# of metonymic and metaphoric
relations in adjective–noun pairs

Hanna Weiland-Breckle
University of Cologne
h.weiland-breckle@uni-koeln.de

Petra B. Schumacher
University of Cologne
petra.schumacher@uni-koeln.de

Abstract: Theories on metaphor and metonymy make different claims about the nature of the underlying processes in the computation of these two types of language use, i.e., whether they differ or not. Experimental investigations of metonymy and metaphor have generally not compared these two phenomena in a straightforward manner among others due to structural variability. To overcome this shortcoming, we conducted a study in German that used adjective–noun combinations to contrast metaphor and metonymy directly in an ERP-study during reading for comprehension. By combining three different nouns with one adjective in predicative position we construed adjective–noun pairs with literal (*the baby was lively*), metonymic (*the eyes were lively*) or metaphoric (*the speed was lively*) relations. The data revealed a more pronounced N400 for the metaphoric relations in comparison to the literal controls. We argue that the enhanced cost for metaphors reflects the activation process of two unrelated domains via mapping or extended predication. The metonymic adjective–noun pairs only showed a small trend to differ from the other two conditions. This might indicate that metonymies require mapping processes or shifts only within a single domain or domain matrix. Moreover, in contrast to previous studies, we did not find a Late Positivity. We explain this result with regard to different discourse representational consequences arising during combinatorial processing.

Keywords: metonymy; metaphor; adjectives; N400; experimental pragmatics

1. Introduction

Linguistic expressions that require the language user to go beyond a pure compositional computation of the literal meaning of the single sentence components are frequently used in our daily communication. Among those are metaphoric and metonymic uses, which can be considered to require some kind of meaning extension or adjustment during composition. While metaphor (e.g., *life is a journey*) has been characterized as analogy or cross-domain mapping, metonymy (e.g., the reading of "Proust" in *the girl*

read Proust) is considered as contiguity, within-domain mapping or underspecification. A crucial question in this regard is whether these two phenomena engage different mechanisms or whether they rely on the same underlying mechanism, possibly along a continuum.

Support for the discrete mechanisms view comes from language change. Diachronic approaches argue that semantic change involves metonymic and metaphoric operations, which are described as sense transfer and sense transmission or inference processes at the pragmatic level, and differ in terms of analogy for metaphor and reanalysis for metonymy (cf., e.g., Geeraerts 1997; Hopper & Traugott 2003; Keller & Kirschbaum 2003). Similar suggestions have been put forth by cognitive approaches (e.g., Croft 1993; 2002; Lakoff 1987; Lakoff & Johnson 1980; Lakoff & Turner 1989), which argue for different underlying processes for metaphors and metonymies in terms of cross-domain and within-domain mapping. Accordingly, the conceptual proximity represents the distinctive feature. In contrast, the single mechanism view argues that metaphor and metonymy are subject to the same inferential mechanism or selection process (cf., e.g., Frisson & Pickering 2001; Sperber & Wilson 1985; 2008). Although there has been an ongoing debate about the similarity or diversity of the underlying interpretation processes, studies on metonymy and metaphor have generally not compared the processing of these two phenomena in a way that allows for a direct comparison. The reason for this is that most metaphoric and metonymic uses differ enormously in their structural makeup. To overcome this obstacle, we used adjective–noun combinations to contrast metaphor and metonymy directly. By recording event-related potentials (ERPs) to an adjective that required meaning adjustment as a function of its head noun, we obtained minimally differing triplets to compare the underlying processes in a direct manner.

In the following we will briefly summarize the theoretical background for metonymy and metaphor, focusing on possible similarities and differences, and introduce the critical manipulation of adjective–noun combinations. Subsequently, we will review the research on the processing of metaphor and metonymy before presenting the ERP experiment geared towards contrasting the mechanisms underlying incremental metaphor and metonymy processing.

1.1. Theoretical background

1.1.1. Metonymy and metaphor

Metonymy and metaphor are two kinds of non-literal language use that require the extension of the meaning of an utterance beyond its standard lexical semantics (but see extensions that have become conventionalized over time). The different non-literal phenomena and phenomenon types differ concerning the possible range of interpretations and also regarding the required modification for successful interpretation. Since Aristotle the interpretation of metaphors has been defined in terms of transferring properties form a vehicle (source) to a target to which the vehicle is not directly connected (cf. "transfer" in Aristotle's *Rhetoric*; Black 1962; Lakoff 1993). For Lakoff (1993) and other representatives of the cognitive linguistic approach the interpretation of metaphor involves a mapping process between two unrelated domains. Taking the example *love is a journey*, Lakoff described metaphor as mapping from a source domain (here *journey*) to a target domain (*love*). The notion of domain is not clear at all but one approximation is that within conceptual representation domains include related concepts (semantic structures) (Croft 2002). Since every concept that provides a basis for at least one other concept is a domain, a concept could be included in a domain and at the same time serve itself as a domain. For example, the concept 'circle' is included in the domain 'shape' and contains the concepts 'arc' and 'radius', etc. If a concept is involved in a combination of domains, this is called a domain matrix (cf. Croft 2002). Crucially, on the basis of this notion of domain the cognitive linguistic approaches differentiate between metaphor and metonymy. In contrast to the cross-domain mappings involved in metaphor, the mapping processes for metonymy are suggested to take place within a domain or within a domain matrix (cf. Croft 2002; Lakoff & Turner 1989) or as a conceptual shift (cf. Barcelona 2002; Kövecses & Radden 1998; for an overview, see Panther & Thornburg 2003).[1] In metonymy like *the girl read Proust*, *Proust* assumes a domain that includes the concepts 'person' and 'work of Proust'. Therefore when encountering *Proust*, the meaning within the domain 'Proust' has to be shifted from the 'person' to 'work of Proust' (cf. Croft 2002). Lexical semantic accounts have implemented the domain-inherent relations by for

---

[1] In the following we will utilize the idea of mapping within and between domain to describe the difference between metonymy and metaphor, since a discussion of the different approaches to figurative language use goes beyond the scope of this article, but see, e.g., Carston (2010a), Fauconnier & Turner (2002) or Récanati (1995) for different semantic and pragmatic approaches.

instance specifying qualia features that detail form, function and genesis among others of the respective concepts (cf. Pustejovsky 1995).

Other theoretical approaches claim that the underlying processes in metonymy and metaphor are the same and meaning selection is essentially driven by context. The underspecification account (cf. Bierwisch 1983; Copestake & Briscoe 1995; Frisson & Pickering 2001) argues that during language processing, if the processor encounters a word with multiple senses, an underspecified meaning is initially activated. The idea that meaning and sense need to be differentiated lies at the heart of these approaches: While homonyms like *bank* have two separate meanings (and hence two separate pathways during lexical access), words like *novel* only have one underspecified meaning but two or more senses (e.g., *novel* as object, *novel* as content). During the processing of metaphoric or metonymic expressions, the processor incrementally activates the underspecified meaning of each word that is compatible with all possible senses in the current context. Eventually the processor selects or generates a specific sense derived from contextual cues. As long as the required sense of a word within a figurative expression is already stored in the mental lexicon, as is the case for conventionalized metaphors or metonymies, no processing differences are predicted between them or in comparison to a literal utterance (Frisson & Pickering 2001; Schumacher 2018). The relevance theoretic account by Sperber and Wilson (SperberWilson1985,SperberWilson2008) argue for the same underlying mechanism for interpreting metonymy and metaphor as well (but see newer relevance theoretic approaches, e.g., Carston 2010b, for a different argumentation). Based on the assumption that the content of each utterance, whether literal or non-literal, is underdetermined and requires the addressee to infer the meaning on the basis of some hints given by the speaker and guided by the Principle of Relevance, we interpret this account to expect no differences for metonymy and metaphor in terms of processing.

1.1.2. Adjectives and nouns

Adjectives have been proposed to be subject to simple compositional (literal),[2] metonymic and metaphoric mechanisms depending on the head noun they combine with. In general, adjectives can adopt different senses in combination with different nouns, for instance, *fast* in *fast typist* can

---

[2] We use literal meaning in the sense that the combination of adjective and noun does not require operations that go beyond the combination of individual meaning aspects. An adjective like *lively* modifies per default an animate noun like *child*. Combinations like *lively child* are therefore considered literal.

modify the activity carried out by the typist while in *fast road* it describes the property that one can move quickly on the particular road. As a consequence of the multiple contexts that adjectives can occur in, they have been claimed to be polysemous (cf. Bons 2010; Fritz 1995), underspecified (cf. Frisson & Pickering 2001; Frisson et al. 2011) or vague (cf. Kennedy 2007). The different uses of a particular adjective have been challenging for linguistic theories since it is shown that not all meanings are stored in the mental lexicon but are constructed compositionally as a sentence unfolds (cf., e.g., Frisson et al. 2011; Murphy & Andrew 1993). Therefore, compositional mechanisms have been proposed that typically rely on rich lexical representations and adjustment functions during combinatorial processing (see, e.g., Asher 2011; Jackendoff 1996; Partee 2007; Pustejovsky 1995). For example, *fast* normally modifies events but in *fast typist* the noun does not represent an event and therefore a mismatch accrues. To solve the mismatch, *fast* operates on the telic role (i.e., the purpose and function in qualia representation) of *typist*, namely typing (cf. Pustejovsky 1995).

The proposal that the meaning of adjective–noun pairs can be explained in terms of metonymic and metaphoric relations is based on research by Fritz (1995; 1998). He discussed the German adjective *scharf* 'sharp' and showed that the different manners of use are reducible to metonymic and metaphoric patterns (see also Aarts & Calbert 1979; Ortony 1979). Starting from the meaning of *sharp* in (1), where the adjective modifies an object that can have a sharp blade, Fritz argues that in (2), as an example for a metonymic pattern, the functional aspect of the eye is emphasized, i.e., being keen-eyed. Finally, (3) is an example of a metaphoric pattern: the intended meaning may be derived by inferring that words could be like knifes, namely being sharp.

(1) scharfes Messer
     sharp   knife
     'sharp knife'

(2) scharfes Auge
     sharp   eye
     'keen eye'

(3) scharfe Worte
     sharp   words
     'cutting words'

This proposal is in line with diachronic approaches that argue that metonymic and metaphoric patterns should be distinguished and play a role in semantic change (e.g., Geeraerts 1997; Koch 2016).

The proposal of Fritz (1995) was assessed by Bons (2009; 2010), who showed in a corpus study that these patterns could also be found for other adjectives and described basic structural principles that are essential for the variability in the use of polysemous adjectives. Bons (2010) identified different metonymic patterns, such as indicating the outcome of a particular condition instead of the condition (e.g., *sanfte Augen* 'gentle eyes' for *sanfter Mensch* 'gentle person') or the classical metonymic relation of part–whole relations (e.g., *grobe Masche* 'coarse stitch' for *grobes Netz* 'coarse net') as well as discrete metaphoric patterns, such as talking about tasks or problems as concrete objects (e.g., *harte Aufgabe* 'hard task') or talking about non-physical activities as physico-motorical activities (e.g., *harte Kritik* 'hard criticism'). Crucially, in its manner of use, a particular adjective (e.g., *sanft* 'soft'/'gentle', literally used in *gentle person*) can surface as metonymic (*gentle eyes*) and metaphoric patterns (*gentle review*).

The observations made by Fritz and Bons inspired the following experiment. The fact that the same adjective could be involved in metonymic and metaphoric utterances offered the opportunity to directly compare the two types of figurative language use and their underlying mechanism(s).

1.2. Experimental background

In the following we briefly summarize previous findings from language processing for metaphor and metonymy comprehension and then present some studies that compared these two phenomena with behavioral methods directly.

Previous reading time and eye-tracking studies that investigated different types of metonymy, among others producer-for-product (like *reading Proust*) and place-for-event metonymy (like *protesting during Vietnam*), did not find any processing differences compared to literal controls (cf. Frisson & Pickering 1999; 2007; McElree et al. 2006). Crucially, producer-for-product metonyms with an unfamiliar author, e.g., *reading Needham*, were more costly without supporting context, indicating that common knowledge is required to arrive at the intended meaning (McElree et al. 2006). For another type of metonymy, namely mass/count alternations, Frisson and Frazier (2005) reported distinct eye tracking profiles in terms of latency and size in comparison to literal controls and in the comparison of these two metonymy types. For metaphors many behavioral studies

reported that at least familiar metaphoric sentences were as easy to understand as literal utterances (see Glucksberg 2003 for an overview). However, findings have been mixed with some studies reporting enhanced costs for metaphors (see, e.g., Ashby et al. 2017; Brisard et al. 2001; Noveck et al. 2001). For instance, in an eye-tracking study on nominal metaphors (*this X is Y*), Ashby and colleagues (2017) found a higher proportion of regressions as well as longer reading times for metaphors relative to similes. The different processing profiles of metaphors have been shown to be sensitive to factors like familiarity, appropriateness or context (see Giora 1997; 2002 for an overview).

The method applied here is the recording of event-related brain potentials (ERPs). ERPs are an online method that measures neuronal activity time-locked to, e.g., a cognitive event. The output is a multi-dimensional signal that includes information about temporal aspects relative to a stimulus-onset, the magnitude and polarity of the signal's amplitude and the topographical distribution across the scalp. ERPs are relative measurements, wherefore the material should be construed in minimal pairs in such a way that the difference between conditions can be attributed to a specific mechanism. ERP effects are typically reported in regard to their polarity and time course, e.g., a negative deflection with its maximum peak around 400ms is labeled N400 and a positive deflection with a later peak latency between roughly 500–1000ms is referred to as Late Positivity. For metonymic expressions that were presented without supporting context, Schumacher (2014) and Weiland et al. (2014) reported enhanced processing costs for metonymies of the property-for-person type (non-established metonymies like: *the ham sandwich wants to pay*) and the producer-for-product type (*reading Proust*) respectively, reflected in a biphasic N400 – Late Positivity for the former type and an N400 for the latter type. If however presented with licensing context, these metonymies did not elicit an N400 effect in comparison to literal controls (cf., e.g., Schumacher 2011; Weiland-Breckle & Schumacher 2017). The enhanced processing costs for the context-free presentation modality were linked to difficulties in the lexical access phase, e.g., more demanding lexical processes due to the lack of contextual predictability (cf. Schumacher 2013). Regarding the Late Positivity time window, the results can be attributed to different types of metonymy (cf. Schumacher 2013; 2018). One type requires sense creation and meaning accommodation (like, e.g., property-for-person or animal-for-statue as in *wooden turtle*) and elicits a more pronounced Late Positivity (cf. Schumacher 2011; 2013; 2014). The other type includes more conventionalized sense alternations and requires simple sense selection processes

– as reflected by portioning or producer-for-product metonymy – and did not evoke a Late Positivity effect (cf. Schumacher 2013; Weiland et al. 2014). Overall, the observed ERPs for metonymy have been associated with expectation-based processes (N400) and reconceptualization during the updating of mental representations (Late Positivity).

Turning to metaphors, they reliably elicited a more pronounced N400 in comparison to literal controls across languages, which was primarily attributed to semantic processing difficulties (cf., e.g., Bambini et al. 2016; Coulson & Van Petten 2002; 2007; Gold et al. 2010; Lai et al. 2009; Pynte et al. 1996; Weiland et al. 2014). Bambini et al. (2016) further showed that if presented with supporting context, the N400 effect is no longer observable. Therefore, they argue that the N400 reflects contextual expectation-based lexical access, in analogy to what has been proposed for metonymy (Schumacher 2014). Regarding later ERP effects, the results are mixed. Studies by Pynte et al. (1996), Coulson & Van Petten (2007) and Lai et al. (2009) did not report a Late Positivity. In contrast, Coulson and Van Petten (2002), De Grauwe et al. (2010), Weiland et al. (2014) and Bambini et al. (2016) reported a more pronounced Late Positivity for metaphors. The Late Positivity was linked to the cost of recovering or constructing the appropriate meaning. The fact that some studies did not report a Late Positivity could be explained by the selection of different time windows (Coulson & Van Petten 2007) or differences in the word classes (adjectives and verbs) of the target words (Lai et al. 2009).

Although no direct comparison using ERPs has been carried out yet, Weiland et al. (2014) tested metonyms and metaphors in the same experimental environment using cross-modal masked priming. Their findings suggest different degrees of lexical access effort reflected in the N400, since the presentation of a literal prime word influenced the two types of figurative language use differently: while for novel metaphors the literal prime merely reduced the N400 amplitude, for metonymies the N400 effect even disappeared. This may indicate that metaphor and metonymy differ along a continuum with respect to demands during lexical access. In the later time-window, metaphors elicited a more pronounced Late Positivity in comparison to the literal control. Metonymies of the producer-for-product type however did not differ from the control items. In line with the findings sketched out above, this difference in the Late Positivity can be linked to metaphors – but not producer-for-product metonymies – requiring an updating of the mental representation of the discourse referent. Crucially, however, this difference must then be attributed to the degree of conventionality and not the type of figurative language use.

Finally, only a few studies compared metonymy and metaphor directly using behavioral methods, the results of which are mixed. Gibbs (1990) used reading and reaction times to investigate how easily figurative reinstatements can be used as anaphors for literal referents (Mr. Smith) introduced in a short story. He found faster reading times for metaphors (*the butcher*) in comparison to metonymies (*the scalpel*) and therefore argues that metaphors are easier to process than metonymies. Taken the example given, the metonymic but not the metaphoric reinstatement involves an animacy shift that might exert costs above and beyond simple meaning shift, which may well explain why his findings differ from subsequent studies. Rundblad and Annaz (2010) investigated the development of the comprehension of metonymy and metaphor and found that metonymies are acquired faster and processed more accurately consistently from childhood to adulthood. Therefore, they argued for a more basic type of conceptual operation for metonymy in contrast to metaphor. Similarly, better performance on metonymy over metaphor was observed for children with autism (Rundblad & Annaz 2010) and William's Syndrome (Annaz et al. 2009; Van Herwegen et al. 2013). Bambini et al. (2013) found higher interpretation costs for metaphors than for metonymies in a timed sensicality judgment task.

Although the studies provide first insights into the processing of metaphor and metonymy, a direct comparison using minimal pairs and measuring the time-course of processing is still missing. Therefore, we conducted a visual ERP study using adjective–noun combinations with literal, metonymic and metaphoric relations to answer the question whether these kinds of language use are subserved by similar or different operations.

Based on previous research on metaphor and metonymy, the following two predictions are tested: If metaphors and metonymies are based on the same mechanism, as suggested by, e.g., Sperber & Wilson (1985; 2008) or Frisson & Pickering (2001), we expect to find no differences between adjective–noun pairs with metonymic and metaphoric relations – nor a difference between these two types and a literal relation. For the latter account this crucially depends on the conventionalization level of the used metaphors and metonymies – but this factor is not assessed in the present study (cf. also Schumacher 2018 for a convention-based explanation for different processing patterns observed with metonymy). If metaphors and metonymies are processed differently, as suggested by, e.g., Lakoff & Turner (1989) and Croft (1993; 2002), we should find different ERP patterns. In particular, we expect to find processing differences in the N400 as well as in the Late Positivity time-window: For the N400 time-window we

predict a three-way modulation as a function of domain involvement, i.e., the biggest deflection for the metaphoric (between domains transfer) relation and the smallest deflection for the literal combination. Based on the findings by Weiland et al. (2014), we predict the metonymic (within domain shift) condition to pattern in-between. In the Late Positivity window we expect to observe costs from the updating of mental representations and sense creation, and the most pronounced deflection for the metaphoric relation in comparison to the literal condition (cf., e.g., Bambini et al. 2016; Coulson & Van Petten 2002; De Grauwe et al. 2010; Weiland et al. 2014). The metonymic adjective–noun pairs we tested here are less conventionalized than the producer-for-product metonymies tested by Weiland-Breckle & Schumacher (2017) and therefore require sense creation and updating processes as well.

2. Methods and materials

2.1. Participants

Thirty-two monolingual native speakers of German participated in the ERP-study. All were right-handed, reported normal or corrected-to-normal eyesight and no history of neurological dysfunction and were paid for participation. Participants gave written inform consent and the experimental procedures were performed in accordance with the Declarations of Helsinki. Due to an excessive number of occular artefacts in the critical regions, five participants had to be excluded from the analysis. In total 27 subjects (mean age: 22.2 years, range: 20–29, 19 women) entered the statistical analysis of the ERP data.

2.2. Stimuli preparation

We constructed three conditions with adjective–noun pairs that were assumed to use the same adjective differently (literally, metonymically, metaphorically). A series of pretests were carried out to select the most appropriate triplets of adjective–noun pairings in terms of their classification into the three conditions and the predictability of the sentence-final adjective.

2.2.1. Pretest

Classification test for adjective–noun pairs. We first constructed near minimal triplets consisting of an adjective and three different nouns yielding a literal, metonymic or metaphoric relation each according to the definitions in Table 1, which are based on Bons's (2010) characterization of discrete adjective–noun combinations. These definitions pursue the idea that a potential difference between metonymy and metaphor can be explained with the notion 'domain' (cf. Croft 2002). Metonymies are seen as a conceptual shift within one domain, e.g., via mapping (cf. Gibbs 1990), while metaphors engage cross-domain transfer.

Table 1: Definitions for the three possible relations in adjective-noun combinations. a is the adjective; X and Y nouns, $a(X)$ means that X is a.

A–N relation	Definition
Literal	The relation between an adjective and a noun is literal if the meaning of the adjective–noun combination $a(X)$ is consistent with the meaning it has in the language system. A frequent use (high convention) of a pair does not automatically imply a literal meaning of the relation.
Metonymic	The relation between an adjective and a noun is metonymic if the adjective–noun combination $a(Y)$ is used for the adjective–noun pair $a(X)$; X and Y are both part of the same domain, that means that there is a strong conceptual relation between them. $a(Y)$ for $a(X)$; if X and Y are part of the same domain.
Metaphoric	The relation between an adjective and a noun is metaphoric if the meaning of the adjective a is transferred from X to Y, and X and Y do not belong to the same domain, that means that there is no conceptual relation between them. Adjective a should not be originally applied to the domain of Y .$a(X) \rightarrow a(Y)$; if X and Y are of different domains.

Based on these definitions, we created 83 triplets by combining a single adjective (e.g., *lively*) with three different nouns (e.g., *baby, eyes, speed*) to arrive at a literal, a metonymic and a metaphoric relation. These items were distributed across a questionnaire in a pseudorandomized order. Then we asked eight native speakers of German (mean age: 28.9 years, 6 female) with a linguistic background to classify the adjective–noun pairs on the basis of the definitions above in two sessions of a written questionnaire. They were instructed to read the definitions and to judge the relation of the randomized adjective–noun pairs. For analyses, we calculated the percentage of accordant classification for each adjective–noun combination

and ultimately chose the 34 triplets that had classifications with the highest accordance in all three conditions. For the selected adjective–noun pairs this amounted to 85.3% for the literal relations, 66.9% for the metonymic relations, and 57.7% for the metaphoric relations).

Cloze probability test for adjective–noun combinations. Since cloze probability is known to have an impact on the N400 (Kutas & Hillyard 1984) we elicited this value for the selected 34 triplets. In order to measure processing demands time-locked to the same lexical item (the adjective), stimuli were created in predicative constructions (see Table 2 for an example of the actual stimuli). For the cloze probability test (cf. Bloom & Fischler 1980), we cut the sentences before the adjective and distributed the critical triplets over three lists so that each participant saw each item only in one of the three conditions. We presented each list in two different randomizations. In the instruction, the subjects were asked to read the following sentence beginnings attentively and to complete them afterwards with the adjective that came to their mind first. 73 students, all native speakers of German, participated in this study (mean age: 22.8 years, ranging from 20 to 34, 56 women). To calculate the cloze probability value, we compared the adjectives given by the participants with the actual adjectives from our material. Then we computed the mean value of percentage of accordance for each condition. In all three conditions, this was below 0.5%.

2.2.2. Stimuli

We created the material by using the 34 adjective–noun combinations rated best in the preceding classification test. Each adjective was presented in predicative position and was combined with three different nouns as illustrated in Table 2. This allowed us to measure processing costs at the adjective. A full list of experimental stimuli can be found in Appendix 2.

In the triplet in Table 2, the critical word *lively* represents an adjective that modifies animate entities and is used to refer to a *living being* directly (literal condition) or indirectly via a part–whole relation (metonymic condition) or involves a transfer to another domain (metaphoric condition). In particular, in this example *baby* belongs to the domain of *living being*, which the adjective *lively* is specified to combine with. The combination of *lively* and *baby* hence does not require operations beyond simple combinatory processing. The interpretation of the metonymic condition however requires additional processes. Again, *lively* wants to combine with an animate entity; since *eyes* are not animate but part of the domain matrix *living being* (part–whole relationship), a shift within this domain is available during interpretation. In the metaphoric condition, *speed* belongs to

the domain *physical quantity* and is an abstract concept wherefore it is not enough to shift within a single domain matrix. The processing of *this speed is lively* therefore requires a cross-domain transfer between the domain *physical quantity* and the domain *living being. Lively* in the *living being* domain means something like being active, full of energy, etc. When mapping this to the domain of *speed* one can apply "active" and "full of energy" to the pace and the driving style.

In the metonymic condition we were further limited to the part–whole relation of living beings. This resulted in the repetition of critical nouns (e.g., eyes, gesture, gaze), which crucially were combined with distinct adjectives each.

Table 2: Item sample: the three conditions differ with respect to the first noun which has either a literal, metonymic or metaphoric relation with the following adjective. The critical region to which the ERP was time-locked is marked bold. Segmentation is marked via vertical lines.

Condition	Example stimulus
Literal	*Olivers Baby	war	**munter**,	obwohl	es sich	gerade	von einer schweren Erkältung	erholte.*
	Oliver's baby was lively, although she herself currently from a bad cold was recovering.
	'Oliver's baby was **lively**, although she was currently recovering from a bad cold.'
Metonymic	*Robertos Augen	waren	**munter**,	obwohl	es	gerade mal	6 Uhr	in der Frühe	war.*
	Roberto's eyes were lively, although it just 6 o´clock in the early morning was.
	'Roberto's eyes were **lively**, although it was just 6 o'clock in the early morning.'
Metaphoric	*Fabians Tempo	war	**munter**,	obwohl	er	bei diesem Wetter	vorsichtig	hätte	sein sollen.*
	Fabian's speed was lively, although he in this weather careful have been should.
	'Fabian's speed was **lively**, although he should have been careful in this weather.'

We had two high ranked constraints in material construction: (i) we wanted to measure on identical adjectives across the three conditions, and (ii) we followed the definitions given in Table 1. Within the challenging task of identifying proper experimental triplets, it was not possible to control for

two potentially intervening factors, animacy and frequency. These two factors have been shown to influence the N400 amplitude when measured immediately at the noun (cf., e.g., Weckerly & Kutas 1999 for animacy and Van Petten & Kutas 1990 for frequency). Since there are no data about the impact of noun frequency and animacy on the combinatorial processing on a subsequent adjective, we constructed two types of fillers that varied systematically in animacy and frequency of the noun while maintaining the same adjective (in analogy to our critical stimuli). For the first type we tested adjectives that can be used to refer to animate and inanimate nouns as, for example, color adjectives. They were combined with an animate noun, e.g., *fish*, and an inanimate noun, e.g., *line*. We matched the nouns in terms of number of syllables, word length and frequency of occurrence based on the database of Deutscher Wortschatz, University Leipzig (http://wortschatz.uni-leipzig.de). With the second filler type, we controlled for the frequency of use of particular adjectives with particular nouns. We used COSMAS (Corpus Search, Management and Analysis System) from the IDS Mannheim to find adjective–noun pairs that have frequent use (e.g., *coarse grid*). Thereafter we selected a second noun (e.g., *layout*) with the same number of syllables and letters that does not co-occur with the respective adjective. Examples of both filler types can be found in Appendix 1. For both filler conditions, statistical analyses of the ERPs did not reveal any differences (all Fs < 1). We therefore only focus on the critical contrasts of adjective–noun pairs in subsequent sections.

2.3. Procedure

During the experiment, participants sat in a soundproof booth and had to silently read sentences on a 17-inch computer monitor. The stimuli were presented segmentally (marked by vertical bars in Table 2) in yellow letters (Verdana font, font size 32 pt) against a blue background in the center of the screen. After each sentence the participants had to perform a word detection task. One hundred and two (3×34) critical items and the two types of filler items (2×30 each) were presented in four different pseudo-randomized order. All items were pseudo-randomized in a way that the items of the critical triplets were distributed across different blocks and the order of presentation of critical conditions was balanced across the entire experiment. Each participant saw all critical items but in different orders. In total, the experiment contained 222 items that were separated in six blocks with brief pauses in-between.

Before the experiment, participants gave written consent and read an instruction that described the procedure. Thereafter the subjects performed a training block with eight trials to become familiar with the experimental setting. Each trial started with a 500 ms presentation of a fixation asterisk in the center of the screen. This was followed by a blank screen for 500 ms. Then the segments were presented with an inter-stimulus interval (ISI) of 150 ms. The presentation rate was set to 400 ms for one word, 450 ms for two words, 550 ms for three words and at 600 ms for four-word segments. After a blank screen (500 ms), three question marks (500 ms) signaled the upcoming word detection task. When the probe word was presented, the participants had up to 3000 ms to press the 'yes' or 'no' button on a game pad to indicate whether the word had occurred in the preceding trial. 'Yes' and 'no' responses were evenly distributed across all items and blocks and the assignment of the left or right button to the 'yes' answer was counterbalanced across participants. The probe words never targeted the critical adjectives and were chosen from all positions in the sentences. On average participants answered over 95% of the critical trials correctly and thus showed a high level of attention. A blank screen was presented for 1000 ms between trials.

2.3.1. EEG recording procedure

The EEG was recorded from 26Ag/AgCI scalp electrodes, mounted on the scalp by an elastic cap (*Electro-Cap International*). All channels were sampled at a rate of 500 Hz and amplified by *Brain Vision Brain-Amp* amplifier. All impedances were kept below 4 kΩ. The EEG was re-referenced offline to linked mastoids and referenced online to the left mastoid (ground: AFz). To control for artefacts resulting from eye-movement, three electrodes were placed over, under and at the outer cantus of the participants' right eye and one electrode at the outer cantus of the left eye so that vertical and horizontal ocular movements could be monitored.

2.3.2. Data analysis

The EEG data were bandpass filtered offline (0.3–20 Hz) to avoid slow signal drifts.[3] Automatic (rejection criterion was set to a threshold standard deviation of 40 µV within a 200 ms sliding window for the eye electrodes) and manual rejections were carried out to remove artefacts due

---

[3] We used a bandpass filter instead of baseline correction since we consider it a better method to deal with potential transient signal differences in the baseline region (cf., e.g., Makeig et al. 2002; Wolff et al. 2008).

to eye-movement or amplifier saturation for the critical region. We had to exclude 10.5% of the trials from the statistical analysis due to artefacts and time-outs or incorrect responses to the probe recognition task.

Averaged ERPs were calculated from −100 to 1000 ms per condition, participant and electrode and time-locked to the adjective (in bold in Table 2). Statistical analyses of the ERP data were computed by means of repeated measure analyses of variance (ANOVA) for the mean amplitude value of the ERP data per time window. The ANOVA included the factors RELATION with three levels (literal, metonymic and metaphoric) and ROI (topographical region of interest). The electrodes were grouped by location as follows: Frontal (F7 / F3 / Fz / F4 / F8), fronto-central (FC5 / FC1 / FCz / FC2 / FC6), central (T7 / C3 / Cz / C4 / T8), centro-parietal (CP5 / CP1 / CPz / CP2 / CP6) and parietal (P7 / P3 / Pz / P4 / P8). The analysis was done in a hierarchical manner. We used the Huynh–Feldt procedure (Huynh & Feldt 1970) in order to control for potential type I errors due the violations of sphericity. When computing pair-wise comparisons, we considered an adjusted significance level of $p < .033$ based on the modified Bonferroni procedure by Keppel (1991).

Critical time-windows were determined by performing pre-analyses from 0 to 1000 ms in 50 ms steps. If two or more adjacent windows reached significance, they were combined to a critical time-window for subsequent analysis (cf. Gunter et al. 2000). This yielded a time-window from 450 to 550 ms and no other time-window of interest.

3. Results

The participants performed well in the word detection task and answered correctly in over 95% of the cases in all three conditions, which indicates to us that they were paying attention to the material.

Figure 1 shows the grand-averaged ERPs for the three conditions (literal, metonymic and metaphoric relation). Pre-analysis indicated differences between conditions in the time-window from 450 to 550 ms only. Within this window, the metaphoric condition shows the most negative deflection. A three-way modulation is observable with an increasing negativity for the control over metonymic over metaphoric condition and an anterior maximum.
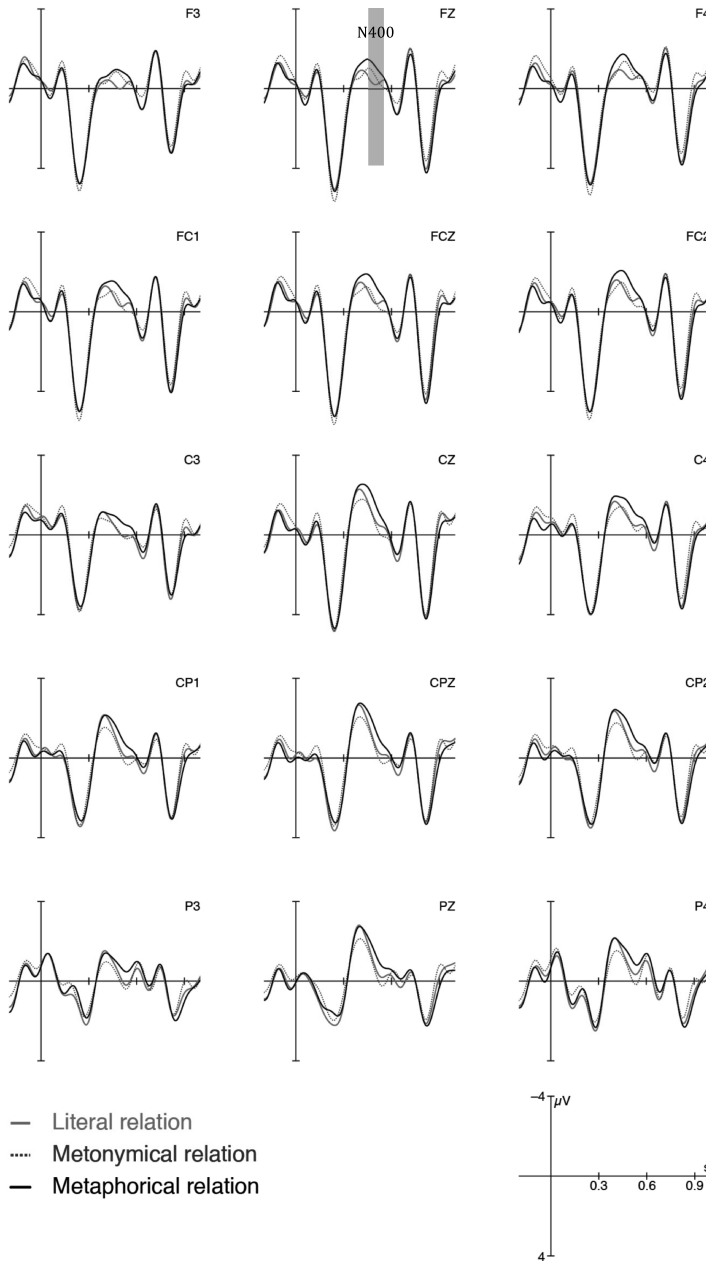
Figure 1: Grand average ERPs for 15 selected electrode sites for the three conditions: literal (gray), metonymic (dashed black) and metaphoric (black) relation. Negativity is plotted up and the vertical bar represents the adjective onset. The critical time-window is marked exemplary with a grey bar at Fz.

In the critical time-window (450–550 ms), ANOVAs revealed a main effect of RELATION ($F(2, 52) = 3.26$, $p < .05$) and an interaction for RELATION × ROI ($F(8, 208) = 2.94$, $p < .01$). The interaction was resolved by ROI and revealed main effects in the frontal ($F(2, 52) = 3.71$, $p < .05$), fronto-central ($F(2, 52) = 3.36$, $p < .05$) and centro-parietal regions ($F(2, 52) = 3.3$, $p < .05$). The other regions yielded no effect of RELATION. Pair-wise comparisons in the three topographical regions that showed reliable effects indicated significant differences between the literal and the metaphoric condition in the frontal ($F(1, 26) = 5.55$, $p < .03$) and in the fronto-central ($F(1, 26) = 5.92$, $p < .03$), but not in the centro-parietal region ($F < 1$). The comparison of the literal with the metonymic condition revealed a marginally significant effect in the frontal region ($F(1, 26) = 4.76$, $p = .038$), for all other regions we found no reliable effect (all Fs < 1). Comparing the metaphoric with the metonymic condition resulted in no reliable effects in the frontal and fronto-central regions (all Fs < 1). However, in the centro-parietal region, the statistical analysis registered a marginally significant difference ($F(1, 26) = 4.62$, $p = .041$).

4. Discussion

The current experiment was adopted to directly compare metaphoric and metonymic utterances by using nouns in combination with adjectives in predicative position. Based on previous ERP-findings on metaphor and metonymy, we expected differences reflected in (i) a three-leveled N400 amplitudinal gradient with the largest deflection for adjective–noun pairs with a metaphoric relation, followed by those with a metonymic relation and the literal control condition and (ii) a two-way Late Positivity pattern for metaphoric and metonymic relations compared to literal adjective–noun combinations. The data revealed a more pronounced negative going wave between 450 ms and 550 ms for the metaphoric relation (*This speed was lively...*) in contrast to the literal condition (*This baby was lively...*). The comparison of the metonymic condition (*These eyes were lively...*) with the literal and the metaphoric condition revealed only marginal differences in both contrasts. Contrary to our expectations, we did not find any Late Positivity effects.

This study revealed two important findings: First, nuanced differences in combinatorial processing of adjective–noun predications can be observed (see the discussion of the N400 differences below). Second, the data shed new light on the nature of the Late Positivity. We discuss these findings in turn.

The processing effort for adjective–noun combinations differs across conditions, at least in predicative adjective–noun comprehension. There was a reliable effect in the 450–550 ms time window between metaphoric and literal conditions and a tendency for the metonymic condition to diverge from the other two conditions. The pronounced negativity for the metaphoric reading may indicate that the computation of two unrelated domains (cf. Dölling 1995; Lakoff & Turner 1989) is more demanding than the combination of adjectives and nouns with matching domains. The metonymic condition, which is considered to require a conceptual shift within a domain (cf. Dölling 1995; Lakoff & Turner 1989), did differ only marginally from the literal control and the metaphoric condition. Although this finding is only a trend, it points towards different processes of mapping within and across domains – see also previous behavioral work from sentence comprehension (Bambini et al. 2013 – but see Gibbs 1990), development (Rundblad & Annaz 2010) and impairment (Annaz et al. 2009; Van Herwegen et al. 2013).

This indicates that the observed N400 effect reflects different degrees of domain activation during combinatorial processing. Based on the absence of effects in the control items, low-level processing features such as frequency or animacy do not have an impact on the observed pattern. The finding for the control items diverges from previous studies that reported differences in the N400 amplitude based on frequency (cf., e.g., Van Petten & Kutas 1990) or animacy (cf., e.g., Weckerly & Kutas 1999). However, this difference may be due to the fact that previous studies measured on words that differed in regard to their animacy or frequency while we measured on an adjective that followed nouns that were either manipulated for animacy or for the frequency of co-occurrence with the tested adjective.

Hence due to the claim that in metaphor interpretation two domains are involved, we interpret these findings as reflecting highest costs for the mapping of these two unrelated domains or extended predication in the metaphoric relation condition and lowest effort for the processing of the literally related adjective–noun pairs. The cost for mapping processes within a domain or domain matrix (metonymically related adjective–noun pairs) only showed a trend to pattern in between. These observations also correspond to the spreading-activation theory, which assumed related concepts to have more links between them (cf. Collins & Loftus 1975). Consequently, related concepts (within a domain matrix) pre-activate each other more and connect faster, while the mutual activation is less strong and the connecting of unrelated concepts (within different unrelated domains) takes longer. Note further that the scalp topography of the observed negativ-

ity showed a frontal maximum, which is untypical for N400 effects (but topographical information is hard to interpret). Research on recognition memory has suggested distinguishing a frontal N400 (FN400) from the N400 but there are reasonable reservations in the literature that the two should be dissociated (cf. Kutas & Federmeier 2011).

Finally, since we did not assess the conventionalization of metonymic and metaphoric meanings in the adjective–noun pairs we cannot entirely exclude the possibility that the difference between these conditions is also affected by discrete levels of lexicalization (cf. Frisson & Pickering 2001). Note however that Schumacher (2013; 2018) argues that differences in conventionalization are reflected in the Late Positivity, such that sense creation involved in the computation of non-conventionalized meanings exerts processing costs (Late Positivity for updating of mental representations) while sense selection in the case of conventionalized interpretations comes at no further cost.

The second important finding was the absence of a Late Positivity effect. A Late Positivity has been reported by many ERP studies on metaphors (cf., e.g., Coulson & Van Petten 2002; Pynte et al. 1996; Weiland et al. 2014) as well as metonymies that involve sense creation (cf. Schumacher 2011; 2013). How can these different findings be explained? A possible answer we would like to entertain is that the processing of non-literal language in adjective–noun predication differs from that in nominal metaphors.[4] Schumacher (2013) recently argued for this view on the basis of experimental findings from different types of meaning adjustment. She proposed that the crucial factor for obtaining a Late Positivity is whether the meaning of the noun (e.g., in *The ham sandwich wanted to pay.*) or the meaning of the adjective (e.g., in *fast road*) is enriched because in the former case the discourse representation structure of the referent has to be modified, but not in the latter case.[5] Accordingly, the

---

[4] An anonymous reviewer suggested an alternative explanation related to task effects. The Late Positivity, as a member of the P300 family, has been observed in response to sensicality ratings or comprehension tasks but has been absent in studies involving probe recognition tasks (which we employed in the current experiment). However, when we turn to studies on compositionality, Coulson and Van Petten (2007) used a comprehension task and did not report a Late Positivity and, more crucially, Schumacher (2013) and Weiland et al. (2014) reported a Late Positivity and used a probe detection task. Task effects, albeit having an effect on the positivity in certain studies, do not seem to be at stake here.

[5] The claim that adjectives behave differently from nouns is also supported by findings from the processing of coercion where an entity must be enriched towards an event reading. In a speed-accuracy tradeoff (SAT) study, McElree et al. (2006) compared

differences between the processing of nominal metaphor or metonymy and adjective–noun predication are associated with operations at the level of discourse representation structure, such that the Late Positivity is linked to discourse updating processes (cf., e.g., Schumacher & Hung 2012; Schumacher 2013). The crucial difference is that the meaning adjustment does not result in a reconceptualization of the nominal head in the current experiment. In a sentence like *his speed was lively*, extra relational information must be retrieved to link the noun's domain 'physical quantity' with the adjective's domain 'living being'. But critically, the representation of the head noun remains unchanged. This can, for instance, be illustrated by subsequent coreference; consider (4), where reconceptualization of the ham sandwich to the ham sandwich eater blocks coreference with the original semantic type. In contrast, in (5) coreference with *speed* is not affected by the domain extension towards living beings.

(4)   After the ham sandwich$_i$ had paid, he$_i$/*it$_i$ left the restaurant.

(5)   Fabian's speed$_i$ was lively. Meg didn't complain about it$_i$/*him$_i$.

Thus if a readjustment and updating of the referential structure has to be performed due to the metaphoric or metonymic reading, extra processing demands are exerted reflected in the Late Positivity.[6] Adjectives by contrast do not act as discourse referents and hence do not force an update of the referential structure. This proposal is supported by research involving combinatorial processing of different word classes (Lai et al. 2009; Schumacher 2013; Schumacher et al. 2018). Lai et al. (2009) measured metaphor processing on verbs and also did not find a Late Positivity. Schumacher (2013) tested among others adjective-induced meaning adjustment of a nominal head (*the wooden dove*); what has to be adjusted in these cases is the representation of the referent (*dove* from animal to statue). Time-locking the ERPs to the noun revealed a more pronounced Late Positivity for *wooden dove* in comparison to *wooden trunk*. Similarly, the processing of so-called privative adjectives (*the fake gun*) calls for the adjustment of

---

coercion processes triggered by event-selecting verbs (e.g., *begin the book*) and event-selecting adjectives (e.g., *survivable mountain*) and found differences in processing speed and accuracy. Although both types, verbs and adjectives, evoke similar coercion effects (control conditions reached higher accuracy than the coercion conditions), coercion triggered by an adjective was processed slower and less accurate.

[6] Late Positivity effects are also reported for reorganization of referential structure triggered by information structural cues (cf. Schumacher & Hung 2012; Schumacher et al. 2015).

the head noun (a fake gun is not a gun in some dimension) and registered a Late Positivity relative to a control condition (Schumacher et al. 2018).

Current models of meaning constitution argue for two stages reflected in the N400 and the Late Positivity respectively: A first processing step which reflects expectation based mechanisms and in which meaning is accessed and a second step in which the mental representation (the discourse model) is updated based on newly encountered information (see Bornkessel-Schlesewsky & Schumacher 2016 for an overview). Our interpretation of the data fits these models by involving a first stage that includes the activation or generation of meaning via mapping processes between domains (N400) and a second stage that may reflect costs for operations on discourse referents. The latter are not required for any of the tested conditions here.

In sum, we presented the first direct comparison of metaphor and metonymy with ERPs by using adjective–noun combinations with adjectives in predicative position. We found that the processing of metaphors is the most demanding which can be associated with mapping processes between two unrelated domains. Metonymies showed a marginally reduced N400 possibly because the mapping operation takes place within a domain. Additionally, we found evidence that these mapping processes do not result in discourse updating (reflected in the Late Positivity for reconceptualization elsewhere).

Acknowledgements

References

Aarts, Jan M. G. and Josepf P. Calbert. 1979. Metaphor and non-metaphor: The semantics of adjective–noun combinations. Tübingen: Niemeyer.

Annaz, Dagmara, Jo Van Herwegen, Michael Thomas, Roza Fishman, Annette Karmiloff-Smith and Gabriella Rundblad. 2009. Comprehension of metaphor and metonymy in children with Williams syndrome. International Journal of Language and Communication Disorders 44. 962–978.

Ashby, Jane, Carlos Roncero, Roberto G. de Almeida and Stephen J. Agauas. 2017. The early processing of metaphors and similes: Evidence from eye movements. The Quarterly Journal of Experimental Psychology 71. 161–168.

Asher, Nicholas. 2011. Lexical meaning in context: A web of words. Cambridge: Cambridge University Press.

Bambini, Valentina, Chiara Bertini, Walter Schaeken, Alessandra Stella and Francesco Di Russo. 2016. Disentangling metaphor from context: An ERP study. Frontiers in Psychology 7. 559.

Bambini, Valentina, Marta Ghio, Andrea Moro and Petra B. Schumacher. 2013. Differentiating among pragmatic uses of words through timed sensicality judgments: Metaphor, metonymy and approximation. Frontiers in Psychology 4. 938.

Barcelona, Antonio. 2002. Clarifying and applying the notions of metaphor and metonymy within cognitive linguistics: An update. In Dirven & Pörings (2002, 207–277).

Bierwisch, Manfred. 1983. Semantische und konzeptuelle Repräsentation lexikalischer Einheiten. In R. Růžička and W. Motsch (eds.) Untersuchungen zur Semantik (Studia Grammatica 22). Berlin: Akademie-Verlag. 61–99.

Black, Max. 1962. Models and metaphors. Ithaca, NY: Cornell University Press.

Bloom, Paul A. and Ira Fischler. 1980. Completion norms for 329 sentence contexts. Memory and Cognition 8. 631–642.

Bons, Iris. 2009. Polysemie und Distribution: Zur Theorie und Methode einer korpusbasierten Semantik deutscher Adjektive. In I. Bons, G. Fritz and T. Gloning (eds.) Linguistische untersuchunge (vol. 1). Gießen: Gießener Elektronische Bibliothek.

Bons, Iris. 2010. Zur Polysemie von Adjektiven. In I. Bons, T. Gloning and D. Kaltwasser (eds.) Fest-Platte für Gerd Fritz. Gießen: Gießener Elektronische Bibliothek.

Bornkessel-Schlesewsky, Ina and Petra B. Schumacher. 2016. Towards a neurobiology of information structure. In C. Féry and S. Ishihara (eds.) The Oxford handbook of information structure. Oxford: Oxford University Press. 581–598.

Brisard, Frank, Steven Frisson and Dominiek Sandra. 2001. Processing unfamiliar metaphors in a self-paced reading task. Metaphor and Symbol 16. 87–108.

Carston, Robyn. 2010a. Lexical pragmatics, ad hoc concepts and metaphor: From a relevance theory perspective. Italian Journal of Linguistics 22. 153–180.

Carston, Robyn. 2010b. XIII-metaphor: Ad hoc concepts, literal meaning and mental images. Proceedings of the Aristotelian Society 110. 295–321.

Collins, Allan M. and Elizabeth F. Loftus. 1975. A spreading-activation theory of semantic processing. Psychological Review 82. 407–428.

Copestake, Ann and Ted Briscoe. 1995. Semi-productive polysemy and sense extension. Journal of Semantics 12. 15–67.

Coulson, Seana and Cyma van Petten. 2002. Conceptual integration and metaphor: An event-related potential study. Memory and Cognition 30. 958–968.

Coulson, Seana and Cyma van Petten. 2007. A special role for the right hemisphere in metaphor comprehension? ERP evidence from hemifield presentation. Brain Research 1146. 128–145.

Croft, William. 1993. The role of domains in the interpretation of metaphors and metonymies. Cognitive Linguistics 4. 335–370.

Croft, William. 2002. The role of domains in the interpretation of metaphors and metonymies. In Dirven & Pörings (2002, 161–206).

De Grauwe, Sophie, Abigail Swain, Phillip J. Holcomb, Tali Ditman and Gina R. Kuperberg. 2010. Electrophysiological insights into the processing of nominal metaphors. Neuropsychologia 48. 1965–1984.

Dirven, R. and R. Pörings (eds.). 2002. Metaphor and metonymy in comparison and contrast. Berlin & New York: Mouton de Gruyter.

Dölling, Johannes. 1995. Ontological domains, semantic sorts and systematic ambiguity. International Journal of Human-Computer Studies 43. 785–807.

Fauconnier, Gilles and Mark Turner. 2002. The way we think: Conceptual blending and the mind's hidden complexities. New York: Basic Books.

Frisson, Steven and Lyn Frazier. 2005. Carving up word meaning: Portioning and grinding. Journal of Memory and Language 53. 277–291.

Frisson, Steven and Martin J. Pickering. 1999. The processing of metonymy: Evidence from eye movements. Journal of Experimental Psychology: Learning, Memory, and Cognition 25. 1366–1383.

Frisson, Steven and Martin J. Pickering. 2001. Figurative language processing in the underspecification model. Metaphor and Symbol 16. 149–171.

Frisson, Steven and Martin J. Pickering. 2007. The processing of familiar and novel senses of a word: Why reading Dickens is easy but reading Needham can be hard. Language and Cognitive Processes 22. 595–613.

Frisson, Steven, Martin J. Pickering and Brian McElree. 2011. The difficult mountain: Enriched composition in adjective–noun phrases. Pychonomic Bulletin and Review 18. 1172–1179.

Fritz, Gerd. 1995. Metonymische Muster und Metaphernfamilien. Bemerkungen zur Struktur und Geschichte der Verwendungsweisen von „scharf". In G. Hindelang, E. Rolf and W. Zillig (eds.) Der Gebrauch der Sprache. Festschrift für Franz Hundsnurscher zum 60. Geburtstag. Münster: Münster Lit. 77–107.

Fritz, Gerd. 1998. Historische Semantik. Stuttgart & Weimar: JB Metzler.

Geeraerts, Dirk. 1997. Diachronic prototype semantics: A contribution to historical lexicology. Oxford: Oxford University Press.

Gibbs, Raymond W. Jr. 1990. Comprehending figurative referential descriptions. Journal of Experimental Psychology: Learning, Memory and Cognition 16. 56–66.

Giora, Rachel. 1997. Understanding figurative and literal language: The graded salience hypothesis. Cognitive Linguistics 8. 183–206.

Giora, Rachel. 2002. Literal vs. figurative language: Different or equal? Journal of Pragmatics 34. 487–506.

Glucksberg, Sam. 2003. The psycholinguistics of metaphor. Trends in Cognitive Sciences 7. 92–96.

Gold, Rinat, Miriam Faust and Abraham Goldstein. 2010. Semantic integration during metaphor comprehension in Asperger syndrome. Brain and Language 113. 124–134.

Gunter, Thomas C., Angela D. Friederici and Herbert Schriefers. 2000. Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. Journal of Cognitive Neuroscience 12. 556–568.

Hopper, Paul J. and Elisabeth Closs Traugott. 2003. Grammaticalization (Second edition). Cambridge: Cambridge University Press.

Huynh, Huynh and Leonard S. Feldt. 1970. Conditions under which mean square ratios in repeated measurements designs have exact F-distributions. Journal of the American Statistical Association 65. 1582–1589.

Jackendoff, Ray. 1996. The architecture of the language faculty (Linguistic Inquiry Monograph 28). Cambridge, MA: MIT Press.

Keller, Rudi and Ilja Kirschbaum. 2003. Bedeutungswandel: Eine Einführung. Berlin: Walter de Gruyter.

Kennedy, Christopher. 2007. Vagueness and grammar: The semantics of relative and absolute gradable adjectives. Linguistics and Philosophy 30. 1–45.

Keppel, Geoffrey. 1991. Design and analysis: A researcher's handbook (3rd edition). Englewood Cliffs, NJ: Prentice-Hall.

Koch, Peter. 2016. Meaning change and semantic shifts. In P. Juvonen and M. Koptjevskaja-Tamm (eds.) The lexical typology of semantic shifts. Berlin & Boston: Walter de Gruyter. 21–66.

Kövecses, Zoltán and Günter Radden. 1998. Metonymy: Developing a cognitive linguistic view. Cognitive Linguistics 9. 37–77.

Kutas, M. and S. A. Hillyard. 1984. Brain potentials during reading reflect word expectancy and semantic association. Nature 307. 161–163.

Kutas, Marta and Kara D. Federmeier. 2011. Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). The Annual Review of Psychology 62. 621–647.

Lai, Vicky Tzuyin, Tim Curran and Lise Menn. 2009. Comprehending conventional and novel metaphors: An ERP study. Brain Research 1284. 145–155.

Lakoff, George. 1987. Women, fire, and dangerous things: What categories reveal about the mind. Chicago: The University of Chicago Press.

Lakoff, George. 1993. The contemporary theory of metaphor. In A. Ortony (ed.) Metaphor and thought. Cambridge: Cambridge University Press. 202–52.

Lakoff, George and Mark Johnson. 1980. Metaphors we live by. Chicago: The University of Chicago Press.

Lakoff, George and Mark Turner. 1989. More than cool reason: A field guide to poetic metaphors. Chicago: University of Chicago Press.

Makeig, Scott, Marissa Westerfield, Tzyy-Ping Jung, Sonia Enghoff, Jeanne Townsend, Eric Courchesne and Terrence J. Sejnowski. 2002. Dynamic brain sources of visual evoked responses. Science 295. 690–694.

McElree, Brian, Steven Frisson and Martin J. Pickering. 2006. Deferred interpretations: Why starting Dickens is taxing but reading Dickens isn't. Cognitive Science 30. 181–192.

Murphy, Gregory L. and Jane. M. Andrew. 1993. The conceptual basis of antonymy and synonymy in adjectives. Journal of Memory and Language 32. 301–319.

Noveck, Ira A., Maryse Bianco and Alain Castry. 2001. The costs and benefits of metaphor. Metaphor and Symbol 16. 109–121.

Ortony, Andrew (ed.). 1979. Metaphor and thought. Cambridge: Cambridge University Press.

Panther, Klaus-Uwe and Linda L. Thornburg (eds.). 2003. Metonymy and pragmatic inferencing. Amsterdam & Philadelphia: John Benjamins.

Partee, Barbara H. 2007. Compositionality and coercion in semantics: The dynamics of adjective meaning. In I. K. G. Bouma and J. Zwarts (eds.) Cognitive foundations of interpretation. Amsterdam: Royal Netherlands Academy of Arts and Sciences. 145–161.

Pustejovsky, James. 1995. The generative lexicon. Cambridge, MA: MIT Press.

Pynte, Joël, Mirelle Besson, Fabrice-Henri Robichon and Jézabel Poli. 1996. The time-course of metaphor comprehension: An event-related potential study. Brain and Language 55. 293–316.

Récanati, François. 1995. The alleged priority of literal interpretation. Cognitive Science 19. 207–232.

Rundblad, Gabriella and Dagmara Annaz. 2010. The atypical development of metaphor and metonymy comprehension in children with autism. Autism 14. 29–46.

Schumacher, Petra B. 2011. The hepatitis called...: Electrophysiological evidence for en-riched composition. In J. Meibauer and M. Steinbach (eds.) Experimental pragmat-ics/semantics. Amsterdam & Philadelphia: John Benjamins. 199–219.

Schumacher, Petra B. 2013. When combinatorial processing results in reconceptualization: Toward a new approach of compositionality. Frontiers in Psychology 4. 677.

Schumacher, Petra B. 2014. Content and context in incremental processing: The "ham sandwich" revisited. Philosophical Studies 168. 151–165.

Schumacher, Petra B. 2018. Metonymy. In C. Cummins and N. Katsos (eds.) The Ox-ford handbook of experimental semantics and pragmatics. Oxford: Oxford University Press.

Schumacher, Petra B., Jana Backhaus and Manuel Dangl. 2015. Backward- and forward-looking potential of anaphors. Frontiers in Psychology 6. 1746.

Schumacher, Petra B. and Yu-Chen Hung. 2012. Positional influences on information pack-aging: Insights from topological fields in German. Journal of Memory and Language 67. 295–310.

Sperber, Dan and Deirdre Wilson. 1985. Loose talk. Proceedings of the Aristotelian Society 86. 153–171.

Sperber, Dan and Deirdre Wilson. 2008. A deflationary account of metaphors. In R. W. J. Gibbs (ed.) The Cambridge handbook of metaphor and thought. Cambridge: Cam-bridge University Press. 84–105.

Van Herwegen, Jo, Dagmara Dimitriou and Gabriella Rundblad. 2013. Development of novel metaphor and metonymy comprehension in typically developing children and Williams syndrome. Research in Developmental Disabilities 34. 1300–1311.

Van Petten, Cyma and Marta Kutas. 1990. Interactions between sentence context and word frequency in event-related brain potentials. Memory and Cognition 18. 380–393.

Weckerly, Jill and Marta Kutas. 1999. An electrophysiological analysis of animacy effects in the processing of object relative sentences. Psychophysiology 36. 559–570.

Weiland, Hanna, Valentina Bambini and Petra B. Schumacher. 2014. The role of literal meaning in figurative language comprehension: Evidence from masked priming ERP. Frontiers in Human Neuroscience 8. 583.

Weiland-Breckle, Hanna and Petra B. Schumacher. 2017. Artist-for-work metonymy. Type clash or underspecification? The Mental Lexicon 12. 219–233.

Wolff, Susann, Matthias Schlesewsky, Masako Hirotani and Ina Bornkessel-Schlesewsky. 2008. The neural mechanisms of word order processing revisited: Electrophysiological evidence from Japanese. Brain and Language 107. 133–157.

Appendix 1: Filler item sample

For filler type animacy the two conditions differ with respect to the animacy of the first noun (animate vs. inanimate). For filler type frequency the two conditions differ regarding the frequency of the combination of noun and adjective (frequent combination vs. new combination). The critical region to which the ERP was time-locked is marked bold.

Filler: Animacy	Example stimulus
animate	*Der Fisch war **rot**, weshalb er wunderbar in Ulrikes Aquarium passte.*
	the fish was red, which is why it wonderfully in Ulrike's aquarium fitted
	'The fish was red, which is why it fitted wonderfully in Ulrike's aquarium.'
inanimate	*Der Strich war **rot**, weshalb man ihn nachts schlecht sehen konnte.*
	the line was red, which is why you it at night poorly see could
	'The line was red, which is why it could be seen poorly at night.'

Filler: Frequency	Example stimulus
frequent	*Das Raster war **grob**, damit nicht zu viele Proben aussortiert werden mussten.*
	the grid was coarse, so that not too many samples eliminated had to be
	'The grid was coarse, so that not too many samples had to be eliminated.'
new	*Das Layout war **grob**, damit der Kunde noch eigene Ideen einbringen konnte.*
	the layout was rough, that the client still his own ideas bring forward could
	'The layout was rough, so that the client could still bring forward his own ideas.'

Appendix 2: List of critical stimuli

Item	Cond	Material
1	Lit	Lenas Kaninchen war ängstlich, als es von ihr in einen Transportkäfig gesetzt wurde.
1	Meto	Sarahs Blick war ängstlich, als sie von ihrem Englischlehrer an die Tafel gerufen wurde.
1	Meta	Philipps Spielzeug war ängstlich, als er den Ball zum Torwart zurückschoß.
2	Lit	Dieser Mönch war barmherzig, weil er den Armen und Bedürftigen geholfen hatte.
2	Meto	Diese Geste war barmherzig, weil die Kinder ohne sie verhungert wären.
2	Meta	Diese Entscheidung war barmherzig, weil so die Flüchtlinge vor dem Ertrinken gerettet wurden.

Item	Cond	Material
3	Lit	Diese Sammler waren begierig auf das Katalogisieren der jüngst erworbenen Schätze.
3	Meto	Diese Augen waren begierig auf die neusten Modefotografien in der Vogue.
3	Meta	Dieses Lesen war begierig auf den Ausgang des spannenden Familiendramas.
4	Lit	Dieser Junge war dreist, weshalb er die Autos vor aller Augen zerkratzte.
4	Meto	Dieses Lächeln war dreist, weshalb dem Erzieher keine angemessene Reaktion dafür einfiel.
4	Meta	Dieses Wetter war dreist, weshalb Klara den ganzen Tag schlechte Laune hatte.
5	Lit	Simones Freund war dumm, auch wenn er angeblich ein abgeschlossenes Studium hatte.
5	Meto	Sylvias Blick war dumm, auch wenn sie versucht hatte ihre Unwissenheit zu verstecken.
5	Meta	Sybilles Situation war dumm, auch wenn ihr alle Freunde zu helfen versucht hatten.
6	Lit	Monikas Helfer waren ehrlich, wenn man der Vermittlungsagentur glaubt.
6	Meto	Mareikes Gesten waren ehrlich, wenn man ihr als Spenderin glaubt.
6	Meta	Luisas Produkte waren ehrlich, wenn man der Werbekampagne Glauben schenkt.
7	Lit	Stefans Tante war elegant, auch wenn sie keine gehobene Herkunft hatte.
7	Meto	Alberts Geste war elegant, auch wenn man sie einem Bauarbeiter nie zugetraut hätte.
7	Meta	Eugens Argumentation war elegant, auch wenn er das Streitgespräch nicht für sich entschied.
8	Lit	Peters Lehrer war energisch, wenn es um die Bestrafung eines Täuschungsversuches ging.
8	Meto	Haralds Blick war energisch, wenn er mit seiner Frau gemeinsam einen Tango tanzte.
8	Meta	Heikos Bericht war energisch, wenn es um seine gefährlichen Klettertouren ging.
9	Lit	Jener Bursche war feige, während die Anderen die Mutprobe meisterten.
9	Meto	Jene Geste war feige, während der restliche Kampf relativ fair verlief.
9	Meta	Jener Angriff war feige, während des Waffenstillstandes ausgeführt worden.
10	Lit	Diese Handwerker waren geschickt, weil sie Manuels Bad geschmackvoll renovierten.
10	Meto	Diese Hände waren geschickt, weil sie eine wunderbare Stickerei anfertigten.
10	Meta	Diese Verhandlungen waren geschickt, weil sie das gewünschte Ergebnis brachten.
11	Lit	Karls Frau war glücklich, weil er ihr einen roten Sportwagen gekauft hatte.
11	Meto	Mikes Gesicht war glücklich, weil er zum dritten Mal Vater geworden war.
11	Meta	Ulfs Umstände waren glücklich, weil seine Firma endlich schwarze Zahlen schrieb.
12	Lit	Der Boxer war grimmig, weil er nach Punkten gegen seinen Erzfeind verloren hatte.
12	Meto	Das Gesicht war grimmig, weil alle Karten bereits ausverkauft waren.
12	Meta	Der Humor war grimmig, weil der alte Mann ins Altersheim geschickt worden war.
13	Lit	Dieser Mensch war höflich, obwohl ihn sein Nachbar aufs Schlimmste beleidigt hatte.
13	Meto	Diese Stimme war höflich, obwohl sich der Anrufer lautstark und unangemessen beschwert hatte.
13	Meta	Dieser Brief war höflich, obwohl sein Schreiber bereits die vierte Anfrage versendet hatte.
14	Lit	Jener Greis war irre, wenn man dem Gerede im Dorf Glauben schenkt.
14	Meto	Jener Blick war irre, wenn man der Modelagenturchefin glaubt.
14	Meta	Jenes Vorhaben war irre, wenn man der Beurteilung von Michaels Eltern glaubt.

Item	Cond	Material
15	Lit	Jener Student war jugendlich, obwohl er definitiv zu den Ältesten seines Studiengangs gehörte.
15	Meto	Jenes Gesicht war jugendlich, obwohl es oft zu intensiv der Sonne ausgesetzt worden war.
15	Meta	Jene Ansichten waren jugendlich, obwohl sie von Johannas Oma vertreten worden waren.
16	Lit	Sophies Junge war kindlich für einen Schüler der siebten Klasse.
16	Meto	Marias Blick war kindlich für eine erwachsene Frau, die bereits zwei Kinder hatte.
16	Meta	Sinas Eifer war kindlich für eine Jugendliche, die mitten in der Pubertät steckte.
17	Lit	Melanies Fohlen war lebhaft, seit es jeden Tag mit seiner Mutter auf die Weide durfte.
17	Meto	Miriams Bick war lebhaft, seit es in der Pause für alle Mitarbeiter einen Kaffee gab.
17	Meta	Susannes Argumentation war lebhaft, seit sich die Diskussion um ein persönliches Thema drehte.
18	Lit	Der Welpe war munter, obwohl er mehrere Impfungen bekommen hatte.
18	Meto	Der Blick war munter, obwohl nach der durchgemachten Nacht Schlafmangel herrschte.
18	Meta	Der Alexanderplatz war munter, obwohl es mitten in der Nacht war.
19	Lit	Das Mädchen war sanft, wenn sie half ihren kleinen Bruder zu wickeln.
19	Meto	Der Augenaufschlag war sanft, wenn Lara ihren Angebeteten zu bezirzen versucht hatte.
19	Meta	Das Herz war sanft, wenn man Sabinas Ehemann Glauben schenkt.
20	Lit	Sissis Hase war scheu, obwohl sie ihn jeden Tag durch den Garten trug.
20	Meto	Susis Augen waren scheu, obwohl sie versucht hatte, ihre Schüchternheit zu überspielen.
20	Meta	Sonjas Fragen waren scheu, obwohl sie ihre Lehrerin bereits seit zwei Jahren kannte.
21	Lit	Diese Frau war sinnlich, wie man ohne Zweifel feststellen musste.
21	Meto	Dieser Mund war sinnlich, wie man gut auf den Portraitfotos erkennen konnte.
21	Meta	Diese Atmosphäre war sinnlich, wie die Clubbesucher während der Party festgestellt hatten.
22	Lit	Dieser Junge war stolz, wenn man seiner Erzieherin glaubt.
22	Meto	Dieses Gesicht war stolz, wenn man dem Redakteur der Zeitschrift glaubt.
22	Meta	Diese Leistung war stolz, wenn man bedenkt, wie wenig Trainingseinheiten absolviert wurden.
23	Lit	Das Mädchen war traurig, weil ihre weiße Katze gestern entlaufen war.
23	Meto	Der Blick war traurig, weil gerade bekannt wurde, dass Loriot gestorben war.
23	Meta	Die Topfpflanze war traurig, weil ihre Besitzerin sie seit Tagen nicht gegossen hatte.
24	Lit	Jener Verkäufer war unfreundlich, seit Nele etwas reklamieren wollte.
24	Meto	Jenes Gesicht war unfreundlich, seit bekannt wurde, dass es kein Hitzefrei geben würde.
24	Meta	Jene Atmosphäre war unfreundlich, seit es zum Streit mit dem Gastgeber kam.
25	Lit	Dieser Bursche war kräftig, wenn man sieht, was er alles auf einmal tragen konnte.
25	Meto	Dieser Arm war kräftig, wenn man bedenkt, dass Tim eigentlich keinen Sport machte.
25	Meta	Diese Argumentation war kräftig, wenn man bedenkt, dass Rachel erst sieben Jahre alt war.

Item	Cond	Material
26	Lit	Diese Mädchen waren entzückend, wenn man ihrer Ballettlehrerin glaubt.
26	Meto	Diese Blicke waren entzückend, wenn man dem Modefotografen Glauben schenkt.
26	Meta	Diese Ideen waren entzückend, wenn man dem Chefdesigner der Autofirma glaubt.
27	Lit	Manfreds Richter war streng, weshalb er nicht mit einer Minderung des Strafmaßes rechnen konnte.
27	Meto	Martins Blick war streng, weshalb er von seinen Kollegen für unsympathisch gehalten wurde.
27	Meta	Patricks Vorschriften waren streng, weshalb sie kein Kind auf dem Abenteuer-spielplatz missachtete.
28	Lit	Marions Baby war fröhlich, weshalb es alle Menschen in seiner Umgebung zum Lächeln brachte.
28	Meto	Lydias Augen waren fröhlich, weshalb man sofort gute Laune bekam.
28	Meta	Lucias Blumenstrauß war fröhlich, weshalb er den ersten Platz beim Wettbewerb gewonnen hatte.
29	Lit	Sein Kind war ruhig, nachdem es endlich etwas zum Spielen bekommen hatte.
29	Meto	Sein Blick war ruhig, nachdem er seine Freundin in der Menschenmenge gefunden hatte.
29	Meta	Sein Gewissen war ruhig, nachdem er bei seinem Priester gebeichtet hatte.
30	Lit	Paulas Hamster war lebhaft, weshalb er mehrmals aus seinem Rad gefallen war.
30	Meto	Majas Augen waren lebhaft, weshalb jeder einen wachen Verstand bei ihr vermutete.
30	Meta	Noras Fantasie war lebhaft, weshalb sie abends oft schlecht einschlafen konnte.
31	Lit	Jenes Mädchen war kitzlig, wenn man Melanie glaubt.
31	Meto	Jene Fußsohlen waren kitzlig, wenn man ihrem Besitzer glaubt.
31	Meta	Jene Situation war kitzlig, wenn man dem Leiter des Einsatzkommandos glaubt.
32	Lit	Ihr Kind war wach, obwohl es die Nacht über viel zu wenig geschlafen hatte.
32	Meto	Ihre Augen waren wach, obwohl die Feier bis spät in die Nacht ging.
32	Meta	Ihr Interesse war wach, obwohl es um ein ziemlich langweiliges Thema ging.
33	Lit	Olivers Baby war munter, obwohl es sich gerade von einer schweren Erkältung er-holte.
33	Meto	Robertos Augen waren munter, obwohl es gerade mal 6 Uhr in der Frühe war.
33	Meta	Fabians Tempo war munter, obwohl er bei diesem Wetter vorsichtig hätte sein sollen.
34	Lit	Dieses Kind war traurig, als ihm sein Lieblingsspielzeug kaputt gegangen war.
34	Meto	Diese Augen waren traurig, als Deutschland bei der Weltmeisterschaft ausschied.
34	Meta	Dieser Baum war traurig, als es diesen heißen Sommer ohne Regen gab.