates of processing metaphoric sentences and their literal counterparts are indistinguishable when embedded in a narrative: both conditions activate bilateral areas along the anterior temporal poles, middle temporal gyri, superior temporal sulci, and the angular gyri. Metaphors embedded in a narrative as compared to their identical counterparts embedded in jabberwocky show increased responses in sensorimotor areas that correspond to the modality of the literal meaning of the target word, perhaps reflecting deeper semantic processing. Our results confirm that context affects neural mechanisms for understanding creative ideas.

1. Introduction

Creativity is characterized by both novelty and appropriateness (or effectiveness or usefulness; Runco and Jaeger, 2012). New ideas do not arise spontaneously in a vacuum. Instead, they build on existing knowledge and recombinations of knowledge to create new and meaningful associations (Beatty et al., 2016; Beatty et al., 2017; Benedek et al., 2014; Benedek and Fink, 2019; Kenett, 2018b; Kenett and Faust, 2019; Volle, 2018). Creating new variations of existing concepts and establishing new relations between distant concepts is one way to yield creative ideas (Kenett, 2018a; Kenett and Faust, 2019; Kenett et al., 2018; Mednick, 1962). In this study, we use novel metaphors as an example of creative language and investigate the neural effects of the context in which they are understood.

Because metaphors use semantic features flexibly to express new meaning, metaphors represent an important case of the creative use of language (Faust, 2012; Vartanian, 2012). Creative people have a more flexible memory structure (Kenett and Faust, 2019; Mednick, 1962), a

characteristic that might explain why they understand and express novel ideas with ease and why they comprehend novel metaphors better than their less creative peers (Gold et al., 2012; Kenett et al., 2018). Unfamiliar metaphors are, by definition, novel, but not all novel metaphors are appropriate (Mcquire et al., 2017). One way to characterize appropriateness is the context in which metaphors are apprehended. The role of context in metaphor processing is relatively unexplored. In this study, we investigate the contribution of meaningful context to the neural underpinnings of comprehension of novel metaphors.

Metaphors often use a familiar concept to highlight specific semantic features with partially overlapping meaning to express another concept. The process of comprehending metaphors was conceived initially as having two stages that involve an initial literal interpretation of a word and a reinterpretation after consideration that the literal meaning does not make sense (Ortony, 1993; Schmidt and Seger, 2009). Empirical observations suggest that metaphoric interpretations can be just as automatic as literal interpretations (Glucksberg, 2003). When embedded in a meaningful context, non-literal expressions are processed with equal

* Corresponding author.

E-mail address: fhartung@pennmedicine.upenn.edu (F. Hartung).

¹ Franziska Hartung and Yoed Kenett share first authorship.

speed (Glucksberg, 2003), indicating that there is no initial bias for literal meanings, and that the predictability of semantic features can outweigh the predictability of specific words. Conventional metaphors do not seem to be processed differently from literal meanings even when encountered in isolation (Faust, 2012). The Career of Metaphor model argues that with increased conventionality, metaphoric meanings of words are lexicalized as additional senses of a word (Bowdle and Gentner, 2005; Cardillo et al., 2012).

Meta-analyses of neuroimaging studies on metaphors and figurative language more generally compared to literal meanings find that while some brain regions show increased activation to figurative language compared to literal language, the reverse contrast does not reveal activations specific to literal language (Bohrn et al., 2012; Rapp, 2012; Rapp et al., 2012). This pattern suggests that literal and non-literal language rely on similar neural structures—and by implication similar cognitive processes—but that figurative language may require more processing than literal language (Bohrn et al., 2012). Similarly, generating new ideas—for instance in problem solving or thinking about new uses for objects—recruits similar brain structures as recalling familiar solutions (Benedek et al., 2014).

While context profoundly affects semantic processing, most metaphor and creative language studies use decontextualized phrases, or sentences with minimal context; conditions that are not typical of how we encounter metaphors (Abraham, 2013, 2018). Supporting context can make metaphors easier to understand (Prat et al., 2012). Context facilitates prediction, limits possible interpretation, and likely facilitates co-activating relevant information and semantically related concepts to ease processing of figurative language (Fig. 1). Prat et al. (2012) reported that figurativeness itself had no effect on comprehension when metaphors are embedded in a minimal context (one or two preceding sentences), regardless of difficulty (see also Diaz and Hogstrom, 2011).

Coherent discourse elicits more robust activations than isolated sentences in left hemisphere language areas as well as in bilateral anterior temporal cortex, superior temporal sulci, and medial prefrontal regions (Diaz and Hogstrom, 2011; Ferstl et al., 2008). Similarly, metaphors compared to literal control sentences robustly activate bilateral inferior frontal and temporal regions (Bohrn et al., 2012; Cardillo et al., 2012; Rapp, 2012; Rapp et al., 2012). The anterior temporal poles may play an important role in discourse processing and semantic memory (Diaz and Hogstrom, 2011; Patterson et al., 2007; Simmons and Martin, 2009), and may be ideally suited to creating a larger abstract semantic representation from individual words (Ferstl et al., 2008).

Creative thinking often involves the right hemisphere (Aberg et al., 2016; Faust, 2012), which historically has been hypothesized to play a major role in processing metaphors (Anaki et al., 1998; Brownell et al., 1990; Faust, 2012; Fiore and Schooler, 1998; Schmidt et al., 2007). However, the experimental evidence on lateralization of non-literal language comprehension is inconsistent (Anaki et al., 1998; Bohrn et al., 2012; Cardillo et al., 2018; Diaz and Hogstrom, 2011; Giora, 2007; Ianni et al., 2014; Rapp et al., 2007; Rapp et al., 2012; Schmidt and Seger, 2009; Winner and Gardner, 1977; Yang, 2014). Instead, studies consistently report left fronto-temporal activations and bilateral activations,

particularly when stimuli are novel (Bohrn et al., 2012; Diaz et al., 2011; Faust, 2012; Rapp et al., 2012; Yang, 2014). Diaz, Barrett and Hogstrom (2011) showed that contextual congruence had a greater influence than figurativeness on bilateral hemisphere recruitment. They reported bilateral activations in inferior frontal cortex and the anterior temporal lobes when figurative expressions were preceded by a congruent sentence. Previously reported effects of figurativeness studies may actually be incongruity effects of an unexpected phrase (prediction violation), a violation experienced when first encountering a novel metaphor.

Some suggest that the right hemisphere recruitment for contextualized complex language might serve as additional resources to support comprehension when demands on working memory increase (Just et al., 1996; Schmidt and Seger, 2009). In complex contexts, semantic processing requires integrating broad semantic fields and concepts that are related to each other (Yang, 2014). An alternative hypothesis proposed by Federmeier suggests that even if both hemispheres are sensitive to language and sentence-level processes, the right hemisphere integrates language components whereas the left hemisphere predicts language sequences (Federmeier, 2007). In either account, the right hemisphere contribution to sentence processing does not reflect figurativeness.

In summary, previous research has found that metaphors are processed in the brain differently than their literal counterparts. However, most studies have used decontextualized, isolated sentences as stimuli. In real-life language comprehension, which typically takes place over temporally developing (narrative) contexts rather than over disconnected sentences, the human brain is constantly generating predictions about what will come next. While metaphors may be unexpected when heard in isolation, they may be processed fluently when they are more predictable in the context of a supporting narrative. Thus, previously reported differences in the neural processes underlying literal and metaphorical language processing may be attributable to using isolated sentences as stimuli. We examined whether there are neural differences in the processing of literal and metaphorical language when embedded in a meaningful narrative context, allowing the brain to engage in normal predictive language processing.

2. Aim of the study

This study investigates the impact of meaningful narrative contexts on the neural processing of metaphoric and literal sentences. We test the hypothesis that an appropriate narrative context changes how metaphors are processed. To test this hypothesis, we measured blood-oxygen level dependent (BOLD) brain activity with functional MRI while participants listened to metaphors and literal sentences embedded in narrative and matched jabberwocky contexts. Jabberwocky streams are grammatically well-formed sentences in which content words are replaced by pseudo-words that are phonotactically and syllabically possible in the language but bear no semantic meaning. We compared brain activity to literal and novel metaphorical target sentences embedded in narrative and jabberwocky contexts to each other and to a jabberwocky baseline. We predicted that target sentences in meaningful contexts would produce greater activations than jabberwocky embedded sentences in brain areas

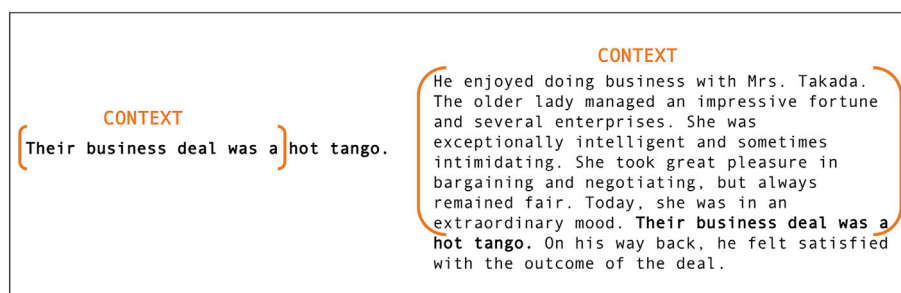


Fig. 1. Contextual embedding significantly reduces ambiguity and prediction violations for non-literal expressions.

(e.g., anterior temporal cortex) that integrate semantics. In addition, we predict that neural processing of metaphoric and literal sentences would not significantly differ when embedded in meaningful, supportive contexts.

3. Methods

3.1. Participants

Thirty right-handed individuals were recruited from the University of Pennsylvania participant database (17 women, 12 men, 1 nonbinary person). Data from three participants were partially excluded because of excessive movement (>3 mm within one run, 25% of total data from 1 participant), headphone malfunction (25% of total data from 2 participants), or scanner synchronization issues (25% of total data from 1 participant). In sum, the exclusion amounted to one participant's entire dataset, 50% of data from another participant, and 25% of data excluded from two other participants. Data from one participant was excluded because of a technical problem with the headphones.

From the participants whose data was included, age of participants ranged from 19 to 36 (mean age = 24.73 years, SD = 4.26, mean education = 16.41 years, SD = 1.35). Participants had normal or corrected to normal vision and no prior history of psychiatric or neurological disease. Before participation in the study, each individual gave informed consent (IRB approval number 806447 of the University of Pennsylvania). Participants were compensated monetarily for their time.

3.2. Stimuli

Stimuli were short narratives constructed around target sentences. Target sentences were sentence-pairs selected from Cardillo et al. (2010). Each metaphor-literal sentence pair had the same syntactic form (The A is a B) and used the same nominalized motion verb for the second term (B). Using the published norms of Cardillo et al. (2010), we ensured metaphors were novel by limiting our selection to metaphors with a maximum value for familiarity (<3.5). This selection criteria provided an initial sample of 23 sentence-pairs. For each of these initial 46 target sentences, we created a unique, short narrative (70–100 words) in which the target sentences provided a meaningful closure of the narrative (Table 1, Supplementary Materials S1). After norming (see below), we excluded three narrative pairs, the final stimulus set consisted of 20 narrative pairs (see below for norming details).

Table 1

Target sentence pairs.

Metaphor	Literal sentence
Her stare was a bull charge.	The battle plan was a charge.
The review was a karate chop.	His gesture was a quick chop.
His work experience was a clumsy clamber.	The final ascent was an exhausting clamber.
His smile was a charming dodge.	His move was a quick dodge.
The art major was a glide.	The skater's entrance was a glide.
The test review was a quick jog.	The race course was an easy jog.
The home purchase was a bungee jump.	The fence was a high jump.
The winter was a heartbroken limp.	The gait was a mild limp.
The friendship was a crazy polka.	The dance was a fast polka.
The purchase was a tiger pounce.	The cat's attack was a pounce.
The road was an irresistible pull.	The magnet was a weak pull.
Her inquiries were a nervous scamper.	Her exit was a nervous scamper.
The home purchase was a skydive.	The prize was a free skydive.
The poetry was a teenage slouch.	His posture was a lazy slouch.
The numbers were a brain swarm.	The bees were a black swarm.
The eviction was a mean sweep.	The chore was a quick sweep.
The reception was an icy swim.	The competitive relay was a swim.
The business deal was a hot tango.	The last romp was a Latin tango.
The ceremony was a swamp trudge.	The way back was a trudge.
The shop display was a gentle tug.	The puppy's grasp was a firm tug.
The letter was a goodbye wave.	The tsunami was a giant wave.

While all target sentences were structured around a motion event noun, the correct interpretation of the noun differed in the literal and metaphoric conditions. In literal sentences, the correct interpretation evoked action semantics. The correct interpretation of the same motion event noun in the matched metaphorical sentence evoked emotion and social semantics. This pattern of associations is typical for metaphors. Metaphors can express abstract and emotional ideas by accessing concrete semantic domains like action semantics (Edwards, 1999; Fainsilber and Ortony, 1987; Jamrozik et al., 2016; Lakoff and Johnson, 2008; Lubart and Getz, 1997), and metaphors can be experienced as more emotional than their literal counterparts (Citron and Goldberg, 2014; Glucksberg, 2003).

3.3. Norming

To norm the narratives, 50 raters were recruited using Amazon Mechanical Turk (AMT; Buhrmester et al., 2011). Rater selection was restricted to IP addresses located in the US. All raters (27 women, 23 men) indicated they were native speakers of English. Age of AMT raters ranged from 23 to 70 years (mean age = 37.85 years, SD = 11.23). Education of AMT raters ranged from 12 to 20 years (mean education = 15.35 years, SD = 1.99). In the norming task, AMT raters read all 46 narratives in random order and rated each of them on 5-point Likert scales for three qualities: 1) ease of comprehension, 2) interestingness of the story, and 3) naturalness of the writing. The norming experiment took approximately 40 min and AMT raters were compensated monetarily. Norming revealed no striking differences in comprehensibility, interestingness, and naturalness leading to no exclusions. We excluded two story pairs based on the norming of the original sentences with a minimum criterion for interpretability of the metaphor (minimum 75%) which led to exclusion of 2 pairs (*balloon ride*, *trickle*). In addition, because of concerns that one literal item could also be interpreted figuratively (*trudge*, figurativeness rating > 3.01), the two narratives based on this sentence-pair was excluded from the stimulus sample. These exclusions left 20 sentence pairs in the final stimulus set (For details on the norming results see Supplementary material S2, see Supplementary Material S3 for a complete story overview and auditory recordings). No data from the norming task were excluded from analysis.

3.4. Stimulus presentation

The experiment was built in OpenSesame (Mathôt et al., 2012) using PsychoPy as a backend. Stimuli recordings were presented through MR-compatible headphones combined with hearing protection. Volume was adjusted to the optimal level based on feedback from participants. All visual stimuli (questionnaires, instructions, etc.) were projected onto a screen using a projector outside the MR scanner room, which could be seen by participants through a mirror mounted above the head coil. Responses to the comprehension questions were recorded with a 4-button response device.

3.5. Procedure

The experiment consisted of one session. After signing the consent form, each participant was placed in the MRI scanner. Padding was used to minimize movement. The experiment began with a volume test in which participants heard a story which was not used in the experiment; the scanner was switched on to adjust the audio presentation to individual optimal volume. Participants listened to the 40 stories presented randomly, of which 20 were meaningful narratives and 20 were jabberwocky stories. In the jabberwocky condition, the syntactic structure and syllable number were identical to their narrative counterparts, but only the target sentence was sensible. In both the jabberwocky and the story condition, half the target sentences were metaphorical and half were literal. Stimulus presentation was randomized across four scanner runs. Participants listened for comprehension. In 25% of trials, a

comprehension probe question followed the audio presentation to ensure participants were paying attention to the task. In comprehension probe task, the target sentence from the narrative they last listened to was presented visually along with four answer choices (metaphoric interpretation of the action event, literal interpretation of the action event, semantically related but incorrect interpretation, random interpretation). Participants were asked to select the correct response to the comprehension probes with a 4-button response device with their right hand. The 4 buttons corresponded to the randomly assigned numbers of the answer choices. There was no time limit to answering the comprehension question. After the experiment, a high-resolution anatomical scan (~7min) was taken. After the scanning session, participants were debriefed.

3.6. fMRI data acquisition and preprocessing

Images of blood-oxygen level dependent (BOLD) changes were acquired with a 3T Siemens Magnetom Prisma scanner (Erlangen, Germany) with a 64-channel head coil. We used cushions to minimize participants' head movement. We used two localizing scans and auto-alignment (AAHead_Scout). Functional images were acquired using a fast multiband sequence, with a multiband factor of 6, and a high temporal resolution (TR: 500 ms, TE: 25 ms, flip angle: 30°, voxel size: 3.0 × 3.0 × 3.0 mm, 48 slices) minimally modified from the Human Connectome Project protocols (Van Essen et al., 2012). High resolution (0.8 × 0.8 × 0.8 mm) structural (anatomical) images were acquired using an SPC T1 GRAPPA sequence (Milchenko and Marcus, 2013). Data was preprocessed using the Matlab toolbox SPM12 (<http://www.fil.ion.ucl.ac.uk/spm>). Images were converted to NIFTI files, motion corrected and registered to the mean image of each scanner run. The motion-corrected images were co-registered with the individual participants' anatomical scan. After segmentation of the anatomical T1 image, we created a mask from the individual participant's normalized tissue probability maps and used the forwards deformations to spatially normalize the functional scans to the standard MNI template. Finally, all images were spatially smoothed using a voxel-wise isotropic 6 mm full width at half maximum (FWHM) Gaussian kernel.

3.7. fMRI data analysis

In order to create regressors for the analysis, the auditory recordings of the stories were scored for the time windows in which the target sentences were presented (Hartung et al., 2017). Additionally, we scored the time windows for the story and jabberwocky contexts preceding the target sentences for a control analysis. At the single subject level, statistical analysis was performed using a general linear model, in which beta weights for the regressor of interest were estimated for the time course of each voxel, using multiple regression analysis (Friston et al., 1994). Regressors were modelled as their true durations, and convolved with a canonical hemodynamic response function (Friston et al., 1996). There were six convolved regressors in total that had no overlap in their onset or duration (4 regressors for sentence type by context, and two regressors for the target sentence preceding context types). The motion estimates of the motion correction algorithm were modelled as regressors of no interest to account for head motion.

For the group level analyses, we used non-parametric statistical inference testing with SnPM (Holmes et al., 1996; Nichols and Holmes, 2002; Winkler et al., 2014). No variance smoothing was applied. Results were corrected for multiple comparisons familywise error correction with a combined cluster forming voxel threshold of $p < .001$ and cluster threshold of $p < .05$ with Monte Carlo permutation testing in SnPM (7000 permutations per comparison, cluster defining $T_{df=1,28} = 3.42$).

We compared activations of literal and metaphoric target sentences within and between the narrative and jabberwocky conditions. We used the jabberwocky context preceding the target sentences as a baseline to compare target sentences in each condition. The jabberwocky baseline

maintains many aspects of language (syntax, number of syllables, sentence intonation, and narrative structure) but lacks semantics. Since metaphors draw on semantics, this baseline was ideal for the research question. We performed an additional control analysis comparing activations to the narrative and jabberwocky contexts preceding the target sentences.

4. Results

4.1. Control analysis

As a control analysis, we compared jabberwocky and narrative contexts preceding the target sentences. Narratives compared to jabberwocky contexts revealed activated regions in the anterior and middle temporal gyrus bilaterally, and a region in the left posterior middle temporal gyrus (Table 2; Fig. 2). Jabberwocky compared to narratives revealed increased activation in a bilateral fronto-parietal network including middle frontal gyri and middle and superior orbital gyri, as well as right inferior frontal gyrus, left superior medial gyrus and medial cingulate cortex, left inferior orbital gyrus extending into left anterior insula, left superior frontal gyrus, and left cuneus (Table 3; Fig. 2).

4.2. Metaphor vs. literal sentences

Comparing metaphorical and literal target sentences in jabberwocky and narrative contexts did not result in BOLD activations surviving statistical thresholding in either condition.

Metaphorical sentences in narrative contexts compared to the jabberwocky baseline revealed increased activation in regions bilaterally along the middle temporal gyrus and superior temporal sulcus, as well as the angular gyrus (Table 4; Fig. 3). Literal sentences in narrative contexts compared to the jabberwocky baseline similarly revealed increased activation in regions bilaterally along the middle temporal gyrus and superior temporal sulcus, as well as the angular gyrus. In addition, we found increased activation in the right superior frontal gyrus (Table 5; Fig. 3).

Metaphoric target sentences embedded in jabberwocky contexts compared to the jabberwocky baseline did not reveal statistically significant BOLD activations differences. Literal sentences in Jabberwocky contexts compared to the jabberwocky baseline revealed increased activation in a region in the left precuneus (Table 6; Fig. 3).

4.3. Effect of context on processing literal and metaphoric target sentences

Metaphoric sentences in narrative compared to metaphoric sentences in jabberwocky revealed increased activation in right premotor cortex as (min $k = 103$; see Table 7 & Fig. 4). Such an effect for literal sentences was not present.

5. Discussion

Metaphors, particularly novel ones, represent an important linguistic tool for creativity; metaphors offer new and useful ways of thinking (Faust, 2012). We tested the hypothesis that the appropriateness of metaphors – i.e., the context in which they are encountered – changes how they are processed.

We found that context matters in metaphor comprehension. Reading metaphors in a meaningful narrative, compared to when encountered

Table 2
Narrative > jabberwocky.

Location	k	T max	X	Y	Z
Left anterior middle temporal gyrus	48	4.91	-54	6	-26
Right anterior middle temporal gyrus	63	4.64	60	4	-18
Left posterior middle temporal gyrus	8	3.98	-46	-46	6

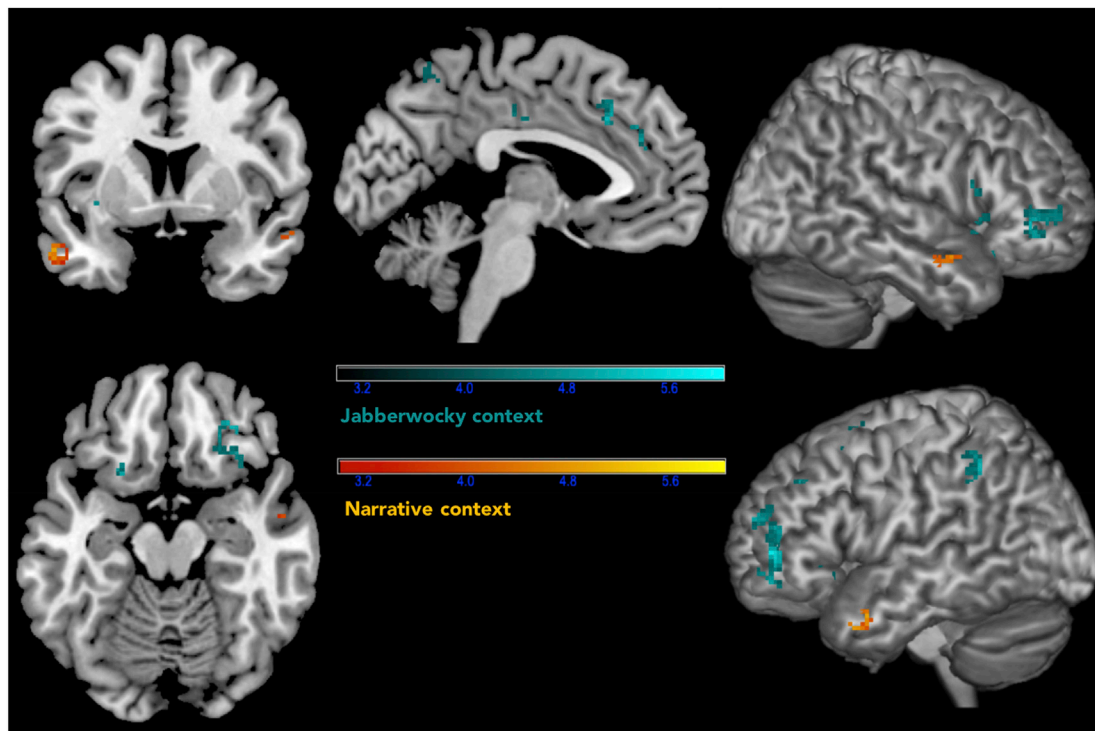


Fig. 2. Results comparing the context types preceding the target sentences. Averaging over narrative contexts as compared to jabberwocky contexts showed bilateral activations in anterior middle temporal areas, whereas jabberwocky contexts as compared to narrative contexts activated areas associated with the fronto-parietal attention network.

Table 3
Jabberwocky > narrative.

Location	k	T max	X	Y	Z
Left middle frontal gyrus, left middle orbital gyrus	87	5.61	-42	52	5
Left superior medial gyrus, Left middle cingulate cortex	141	5.15	-8	32	34
L middle cingulate cortex	9	5.00	-6	-18	36
R middle/superior Orbital gyrus, R inferior frontal gyrus	74	4.83	30	38	-16
R inferior frontal gyrus, R middle frontal gyrus	77	4.82	42	42	-2
Left cuneus	78	4.80	-46	-48	46
Left inferior orbital gyrus/left anterior insula	8	4.56	-22	16	-16
Left superior frontal gyrus	28	4.55	-24	-60	20

Table 4
Metaphorical sentences in narrative contexts > jabberwocky baseline.

Location	k	T max	X	Y	Z
Left angular gyrus	148	5.71	-52	-66	30
Left anterior middle temporal gyrus	243	5.54	-52	10	-26
Right anterior middle temporal gyrus & superior temporal sulcus	149	5.17	52	14	-26
Right posterior middle temporal gyrus	278	4.55	58	-46	14

embedded in jabberwocky, robustly activated a region in the right superior frontal (premotor) cortex. Metaphors and matched literal sentences both activate anterior and posterior temporal regions bilaterally when embedded in a narrative. Listening to a meaningful narrative, compared to jabberwocky, robustly activated (anterior) middle temporal regions bilaterally whereas jabberwocky activated a fronto-parietal network.

In narrative contexts, both literal and metaphoric target sentences showed overlapping activations in the anterior temporal poles, anterior and posterior middle temporal gyri and superior temporal sulci bilaterally. This observation is consistent with previous findings (Diaz and Hogstrom, 2011; Glucksberg, 2003; Prat et al., 2012). Why might metaphors and literal sentences share neural instantiations when presented within a narrative? One possibility is that the narrative provides a context in which the reader expects an interpretation that is congruent with some semantic features of the word used metaphorically and facilitates prediction of such features. These semantic features are likely already co-activated (predicted) by contextual constraints, which facilitate mapping and minimize any violation of expectation that might otherwise arise. It is possible that findings from previous research with isolated sentences or phrases found differences in brain activation to literal and metaphor expression as a result of asymmetry in the probability of the correct interpretation of the target word, and these previously reported findings might hence reflect prediction violation rather than figurative processing.

We found bilateral semantic processing of metaphors across all our conditions. These observations add to previous reports of bilateral processing of metaphors when encountered in isolation using fMRI (Cardillo et al., 2012) and in patients with focal brain damage (Cardillo et al., 2018; Ianni et al., 2014).

Notably, we did not find differences between literal and metaphoric sentences in the jabberwocky condition. This might be a result of the jabberwocky context: listening to jabberwocky already activated many areas previous research linked to metaphor processing, which are in the same areas as the activations we find for the jabberwocky contexts. One area frequently associated with metaphor and complex semantic processing that we did not observe is the inferior prefrontal cortex in either hemisphere. Because the jabberwocky condition showed increased activation in the inferior frontal gyri, perhaps the statistical comparison was less likely to exceed threshold activations with background activations produced by the jabberwocky condition. Our design might thus be

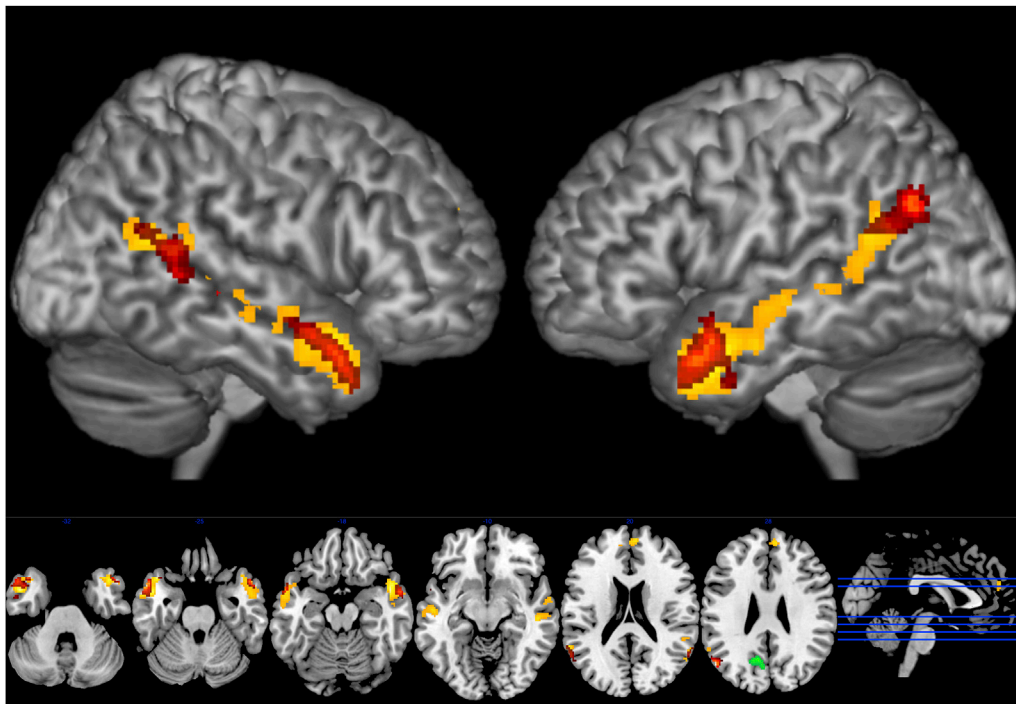


Fig. 3. Activations to literal (yellow) and metaphoric (red) target sentences in narrative contexts and literal target sentences in jabberwocky contexts (green).

Table 5

Literal sentences in narrative contexts > jabberwocky baseline.

Location	k	T max	X	Y	Z
Right superior frontal gyrus	118	4.94	6	54	24
Left anterior middle temporal gyrus	1043	4.76	-46	6	-30
Right anterior middle temporal gyrus & superior temporal sulcus	552	4.70	48	6	-18
Right angular gyrus	114	4.17	60	-56	16
Right angular gyrus	83	4.16	46	-38	12
Right posterior superior temporal sulcus	91	4.09	50	-22	-10

Table 6

Literal sentences in jabberwocky contexts > jabberwocky baseline.

Location	k	T max	X	Y	Z
Left precuneus	134	4.85	-10	-64	28

Table 7

Metaphors in narrative context > metaphors in jabberwocky context.

Location	k	T max	X	Y	Z
Right superior frontal gyrus/premotor cortex, BA6	119	4.44	20	-2	56

unintentionally biased against finding differences in those areas for the target sentence. It is also possible that the sentences embedded in jabberwocky context have a slightly different time progression in cognitive processing since the task is different from comprehending continuous narrative or single sentences in isolation. We cannot exclude the possibility that despite the natural jitter in the audio recordings there might be some bleeding of the context effects into the target sentence presentation from a data analysis perspective. The lack of differences in activation between literal and metaphor sentences in the absence of supporting context may also be the result of statistical thresholding or

because the target sentences are so closely matched on a wide range of linguistic properties. In the narrative condition, the absence of increased inferior frontal gyrus activation is likely because narrative contexts minimize ambiguity, facilitating comprehension, whereas the jabberwocky condition likely involves a search for meaning in an auditory stream that resembles linguistic properties. Narrative contexts facilitate semantic integration and potentially decrease working memory demands by predicting and pre-activating relevant semantic information and aiding the resolution of a novel metaphor's semantic ambiguity.

Beyond these claims about our major hypotheses, we make other observations about the neural processing of language. Our target sentences in both the metaphor and literal condition compared to baseline also showed increased activation in the angular gyri. This region likely serves as a supramodal semantic processing hub (Binder and Desai, 2011; Binder et al., 2009).

Metaphor sentences in narrative contexts showed increased activations in premotor cortex compared to the identical sentences in jabberwocky contexts. Novel metaphors are thought to initially draw strongly on the sensori-motor properties of the literal word meaning (Jamrozik et al., 2016). All metaphorical expressions used in this study were constructed with a noun referring to an action event (e.g. 'Their business deal was a hot tango'). Even when used metaphorically – e.g. to express emotion, metaphors can activate the sensorimotor properties of the original action word meaning. Saygin et al. (2009) suggests that motion semantics are accessed when we read about motion in a fictive manner (Boulenger et al., 2008; Desai et al., 2013). In line with this hypothesis, Samur et al. (2015) found that metaphors in highly emotional contexts activate visual motion processing areas and primary motor cortex, and concluded that 'sensory simulation gets a boost from emotionality in contexts where a figurative interpretation is appropriate, but not in contexts where a literal interpretation is appropriate (Pomp et al., 2018). This observation is congruent with the increased activation in premotor cortex for metaphors in narrative context (which were all emotional) as compared to jabberwocky contexts -that were void of semantics and hence emotional meaning- and suggests deeper semantic processing in narrative contexts. Moreover, the meaningful preceding context supports

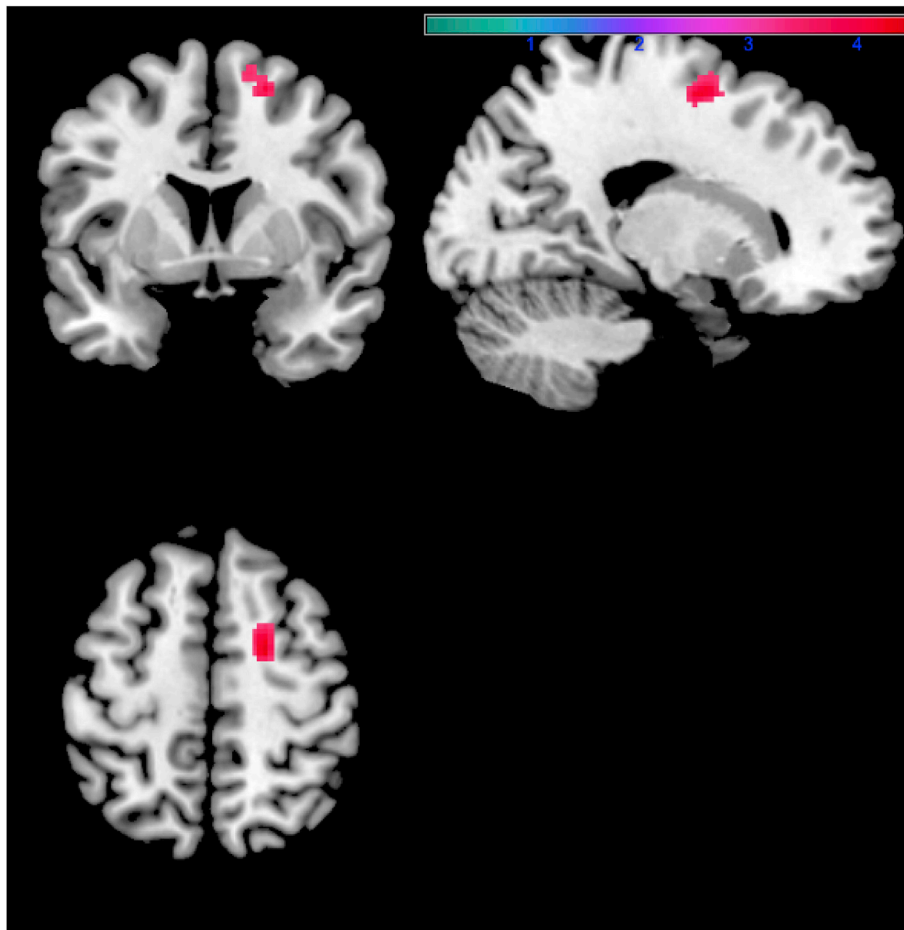


Fig. 4. Increased activations to metaphoric target sentences in narrative as compared to jabberwocky contexts.

prediction of appropriate semantic features, and hence pre- or co-activation of brain areas sensitive to specific types of information.

Literal sentences in jabberwocky contexts activated a region in the precuneus. Bohrn et al.'s (2012) meta-analysis reported precuneus activation for isolated literal sentences, but it is unclear what this increased activation represents. We speculate that this activation could either reflect mental imagery for literal sentences since the literal sentences describe more concrete concepts and have higher-sentence level imageability ratings according to a prior norming study (Cardillo et al., 2018; see also Mashal et al., 2014). Alternatively, it might be an effect of memorizing the sentence for the comprehension probe.

More generally, our observations emphasize the role of medial and superior temporal regions in this comprehension, consistent with a recent theoretical model that highlights the role of temporal regions in creativity (Shen et al., 2017). According to this proposed model, the role of the temporal lobes in the creative process include semantic processing, inhibiting prototypical responses (anterior and posterior MTG), and integrating semantic information (anterior and posterior STG) – all processes involved in comprehending novel metaphors (Faust, 2012). This framework aligns with behavioral and computational findings relating novel metaphor processing to individual differences in creativity (Gold et al., 2012; Kenett et al., 2018). Furthermore, a large meta-analysis of imaging research of creativity conducted activation likelihood estimation analysis to identify key brain regions that relate to various types of creativity related tasks (Gonen-Yaacovi et al., 2013). Among other key regions, this meta-analysis highlighted the role of the left posterior MTG in relation to meaning integration during creativity tasks. Importantly, the left posterior MTG region the authors identified roughly corresponds to our current finding. We cannot directly relate—via reverse inference

(Hong et al., 2019; Poldrack, 2006)—neural mechanisms that have been implicated in creativity and our findings. However, both types of tasks require similar cognitive functions—semantic inhibition and integration—and relate to activation in overlapping brain regions.

Several limitations exist in our study design. First, we did not control for potential inhomogeneity of the magnetic field during our data acquisition (Van Essen et al., 2012). Such magnetic inhomogeneity may introduce noise and weaken statistical power of our analysis. Nevertheless, despite such potential added noise, we found robust differences in our conditions. Further research is needed to replicate and extend our findings while better controlling for such confounds rising from fMRI data acquisition. Second, in the current study we examine the effect of context on performing a linguistic task that involves a creative aspect to it, namely novel metaphors. However, we do not examine how individual differences in creative ability relate to the context effects that were found. Since previous studies have highlighted the link between individual differences and novel metaphor comprehension (Gold et al., 2012; Kenett et al., 2018), future studies are needed to relate individual differences in creative ability to creative language comprehension. Third, our target sentences were selected for homogeneity. Future studies on creative language comprehension should additionally test parametric modulations of activations with semantic aspects such as familiarity, imageability, figurativeness, or other variables that affect semantic processing to disentangle what possible underlying processes are reflected in increased activation in different regions, e.g. whether the precuneus activation for literal sentences in jabberwocky contexts reflects mental imagery or memory processes or both. Fourth, while the jabberwocky contexts are ideal as a comparison condition to test for effects of semantically meaningful context, it is unclear whether the processing of

the embedded target sentences is qualitatively different or follows a similar time course. Future studies should consider different types of contexts, including different language contexts (e.g. expository text, semantically unhelpful context), as a comparison.

In conclusion, novel metaphors represent a prime example of the creative use of language by virtue of being novel and—ideally—useful or appropriate. However, they are only useful if understood. Our findings demonstrate that context affects neural mechanisms that promote the usefulness of novel language.

Data availability statement

All data will be made available without restriction at <https://osf.io/quxk4/>.

Author contribution

FH developed the study concept. All authors contributed to the study design. Testing and data collection were performed by FH. FH performed the data analysis and interpretation under the supervision of AC. FH and YK drafted the manuscript, and AC provided critical revisions. All authors approved the final version of the manuscript for submission.

Acknowledgements

This study was supported by NIH grant 5R01DC012511 awarded to AC. Mark Elliot developed the BOLD scanning sequence. We thank Geoff Aguirre for helpful guidance in designing the study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.neuroimage.2020.116645>.

References

- Aberg, K.C., Doell, K.C., Schwartz, S., 2016. The “creative right brain” revisited: individual creativity and associative priming in the right hemisphere relate to hemispheric asymmetries in reward brain function. *Cerebr. Cortex*.
- Abraham, A., 2013. The promises and perils of the neuroscience of creativity. *Front. Hum. Neurosci.* 7, 246.
- Abraham, A., 2018. *The Neuroscience of Creativity*. Cambridge University Press, Cambridge, UK.
- Anaki, D., Faust, M., Kravetz, S., 1998. Cerebral hemispheric asymmetries in processing lexical metaphors. *Neuropsychologia* 36, 353–362.
- Beaty, R.E., Benedek, M., Silvia, P.J., Schacter, D.L., 2016. Creative cognition and brain network dynamics. *Trends Cognit. Sci.* 20, 87–95.
- Beaty, R.E., Silvia, P.J., Benedek, M., 2017. Brain networks underlying novel metaphor production. *Brain Cognit.* 111, 163–170.
- Benedek, M., Beaty, R., Jaak, E., Koschutnig, K., Fink, A., Silvia, P.J., Dunst, B., Neubauer, A.C., 2014. Creating metaphors: the neural basis of figurative language production. *Neuroimage* 90, 99–106.
- Benedek, M., Fink, A., 2019. Toward a neurocognitive framework of creative cognition: the role of memory, attention, and cognitive control. *Curr. Opin. Behav. Sci.* 27, 1116–1122.
- Binder, J.R., Desai, R.H., 2011. The neurobiology of semantic memory. *Trends Cognit. Sci.* 15, 527–536.
- Binder, J.R., Desai, R.H., Graves, W.W., Conant, L.L., 2009. Where Is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebr. Cortex* 19, 2767–2796.
- Bohri, I.C., Altmann, U., Jacobs, A.M., 2012. Looking at the brains behind figurative language—a quantitative meta-analysis of neuroimaging studies on metaphor, idiom, and irony processing. *Neuropsychologia* 50, 2669–2683.
- Boulenger, V., Hauk, O., Pulvermüller, F., 2008. Grasping ideas with the motor system: semantic somatotopy in idiom comprehension. *Cerebr. Cortex* 19, 1905–1914.
- Bowlby, B.F., Gentner, D., 2005. The career of metaphor. *Psychol. Rev.* 112, 193–216.
- Brownell, H.H., Simpson, T.L., Bihle, A.M., Potter, H.H., Gardner, H., 1990. Appreciation of metaphoric alternative word meanings by left and right brain-damaged patients. *Neuropsychologia* 28, 375–383.
- Buhrmester, M., Kwang, T., Gosling, S.D., 2011. Amazon’s mechanical turk a new source of inexpensive, yet high-quality, data? *Perspect. Psychol. Sci.* 6, 3–5.
- Cardillo, E.R., McQuire, M., Chatterjee, A., 2018. Selective metaphor impairments after left, not right, hemisphere injury. *Front. Psychol.* 9.
- 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. *Behav. Res. Methods* 42, 651–664.
- Cardillo, E.R., Watson, C.E., Schmidt, G.L., Kranjec, A., Chatterjee, A., 2012. From novel to familiar: tuning the brain for metaphors. *Neuroimage* 59, 3212–3221.
- Citron, F.M.M., Goldberg, A.E., 2014. Metaphorical sentences are more emotionally engaging than their literal counterparts. *J. Cognit. Neurosci.* 26, 2585–2595.
- Desai, R.H., Conant, L.L., Binder, J.R., Park, H., Seidenberg, M.S., 2013. A piece of the action: modulation of sensory-motor regions by action idioms and metaphors. *Neuroimage* 83, 862–869.
- Diaz, M.T., Barrett, K.T., Hogstrom, L.J., 2011. The influence of sentence novelty and figurativeness on brain activity. *Neuropsychologia* 49, 320–330.
- Diaz, M.T., Hogstrom, L.J., 2011. The influence of context on hemispheric recruitment during metaphor processing. *J. Cognit. Neurosci.* 23, 3586–3597.
- Edwards, D., 1999. Emotion discourse. *Cult. Psychol.* 5, 271–291.
- Fainsilber, L., Ortony, A., 1987. Metaphorical uses of language in the expression of emotions. *Metaphor Symbolic Act.* 2, 239–250.
- Faust, M., 2012. Thinking outside the left box: the role of the right hemisphere in novel metaphor comprehension. In: Faust, M. (Ed.), *Advances in the Neural Substrates of Language: toward a Synthesis of Basic Science and Clinical Research*. Wiley Blackwell, Malden, MA, pp. 425–448.
- Federmeier, K.D., 2007. Thinking ahead: the role and roots of prediction in language comprehension. *Psychophysiology* 44, 491–505.
- Ferstl, E.C., Neumann, J., Bogler, C., Von Cramon, D.Y., 2008. The extended language network: a meta-analysis of neuroimaging studies on text comprehension. *Hum. Brain Mapp.* 29, 581–593.
- Fiore, S.M., Schooler, J.W., 1998. Right hemisphere contributions to creative problem solving: converging evidence for divergent thinking. In: Beeman, M., Chiarello, C. (Eds.), *Right Hemisphere Language Comprehension: Perspectives from Cognitive Neuroscience*. Erlbaum, Mahwah, NJ, pp. 349–372.
- Friston, K.J., Holmes, A.P., Worsley, K.J., Poline, J.P., Frith, C.D., Frackowiak, R.S.J., 1994. Statistical parametric maps in functional imaging: a general linear approach. *Hum. Brain Mapp.* 2, 189–210.
- Friston, K.J., Williams, S., Howard, R., Frackowiak, R.S.J., Turner, R., 1996. Movement-related effects in fMRI time-series. *Magn. Reson. Med.* 35, 346–355.
- Giora, R., 2007. Is metaphor special? *Brain Lang.* 100, 111–114.
- Glucksberg, S., 2003. The psycholinguistics of metaphor. *Trends Cognit. Sci.* 7, 92–96.
- Gold, R., Faust, M., Ben-Artzi, E., 2012. Metaphors and verbal creativity: the role of the right hemisphere. *Asymmetries Body Brain Cognit.* 17, 602–614.
- Gonen-Yaacovi, G., De Souza, L.C., Levy, R., Urbanski, M., Josse, G., Volle, E., 2013. Rostral and caudal prefrontal contribution to creativity: a meta-analysis of functional imaging data. *Front. Hum. Neurosci.* 7, 465.
- Hartung, F., Withers, P., Hagoort, P., Willems, R.M., 2017. When fiction is just as real as fact: No differences in reading behavior between stories believed to be based on true or fictional events. *Front. Psychol.* 8, 1618.
- Holmes, A.P., Blair, R.C., Watson, J.D.G., Ford, I., 1996. Nonparametric analysis of statistic images from functional mapping experiments. *J. Cerebr. Blood Flow Metabol.* 16, 7–22.
- Hong, Y.-W., Yoo, Y., Han, J., Wager, T.D., Woo, C.-W., 2019. False-positive neuroimaging: undisclosed flexibility in testing spatial hypotheses allows presenting anything as a replicated finding. *Neuroimage* 195, 384–395.
- Ianni, G.R., Cardillo, E.R., McQuire, M., Chatterjee, A., 2014. Flying under the radar: figurative language impairments in focal lesion patients. *Front. Hum. Neurosci.* 8, 871.
- Jamrozik, A., McQuire, M., Cardillo, E.R., Chatterjee, A., 2016. Metaphor: bridging embodiment to abstraction. *Psychon. Bull. Rev.* 23, 1080–1089.
- Just, M.A., Carpenter, P.A., Keller, T.A., Eddy, W.F., Thulborn, K.R., 1996. Brain activation modulated by sentence comprehension. *Science* 274, 114–116.
- Kenett, Y.N., 2018a. Going the extra creative mile: the role of semantic distance in creativity – theory, research, and measurement. In: Jung, R.E., Vartanian, O. (Eds.), *The Cambridge Handbook of the Neuroscience of Creativity*. Cambridge University Press, New York, NY, pp. 233–248.
- Kenett, Y.N., 2018b. Investigating creativity from a semantic network perspective. In: Kapoula, Z., Volle, E., Renoult, J., Andreatta, M. (Eds.), *Exploring Transdisciplinarity in Art and Sciences*. Springer International Publishing, Cham, pp. 49–75.
- Kenett, Y.N., Faust, M., 2019. A semantic network cartography of the creative mind. *Trends Cognit. Sci.* 23, 271–274.
- Kenett, Y.N., Gold, R., Faust, M., 2018. Metaphor comprehension in low and high creative individuals. *Front. Psychol.* 9.
- Lakoff, G., Johnson, M., 2008. *Metaphors We Live by*. University of Chicago press, Chicago, IL.
- Lubart, T.I., Getz, I., 1997. Emotion, metaphor, and the creative process. *Creativ. Res. J.* 10, 285–301.
- Mashal, N., Vishne, T., Laor, N., 2014. The role of the precuneus in metaphor comprehension: evidence from an fMRI study in people with schizophrenia and healthy participants. *Front. Hum. Neurosci.* 8.
- Mathôt, S., Schrei, D., Theeuwes, J., 2012. OpenSesame: an open-source, graphical experiment builder for the social sciences. *Behav. Res. Methods* 44, 314–324.
- McQuire, M., Mccollum, L., Chatterjee, A., 2017. Aptness and beauty in metaphor. *Lang. Cognit.* 9, 316–331.
- Mednick, S.A., 1962. The associative basis of the creative process. *Psychol. Rev.* 69, 220–232.
- Milchenko, M., Marcus, D., 2013. Obscuring surface anatomy in volumetric imaging data. *Neuroinformatics* 11, 65–75.
- Nichols, T.E., Holmes, A.P., 2002. Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum. Brain Mapp.* 15, 1–25.

- Ortony, A., 1993. *Metaphor and Thought*. Cambridge University Press.
- Patterson, K., Nestor, P.J., Rogers, T.T., 2007. Where do you know what you know? The representation of semantic knowledge in the human brain. *Nat. Rev. Neurosci.* 8, 976–987.
- Poldrack, R.A., 2006. Can cognitive processes be inferred from neuroimaging data? *Trends Cognit. Sci.* 10, 59–63.
- Pomp, J., Bestgen, A.-K., Schulze, P., Müller, C.J., Citron, F.M.M., Suchan, B., Kuchinke, L., 2018. Lexical olfaction recruits olfactory orbitofrontal cortex in metaphorical and literal contexts. *Brain Lang.* 179, 11–21.
- Prat, C.S., Mason, R.A., Just, M.A., 2012. An fMRI investigation of analogical mapping in metaphor comprehension: the influence of context and individual cognitive capacities on processing demands. *J. Exp. Psychol. Learn. Mem. Cognit.* 38, 282–294.
- Rapp, A.M., 2012. The brain behind nonliteral language : insights from brain imaging. In: Faust, M. (Ed.), *Advances in the Neural Substrates of Language: toward a Synthesis of Basic Science and Clinical Research*. Wiley Blackwell, Malden, MA, pp. 406–424.
- Rapp, A.M., Leube, D.T., Erb, M., Grodd, W., Kircher, T.T.J., 2007. Laterality in metaphor processing: lack of evidence from functional magnetic resonance imaging for the right hemisphere theory. *Brain Lang.* 100, 142–149.
- Rapp, A.M., Mutschler, D.E., Erb, M., 2012. Where in the brain is nonliteral language? A coordinate-based meta-analysis of functional magnetic resonance imaging studies. *Neuroimage* 63, 600–610.
- Runco, M.A., Jaeger, G.J., 2012. The standard definition of creativity. *Creativ. Res. J.* 24, 92–96.
- Samur, D., Lai, V.T., Hagoort, P., Willems, R.M., 2015. Emotional context modulates embodied metaphor comprehension. *Neuropsychologia* 78, 108–114.
- Saygin, A.P., McCullough, S., Alac, M., Emmorey, K., 2009. Modulation of BOLD response in motion-sensitive lateral temporal cortex by real and fictive motion sentences. *J. Cognit. Neurosci.* 22, 2480–2490.
- Schmidt, G.L., DeBuse, C.J., Seger, C.A., 2007. Right hemisphere metaphor processing? Characterizing the lateralization of semantic processes. *Brain Lang.* 100, 127–141.
- Schmidt, G.L., Seger, C.A., 2009. Neural correlates of metaphor processing: the roles of figurativeness, familiarity and difficulty. *Brain Cognit.* 71, 375–386.
- Shen, W., Yuan, Y., Liu, C., Luo, J., 2017. The roles of the temporal lobe in creative insight: an integrated review. *Think. Reas.* 23, 321–375.
- Simmons, W.K., Martin, A., 2009. The anterior temporal lobes and the functional architecture of semantic memory. *J. Int. Neuropsychol. Soc.* 15, 645–649.
- Van Essen, D.C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T.E.J., Bucholz, R., Chang, A., Chen, L., Corbetta, M., Curtiss, S.W., 2012. The Human Connectome Project: a data acquisition perspective. *Neuroimage* 62, 2222–2231.
- Vartanian, O., 2012. Dissociable neural systems for analogy and metaphor: implications for the neuroscience of creativity. *Br. J. Psychol.* 103, 302–316.
- Volle, E., 2018. Associative and controlled cognition in divergent thinking: theoretical, experimental, neuroimaging evidence, and new directions. In: Jung, R.E., Vartanian, O. (Eds.), *The Cambridge Handbook of the Neuroscience of Creativity*. Cambridge University Press, New York, NY, pp. 333–362.
- Winkler, A.M., Ridgway, G.R., Webster, M.A., Smith, S.M., Nichols, T.E., 2014. Permutation inference for the general linear model. *Neuroimage* 92, 381–397.
- Winner, E., Gardner, H., 1977. The comprehension of metaphor in brain-damaged patients. *Brain* 100, 717–729.
- Yang, J., 2014. The role of the right hemisphere in metaphor comprehension: a meta-analysis of functional magnetic resonance imaging studies. *Hum. Brain Mapp.* 35, 107–122.