NeuroImage 63 (2012) 1432–1442

Contents lists available at SciVerse ScienceDirect

NeuroImage

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

Neural correlates of combinatorial semantic processing of literal and figurative noun noun compound words

Bálint Forgács [a,b,*], Isabel Bohrn [c], Jürgen Baudewig [b,c], Markus J. Hofmann [d], Csaba Pléh [a], Arthur M. Jacobs [b,c,d]

[a] Department of Cognitive Science, Budapest University of Technology and Economics (BME), Egry József utca 1., T building, V. 506, 1111, Budapest, Hungary
[b] Dahlem Institute for Neuroimaging of Emotion (D.I.N.E.), Cluster Languages of Emotion, Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany
[c] Languages of Emotion Cluster of Excellence, Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany
[d] Department of Experimental and Neurocognitive Psychology, Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany

ARTICLE INFO

Article history:
Accepted 15 July 2012
Available online 23 July 2012

Keywords:
fMRI
Metaphor
Right hemisphere
LIFG
Coarse semantic coding
Semantic distance

ABSTRACT

The right hemisphere's role in language comprehension is supported by results from several neuropsychology and neuroimaging studies. Special interest surrounds right temporoparietal structures, which are thought to be involved in processing novel metaphorical expressions, primarily due to the coarse semantic coding of concepts. In this event related fMRI experiment we aimed at assessing the extent of semantic distance processing in the comprehension of figurative meaning to clarify the role of the right hemisphere. Four categories of German noun noun compound words were presented in a semantic decision task: a) conventional metaphors; b) novel metaphors; c) conventional literal, and; d) novel literal expressions, controlled for length, frequency, imageability, arousal, and emotional valence. Conventional literal and metaphorical compounds increased BOLD signal change in right temporoparietal regions, suggesting combinatorial semantic processing, in line with the coarse semantic coding theory, but at odds with the graded salience hypothesis. Both novel literal and novel metaphorical expressions increased activity in left inferior frontal areas, presumably as a result of phonetic, morphosyntactic, and semantic unification processes, challenging predictions regarding right hemispheric involvement in processing unusual meanings. Meanwhile, both conventional and novel metaphorical expressions induced BOLD signal change in left hemispherical regions, suggesting that even novel metaphor processing involves more than linking semantically distant concepts.

Introduction

Although brain regions traditionally held responsible for language processing, like Broca's and Wernicke's areas, are located in the left hemisphere (LH), a growing number of studies are reporting evidence for linguistic functions localized in the right hemisphere (RH). The hemisphere historically often treated as the "mute" one apparently takes part in a number of linguistic functions, especially in the processing of meaning below the surface, as in indirect requests (Weylman et al., 1989), lexical ambiguity resolution (Faust and Chiarello, 1998), understanding jokes (Coulson and Williams, 2005; Coulson and Wu, 2005), irony (Eviatar and Just, 2006), and metaphors (Ahrens et al., 2007; Anaki et al., 1998; Arzouan et al., 2007; Bottini et al., 1994; Mashal et al., 2005, 2007; Pobric et al., 2008; Sotillo et al., 2005; Stringaris et al., 2006). The message level meaning seems to be an important factor in the interpretation of such linguistic materials, what is in line with the RH's sensitivity to contextual effects (Grindrod and Baum, 2003; Van Lancker Sidtis, 2006; Vigneau et al., 2011). Several studies have found evidence for a RH involvement also in the processing of short, out of context, two word expressions, such as novel metaphors (Anaki et al., 1998; Mashal et al., 2005, 2007; Pobric et al., 2008), or during the semantic combination of two nouns into a highly meaningful phrase (Graves et al., 2010).

The aim of the present study was to assess the impact of two often confounded factors on RH language processing, familiarity and figurativeness, while controlling for context, imageability, emotional valence, and arousal, thought to be posing higher processing demands on the RH. Specifically, the goal was to see whether there are neural processes associated with novel metaphor comprehension independently of processing semantic distance, namely could the selection and suppressions of certain semantic features play a separate role.

* Corresponding author at: Department of Cognitive Science, Budapest University of Technology and Economics (BME), Egry József utca 1., T building, V. 506, 1111, Budapest, Hungary.
E-mail addresses: forgacsb@cogsci.bme.hu (B. Forgács), isabel.bohrn@fu-berlin.de (I. Bohrn), juergen.baudewig@fu-berlin.de (J. Baudewig), mhof@zedat.fu-berlin.de (M.J. Hofmann), pleh@cogsci.bme.hu (C. Pléh), ajacobs@zedat.fu-berlin.de (A.M. Jacobs).

1053-8119/$ – see front matter
doi:10.1016/j.neuroimage.2012.07.029

Neural processing of metaphors

While there had been extensive previous work on metaphors (e.g., Miller, 1979; Ortony, 1979; Richards, 1936; Searle, 1979; Tversky, 1977), the cognitive metaphor theory by Lakoff and Johnson (1980a, 1980b; also Lakoff, 1987) brought the issue real popularity in the field. Breaking away from the classical view of metaphors regarded as poetic or rhetorical tools, basically ornaments of language (Aristotle, 335 BC/1952), and primarily violations (Grice, 1975), they pointed out that metaphors are widely used in everyday language, and proposed that even the conceptual system is metaphorical in nature (Lakoff and Johnson, 1980b). Abstract concepts are understood via the systematic mapping of more concrete concepts onto them, which are based on the experiential gestalts of bodily perceptions in the case of primary metaphors, or on the recombination of the latter in the case of complex metaphors (Grady, 1997). For instance, the metaphorical expression "I can see your point" is an example of the conceptual metaphor SEEING IS UNDERSTANDING, where a concrete experience, *seeing* is the source domain mapped onto the abstract target domain, *understanding* (Lakoff and Johnson, 1980a). During these mappings certain features of the source domain are selected and others are filtered, hence there is no complete correspondence between the two conceptual domains (Kövecses, 2005). Cognitive metaphor theory, though, has been criticized (e.g., Jackendoff and Aaron, 1991; McGlone, 2007; Murphy, 1996, 1997), and there are alternative theories, like the class inclusion theory (Glucksberg and Keysar, 1990), the structure mapping theory (Gentner, 1983), or the conceptual blending theory (Fauconnier and Turner, 1998).

The classical linguistic approach proposed a sequential processing for metaphors, requiring a re-analysis of the literally false meaning (Grice, 1975), but the parallel view suggests that literal meaning has no advantage, as figurative language (an indirect request or an idiom) does not take more time to comprehend in a supportive context (Gibbs, 1994). At the same time some ERP studies suggest that there is a gradual component to metaphor processing, conventional metaphors requiring a slightly higher effort than literal expressions, while novel metaphors posing even more demand on comprehension (Arzouan et al., 2007; Lai et al., 2009), perhaps because of the selection and filtering of specific conceptual features.

There seems to be a systematic division of labor between the two cerebral hemispheres regarding words and concepts (Beeman, 1998; Chiarello, 1991), but more broadly the LH is thought to expect and actively predict likely upcoming material, while the RH is assumed to integrate and assemble meaning directly from the ongoing information (Federmeier, 2007; Federmeier and Kutas, 1999; Federmeier et al., 2005). The RH theory of metaphor processing suggests a division for literal and figurative language. It evolved from studies with RH damaged patients (Winner and Gardner, 1977) and was strengthened by a landmark PET study with healthy individuals (Bottini et al., 1994). However, there are several studies that could not confirm a special role of the RH, and reported bilateral processing (Coulson and Van Petten, 2007; Schmidt and Seger, 2009), while still others found mainly LH involvement (Chen et al., 2008; Eviatar and Just, 2006; Lee and Dapretto, 2006; Rapp et al., 2004, 2007; Stringaris et al., 2007). Nevertheless, as Schmidt and Seger (2009) pointed out, studies that have reported RH activations for figurative language have been involving novel metaphorical expressions and unusual semantic relations (Ahrens et al., 2007; Arzouan et al., 2007; Bottini et al., 1994; Mashal et al., 2005, 2007; Pobric et al., 2008; Sotillo et al., 2005; Stringaris et al., 2006).

With frequent use novel metaphors lose their novelty, and as eventually they become conventionalized, fixed, and familiar expressions, there is no need to create the conceptual mappings, as proposed by the career of metaphor hypothesis (Bowdle and Gentner, 2005). When compared directly, such "dead" metaphors were found to be processed similarly to literal expressions, mainly by LH areas (Mashal et al., 2005, 2007; Pobric et al., 2008). This could account for parts of the diverse results found in previous studies. However, the re-activation of the mapping can trigger RH processing again, for example when the literal meaning of idioms is evoked (Mashal et al., 2008); for this reason the term "sleeping" metaphor seems to be a useful refinement (Müller, 2008).

Semantic distance

Most of the time RH involvement is not attributed to metaphorical meaning per se, but to the bridging of unusual semantic relations in novel expressions. The graded salience hypothesis (Giora, 1997, 1999, 2002, 2003) suggests that the figurative–literal distinction is not a good predictor of processing. Highly salient meanings, both literal and figurative (e.g. conventional metaphors) are always activated directly and processed first, regardless of context. Even contexts favoring less salient meanings (e.g. literal interpretation of conventional metaphors) do not inhibit the activation of salient meanings (Giora, 1999). If the context supports an alternative interpretation that is similarly salient, parallel processes are activated, whereas novel metaphors require a serial processing where the intended figurative meaning is derived following the more salient literal meaning (Giora, 1997).

The salience of meaning is determined by a number of factors, such as being coded in the mental lexicon, prominence due to conventionality, frequency, familiarity, and prototypicality (Giora, 2002). In terms of hemispheric processing the graded salience hypothesis predicts (Giora, 2003), regardless of figurativeness, a selective LH processing during the comprehension of salient meanings (e.g. even conventional metaphors), and a selective RH activation for non-salient meanings (e.g., novel metaphors).

Another important framework focuses more on the neural attributes of the hemispheres. The coarse semantic coding theory (Beeman, 1998; Beeman et al., 1994; Jung-Beeman, 2005) proposes that the LH is coding narrow semantic fields in a fine grained manner, including word representations, synonyms, the word's semantic features, and first-order associates. The RH is coding broad semantic fields coarsely, including distant meanings too, allowing for the semantic integration of otherwise non-overlapping concepts. When Beeman et al. (1994) presented subjects the prime words "foot", "cry", and "glass", none of which is closely associated with the target word "cut", the RH benefitted more from the sum of the priming effects than the LH. In a second experiment they showed that the RH benefits equally from direct and summation primes, while the LH only from direct primes.

According to Beeman's model, the critical factor that determines which hemisphere is more sensitive to a given semantic relation is closeness of association or in other words, semantic distance. For example, when two words are strongly associated and are category co-exemplars ("arm"–"leg") priming is equivalent in the two hemispheres, but when they are nonassociated category members ("arm"–"nose"), priming is observed only in the RH (Chiarello et al., 1990). Even though this is rather due to semantic feature overlap than association per se, the higher the number and the more central the shared features of the concepts are, the more strongly they are associated. This suggests that even though category members also share some features, only strongly associated ones share enough to prime the LH (Beeman, 1998).

On the one hand, these theories provide an elegant account for the LH processing of most conventional metaphors, where narrow semantic field processing and high salience go hand in hand, and figurative meaning is accessed directly. On the other hand, it is still not exactly clear what role the processing of large semantic distances play in the processing of figurative meaning in novel metaphors. The question whether low salience and/or coarse coding by itself can account for RH processing of novel metaphors has been scarcely addressed directly.

In a divided visual field experiment Schmidt et al. (2007) found RH effects for unfamiliar metaphorical and unfamiliar literal sentences too, although there were no LH effects even for familiar literal sentences. It is possible that the RH processing dominance for unfamiliar conditions was not induced by semantic distance, but by context. In their fMRI study, also involving sentences, Schmidt and Seger (2009) found the right insula involved in the processing of unfamiliar vs. familiar metaphors, but the opposite contrast revealed right hemispheric regions also (inferior and middle frontal gyrus). In an experiment employing the same conditions as the present study, but using sentences, Diaz et al. (2011) found both the two novel and the two figurative conditions activating right inferior frontal gyrus (IFG). However, familiar and novel literal sentences, and familiar metaphors all evoked RH regions; novel metaphors did not differ from familiars or novel literals at all; and when contrasting the two literal conditions only LH regions showed up for novel ones. All in all, as the authors also point out, the complexity of stimulus construction could have played a role. In further neuroimaging studies semantic distance, context, and figurativeness all could have been similarly tangled with each other: Intriguingly there were no RH activations for novel metaphors embedded in sentences (Mashal and Faust, 2010; Mashal et al., 2009; Shibata et al., 2007). As sentences put a higher processing demand on the RH via pragmatics (Van Lancker, 1997; Van Lancker Sidtis, 2006), the RH effects could have been canceled in the analysis. As metaphorical contexts' numerous linguistic dimensions (Steen, 2004) can mask RH effects, isolated word pairs or compound words could help reduce the computational load on the RH.

Noun noun compound processing

Compound words belong to a special linguistic realm being combinations of nouns (or adjectives and nouns, not considered from now on): more complex than single words, governed by morphology, but simpler than propositions or sentences, governed by syntax. Their morphological complexity does not stem from pre- or suffixes, but from their constituents' internal hierarchical structure. In German (and in English) noun noun compound words (NNCs) are right headed, meaning that the second constituent, the *head* determines the semantic category and the morphosyntactic features of the whole compound, while its meaning is altered by the first noun, the *modifier* (Downing, 1977). This idea is by an eye tracking study showing strong second lexeme frequency effects (Juhász et al., 2003). Compounds can be endocentric/transparent (e.g., “snowball”) where the meaning is constructed from the parts, or can be exocentric/opaque with no head (as in “humbug”, which is not a kind of bug) where the meaning does not emerge as the result of a semantic combination (Spencer, 1991).

Compounds are processed slower when separated by a space, suggesting that they are represented as lexical units, at least to a certain extent, however both constituents can have some priming effect, even in opaque compounds (Libben et al., 2003), which are nevertheless processed more slowly than matched transparent ones (Ji, 2008).

Eye-tracking studies suggest that there are two separate processing steps both in German (Inhoff et al., 2000), and in English (Juhász et al., 2005): a decomposition and a reintegration stage. The second stage seems to be a semantic composition, determined by the relational structure of the constituents, like head FOR modifier (e.g., “cheese-knife”), or modifier HAS head (e.g. “coat-button”). This conceptually driven integration is true not only for novel compounds (Gagné and Spalding, 2004), but apparently for familiar ones too (Gagné and Spalding, 2009). According to a picture naming experiment, relations are represented independently of the parts, and relational priming might be similar to syntactic priming (Raffray et al., 2007).

The above results are best accounted for by the structured storage theory of compounds (Bien et al., 2005), which suggests that compounds are decomposed and reassembled along the stored structural position of the constituents: The structural position is part of the representation, allowing a differentiation between “doghouse” and “housedog”. The theory thereby lies somewhere in-between nondecompositional and fully decompositional views, the former proposing a complete list of compounds in the mental lexicon, while the latter taking the position that all of them are decomposed and reassembled at every instance.

Event-related potential (ERP) studies also support a semantic integration account. The N400 component, a response often associated with semantic processing (Hillyard and Kutas, 1983; Kutas and Federmeier, 2000), has been found sensitive to the lexical-semantic integration, and the late anterior negativity (LAN) suggests morphosyntactic decomposition (Chiarelli et al., 2007; Koester et al., 2004).

In an fMRI experiment the production of Dutch NNCs has been primed via the presentation of the picture of the first constituent (the modifier). This morphological process activated BA 47 in left inferior frontal gyrus (LIFG) independently of phonological and semantic processes (Koester and Schiller, 2011).

Taken together these results support the idea of a hierarchical representation of the internal structure of NNCs, suggesting that morphosyntactic and semantic features are integrated primarily at a conceptual level.

Combinatorial semantic processing

In some special cases it is possible to dissociate the almost always overlapping dimensions: the salience of an expression, referring mainly to familiarity, frequency, etc., and the coarseness of coding, referring mainly to associatedness and semantic feature overlap.

In an experiment aimed directly at the processing of noun noun phrases the constituents were not unfamiliar, and were co-occurring, but they were not closely associated either. Stronger activations were found in angular gyrus (AG), adjacent supramarginal gyrus (SMG), and middle temporal gyrus (MTG), but unexpectedly in the RH for highly meaningful phrases (e.g., “lake house”) as compared to their less meaningful reversals (e.g., “house lake”). The latter in turn evoked a stronger activation of the left inferior frontal junction (IFJ) and LIFG (Graves et al., 2010). According to the authors the phrases required coarse semantic coding (Beeman et al., 1994) that allowed more space for the constructive combinatorial semantic processing of compatible concepts, even though they were not novel.

Conventional German NNCs' are also unique linguistic constructs: Two lemmas are joined together to form a compound with a salient meaning, however the second constituents (the heads) are neither closely associated, nor do they share several semantic features with the first constituents (the modifiers). Unlike highly familiar, conventional adjective–noun word pairs that are strongly associated and highly co-occur, NNC constituents do not go together often. They most likely appear together in specific NNC combinations, but NNCs even have a relatively low frequency in general (as compared to non-compound words, which is actually a methodological concern for compound research, see Juhász and Rayner, 2003).

As even conventional NNCs are processed via a semantic decomposition and reintegration of not strongly associated elements, they could require coarse semantic coding (despite their salient meaning). Their constituents are definitely compatible, and so their processing is expected to resemble the RH combinatorial semantic processing of highly meaningful noun noun phrases observed by Graves et al. (2010).

However, according to the graded salience hypothesis (Giora, 2003) it is salience that determines hemispheric processing, both metaphorical and literal novel NNCs, regardless of figurativeness should increase BOLD signal change in RH regions more than conventional metaphorical and literal expressions. At the same time the latter two should increase BOLD signal change in LH regions that are thought to process salient meanings.

Taking both theories into consideration novel and conventional NNCs should not be processed identically. Novel NNCs also should require coarse coding, but most probably on a much more thorough level than conventional NNCs. Nevertheless, based on previous findings novel metaphors are expected to evoke a stronger BOLD signal change in the RH. Contrasting them to novel literal expressions could shed light on metaphor processing independent of semantic distance processing.

Conventional and novel NNCs allow a gradual testing of the interaction between semantic relatedness and figurativeness. Novel and conventional compounds, regardless of figurativeness, should require a very similar level of semantic combination, and could be indistinguishable in terms of behavioral measures. Meanwhile, as metaphors require the selection and suppression of certain features of one of the constituents, metaphorical NNCs could pose an overall higher computational demand on the system than literal NNCs, above the semantic combination they both require. For this reason a gradually increasing processing demand was predicted for our four categories of NNCs: because of their salient meaning *conventional literal* NNCs should pose the lowest computational demand, followed by *conventional metaphorical* NNCs, with an extra meaning selection step. *Novel literal* NNCs should be even more demanding, because of the non-salient nature of the unfamiliar combination of the nouns, whereas *novel metaphorical* NNCs should put the highest computational load on the system being non-salient, and because of the required meaning selection procedure.

Methods

Participants

Forty healthy adult volunteers (20 females, mean age: 24.2 years, range: 19–30) participated in the study for cash or course credit. All were native speakers of German, right handed, as determined by the Edinburgh Handedness Inventory (Oldfield, 1971), $M = 89.7$, $SD = 12.5$, had normal or corrected to normal vision, and had no history of neurological or psychiatric disorders. Approval of the ethics committee of the Freie Universität, Berlin, and informed consent of participants were obtained.

Stimuli

The stimuli consisted of 200 German noun noun compound words (NNCs), divided equally among four conditions: conventional metaphors (CM) e.g. “Stuhlbein” (“chair-leg”), novel metaphors (NM) e.g. “Plastikschwur” (“plastic-oath”), conventional literal (CL) e.g. “Alarmsignal” (“alarm-signal”), and novel literal expressions (NL) e.g. “Stahlhemd” (“steel-shirt”). The criterion for metaphors was that they should make no sense when read strictly literally, whereas novel literal expressions should have literally possible, but unusual meaning. NNCs also allow for the simplest possible (single word) presentation for metaphorical expressions.

For each condition 100 items were produced by three German native speaker research assistants at the Freie Universität, Berlin. NNCs were controlled for length (number of letters); the sum of the frequencies of the constituents' lemma form, and the sum of their lemma frequency class (e.g., the German word “*der*” *(“the”)* has got about $2^{(\text{frequency class})}$ the number of occurrences than the selected word), based on University of Leipzig's Wortschatz Lexikon: http://wortschatz.uni-leipzig.de/; and factors of the Berlin Affective Word List/BAWL (Võ et al., 2006): emotional valence, arousal, and imageability. For compounds not listed in the BAWL (e.g., novel ones) ratings were obtained in line with the original procedure from 19 volunteer university students, who received course credit, and did not participate in the fMRI experiment. In the next step, three linguist experts ranked the words for being plausible examples of their category or not, and selected the 50 best representatives.[1] Since the most important goal was to keep the key qualitative differences between conditions, while using the best examples, it was not possible to match all the above factors completely across all conditions (e.g., novel compounds naturally being less imageable or meaningful than conventional compounds). Still, differences were reduced as much as possible, and factors were controlled for during the final data analysis. An additional 26 volunteer university students, who also did not participate in the fMRI experiment, rated the compound words also for how meaningful and how literal they are on a 7 point Likert scale. The values of all the factors are presented in Table 1.

Table 1
Mean (SD) values of linguistic factors of the compounds.

	Conventional metaphor	Conventional literal	Novel metaphor	Novel literal
Nr. of letters	10.78 (1.67)	10.52 (1.64)	11.16 (1.67)	10.82 (1.45)
Lemma frequency sum	39081 (52202)	86190 (150649)	42181 (62895)	27827 (35758)
Lemma frequency class sum	21.02 (5.28)	19.44 (5.16)	23.36 (3.02)	23.98 (3.41)
Valence (between −3 and +3)	−0.379 (1.176)	0.122 (0.881)	−0.559 (1.151)	−0.213 (0.780)
Arousal (1 = none)	3.268 (0.619)	2.968 (0.656)	3.325 (0.636)	3.095 (0.459)
Imageability (1 = none)	4.835 (1.232)	5.356 (1.189)	3.075 (0.866)	4.579 (0.832)
Meaningfulness (1 = highest)	2.434 (0.581)	1.813 (0.314)	4.395 (0.646)	4.122 (0.716)
Literalness (1 = highest)	3.863 (0.742)	2.174 (0.244)	3.917 (0.417)	2.582 (0.416)

Although novelty and unfamiliarity refer to large semantic distances by definition, it is possible that some unfamiliar items are in fact existing but outdated expressions, or some novel items are not truly distant semantically (e.g. according to co-occurrence measures). Based on the above concerns, semantic relatedness for the novel NNCs (NM and NL) was controlled by excluding all compounds for which the constituent lemmas were significantly co-occurring in the Wortschatz corpus of 43 million German sentences (Quasthoff et al., 2006), and conventional compounds had to occur in the corpus of contemporary German. A recent computational and behavioral analysis has provided evidence that this measure of semantic distance accounts well for semantic relations between words (Hofmann et al., 2011). Familiar NNCs (CM and CL), being already existing words, all have a frequency value of their own, and a salient meaning — despite the fact that they are neither sharing many semantic features, nor are they closely associated. Although a portion of them was found significantly co-occurring, none of the second constituents was a significant right neighbor of the first constituents.

Experimental procedure

After reading the instructions and completing a 20 item practice task, participants were scanned in 5 imaging runs, each consisting of 40 trials. In each trial a compound word was presented centrally for 2000 ms on a black background, using white, 16 pt Arial capital letters, followed by a fixation cross jittered between 4000 and 8000 ms. Participants were instructed to read the items silently, and to indicate via button press as fast and as accurately as possible whether the word appearing on the screen seemed familiar or unfamiliar to them. Participants were required to respond with their right thumb using an MR-compatible button box.

fMRI data acquisition

Neuroimaging data was collected by a 3 T Siemens Tim Trio MRI scanner fitted with a 12-channel head coil (Siemens Erlangen, Germany), at the laboratory of the Dahlem Institute for Neuroimaging of Emotion

[1] An attempt to have doctoral students of linguistics categorize the words according to Lakoffian theoretical concerns failed, as the results were few in number and strongly inconsistent.

(D.I.N.E.), Freie Universität, Berlin. Initially, a high-resolution 3D T1-weighted dataset was acquired from each subject (176 sagittal sections, $1\times1\times1$ mm^3). During every run 200 whole-brain functional T2*-weighted echo planar images (EPI) were taken with the parameters as follows: $3.0\times3.0\times3.0$ mm voxels, TR 2 s, TE 30 ms, flip angle 90°, matrix size 64×64, FOV 192 mm, slice thickness 3 mm, no gap, 37 slices.

Data analysis

The behavioral data were analyzed using SPSS 13 (IBM SPSS Statistics). To analyze the recorded fMRI data BrainVoyager QX 2.2 (Brain Innovation, Maastricht, The Netherlands) was used. The data were motion and slice-scan time corrected (cubic spline interpolation). Intra-session image alignment to correct for motion across runs was performed using the first image of the first functional run as the reference image. Following linear trend removal, data was filtered temporally in 3D with a high pass Fourier filter of 2 cycles in time course to remove low frequency drifts. Preprocessed data were spatially smoothed using an 8 mm full-width-half maximum Gaussian kernel to reduce noise. Statistical analyses were performed in Talairach space (Talairach and Tournoux, 1988) (Table 2). The T1 images were first rotated into the AC–PC plane, transformed into Talairach space, and then used to register the functional data to the subjects' 3D images. Anatomical regions were identified by manual inspection using the Talairach atlas and the Talairach demon (http://www.talairach.org).

The statistical analyses were carried out using a voxel-wise General Linear Model (GLM) at the single-participant-level first, based on design matrices built from the four conditions (CM, CL, NM, NL). BOLD responses were separately modeled using a boxcar function, which was convolved with a theoretical two gamma hemodynamic response function (Friston et al., 1998) for each experimental condition, and the model was independently fitted to the signal of each voxel. Subsequently these parameter fits were evaluated in the second level analysis applying the Random Effects Model. To examine the effects of familiarity and metaphoricity direct contrasts of the conditions were calculated, using a threshold of $p<.00001$ and a cluster size >4. This cluster threshold was determined by running an AlphaSim analysis with NeuroElf v0.9c (http://neuroelf.net/) to correspond to an FWE-correction of $p<.05$.

To detect brain areas responding to the degree of valence, arousal, imageability and meaningfulness parametric analyses were carried out. The former linguistic factors were separately modeled as parametric regressors. Additionally, as measurement of the BOLD response beta-values were extracted from the LIFG for each single word and correlation coefficients were calculated from these values with meaningfulness in order to visualize the results of the afore mentioned parametric modulation analysis. Emotional valence, arousal, and imageability were included as covariates in one, and the sum of the logarithm of the constituent's word frequency and reaction times (as an extra control for difficulty) in another analysis. These regressors were generated in the following way: the previously modeled BOLD responses (evoked by the four main conditions) were modulated by multiplying them with normalized values (from −1 to +1) of individual reaction times (and other variables) for each single word. Hereby the response to each condition was split into 2 parts: the condition itself and the parametric modulation of the specific effect. Then General Linear Models were calculated including these additional regressors to create an extended model.

Table 2
Talairach coordinates of BOLD signal change peaks.

Contrast	x	y	z	k	Max	Diameter
(CM + CL) > (NM + NL)						
Precuneus	0	−58	31	1572	10.491110	d = 1.0 mm
	0	−58	31	L	10.491110	d = 1.0 mm
	0	−34	33	L	9.673079	d = 1.0 mm
	−3	−64	50	L	6.259244	
L inferior temporal gyrus	−56	−14	−11	237	10.207053	d = 1.0 mm
Medial frontal gyrus	−11	51	2	1794	10.024557	
	−11	51	2	L	10.024557	
	7	34	2	L	9.460264	d = 1.0 mm
	0	64	12	L	9.067936	d = 3.6 mm
	0	6	4	L	6.866185	d = 4.0 mm
	−19	65	21	L	6.201931	d = 1.0 mm
	0	56	40	L	5.525948	d = 4.1 mm
R SMG	50	−47	35	1805	9.406252	d = 3.7 mm
	50	−47	35	L	9.406252	d = 3.7 mm
R MTG	59	−20	−6	L	8.942244	
R angular gyrus	45	−67	29	L	8.801411	d = 2.4 mm
R posterior STS	61	−55	1	L	7.559316	
	55	−35	−4	L	7.452419	
	61	−31	16	L	6.512668	d = 2.2 mm
	64	−44	12	L	6.374903	d = 1.0 mm
R SFG	25	24	46	282	9.344021	
L angular gyrus	−47	−64	33	542	8.982792	d = 1.4 mm
L SFG	−40	26	45	92	6.594738	
	−40	26	45	L	6.594738	
	−34	15	46	L	6.459573	d = 2.0 mm
	−22	31	40	14	6.228384	d = 1.0 mm
R STG	50	4	−13	37	6.188431	
	−29	31	22	6	5.756945	d = 5.0 mm
(CM + CL) < (NM + NL)						
LIFG	−43	−2	28	921	10.096503	d = 1.0 mm
	−43	−2	28	L	10.096503	d = 1.0 mm
	−46	22	21	L	9.711818	d = 1.4 mm
L insular cortex	−34	21	8	L	8.039451	
	−50	−8	43	L	7.628479	
Pre-SMA	−8	−1	53	265	9.481983	d = 2.8 mm
	−8	−1	53	L	9.481983	d = 2.8 mm
	11	22	38	L	7.877074	
R insular cortex	30	20	10	139	8.968314	d = 1.0 mm
L fusiform gyrus	−43	−55	−6	77	7.122552	d = 5.8 mm
	−43	−55	−6	L	7.122552	d = 5.8 mm
	−39	−40	−8	L	6.071026	d = 4.0 mm
(CM + NM) > (CL + NL)						
LIFG and LIFJ	−46	19	14	609	8.258172	d = 1.0 mm
	−46	19	14	L	8.258172	d = 1.0 mm
	−50	6	24	L	7.126366	d = 2.2 mm
L temporal pole (aSTS)	−49	3	−6	L	6.776460	d = 1.0 mm
	−50	14	−1	L	6.755461	d = 2.2 mm
L posterior STS	−53	−41	8	46	6.565135	
L amygdala	−21	−11	1	8	5.759394	
L anterior STS	−54	−10	1	5	5.671626	d = 2.0 mm
CM > CL						
LIFG and LIFJ	−46	25	11	731	12.040571	d = 2.4 mm
	−46	25	11	L	12.040571	d = 2.4 mm
	−42	7	29	L	10.068727	d = 3.0 mm
Pre-SMA	−8	10	48	16	6.507284	
L posterior STS	−53	−36	8	12	5.919095	d = 3.2 mm
L hippocampus	−33	−11	−14	4	5.458577	d = 2.2 mm
NM > NL						
L temporal pole (aSTS)	−52	3	−3	85	7.105155	d = 1.7 mm
L posterior STS	−55	−41	10	11	5.746142	d = 1.0 mm
NM < NL						
L parahippocampal gyrus	−27	−36	−8	9	5.709272	d = 1.4 mm

Results

Behavioral results

During the outlier procedure 4.7% of all the recorded data were removed. Reaction time and error rate data were submitted for both a subject (F_1) and an item (F_2) based one-way ANOVA analysis, and post-hoc tests were performed to determine the differences between categories (Fig. 1).

Subject based analysis revealed significant main effect of the categories for error rates, $F_1(3,156)=17.598$, $p<.001$. Levene's test for the

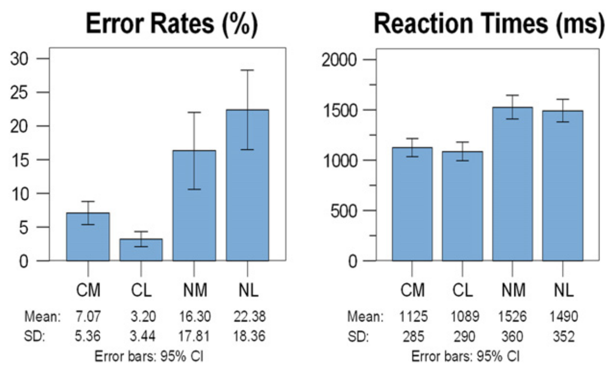

Fig. 1. Mean error rates and reaction times according to the F_1 analysis.

homogeneity of variances proved significant, $L(3,156) = 24.902$, $p < .001$, hence Tamhane post hoc test was performed, revealing significant differences between all but the NM and NL categories. Reaction time differences were calculated only for accurate responses, and were also significantly different between categories, $F_1(3,156) = 20.865$, $p < .001$. Tukey post hoc test showed differences between all categories, except for the CM–CL and the NM–NL comparisons.

Item based F_2 analysis provided similar results. Word category had a significant main effect on error rates $F2(3,196) = 28.909$, $p < .001$. As Levene's test proved to be significant, $L(3,196) = 8.522$, $p < .001$, the homogeneity of variances was not assumed; Tamhane post hoc test revealed significant differences between all categories except for CM and CL, and for NM and NL. Item based analysis of reaction times also showed a significant main effect of categories $F_2(3,196) = 119.466$, $p < .001$, and as the variances were not homogenous ($L(3,196) = 3.083$, $p < .028$), Tamhane post hoc test was applied, showing differences for all comparisons, but CM–CL and NM–NL.

Results were calculated for the uncorrected data set also, but the differences between categories remained exactly the same.

Neuroimaging results

Familiarity

To examine familiarity effects, the two conditions with salient meaning (CM and CL), were joined and contrasted against the two novel conditions with non-salient meaning (NM and NL): (CM + CL) > (NM + NL). Conventional compounds significantly increased the BOLD signal in right MTG (BA 21), right SMG (BA 40), bilateral AG (BA 39), right superior frontal gyrus (SFG: BA 8), left inferior temporal gyrus (ITG: BA 20) and in broad bilateral midline structures, as the ventromedial prefrontal cortex (VMPFC: BA 10, 12), the dorsal anterior cingulate cortex (dACC: BA 32), and subgenual cingulate area (BA 25), the posterior cingulate cortex (PCC: BA 23, 31), and the precuneus (BA 7). Novel NNCs increased BOLD responses in left IFJ (ventral BA 6) and LIFG (BA 44, 45), left fusiform gyrus (BA 37), bilateral insula, and pre-SMA (BA 6), as illustrated in Fig. 2.

Emotional valence, arousal, and imageability included in the analysis as covariates did not change the findings. The sum of the logarithm of the constituent's word frequency, and reaction times (to control for difficulty) also have been included as covariates in a separate analysis, and were found not to affect our main results either. As these factors cannot explain our findings we included an image and coordinates of activation peaks for this extended model in the Supplementary material. Results of parametric analyses and corresponding coordinates also can be observed in the Supplementary material.

Figurativeness

Brain areas associated with metaphor processing were found active by contrasting the two metaphorical against the two literal conditions (CM + NM) > (CL + NL). BOLD responses increased in LIFG (BA 44, 45), left IFJ (ventral BA 6), left temporal pole (BA 38), left posterior STS (BA 22), and left amygdala. As the LIFG was found involved

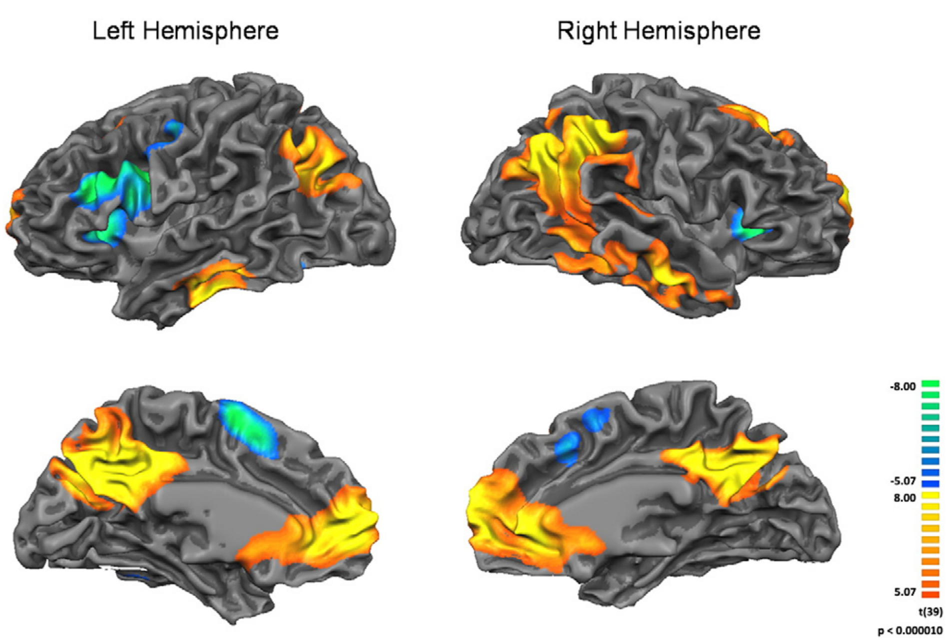

Fig. 2. BOLD signal change for contrasting conventional NNCs (warm colors) against novel NNCs (cold colors): (CM + CL) > (NM + NL). Conventional metaphorical and literal NNCs increased BOLD signal change in right temporoparietal areas, suggesting combinatorial semantic processing, in line with the coarse semantic coding theory, as their constituents are not closely associated. Novel metaphors and novel literal expressions induced BOLD signal increase in LIFG, presumably as a result of meaning making: unifying phonetic, morphosyntactic, and semantic features of novel words, via fine grained semantic coding.

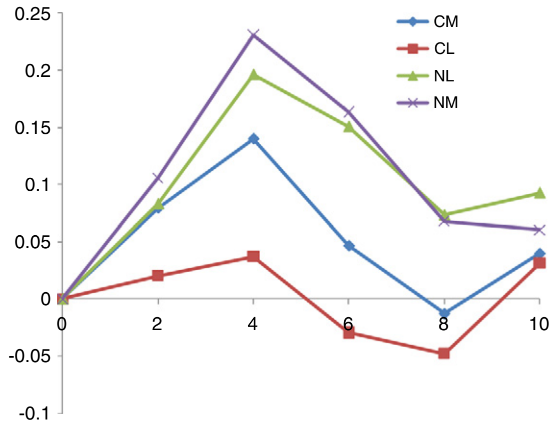

Fig. 3. BOLD responses in the LIFG in the (CM + NM) > (CL + NL) contrast. X-axis: percent BOLD signal change, Y-axis: time (s). The gradual BOLD signal increase of the four conditions suggests a gradual semantic processing demand for conventional literal expressions being the least complex, followed by conventional metaphors, requiring the selection and suppression of certain semantic features to construct figurative meaning; then came novel literal NNCs, where a new meaning has to be constructed from the two constituents, and finally by novel metaphors, where the novel figurative meaning has to be established via the selection and suppression of certain semantic features.

in several different contrasts, dynamics of the BOLD response in this region can be observed in Fig. 3.

To disentangle the effect of metaphoricity from the effect of familiarity, conventional and novel metaphors were separately contrasted against the corresponding literal condition with comparable salience. Conventional metaphors (CM > CL) increased the BOLD signal in left IFJ (ventral BA 6), LIFG (BA 44, 45), and pre-SMA (medial BA 6), and left posterior STS (BA 22). Novel metaphors (NM > NL) activated left temporal pole (BA 38) and left posterior STS (BA 22); this latter contrast revealed that NLs increased activation in left parahippocampal gyrus. The above results are shown in Fig. 4.

Discussion

The present experiment examined figurative language processing with special emphasis on semantic relatedness in an effort to separate these factors. Since all four categories of NNCs require some, albeit different kinds of semantic combination the question was: how does the computational load change as these factors interact, and specifically how much does semantic distance processing contribute to the processing of novel metaphors?

Familiarity

According to the graded salience hypothesis (Giora, 2003), non-salient (not coded, not co-occurring, not conventional, and not familiar) novel items seemed to be better candidates for activating the RH, while salient (coded, familiar, conventional, etc.) items were expected to more likely activate the LH. However, just the opposite pattern was observed: Despite being salient, conventional items (CM and CL) elicited higher BOLD signal increase in right temporoparietal regions, specifically in the SMG. Although the AG was activated bilaterally, the signal increase was lateralized to the right side in the SFG and MTG too. Nevertheless, these results can be interpreted according to Beeman's (1998) coarse semantic coding theory, as there was no close semantic relation even between the constituent words of familiar NNCs. They also fit well with the results of Graves et al. (2010) who also found right SMG activation. They attributed this to combinatorial semantic processing of the highly meaningful noun noun phrases, where the constituents are weakly associated with no overlapping semantic fields. Nonetheless, the right temporoparietal cortex also plays an important role in recognition memory (Cabeza et al., 2008): it is thus possible that memory processes modulated the familiarity effect in both studies.

Temporal areas are traditionally associated with the mental lexicon and are thought to store information about objects and their attributes, while right SFG seems to play an important role in goal-directed

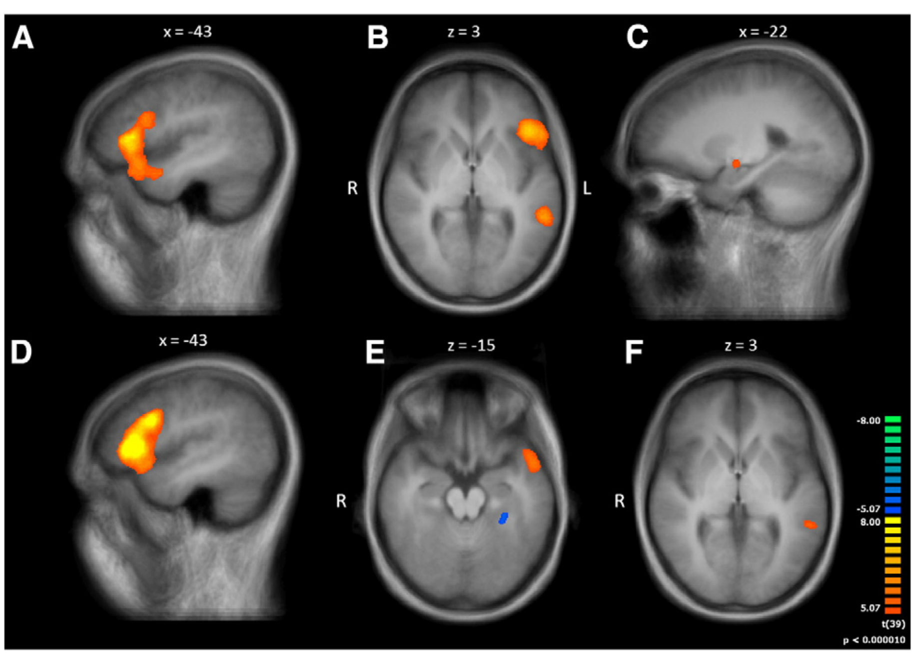

Fig. 4. BOLD signal increase for metaphorical NNCs. A, B, C: (CM + NM) > (CL + NL); metaphors activated LIFG, left IFJ, left temporal pole, left posterior superior temporal sulcus. Activations elicited by metaphors are constituted almost entirely of regions that showed an increased BOLD signal either for the CMs in the CM > CL, or for the NMs in the NM > NL contrast, suggesting that these conditions could have played decisive role in activating corresponding regions in the general figurative contrast. D: (CM > CL); conventional metaphors (compared to matched conventional literal expressions) activated LIFG and left IFJ, suggesting semantic selection and unification procedures, and left posterior superior temporal sulcus as a result of stronger semantic activation. E, F: (NM > NL); novel metaphors (relative to matched novel literal NNCs) increased BOLD responses in left temporal pole, perhaps as a result of higher demands on semantic integration, and in left posterior superior temporal sulcus also, again for stronger semantic activation. According to radiological convention the left side of the brain is on the right side of the figure.

semantic retrieval, especially when a large set of responses is possible (Binder et al., 2009). Together with the above mentioned regions and the broadly activated medial structures such as the VMPFC, the dACC, the PCC, and the central region of the precuneus, these areas are all part of a large semantic network identified by a comprehensive meta-analysis of the semantic system by Binder et al. (2009). The medial activations completely overlap with the default state network, which could reflect the ease of processing, but most probably they took part in comprehension too, as this network is thought to be a virtually inwardly oriented conceptual system, being responsible for semantic processing (Binder et al., 2009). As even conventional NNCs have complex relational structure, RH activations might be reflecting more than a mere linking, but a non-syntactic semantic combination of the two elements. Apparently coarse semantic coding does not necessarily pose a higher processing demand and can be effortless, reflected in short reaction times and default state network activations.

Novel NNCs (NM and NL) elicited strong activations in LH prefrontal areas, which seems to be at odds with the graded salience hypothesis, and at first glance even with the coarse semantic coding theory, since the lemmas did not share narrow semantic fields. However, when it came to the semantic composition of non-associated, non-salient, and not even significantly co-occurring lemmas into truly novel NNCs, processing requirements might change. Beeman (1998) suggests that hemispheric activation primarily depends on semantic feature overlap. The system could have required a more focused, fine grained conceptual analysis, and narrower semantic feature selection to establish the meaning, as it is forced to come up with a single solution during retrieval, and competing candidate concepts need to be filtered during the selection of an appropriate one. Longer reaction times could also reflect a higher processing demand and hence a more thorough analysis of novel items.

Left inferior frontal areas were found responsible for both linguistic and non-linguistic processes. According to a meta-analysis (Owen et al., 2005) the IFG plays an important role in working memory and attention, while the IFJ was found to be involved in cognitive control and task switching by another meta-analysis (Derrfuss et al., 2005). However, the LIFG is associated with the processing of morphological complexity in general (Bozic et al., 2007; Marslen-Wilson and Tyler, 2007), morphosyntactic compounding (Koester and Schiller, 2011), but even with the processing of difficult unfamiliar metaphors as compared to easy unfamiliar metaphors (Schmidt and Seger, 2009). In fact different subregions may actually play different roles: In their meta-analysis Liakakis et al. (2011) found left BA 44 involved in working memory, whereas left BA 45 and BA 46 associated with semantic, and phonological processing. This latter area, the anterior portion of the IFG, is identical to the cluster identified by an earlier meta-analysis, found to be activated by semantic processing (Bookheimer, 2002). These results partly serve as the basis of Hagoort's (2005) neurobiological language model, the Memory, Unification, Control (MUC) framework, where the LIFG is responsible for the Unification gradient: the interactive and concurrent integration of phonology, syntax, and semantics into a complex whole. Importantly working memory is an integral part of the system, as the neural requirements of the unification include keeping the lexical building blocks activated.

Jung-Beeman's (2005) Bilateral Activation, Integration, and Selection (BAIS) framework assigns a slightly different role to the LIFG. As bilateral language areas are interacting in line with task demands, fine grained coding taking place in LH, and coarse coding in RH areas, this model suggests that the LIFG is responsible for the meaning Selection component within narrow semantic fields.

Although these theories propose somewhat different procedures to the LIFG, presenting novel NNCs could easily impose higher processing demands on this region, as the main challenge is the precise selection and/or complex unification of the phonetic, syntactic, and semantic features of the parts into novel units.

The left fusiform gyrus showed a negative correlation with association values in the study of Graves et al. (2010), hence the activation found in the present experiment fits well with the processing of novel NNCs, with no significant co-occurrence. The anterior insula was found activated for novel metaphors previously (Ahrens et al., 2007; Mashal et al., 2007), but it could be a marker of the selective ventral attention system (Eckert et al., 2009). Pre-SMA also expressed higher BOLD signals, an area taking part in working memory tasks, such as sequence learning (Owen et al., 2005), hence this neural response could reflect the sequential ordering aspect of processing novel NNCs.

This complex pattern of phonetic, morpho-syntactic, and semantic unification, meaning selection, processing and sequencing of non-associated lemmas, cognitive control, and working memory load could reflect a more demanding (and more compelling) meaning-making procedure (cf. Bruner, 1990), where meaning is actively constructed, rather than passively comprehended. Such a productive effort would not be unusual for poetic, non-everyday language that does not necessarily always give in easily to understanding, and requires interpretation.

Figurativeness

The activations elicited by metaphorical (CM + NM) vs. literal (CL + NL) NNCs are constituted almost entirely of regions that showed an increased BOLD signal either for the CMs in the CM > CL, or for the NMs in the NM > NL contrast. This suggests that activations showing up in the combined figurative contrast could have been mainly the sum of the activations of the two otherwise not overlapping metaphorical conditions (except for left anterior STS).

Contrasting CMs and CLs (that are indistinguishable by reaction times) revealed a BOLD signal increase in LIFG and left IFJ for the CMs, probably as a result of their higher complexity. LIFG was found activated also by Diaz et al. (2011) for an identical contrast. Metaphors require the listener to select non-concrete features of the figurative constituent words — a "chair-leg" is not a leg in the literal, physical sense. Hence they could have imposed higher computational demand on meaning selection processes, and required a more thorough unification procedure. In general these results are in line with conventional metaphors evoking stronger LH activations in fMRI studies, and posing a slightly higher effort relative to literal expressions in ERP experiments (Arzouan et al., 2007; Lai et al., 2009).

The contrast between the behaviorally also indistinguishable NM and NL categories showed activations for NMs in the left posterior STS (BA 22), probably as a result of the higher conceptual complexity of figurative language, and in the left anterior STS, an area suggested to be responsible for verbal as compared to perceptual knowledge by Binder et al. (2009). The region included the temporal pole, also found activated by Schmidt and Seger (2009) for figurative language in general, and by Ahrens et al. (2007) for novel (anomalous) metaphors. According to the MUC model (Hagoort, 2005) temporal regions play a role in memory retrieval, while according to the BAIS model (Jung-Beeman, 2005) they are responsible for two separate functions: posterior STS is where semantic information is supposed to be activated (required by both metaphorical conditions), while anterior STS and temporal pole are held responsible for semantic integration. Based on the observed pattern of activations of brain regions associated with semantic functions, our results suggest that novel metaphorical expressions required higher conceptual processing than similarly novel, unfamiliar, but literal NNCs. This is most probably not due to coarse coding, but to the fine grained activation of an appropriate, not dominant, and not literal sense of one of the constituents, and the following, more complex integration of the two parts into a novel figurative meaning. Up to this date, to our knowledge, this is the first study reporting LH activations for novel metaphorical stimulus material out of context. Previous studies might have found RH activations mainly because of semantic distance processing, but since in the present experiment semantic relatedness was carefully controlled for, it was possible to parse it out from the neural processing correlates of novel metaphorical expressions.

Posterior STS (BA 22) and LIFG, regions found expressing BOLD signal increase in the metaphorical vs. literal contrast, are located at the overlap of areas activated by both internal-conceptual, and external-perceptual information (Binder et al., 2009), suggesting that the integration of both knowledge domains is important for metaphor comprehension. Metaphorical items apparently required a thorough processing, involving stronger neural markers for activating, selecting and integrating semantic information.

Finally, a gradually increasing processing demand was proposed for the four conditions, and has been confirmed according to the LIFG activation patterns (Fig. 3). Familiar CLs induced the least BOLD signal change, followed by CMs, requiring the selection and filtering of certain semantic features in order to construct the figurative meaning; reflected also in behavioral results, NLs posed even higher processing demand, as a result of integrating semantically distant concepts, and finally NMs evoked the highest BOLD signal change, requiring both bridging semantic distance, and establishing metaphorical meaning.

Conclusions

The present study examined the neural correlates of processing familiar and unfamiliar, literal and figurative NNCs. On the one hand, at odds with the graded salience hypothesis (Giora, 2003), but in line with the coarse semantic coding theory (Beeman, 1998), distantly related familiar NNCs activated right temporoparietal regions (e.g., SMG) probably reflecting combinatorial semantic processing (Graves et al., 2010). On the other hand, unfamiliar conditions increased BOLD signal change in LH regions, such as the LIFG, which could be accounted for by the coarse semantic coding theory, since novel items could have required fine grained conceptual analysis, and narrow semantic feature selection (Jung-Beeman, 2005) for the unification of phonological, (morpho-)syntactic and semantic information (Hagoort, 2005), presumably due to meaning-making (Bruner, 1990). The comprehension of figurative language was successfully separated from semantic distance processing, and LH regions were found activated even for novel metaphorical expressions, suggesting a fine grained conceptual analysis during the selection and suppression of certain conceptual features in order to establish figurative meaning.

Acknowledgments

We would like to thank the highly valuable suggestions of anonymous reviewers. This research could not have been realized without the generous help of Markus Conrad, Ph.D., Mario Braun, Ph.D., Maren Luitjens, Isabel Amberger, Johannes Ecker, and Hauke Blume.

Appendix A

Conventional metaphorical	Conventional literal	Novel metaphorical	Novel literal
ANGSTHASE	ALARMSIGNAL	ALPENMATTEN	AUTOFILM
ARMLEUCHTER	ALTARBILD	ÄRGERBERG	BLECHGLOCKE
ARMUTSZEUGNIS	BAUFIRMA	BEWERBERPEST	DACHFEUCHTE
ARSCHKRIECHER	BETTRUHE	BLUMENBLICK	DAUMENVERBAND
AUGAPFEL	BLEISTIFT	BUMERANGLAUNE	EINSATZANTRAG
BÄRENHUNGER	BOXSACK	DANKEBBE	ERZTROMPETE
BAUMKRONE	BRIEFMARKE	DUFTGESANG	EULENFALLE
BEIFALLSSTURM	BÜRGERAMT	EREIGNISPULS	FASANMÖRDER
BÖRSENHAI	EHEPARTNER	ESSIGHUMOR	FELLINSEKT
DONNERBALKEN	FENSTERGRIFF	FANTASIEPAPST	FILZKANNE
DRAHTESEL	GASHEIZUNG	FETTGÜRTEL	FLÖTENKISTE
ERFOLGSREZEPT	GEBETSSTUNDE	FLAMMENSCHRIFT	FLUCHTLUKE

Appendix A (*continued*)

Conventional metaphorical	Conventional literal	Novel metaphorical	Novel literal
FLUGHAFEN	GERICHTSSAAL	FLUMMIVERSTAND	FUGENCREME
FLUSSLAUF	GESCHÄFTSMANN	GEIERBERUF	FUNDREGAL
FRAUENHELD	GITTERSTAB	GELDDURST	GALGENLEITER
GEBÄRMASCHINE	HILFSARBEITER	GLAUBENSSÄER	GÄNSEPFEIFE
GEDANKENGANG	HUNDELEINE	HAUSDIKTATUR	GEWITTERSPUR
GEHIRNWÄSCHE	IMBISSSTAND	HENNENHYSTERIE	GLASAFFE
GEWISSENSBISSE	KAFFEESAHNE	IDEENHAGEL	GLITZERTELEFON
GNADENBROT	KIRCHTURM	IGELFROST	GURKENTÜTE
HAARWURZEL	KLEIDERHAKEN	KAKTUSBART	HIRTENTROMMEL
HANDSCHUH	KREDITKARTE	KATERBREMSE	HOLZFLASCHE
HEIZKÖRPER	KÜCHENMESSER	KIRSCHWANGE	KARTONFLIEGE
HERZKAMMER	LASTWAGEN	LEIDWOGE	KEKSVERSTECK
JAMMERLAPPEN	LEBENSFREUDE	LISTENTEUFEL	KEROSINDOSE
KABELSALAT	LEHRJAHR	LÜGENBRATEN	KÖRNERKUCHEN
KADERSCHMIEDE	MARINESOLDAT	MATRATZENRUHM	KRANSCHRAUBEN
KINDERGARTEN	MIETZINS	MEINUNGSKÄFIG	KRÄUTERHEFE
KUMMERKASTEN	MOTTOPARTY	MENSAKOMA	KUNSTSCHWAN
LUFTBRÜCKE	NATURSCHUTZ	MUSIKSUPPE	LABORTABLETTE
LUNGENFLÜGEL	OFENROHR	MUTTROPFEN	LEHRERORDNER
LUSTMOLCH	PFIRSICHKERN	MÜCKENKUSS	LIPPENFALTEN
MEERBUSEN	PFLEGEVATER	NASENSCHAUER	LÖWENNETZ
MONDGESICHT	POLIZEIBEAMTE	NEIDFIEBER	MODELLGELENK
MOTORHAUBE	POSTFACH	ORDNUNGSBIENE	MÖNCHSJACKE
NOTNAGEL	RASIERAPPARAT	PARADIESMORAL	PAPIERSÄGE
ORANGENHAUT	REGENSCHIRM	PHRASENVULKAN	PARKTONNE
PANZERFAUST	REISEPASS	PLASTIKSCHWUR	RASENKREUZ
PECHVOGEL	SCHLACHTFELD	SCHMUTZMAGNET	SANDTISCH
RABENELTERN	SEGELBOOT	SEELENFARBEN	SCHILFGEHEGE
RAMPENSAU	STEUERZAHLER	SPRACHKANONE	SILBERSAITE
SCHÜRZENJÄGER	STIERKAMPF	STADTNARBE	SOFALADEN
SKANDALNUDEL	STREIKFÜHRER	STAUBKOST	STAHLHEMD
SPIELHÖLLE	STROHBALLEN	TASCHENBAUCH	STEINLACK
STUHLBEIN	TAUFBECKEN	TRAUMACHIRURG	TABLETTFOLIE
TALSOHLE	TAXIFAHRER	WALFIGUR	TANZHOSE
WASSERSPIEGEL	WEINKELLER	WANDKOSTÜM	TIGERPYJAMA
WESPENTAILLE	WETTERBERICHT	WEISHEITSSEE	TRAKTORBUCH
WOLKENKRATZER	WOCHENENDE	WISSENSPIRAT	ZIEGELMASSE
WÜSTENSCHIFF	ZIMMERPFLANZE	WORTHAUFEN	ZUFALLTASTE

Appendix B. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2012.07.029.

References

Ahrens, K., Liu, H., Lee, C., Gong, S., Fang, S., Hsu, Y., 2007. Functional MRI of conventional and anomalous metaphors in Mandarin Chinese. Brain Lang. 100, 163–171.
Anaki, D., Faust, M., Kravetz, S., 1998. Cerebral hemispheric asymmetries in processing lexical metaphors. Neuropsychologia 36, 691–700.
Aristotle, 335 BC/1952. Rhetoric, in: Ross, W.D. (Ed.) & Roberts, W.R. (Trans.), The works of Aristotle (vol. 11): Theoretica, de rhetorica ad alexandrum, poetica. Clarendon Press, Oxford, England.
Arzouan, Y., Goldstein, A., Faust, M., 2007. Dynamics of hemispheric activity during metaphor comprehension: electrophysiological measures. Neuroimage 36, 222–231.
Beeman, M.J., 1998. Coarse semantic coding and discourse comprehension. In: Beeman, M., Chiarello, C. (Eds.), Right Hemisphere Language Comprehension: Perspectives from Cognitive Neuroscience. Erlbaum, Mahwah, NJ, pp. 255–284.
Beeman, M., Friedman, R.B., Grafman, J., Perez, E., Diamond, S., Lindsay, M.B., 1994. Summation priming and coarse semantic coding in the right hemisphere. J. Cogn. Neurosci. 6, 26–45.
Bien, H., Levelt, W.J., Baayen, R.H., 2005. Frequency effects in compound production. Proc. Natl. Acad. Sci. U. S. A. 102, 17876–17881.
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. Cereb. Cortex 19, 2767–2796.
Bookheimer, S., 2002. Functional MRI of language: new approaches to understanding the cortical organization of semantic processing. Annu. Rev. Neurosci. 25, 151–188.
Bottini, G., Corcoran, R., Sterzi, R., Paulesu, E.S.P., Scarpa, P., Frackoviak, R.S.J., 1994. The role of the right hemisphere in the interpretation of the figurative aspects of language: a positron emission tomography activation study. Brain 117, 1241–1253.
Bowdle, B., Gentner, D., 2005. The career of metaphor. Psychol. Rev. 112, 193–216.
Bozic, M., Marslen-Wilson, W.D., Stamatakis, E.A., Davis, M.H., Tyler, L.K., 2007. Differentiating morphology, form, and meaning: neural correlates of morphological complexity. J. Cogn. Neurosci. 19 (9), 1464–1475.

Bruner, J.S., 1990. Acts of Meaning. Harvard University Press.
Cabeza, R., Ciaramelli, E., Olson, I., Moscovitch, M., 2008. The parietal cortex and episodic memory: an attentional account. Nat. Rev. Neurosci. 9 (8), 613–625.
Chen, E., Widick, P., Chatterjee, A., 2008. Functional–anatomical organization of predicate metaphor processing. Brain Lang. 107, 194–202.
Chiarelli, V., El Yagoubi, R., Mondini, S., Danieli, M., Perrone, G., Semenza, C., 2007. The electrophysiological correlates of noun–noun compounds. Brain Lang. 103, 8–249.
Chiarello, C., 1991. Interpretation of word meanings by the cerebral hemispheres: one is not enough. In: Schwanenflugel, P. (Ed.), The Psychology of Word Meaning. Lawrence Erlbaum Associates, Hillsdale, NJ.
Chiarello, C., Burgess, C., Richards, L., Pollock, A., 1990. Semantic associative priming in the cerebral hemispheres: some words do, some words don't… sometimes, some places. Brain Lang. 38, 75–104.
Coulson, S., Van Petten, C., 2007. A special role for the right hemisphere in metaphor comprehension? ERP evidence from hemifield presentation. Brain Res. 1146, 128–145.
Coulson, S., Williams, R.W., 2005. Hemispheric asymmetries and joke comprehension. Neuropsychologia 43, 128–141.
Coulson, S., Wu, Y.C., 2005. Right hemisphere activation of joke related information: an event-related potential study. J. Cogn. Neurosci. 17 (3), 494–506.
Derrfuss, J., Brass, M., Neumann, J., von Cramon, D.Y., 2005. Involvement of the inferior frontal junction in cognitive control: meta-analyses of switching and Stroop studies. Hum. Brain Mapp. 25, 22–34.
Diaz, M.T., Barrett, K.T., Hogstrom, L.J., 2011. The influence of sentence novelty and figurativeness on brain activity. Neuropsychologia 49, 320–330.
Downing, P., 1977. On the creation and use of English compound nouns. Language 53, 810–842.
Eckert, M.A., Menon, V., Walczak, A., Ahlstrom, J., Denslow, S., Horwitz, A., Dubno, J.R., 2009. At the heart of the ventral attention system: the right anterior insula. Hum. Brain Mapp. 30 (8), 2530–2541.
Eviatar, Z., Just, M.A., 2006. Brain correlates of discourse processing: an fMRI investigation of irony and conventional metaphor comprehension. Neuropsychologia 44, 2348–2359.
Fauconnier, G., Turner, M., 1998. Conceptual integration networks. Cognit. Sci. 22, 133–187.
Faust, M., Chiarello, C., 1998. Sentence context and lexical ambiguity resolution by the two hemispheres. Neuropsychologia 3, 827–836.
Federmeier, K.D., 2007. Thinking ahead: the role and roots of prediction in language comprehension. Psychophysiology 44, 491–505.
Federmeier, K.D., Kutas, M., 1999. Right words and left words: electrophysiological evidence for hemispheric differences in meaning processing. Cogn. Brain Res. 8, 373–392.
Federmeier, K.D., Mai, H., Kutas, M., 2005. Both sides get the point: hemispheric sensitivities to sentential constraint. Mem. Cognit. 33 (5), 871–886.
Friston, K.J., Fletcher, P., Josephs, O., Homes, A., Rugg, M.D., Turner, R., 1998. Event related fMRI: characterizing differential responses. Neuroimage 7 (1), 30–40.
Gagné, C.L., Spalding, T.L., 2004. Effect of relation availability on the interpretation and access of familiar noun–noun compounds. Brain Lang. 90, 478–486.
Gagné, C.L., Spalding, T.L., 2009. Constituent integration during the processing of compound words: does it involve the use of relational structures? J. Mem. Lang. 60, 20–35.
Gentner, D., 1983. Structure-mapping: a theoretical framework for analogy. Cognit. Sci. 7 (2), 155–170.
Gibbs, R.W., 1994. The Poetics of Mind: Figurative Thought, Language, and Understanding. Cambridge University Press, Cambridge, UK.
Giora, R., 1997. Understanding figurative and literal language: the graded salience hypothesis. Cogn. Linguist. 8, 183–206.
Giora, R., 1999. On the priority of salient meanings: studies of literal and figurative language. J. Pragmat. 31, 919–929.
Giora, R., 2002. Literal vs. figurative language: different or equal? J. Pragmat. 34, 487–506.
Giora, R., 2003. On Our Mind: Salience, Context and Figurative Language. Oxford University Press, New York.
Glucksberg, S., Keysar, B., 1990. Understanding metaphorical comparisons: beyond similarity. Psychol. Rev. 97 (1), 3–18.
Grady, J.E., 1997. Foundations of meaning: primary metaphors and primary scenes. PhD Dissertation, University of California, Berkeley.
Graves, W.W., Binder, J.R., Desai, R.H., Conant, L.L., Seidenberg, M.S., 2010. Neural correlates of implicit and explicit combinatorial semantic processing. Neuroimage 53, 638–646.
Grice, H.P., 1975. Logic and conversation. In: Cole, P., Morgan, J. (Eds.), Syntax and Semantics 3: Speech Acts. Academic Press, New York, pp. 41–58.
Grindrod, C.M., Baum, S.R., 2003. Sensitivity to local sentence context information in lexical ambiguity resolution: Evidence from left- and right-hemisphere-damaged individuals. Brain Lang. 85, 503–523.
Hagoort, P., 2005. On Broca, brain, and binding: a new framework. Trends Cogn. Sci. 9 (9), 416–423.
Hillyard, S.A., Kutas, M., 1983. Electrophysiology of cognitive processing. Annu. Rev. Psychol. 34, 33–61.
Hofmann, M.J., Kuchinke, L., Biemann, C., Tamm, S., Jacobs, A.M., 2011. Remembering words in context as predicted by an associative read-out model. Front. Psychol. 2, 252.
Inhoff, A.W., Radach, R., Heller, D., 2000. Complex compounds in German: interword spaces can facilitate segmentation but hinder assignment of meaning. J. Mem. Lang. 42, 23–50.
Jackendoff, R., Aaron, D., 1991. Review article. More than cool reason: A field guide to poetic metaphor by George Lakoff and Mark Turner. Language 67 (2), 320–338.
Ji, H., 2008. The influence of morphological complexity on word processing. Doctoral Dissertation, University of Alberta.
Juhász, B.J., Rayner, K., 2003. Investigating the effects of a set of intercorrelated variables on eye fixation durations in reading. J. Exp. Psychol. Learn. Mem. Cogn. 29 (6), 1312–1318.
Juhász, B.J., Starr, M.S., Inhoff, A.W., Placke, L., 2003. The effects of morphology on the processing of compound words: evidence from naming, lexical decisions and eye fixations. Br. J. Psychol. 94, 223–244.
Juhász, B.J., Inhoff, A.W., Rayner, K., 2005. The role of interword spaces in the processing of English compound words. Lang. Cognit. Process. 20, 291–316.
Jung-Beeman, M., 2005. Bilateral brain processes for comprehending natural language. Trends Cogn. Sci. 9 (11), 512–518.
Koester, D., Schiller, N.O., 2011. The functional neuroanatomy of morphology in language production. Neuroimage 55, 732–741.
Koester, D., Gunter, T.C., Wagner, S., Friederici, A.D., 2004. Morphosyntax, prosody, and linking elements: the auditory processing of German nominal compounds. J. Cogn. Neurosci. 16, 1647–1668.
Kövecses, Z., 2005. A metafora (The Metaphor). Typotex, Budapest.
Kutas, M., Federmeier, K.D., 2000. Electrophysiology reveals semantic memory use in language comprehension. Trends Cogn. Sci. 4, 463–470.
Lai, V.T., Curran, T., Menn, L., 2009. Comprehending conventional and novel metaphors: an ERP study. Brain Res. 1284, 145–155.
Lakoff, G., 1987. Women, Fire, and Dangerous Things. University of Chicago Press, Chicago.
Lakoff, G., Johnson, M., 1980a. Metaphors We Live By. University of Chicago Press, Chicago.
Lakoff, G., Johnson, M., 1980b. The metaphorical structure of the human conceptual system. Cognit. Sci. 4, 195–208.
Lee, S.S., Dapretto, M., 2006. Metaphorical vs. literal word meanings: fMRI evidence against a selective role of the right hemisphere. NeuroImage 29, 536–544.
Liakakis, G., Nickel, J., Seitz, R.J., 2011. Diversity of the inferior frontal gyrus — a meta-analysis of neuroimaging studies. Behav. Brain Res. 225 (1), 341–347.
Libben, G., Gibson, M., Yoon, Y.B., Sandra, D., 2003. Compound fracture: the role of semantic transparency and morphological headedness. Brain Lang. 84, 50–64.
Marslen-Wilson, W.D., Tyler, L.K., 2007. Morphology, language and the brain: the decompositional substrate for language comprehension. Phil. Trans. R. Soc. B 362, 823–836.
Mashal, N., Faust, M., 2010. The effects of metaphoricity and presentation style on brain activation during text comprehension. Metaphor. Symb. 25, 19–33.
Mashal, N., Faust, M., Hendler, T., 2005. The role of the right hemisphere in processing nonsalient metaphorical meanings: application of principal components analysis to fMRI data. Neuropsychologia 43, 2084–2100.
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 Lang. 100, 115–126.
Mashal, N., Faust, M., Hendler, T., Jung-Beeman, M., 2008. Hemispheric differences in processing the literal interpretation of idioms: converging evidence from behavioral and fMRI studies. Cortex 44, 848–860.
Mashal, N., Faust, M., Hendler, T., Jung-Beeman, M., 2009. An fMRI study of processing novel metaphoric sentences. Laterality 14, 30–54.
McGlone, M.S., 2007. What is the explanatory value of a conceptual metaphor? Lang. Commun. 27, 109–126.
Miller, G.A., 1979. Images and models: similes and metaphors. In: Ortony, A. (Ed.), Metaphor and Thought. Cambridge University Press, Cambridge, England, pp. 202–250.
Müller, C., 2008. Metaphors. Dead and Alive, Sleeping and Waking. A Dynamic View. Chicago University Press, Chicago.
Murphy, G.L., 1996. On metaphoric representation. Cognition 60, 173–204.
Murphy, G.L., 1997. Reasons to doubt the present evidence for metaphoric representation. Cognition 62, 99–108.
Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh Inventory. Neuropsychologia 9, 97–113.
Ortony, A., 1979. Beyond literal similarity. Psychol. Rev. 86, 161–180.
Owen, A.M., McMillan, K.M., Laird, A.R., Bullmore, E., 2005. N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. Hum. Brain Mapp. 25, 46–59.
Pobric, G., Mashal, N., Faust, M., Lavidor, M., 2008. The role of the right cerebral hemisphere in processing novel metaphoric expressions: a transcranial magnetic stimulation study. J. Cogn. Neurosci. 20, 170–181.
Quasthoff, U., Richter, M., Biemann, C., 2006. Corpus portal for search in monolingual corpora. Proceedings of LREC-06. Genoa, Italy.
Raffray, C.N., Pickering, M.J., Branigan, H.P., 2007. Priming the interpretation of noun–noun combinations. J. Mem. Lang. 57, 380–395.
Rapp, A.M., Leube, D.T., Erb, M., Grodd, W., Kircher, T.T.J., 2004. Neural correlates of metaphor processing. Cogn. Brain Res. 20, 395–402.
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.
Richards, I.A., 1936. The Philosophy of Rhetoric. Oxford University Press, London.
Schmidt, G.L., Seger, C.A., 2009. Neural correlates of metaphor processing: the roles of figurativeness, familiarity and difficulty. Brain Cogn. 71 (3), 375–386.
Schmidt, G.L., DeBuse, C., Seger, C.A., 2007. Right hemisphere metaphor processing? Characterizing the lateralization of semantic processes. Brain Lang. 100, 127–141.
Searle, J., 1979. Expression and Meaning. Cambridge University Press, Cambridge, England.
Shibata, M., Abe, J-i., Terao, A., Miyamoto, T., 2007. Neural mechanisms involved in the comprehension of metaphoric and literal sentences: An fMRI study. Brain Res. 1166, 92–102.

Sotillo, M., Carretié, L., Hinojosa, J.A., Manuel, M., Mercado, F., López- Martín, S., Albert, J., 2005. Neural activity associated with metaphor processing: spatial analysis. Neurosci. Lett. 373, 5–9.
Spencer, A., 1991. Morphological Theory: an Introduction to Word Structure in Generative Grammar. Blackwell Publishers, Oxford, UK.
Steen, G., 2004. Can discourse properties of metaphor affect metaphor recognition? J. Pragmat. 36, 1295–1313.
Stringaris, A.K., Medford, N.C., Giora, R., Giampietro, V.C., Brammer, M.J., David, A.S., 2006. How metaphors influence semantic relatedness judgments: the role of the right frontal cortex. Neuroimage 33, 784–793.
Stringaris, A.K., Medford, N.C., Giampietro, V.C., Brammer, M.J., David, A.S., 2007. Deriving meaning: distinct neural mechanisms for metaphoric, literal, and nonmeaningful sentences. Brain Lang. 100, 150–162.
Talairach, J., Tournoux, P., 1988. Co-planar Stereotaxic Atlas of the Human Brain. Thieme, New York.
Tversky, A., 1977. Features of similarity. Psychol. Rev. 85, 327–352.
Van Lancker, D., 1997. Rags to riches: our increasing appreciation of cognitive and communicative abilities of the human right cerebral hemisphere. Brain Lang. 57, 1–11.
Van Lancker Sidtis, D., 2006. Does functional neuroimaging solve the questions of neurolinguistics? Brain Lang. 98, 276–290.
Vigneau, M., Beaucousin, V., Hervé, P.-Y., Jobard, G., Petit, L., Crivello, F., Mellet, E., Zago, L., Mazoyer, B., Tzourio-Mazoyer, N., 2011. What is right-hemisphere contribution to phonological, lexico-semantic, and sentence processing? Insights from a meta-analysis. NeuroImage 54, 577–593.
Võ, M.L.-H., Jacobs, A.M., Conrad, M., 2006. Cross-validating the Berlin Affective Word List. Behav. Res. Methods 38, 606–609.
Weylman, S.T., Brownell, H.H., Roman, M., Gardner, H., 1989. Appreciation of indirect requests by left and right damaged patients. The effects of verbal context and conventionality of wording. Brain Lang. 36, 580–591.
Winner, E., Gardner, H., 1977. The processing of metaphor in brain damaged patients. Brain 100, 717–729.