



Speedy Metonymy, Tricky Metaphor, Irrelevant Compositionality: How Nonliteralness Affects Idioms in Reading and Rating

Diana Michl¹

Published online: 25 July 2019

© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

It is widely acknowledged that fixed expressions such as idioms have a processing advantage over non-idiomatic language. While many idioms are metaphoric, metonymic, or even literal, the effect of varying nonliteralness in their processing has not been much researched yet. Theoretical and empirical findings suggest that metonymies are easier to process than metaphors but it is unclear whether this applies to idioms. Two self-paced reading experiments test whether metonymic, metaphoric, or literal idioms have a greater processing advantage over non-idiomatic control sentences, and whether this is caused by varying nonliteralness. Both studies find that metonymic and literal idioms are read significantly faster than controls, while the advantage for metaphoric idioms is only tenuous. Only experiment 2 finds literal idioms to be read fastest of all. As compositionality of the idioms cannot account for these findings, some effect of nonliteralness is suggested, together with idiomaticity and the sentential context.

Keywords Idiom processing · Nonliteralness · Metaphor · Metonymy · Self-paced reading

Introduction

She holds the whip hand, the truth comes to light, he always has the last word, she changes the subject—are these four idiomatic expressions equally easy and fast to process, provided they are equally well-known, understandable and share a similar syntactic structure? Or does the difference in nonliteralness make some of them easier to process than others? Varying nonliteralness (often referred to as figurativeness) has rarely been the focus of idiom processing studies. Yet, idioms have a strong tendency among both laypeople and linguists to be conceptualized as “nonliteral expressions”, as idiom literature as well as simple informal surveys in the form of “what is an idiom/a figure of speech/turn of expression?” can show. They are sometimes even referred to as “frozen” metaphors (see also Michl 2019), although they are by no means all metaphoric. In fact,

✉ Diana Michl
diana.michl@fu-berlin.de

¹ Department Linguistik, Universität Potsdam/Freie Universität Berlin, Potsdam/Berlin, Germany

idioms differ in semantic structure and can be based on, for example, hyperbole, irony, metaphor, metonymy, and even on literal structures. Expressions such as *to change the subject* or *to have one's own life* are often counted among collocations, formulaic sequences, or idiomatic constructions. However, they strongly resemble nonliteral idioms in that they also vary along many properties such as compositionality, conventionality, familiarity, and others (see also Nunberg et al. 1994). The key difference between them and nonliteral idioms such as *to hold the whip hand* or *to come to light* is that they do not have a nonliteral idiomatic meaning, but instead, their literal meaning is their idiomatic meaning. As such, they are particularly interesting in comparison to the classic nonliteral idioms, because it is unclear how varying nonliteralness affects the processing of idiomatic constructions. To examine this effect, we use a broad definition of idioms as “syntactically complex, fixed expressions” (Sailer 2013). This also includes literal idiomatic constructions which we will term “literal idioms” for the sake of readability and consistency.

A number of contradictory models have been proposed to explain how idioms are processed and stored in the lexicon. Compositional approaches assume that idioms do not have entries as single units but that idiomatic meaning has to be composed directly from the meanings of its component words (Gibbs et al. 1989; Hamblin and Gibbs 1999; Nunberg 1978). Others view idioms as lexical and semantic units represented as single units once learned, like large words (Bobrow and Bell 1973; Gibbs 1980; Holsinger 2013). These views either stem from Construction Grammar accounts or are based on the findings that idiomatic sequences are generally processed faster than literal, non-idiomatic controls (Conklin and Schmitt 2008, 2012; Cronk and Schweigert 1992; Sprenger et al. 2006; Swinney and Cutler 1979) and that priming effects seem to be stronger in idioms than in non-idiomatic sentences (Sprenger et al. 2006), at least in highly familiar and conventional (i.e. salient) ones (Laurent et al. 2006). The reason is that idiomatic expressions are either commonly used or can be recognized as familiar once learned while non-idiomatic, purely compositional language mostly produces novel sentences. Given that the nature of idiom processing seems to be very complex and driven by different properties also dependent on task and specific materials in experimental studies, hybrid accounts of storage and processing were developed. One example is the Configuration Hypothesis (Cacciari and Glucksberg 1991; Cacciari and Tabossi 1988). It holds that idioms are represented by their single constituents and as configurations of meaning which arise from the links between their constituents. Idiomatic meaning is comprised of a distributed representation associated with the idiom. An idiom is first interpreted unidiomatically until it is recognized as an idiom, upon which idiomatic interpretation occurs. The Hybrid Model of idiom comprehension (Titone and Connine 1999) also combines aspects of both noncompositional and compositional accounts by characterizing idiomatic expressions both as single units *and* as compositional word sequences, whose meanings are, however, both directly retrieved and literally analyzed during comprehension.

Varying nonliteralness is not considered a determining factor in commonly known idiom processing models. However, underlying metaphors have been found to be partially accessed in idiom comprehension (Gibbs et al. 1997). This implies that a conceptual level seems to be active in the processing of idioms, which in turn indicates that different semantic structures—such as the very common metonymic and metaphoric structures—may determine idiom processing ease as well. As such, the next section will show why differentiation of idioms based on their semantic structure may be useful.

Metaphor and Metonymy

Metaphor and metonymy are assumed to be fundamental, pervasive and immanent aspects of conceptual thinking and everyday language (Barcelona 1997; Glucksberg and Keysar 1990; Radden 2003, 2005). In a metaphor such as *to swim against the current*, what is said (*current* in a body of water) stands for something literally unrelated (a ‘general trend’ or ‘fashion’). In a metonymy such as *to have an eye for detail*, *eye* refers to something that is literally related to the target concept (i.e. ‘talent to acknowledge detail visually’) (Michl 2019). Despite the fact that definitions of metonymy and metaphor vary in scope, general consent is that metaphors operate between at least two distinct semantic concepts whereas metonymies operate within one semantic concept (see also Fauconnier and Turner 1996; de Mendoza Ibáñez and José 2003; Lakoff and Johnson 1980; Spieß et al. 2015). More generally, metonymy relies on a contiguity relationship between what is said and what is meant (Annaz et al. 2009; Bartsch 2002; Croft 1993; Dirven 2002; Feyaerts 2003; Klepousniotou 2002) while metaphor essentially relies on a similarity or analogy relationship (see also Barnden 2007; Bartsch 2002; Bortfeld and McGlone 2001; Coulson and Matlock 2001; Gentner et al. 2001; Ortony 1979, 1993). Given the closer relationship between what is said and what is meant in metonymy compared to metaphor, one might also say that metaphors tend to be more nonliteral than metonymies. An extensive rating study on idiomatic properties (Michl 2019) has revealed that metonymical idioms are indeed perceived as substantially and significantly more literal than metaphoric idioms.

Very few studies have compared the processing of metaphors and metonymies directly. While many studies exist that examine the processing of metaphors, few studies have examined the processing of metonymies (for the following studies discussed, see also Michl 2019). Metonymies are suggested to be more basic to cognition, easier to learn and easier to comprehend than metaphors (Goossens 1995a, b; Taylor 1995). Experimental research corroborates that assumption: 5-year-olds were found to be less error-prone and faster at processing metonymies in both comprehension and production (Annaz, van Herwegen, Thomas, Fishman, Karmiloff-Smith, and Rundblad 2009), and adults showed a processing advantage for metonymies compared to metaphors, as tested in story-telling tasks (Rundblad and Annaz 2010). In ERP studies, metonymic sentences were found to elicit smaller N400 effects than metaphoric sentences, with and without masked priming (Weiland et al. 2014) which indicates a processing advantage for metonymies.¹ Metonymy, in turn, is processed more slowly when compared to literal language stimuli. Adults produced more errors in comprehension questions on metonymic sentences and exhibited slower responses than for literal sentences (Rapp et al. 2011). Metonymic sentences were found to elicit larger N400 responses compared to literal sentences, which indicates higher processing costs (Weiland et al. 2014).

Metaphors seem to be more variable in their processing demands. Highly familiar and conventional metaphors are processed faster than less conventional ones (Blasko and Connine 1993). Semantic priming, however, was not found to facilitate the processing of conventional

¹ It should be noted that all of these studies used metaphors that either had the form “An X is a Y”, “Some X are Y” or sentences that used this metaphoric form, such as *Life can sometimes be bumpy*, which contains the metaphor *life is a road* which again has the form “X is Y” (example taken from Lai et al. 2009). Given that predicates in the form of “be + N” are not the only kind of metaphor, it is theoretically possible that the findings discussed here are not directly transferable to other forms of metaphors as found in idioms such as *to have the cards stacked against oneself*.

metaphors but only that of novel metaphors (Lai et al. 2009; Lai and Curran 2013). This suggests that the meanings of conventional metaphors are rather automatically retrieved than processed metaphorically. Further evidence from measurements of N400 and P600 indicates that metaphors generally require more processing effort than literal sentences (Coulson and Van Petten 2002; Pynte et al. 1996; Weiland et al. 2014). Metaphors also exhibit slower reaction times in comparison to similes (e.g. *life is like an empty beer can* vs. *life is an empty beer can* (Shibata et al. 2012).

In sum, findings suggest that literal language seems to be processed faster than nonliteral language under certain conditions. Metaphor seems to be more difficult to acquire and comprehend than metonymy which in turn seems to be more difficult to acquire and comprehend than literal language. Rather than a dichotomy, this indeed suggests different nonliteral nuances between the literal and nonliteral poles, as also supported by ERP evidence (Coulson and Van Petten 2002), or possibly a continuum of nonliteral to literal language. It may also suggest that metonymies are easier to process and understand due to their higher degree of literalness, compared to metaphors. But can these differences in metaphor and metonymy also be found in the reading of idioms?

Two self-paced reading experiments are conducted to attempt to unite findings on metaphor, metonymy, and idiom processing and to investigate the effect of nonliteralness on reading times. As idioms differ in their semantic structure and semantic structures differ in their nonliteralness (see introductory examples), the research question is: Do literal, metaphoric, or metonymic structures affect the processing of idioms? Literal idioms are an interesting case in comparison to the classic nonliteral idioms, because if nonliteralness affects idiom processing, then literal idioms should be easier to process than both metonymic and metaphoric idioms. Metonymic idioms should then be easier to process than metaphoric idioms, as also predicted by the findings from metonymy and metaphor research. To test this under different conditions, experiment 1 tested sentences with a neutral context that did not bias the subsequent idiom, while experiment 2 tested sentences with a biasing context that prepared readers for the semantic content of the impending idiom.

In addition to nonliteralness, idioms vary along other properties, such as familiarity, comprehensibility, and compositionality. *Familiarity* refers to how well-known an idiom is and is known to influence processing ease in idioms and fixed expressions in general (Schweigert 1986; Cronk and Schweigert 1992; Gibbs 1980; Libben and Titone 2008; Nippold and Taylor 2002; Tabossi et al. 2009; Titone and Connine 1994). *Comprehensibility* is related to familiarity and refers to the ease with which an idiom can be made sense of or be understood (Katz et al. 1988; also referred to as “meaningfulness” by Titone and Connine 1994). *Compositionality* has been construed as the degree to which the meaning of an idiom can be composed from the meaning of its constituents. Decomposability was shown to have a processing advantage over nondecomposability at least in lexical decision tasks (Caillies and Butcher 2007) and meaningfulness judgments (Gibbs et al. 1989, although their findings have been questioned by Tabossi et al. 2008; Libben and Titone 2008). While a correlation between nondecomposability and higher nonliteralness is expected, the two properties are conceptually different and our goal is to account for them separately in testing for nonliteralness while controlling for compositionality.

Experiment 1

Method

To ensure a natural situation of immediate on-line processing, self-paced reading experiments of German sentences containing literal, metonymic, and metaphoric German idioms were conducted. Sentences were broken down into three chunks and presented as a moving-window paradigm (Just et al. 1982). Idioms were presented as complete chunks to be read at once. This decision was based on Schmitt and Underwood (2004) who conducted a self-paced reading study on formulaic sequences in which idioms were not presented as single units but each word appeared separately. Schmitt and Underwood (2004) found no processing advantage for formulaic sequences in their word-by-word presentation and concluded that formulaic sequences should be more adequately presented as single units. This is corroborated by Conklin and Schmitt (2008) who found a processing advantage for idiomatic expressions when presented as single units.

Whether compositionality effects arise, may depend on the experimental task. While self-paced moving-window word-by-word reading tasks do not show a compositionality effect (Libben and Titone 2008), chunk-by-chunk presentation might require different reading and comprehension mechanisms, thus effects might differ. For this reason, a compositionality rating study was conducted. Compositionality was operationalized as a graded scale from extreme decomposability to extreme nondecomposability. To detect effects, it should be most effective to ask the same participants tested in the reading experiment to also rate the compositionality of the test idioms. In this way, individual ratings can be linked directly to the reading times of each participant for each idiom. To avoid skewing reading times by repetition and memory effects, the online reading task preceded the offline rating task.

Predictions

A general processing advantage is expected for idioms. If only idiomaticity (i.e. being an idiom and recognizable as such) matters, literal or nonliteral structure should not be reflected in processing speed; consequently, all idioms should be processed equally fast, regardless of whether they are literal, metonymic, or metaphoric, as long as length, familiarity, and comprehensibility are controlled for. If, however, the differing nonliteralness is processed or at least accessed in the idiom, literal idioms should be easiest to process, hence read the fastest, metaphors should be read the most slowly, while reading times for metonymies should be in-between.

If compositionality indeed affects processing difficulty in self-paced reading, the idioms rated by participants as rather decomposable should be read faster by those participants, as is predicted by the hybrid models. In line with general definitions of compositionality, nonliteralness should correlate with nondecomposability, so more nonliteral idioms should be more nondecomposable than less nonliteral idioms. Thus metaphoric idioms should be most nondecomposable, literal idioms least nondecomposable, and metonymic idioms in-between.

Participants

A total of 42 adult monolingual native speakers (mostly undergraduates, 12 male) of the Universities of Potsdam and Tübingen (Germany) volunteered to participate in the experiment for either course credit or financial reimbursement. All participants had normal or corrected-to-normal vision and were on average 25.1 years old ($sd=8.8$ years).

Material

The experimental material consisted of 86 idioms (22 literal,² 24 metonymic, and 40 metaphoric idioms, see Table 1 for examples) embedded in sentences. All idioms were taken from an idiom corpus containing 320 idioms (please see Michl 2019 on how exactly these idioms were selected), selected from a modern German idiom dictionary (Schemann 2011). The classification into literal, metonymic, and metaphoric was done by the author and by 17 participants, half of which work in language-related fields. Every idiom received 4 votes of classification. Only idioms whose classification received at least 75% consensus were selected for this study.

Nonliteralness

In addition to the aforementioned rating study on the perception of nonliteralness in metonymic and metaphoric idioms (see Michl 2019), the perception of nonliteralness in literal idioms was tested with adult German monolingual native speakers that would not participate in the subsequent self-paced reading tasks. Prestudy 1 compared literal idioms (22 of which used in the subsequent processing experiments) to metaphoric idioms. 26 participants rated 82 idioms (53 literal, 29 metaphoric) on a 4-point Likert scale (4 being “completely literal”, 1 being “not at all literal”). Prestudy 2 compared literal to metonymic idioms. 28 participants rated 80 idioms (the same 53 literal plus 27 metonymic idioms). Results revealed substantial differences by type of idiom: The two participant groups (54 participants in total) rated literal idioms on average 3.4 and 3.0 (“highly” and “fairly literal”), while the metaphoric idioms were rated 1.27 (“completely nonliteral”) on average and the metonymic idioms 2.34 (“somewhat literal”). Please see Table 2 for an overview of descriptive statistics and significance tests. It was thus ensured that the literal idioms were indeed perceived as literal, the metaphoric idioms as nonliteral, and the metonymic idioms as somewhat literal.

Familiarity and Comprehensibility

The idioms in the original corpus were rated by two different groups of participants on familiarity ($n=96$) and comprehensibility ($n=86$). Participants were asked to rate on a 5-point Likert scale how often they encountered an idiom or how comprehensible they found it. Only idioms scoring “highly” to “extremely familiar/easy to understand” (based

² Literalness itself is a fuzzy concept as all language is ambiguous and meaning construction is always subject to a number of different language- and situation-inherent factors. It is partly structured through homonymy and polysemy. In the case at hand, however, this is unproblematic: Literal idioms differ crucially from nonliteral idioms in that they do not have nonliteral meaning and thus lack all additional complexity.

Table 1 Examples of the material of Experiment 1 with translation

Condition	Pre-idiom chunk	Idiom chunk	Wrap up
Metonymic	Seit zwei Stunden <i>For two hours</i>	knurrt Marie bereits der Magen <i>growls Marie already the stomach</i> <i>Marie has been hungry now</i>	und sie will etwas essen. <i>and she wants to eat something</i>
Control	Seit zwei Stunden <i>For two hours</i>	schmerzt Marie bereits der Magen <i>hurts Marie already the stomach</i> <i>Marie's stomach has been hurting</i>	und sie will nichts essen <i>and she wants to eat nothing</i> <i>and she doesn't want to eat anything</i>
Metonymic	Diese Bücher <i>These books</i>	bezahlt Julia aus eigener Tasche <i>pays Julia out of her own bag</i> <i>Julia pays out of her own pocket</i>	und das ist ärgerlich <i>and that is annoying</i>
Control	Diese Bücher <i>These books</i>	schleppt Eva in der eigenen Tasche <i>drags Eva inside her own bag</i>	und das ist ärgerlich <i>and that is annoying</i>
Metaphoric	Mit der Zeit <i>Given time</i>	kommt die Wahrheit doch noch ans Licht <i>comes the truth after all to light</i> <i>the truth comes to light after all</i>	und alle staunen <i>and everyone is astonished</i>
Control	Mit der Zeit <i>Given time</i>	kommen die Tiere doch noch ans Licht <i>come the animals to light after all</i> <i>the animals come to the light after all</i>	und sind neugierig <i>and are curious</i>
Metaphoric	Wieder einmal <i>Once more</i>	beißt Katrin in den sauren Apfel <i>bites Katrin the sour apple</i> <i>Katrin bites the bullet</i>	und stimmt allem zu <i>and agrees to everything</i>
Control	Wieder einmal <i>Once more</i>	wählt Katrin einen sauren Apfel <i>picks Katrin a sour apple</i>	und beißt hinein <i>and takes a bite</i>
Literal	Beim Training <i>At her training</i>	rührt Lisa sich nicht von der Stelle <i>moves Lisa not from the spot</i> <i>Lisa doesn't move away from the spot</i>	und bleibt still stehen <i>and keeps standing still</i>
Control	Beim Training <i>At her training</i>	rennt Lisa eine Weile auf der Stelle <i>runs Lisa a while on the spot</i> <i>Lisa runs on the spot for a bit</i>	und bleibt dann stehen <i>and then comes to a stop</i>
Literal	Beim dem Ausflug <i>At the trip</i>	bringt Stefan die Kinder in Gefahr <i>brings Stefan the children in danger</i> <i>Stefan puts the children in danger</i>	aber alle haben Spaß <i>but everybody has fun</i>
Control	Beim dem Ausflug <i>At the trip</i>	sieht Stefan die Kinder in Gefahr <i>sees Stefan the children in danger</i> <i>Stefan thinks the children are in danger</i>	aber alle haben Spaß <i>but everybody has fun</i>

on median and mean ratings) were selected for the current reading experiments. Matching on familiarity also provides some matching on cloze probability (as indicated by a high correlation of familiarity and cloze probability found by Titone and Connine 1994).

Table 2 Descriptive statistics: nonliteralness ratings of idioms (1 = completely nonliteral, 4 = completely literal). Lines 1–2 show prestudy 1, lines 3–4 show prestudy 2

	Min	Median	Mean	SD	Max
Literal	1	4	3.42	0.78	4
Metaphoric	1	1	1.27	0.61	4
Literal	1	3	3.02	0.99	4
Metonymic	1	2	2.35	1.02	4

Wilcoxon rank sum tests (the nonparametric equivalent to t-tests) show that the difference in rating between literal and metaphoric idioms ($W=45,290$, $p<0.001$) and between literal and metonymic idioms ($W=361,680$, $p<0.001$) is significant. For further analyses, see Appendix 2. It can be concluded that literal idioms are indeed perceived as much more literal than both metonymic and metaphoric idioms, and that metonymic idioms are perceived as more literal than metaphoric idioms

Structure

The idiomatic sentences were subdivided into three chunks and contained the idiom in the middle chunk. Each idiom chunk consisted of its verb, the subject, and the object of the idiom, in that order. In 43% of the sentences, a function word or adverb such as *again*, *now*, or *this time* was inserted between subject and noun. This was done if necessary to achieve equal chunk length. Each idiom chunk was preceded by a neutral pretext of 2 or 3 words to avoid any expectations of the ensuing content, and followed by a third chunk consisting of a second main clause. This last clause was usually introduced by *und* ('and') and served as an addendum to absorb any sentence wrap-up effects (Aaronson and Scarborough 1976; Just and Carpenter 1980; Rayner et al. 1989) and any possible spill-over, although based on the findings by Just et al. (1982), no spill-over was expected for the moving-window condition.

Matching

Each idiomatic sentence was matched with a control sentence that contained as many identical words in the same order as possible (see Appendix 1 for discussion for a theoretical alternative). All chunk pairs and sentences were matched for length and mean frequency of their words together.

Length

Since self-paced reading times are strongly affected by the length of units (Stowe and Kaan 2006), number of characters per chunk was chosen as a measure for length because it is the smallest and most exact measure and adjusting chunk length by it is relatively feasible. Given that words cannot simply be replaced by others in an idiom and the selected idioms differed in length by character, chunk lengths were aligned by the choice of subject and by inserting function words if necessary.

Frequency

Frequency was matched within each chunk pair and each sentence pair. Normalized log-10 frequency measures were obtained for all lemmas and added within each separate chunk. Then, a mean frequency was calculated for each chunk. Measures were obtained from the DLEXDB³ corpus for German language which was chosen for its modernity, size, and well-balanced composition. It comprises 100 million running tokens, is based on the *Kernkorpus des Digitalen Wörterbuchs der deutschen Sprache (DWDS)* which is a reference corpus for the 20th century, and is composed of 28% fictional, 27% journalistic, 23% scientific, and 21% functional literature.

Plausibility

All stimulus sentences were rated for plausibility on a 5-point scale by 3 adult German native speakers who did not participate in the final experiment. All sentences were modified until they received a mean rating of at least 2 (“quite plausible”).

Fillers

In addition to the idiomatic and control sentences, 86 filler sentences were constructed. Their chunk lengths and syntax varied in order to avoid structural priming and keep up participants’ attention and interest. 22 of the fillers contained an idiom at the beginning, the middle, or the end of the sentence.

Distractor Task

To further keep up attention and quality of participants’ performance, 42 sentence comprehension questions were constructed of which half required a ‘yes’ and half a ‘no’ answer. The questions referred to the sentence immediately preceding and focused on various parts and aspects of the sentence to distract participants from the purpose of the experiment and keep their attention steady across all chunks (Keating and Jegerski 2015).

Presentation

Two separate lists were created so that each participant was exposed to each target type but read only one item of each idiom/control sentence pair. Each list contained 210 items: 43 idioms (11 literal, 12 metonymic, 20 metaphoric), 43 control sentences, 86 filler sentences, and 42 yes/no comprehension questions. ‘Idiom type’ and ‘condition’ (control vs. idiom) were treated as within-participants factors. Within the lists, 4 blocks were created. All blocks were counterbalanced to consist of 21 or 22 test items (4–6 literal, metonymic, or metaphoric or their controls) 21 fillers, and 10 or 11 questions. Blocks were matched for overall frequency and length on means and standard deviations. Items were pseudorandomized for each participant such that no two test items were ever presented consecutively and that no two items of the same idiom type were presented before and after a filler, if possible. The blocks were then scrambled which yielded 4 versions of each list and thus

³ available at www.dlexdb.de.

8 versions of the experimental material. These versions were randomly assigned to the participants.

Procedure

After filling in the consent form, participants were placed about 65 cm before the computer screen and instructed to read sentences for comprehension as quickly as possible. The screen indicated a complete sentence of 3 chunks of generic length. After an interval of 2000 ms, the beginning of the sentence appeared. After reading it, participants had to press the space-bar for the second and then third chunk to appear, and wait the 2000 ms lapse until the beginning of the next sentence appeared. Pressing the space-bar caused the preceding chunk to disappear. Comprehension questions appeared randomly and could be answered by key press. Participants were instructed to leave the index finger of their dominant hand on the space bar at all times while reading. A practice session of 6 sentences and 2 questions preceded the experiment. During the practice session, participants could ask questions. The experimental self-paced reading session lasted about 20 min and offered 3 between-blocks breaks whose length participants could determine.

After the self-paced reading task, participants were directed to the compositionality rating. All 86 idioms were presented in individually randomized order on 4 consecutive pages. Compositionality of idioms was explained as a graded property of how much the single words contribute to the overall meaning. Participants read each idiom and its meaning and rated the idiom with respect to the question “how compositional is each idiom?” using a 7-point Likert scale. The rating task lasted 15 min on average.

Apparatus

The self-paced reading experiment was implemented in the DMASTR computer software (DMDX), Version 5.1.3.3 (Forster and Forster 2003) on Windows 7. The material was presented in marine blue 38-point serif font against a light blue background. The compositionality rating study was designed and presented on the non-commercial social studies platform Sosci Survey, Version 2.5.00-I (Leiner 2014).

Analysis

Linear mixed effects regression was chosen as the most appropriate analysis because it allows accounting for individual variances between participants as well as items within a single model (Baayen et al. 2008), such as remaining differences in length and frequency. Intercepts and slopes for each individual and independent variables can be modelled by groups of data (Field et al. 2014). All analyses were conducted in the R environment and language Version 3.3.2 (R Studio Version 1.0.136; R Core Team 2019a, b). Reading times (RT) data points were excluded if responses to respective comprehension questions were false or timed-out, or if they were under 350 ms or over 4000 ms for the relevant chunk. Thus 3% of RT data points had to be excluded. RT were transformed by their natural logarithm to reduce the very RT-typical right-skew of the distribution while maintaining relevant differences in shorter RT and to achieve better model fit making results more robust and reliable. For comparison, however, final analyses were also performed on raw RT.

Table 3 (a) Log-transformed reading times from self-paced reading experiment 1, lmer coefficients and standard error, t-value, confidence intervals, and *p* value; (b) Raw RT in ms from self-paced reading experiment 1, lmer coefficients and standard error, t-value, confidence intervals, and *p* value

	β (SE)	t-value	CI-lower	CI-upper	Pr(> t)
(a)					
Controls (intercept)	7.26 (0.06)	131.31	7.15	7.37	< 0.001***
Literal	− 0.05 (0.02)	− 3.14	− 0.08	− 0.02	0.002**
Metonymic	− 0.06 (0.01)	− 4.32	− 0.09	− 0.04	< 0.001***
Metaphoric	− 0.02 (0.01)	− 1.38	− 0.04	0.01	0.169
Length	0.03 (0.01)	3.94	0.01	0.04	< 0.001***
(b)					
Controls (intercept)	1564.95 (77.84)	20.1	1409.09	1720.81	< 0.001***
Literal	− 93.38 (26.11)	− 3.58	− 144.59	− 42.18	< 0.001***
Metonymic	− 103.05 (25.27)	− 4.08	− 152.59	− 53.5	< 0.001***
Metaphoric	− 31.39 (19.8)	− 1.59	− 70.21	7.43	0.113
Length	47.1 (11.76)	4	24	70.21	< 0.001***

****p* < 0.001***p* < 0.01**p* < 0.05

Log RT were fit as a function of condition as a categorical fixed effect with the levels literal, metonymic, metaphoric idiom or control, and length as a scaled numeric fixed effect and random intercepts for subjects and items (R packages “lme4”, version 1.1-12; Bates et al. 2015, and “lmerTest”, version 2.0-33; Kuznetsova et al. 2016). Chunk frequency was tested and excluded as a fixed effect as it did not reveal any effect. An initial lmer model comparing idiom to control condition revealed a significant processing advantage for idioms ($b = -0.04$, $t = -4.33$, $p < 0.001$). As control sentences can begin to differ from their idiom sentence at different points (see “Matching”), be it in the initial chunk, the verb or later, this was initially entered as a 3-level covariate. As it revealed no effect, it was excluded from the regression. Length had a significant effect and was thus kept in the final model.

To investigate the relationship between idiom type and compositionality, a correlation test of Kendall’s tau b was run. Kendall’s tau b is a rank-based correlation coefficient and thus more robust and more suitable for non-normal data, discrete data and ties than Pearson’s product-moment correlation. A Pearson product-moment correlation test was run to check the relationship between idiom length and compositionality.

Also, the participant effects of age, handedness, sex, geographical origin, and subject of study of the students were entered as fixed effects, but no effects on reading times could be found ($t < 2.0$, $p > 0.1$).

Results

Results showed significant differences between literal idioms and metonymic idioms compared to their control sentences, but no significant difference between metaphoric idioms and controls (see Table 3). In terms of raw reading times, metonymic idioms were read

roughly 103 ms faster than controls which is a significant difference, literal idioms 93 ms, and metaphoric idioms 31 ms faster. The processing advantage for metonymic idioms was significantly larger than for metaphoric idioms as each compared to their controls ($b=0.05$, $t=-2.4$, $p=0.02$). No significant difference was found between literal and metonymic idioms ($b=0.02$, $t=0.74$, $p=0.46$).

Compositionality rating and idiom type showed a moderate, significant correlation ($r_{\tau}=0.5$, $p<0.001$), expressing that the more literal the idiom, the more decomposable it was rated. There was no correlation between length and compositionality ($r=-0.02$, $p=0.38$). Ratings were then fit as a function of idiom type and length and random intercepts for participants and items. The ratings were highly driven by idiom type ($p<0.001$), yet not by idiom length ($p=0.7$). Literal idioms were rated as much more decomposable than metonymic idioms, while metonymic idioms were rated as much more decomposable than metaphoric ones (see Table 6). Compositionality ratings, in turn, seemed to have no effect on reading times, neither alone ($p=0.25$) nor in combination with idiom type ($p\approx 0.3$). Goodness-of-fit tests showed that ratings as a fixed effect neither significantly increased nor decreased model fit.

Discussion

As no effects of compositionality were found, idiom type and length were the only factors that clearly impacted reading times. Metonymic and literal idioms had a clear processing advantage while metaphoric idioms had only a small, nonsignificant processing advantage compared to controls. Given their status as idioms, this is an interesting finding. While low frequency could cause more processing difficulty and longer reading times (Just et al. 1982), out of the rarer words responsible for low overall chunk frequency—the nouns *Durchreise* ('transit'), *Schrittempo* ('walking speed') and *Fahrerflucht* ('hit and run') and the verbs *knurren* ('growl'), *wickeln* ('wrap'), *erraten* ('guess correctly'), *verschlucken* ('swallow up')—only two occurred in metaphoric idioms. At any rate, it is doubtful whether frequency measures of single words actually matter when part of an idiom is processed, because the idiom is likely to be recognized as an entity rather than as sequential single words.

Instead, one reason metaphoric idioms had no significant processing advantage might be that they can be less apt (Blasko and Connine 1993), given that metaphors are based on a relationship of similarity or analogy. Since similarity and analogy form comparably loose, indirect, and somewhat idiosyncratic links between concepts, they could easily be argued to be ill-fitting, far-fetched, unintuitive, and difficult to understand, at least more so than constructions based on a contiguity of meaning, such as metonymies. For metaphoric idioms, this means that the relationship between the literal and the idiomatic meaning is more distant, less straight-forward, and less obvious than in metonymic idioms (see also Keysar and Bly 1999). Their very small processing advantage may be explained by their familiarity and idiomaticity which make them recognizable as idioms at some point, thus it and its metaphoric meaning has likely been encountered before. Any information activated from the literal meaning and irrelevant to the intended metaphoric meaning might have been blocked in the processing, as has been found for novel metaphors (Glucksberg et al. 2001). Given the small effect size, however, this tendency should not be overestimated, and both confidence interval and a 10,000-replication

bootstrap suggest that it is unreliable. This indicates that a different property—likely nonliteralness—overrides idiomaticity and adds processing load.

A potential reason why metonymies showed a strong advantage appears to be that they are often grounded in physical experience (Radden 2003, 2005). Bodily sensations expressing psychological states are seen as metonymically linked (Kövecses and Radden 1998; Turner and Fauconnier 2003), for example, *to keep an eye on sth.*, *sb.'s heart is pounding like mad*, *sb.'s stomach is churning*. Many metonymic idioms are in fact expressions about body parts (see also Dobrovolskij 1995). Such metonymies are called ‘kinegrams’. It might be argued that these metonymies had the greatest processing advantage because they are inherent to a readers’ natural experience. The metonymic idioms in our experiments consisted of 14 kinegrams and 10 others. If kinegrams were the reason for the processing advantage of metonymic idioms as a group, only the kinegrams would be processed faster than controls in a direct comparison. However, in a mixed regression, kinegrams were not read faster than the other metonymies ($b = -0.002$, $t = -0.15$, $p = 0.88$). This indicates that the processing advantage may rather be due to the metonymic structure or the greater literalness than the idioms’ specific semantic content. The significant difference between metonymic idioms and controls versus metaphoric idioms and controls strongly suggests that the processing system does not treat all idioms equally but differentiates according to the literalness of the semantic structure involved. Along with the similarly large processing advantage of literal idioms, this indicates that nonliteralness or metaphoric structure could override idiomaticity in reading comprehension and that higher nonliteralness is more difficult to process, even in well-known expressions.

The limitation of this experiment is that the context preceding the idiom chunk was neutral and did not create any expectations of the semantic content to follow. Since processing ease is highly influenced by contextual information, the following questions arise: Do the observed patterns also exist when contextual information causes the idiomatic content to be expected? Would nonliteralness show the same effects under a biasing condition? To test this effect, a second self-paced reading was conducted to test the processing of idioms with a biasing context.

Experiment 2

Idioms, procedure, task, apparatus and presentation were the same as in experiment 1, as was matching for length, frequency and plausibility of the materials.

Predictions

If expectedness of an idiom’s content has the same facilitating effect on its reading speed regardless of its nonliteralness, then all idioms should be read equally faster than their controls. Alternatively, if nonliteralness still affects idiom processing when the idiomatic semantic content is expected, then the more nonliteral an idiom, the more slowly it should be read. This would mean that especially metaphors should not be read significantly faster than their controls.

Participants

A total of 93 adult monolingual native speakers (35 male) from several German cities and the Universities of Potsdam and Tübingen volunteered to participate in the experiment for course credit or financial reimbursement. All participants had normal or corrected-to-normal vision and were on average 32 years old ($sd=11.8$ years). The number of participants was increased in an attempt to increase reliability of the findings. Experiment 1 and an interim analysis of 46 participants in Experiment 2 showed that compositionality ratings had no effect on reading times. For this reason, the compositionality rating was discontinued for the remainder of Experiment 2.

Material

The experimental material consisted of the same 86 idioms as in study 1. The only difference was that each idiom chunk was preceded by a biasing context that prepared for the semantic content of the idiom in the middle chunk. The third and last chunk consisted of a subordinate clause to avoid any sentence wrap-up effects (see Table 4 for examples).

Analysis and Results

The analyses conducted were the same as in Experiment 1. Finally, 5% of data points were missing due to recording failure of the program, or had to be excluded due to false or missing responses to comprehension questions or RT over 4000 ms and under 350 ms. RT were again log transformed, but final analyses were also performed on raw RT. Analyses with raw RT yielded slightly more extreme effects with lower significance levels (see Table 5 for comparison). Once again, no participant effects of age, handedness, sex, origin, testing location, subject of study or profession could be found ($t < 2.0$, $p > 0.1$). Length had a significant effect.

All types of idioms were read significantly faster than their non-idiomatic control sentences. Metonymic idioms were read 51 ms faster, literal idioms 70 ms, and metaphoric idioms 39 ms faster. However, no significant differences were found between reading times of literal vs. metonymic ($b = -0.02$, $t = -1.16$, $p = 0.25$) and metonymic vs. metaphoric idioms ($b = -0.01$, $t = 0.45$, $p = 0.62$). The final model was also run without 10 high leverage points which did not change results.

Again, compositionality ratings were highly determined by idiom type ($p < 0.001$), and not by length ($p = 0.1$). A Kendall's tau b correlation between idiom type and compositionality rating ($r_t = 0.39$, $p < 0.001$) was found, while no Pearson product-moment correlation was present between rating and length ($r = 0.00$, $p = 0.99$). As expected, literal idioms were rated most decomposable, metaphoric idioms most nondecomposable, while metonymic idioms were in-between (see Table 6). Compositionality ratings neither had an effect on reading times on their own ($p = 0.45$) nor in combination with idiom type ($p \approx 0.2$). Goodness-of-fit tests showed that ratings as a fixed effect neither significantly increase nor decrease model fit.

Table 4 Examples of the material of Experiment 2 with translation

Condition	Pre-idiom chunk	Idiom chunk	Wrap up
Metonymic	Seit Mittags will Marie etwas essen, denn <i>Since noon wants Marie to eat something as</i>	ihr knurrt längst der Magen <i>her growls already the stomach</i>	und es ist spät <i>and it is late</i>
Control	Seit Mittags will Marie nichts essen, denn <i>Since noon wants Marie to eat nothing as</i>	ihr schmerzt noch der Magen <i>her hurts still the stomach</i>	von gestern <i>from yesterday</i>
Metonymic	Diese Bücher sind zu teuer für Julia, doch <i>These books are too expensive for Julia but</i>	sie zahlt sie aus eigener Tasche <i>she pays them out of her own bag</i>	, was ärgerlich ist <i>which is annoying</i>
Control	Diese Bücher sind zu schwer für Julia, doch <i>These books are too heavy for Julia but</i>	sie trägt sie in der eigenen Tasche <i>she carries them inside her own bag</i>	, was anstrengend ist <i>which is strenuous</i>
Metaphoric	Letztlich kennt jemand die Wahrheit und <i>In the end, someone knows the truth and</i>	sie kommt doch ans Licht <i>it comes after all to light</i>	, früher oder später <i>sooner or later</i>
Control	Letztlich sind die Tiere neugierig und <i>In the end, the animals are curious and</i>	sie kommen doch ans Licht <i>they come after all to the light</i>	, früher oder später <i>sooner or later</i>
Metaphoric	Katrin hat keine andere Wahl mehr und <i>Katrin has no other choice anymore and</i>	sie beißt in den sauren Apfel <i>she bites into the sour apple</i>	, wie so häufig <i>, as so often</i>
Control	Katrin hat kein anderes Obst mehr und <i>Katrin has no other fruit anymore and</i>	sie kostet den sauren Apfel <i>she tries the sour apple</i>	, der nicht schmeckt <i>which doesn't taste good</i>
Literal	Beim Training bleibt Lisa still stehen und <i>At her training Lisa keeps standing still and</i>	sie rührt sich nicht von der Stelle <i>she moves not from the spot</i>	, aus Unlust <i>, out of aversion</i>
Control	Beim Training bleibt Lisa in Bewegung und <i>At the training Lisa keeps moving and</i>	<i>doesn't move away from the spot</i> sie rennt eine Weile auf der Stelle <i>she runs for a while on the spot</i>	, aus Spaß <i>for fun</i>
Literal	Stefan nimmt die Kinder mit zum Jagen, doch <i>Stefan takes the children hunting, but</i>	<i>she runs on the spot for a bit</i> er bringt sie damit in Gefahr <i>he puts them with that in danger</i>	, wie er glaubt <i>, as he thinks</i>
Control	Stefan hält die Kinder ab vom Jagen, denn <i>Stefan prevents the children from hunting, because</i>	er sieht sie dabei in Gefahr <i>he sees them at that in danger</i> <i>he thinks they are in danger</i>	, wie er sagt <i>, as he says</i>

Table 5 (a) Log-transformed reading times from self-paced reading experiment 2, coefficients and standard error, t-value, confidence intervals, and p-value. (b) Reading times in ms from self-paced reading experiment 2, lmer coefficients and standard error, t-value, confidence intervals, and p-value. (c) Bootstrap on 95% confidence intervals for idiom type in ms, based on 10,000 replications

	β (SE)	t-value	CI-lower	CI-upper	Pr(> t)
<i>(a)</i>					
Controls (intercept)	7 (0.04)	31.1	6.93	7.06	<0.001***
Literal	−0.05 (0.01)	−4.18	−0.08	−0.03	<0.001***
Metonymic	−0.03 (0.01)	−3.21	−0.05	−0.01	0.008**
Metaphoric	−0.02 (0.01)	−3.11	−0.04	0	0.017*
Length	0.07 (0.01)	8.45	0.06	0.09	<0.001***
<i>(b)</i>					
Controls (intercept)	1206.3 (38.79)	31.1	1129.588	1282.965	<0.001***
Literal	−70.36 (16.83)	−4.18	−103.523	−37.061	<0.001***
Metonymic	−50.49 (15.74)	−3.21	−81.354	−19.617	0.001**
Metaphoric	−38.61 (12.41)	−3.11	−63.016	−14.226	0.002**
Length	82.94 (9.82)	8.45	63.659	102.792	<0.001***
<i>(c)</i>					
Confidence level 95%	CI-lower		CI-upper		Significant
Literal	−135.92		−62.98		Yes
Metonymic	−89.78		−20.32		Yes
Metaphoric	−45.77		15.07		No

***p < 0.001

**p < 0.01

*p < 0.05

Table 6 Compositionality ratings by idiom type in experiments 1 and 2

Rating/type	Literal		Metonymic		Metaphoric	
	Exp1	Exp2	Exp1	Exp2	Exp1	Exp2
1 (decomp.)	236	181	87	113	44	51
2	118	122	116	115	59	85
3	48	56	113	87	99	110
4	27	37	67	63	83	134
5	14	21	56	54	146	141
6	12	13	39	55	177	187
7 (nondecomp.)	7	23	26	53	232	190

Exp.1: number of participants = 42, Exp2: number of participants = 46
Numbers indicate how many times each idiom type was rated 1, 2 etc. in each experiment

Discussion

Findings show that context clearly impacts how idioms are processed: with a biasing context, all idioms were processed significantly faster than control sentences. Given that compositionality ratings once again showed no effect on reading times, length was controlled and the only other effects were those of idiom type, there seemed to be a tendency for increasing nonliteralness to cause greater processing difficulty in idioms, as long as there was a biasing context preparing for the idiomatic content. Effect sizes differ and there is a tendency for literal idioms to show the greatest and nonliteral idioms the smallest processing advantage, with metonymic idioms in-between. However, these differences were non-significant. This could either indicate that idiomaticity overrides nonliteralness in a biasing context, meaning that idioms are all understood faster as long as they are equally familiar, because their processing is somewhat automatized. Alternatively, it could mean that nonliteralness effects exist but are too subtle to be significant in this study. To check reliability of the nonliteralness effect, confidence-interval bootstraps on the final model with 100, 1000 and 10,000 replications were performed. Results unanimously suggest that the processing advantage for metaphoric idioms is unreliable (see Table 6), whereas it is reliable for metonymic idioms and largest for literal idioms. Still, experiment 2 showed that expectedness of content influences idiom processing, seems to strengthen the effect of idiomaticity and weaken the effect of nonliteralness. Most strikingly, the difference between metonymic versus metaphoric idioms was considerably smaller than in a neutral context.

General Discussion

In both sentence comprehension experiments, idioms as a group had a processing advantage over control sentences, as has been found in most other idiom processing studies. Our results, however, suggest processing ease further depends on nonliteralness, and on the context.

When the semantic content of an idiom is not expected, mainly nonliteralness seems to determine whether an idiom has a processing advantage over a comparable non-idiomatic sentence. Literal and metonymic idioms are processed much faster than controls whereas metaphoric idioms have only a very small processing advantage. Given that metaphoric idioms are clearly perceived as more nonliteral than metonymic and literal idioms, this implies that even in familiar, well-comprehensible idioms, greater nonliteralness is more difficult to process. It also implies that the meanings of idioms, if unexpected and presented as entire chunks embedded in sentences, are not automatically retrieved, but at least partly semantically accessed and constructed which seems to be easier for more literal idioms.

This finding is compatible with hybrid accounts of storage and processing. The Configuration Hypothesis (Cacciari and Tabossi 1988) holds that when an idiom is encountered, it is first processed non-idiomatically, the literal meanings of the individual words are activated until it is recognized as a familiar idiom. The recognition will then cause the processing mode to change, thus cause access of the idiomatic meaning which is only activated once a sufficient portion of the idiom has been encountered (Cacciari and Glucksberg 1991; Cacciari and Tabossi 1988). It is possible that the change to a different interpretation path comes at a cost and leads to a slackening of processing speed, the more nonliteral an idiom is. If comprehension mechanisms have to shift from a literal to a nonliteral path and

inhibit the literal interpretation in a sort of self-correction, this likely requires more processing effort than remaining on the literal path.

In a biasing context, idiom recognition can happen much earlier (Cacciari and Tabossi 1988; Tabossi and Zardon 1993) and switching interpretation paths might thus come at a smaller cost. This is suggested by our findings in a biasing-context condition: While all idioms at first seem to be processed significantly faster than controls, the small differences between literal and metonymic and, respectively, metonymic and metaphoric idioms might indicate that nonliteralness still has a small effect even when the semantic content is expected. This is backed by the bootstraps which also imply a hierarchy of processing ease from literal to nonliteral idioms and the processing advantage for metaphoric idioms to be unreliable. Accordingly, metaphoric idioms have at best a very small processing advantage, possibly even none. At any rate, effect sizes suggest that the differences in processing ease by idiom diminish in a context that prepares for their semantic content, as this preparation likely leads to faster idiom recognition. While the Configuration Hypothesis makes no specific predictions about nonliteralness in idioms, it can account for most of our findings—with the addendum that a shift from a literal to a highly nonliteral interpretation comes at a cost and increases processing load.

Our results are also partially in line with the Hybrid Model (Titone and Connine 1999). Based on their eye tracking results, processing is facilitated when the products of a literal analysis overlap with the idiomatic meaning, while more processing effort is required if the literal interpretation is distinct from the idiomatic meaning. The Hybrid Model, similarly to the Configuration Hypothesis, proposes that all idioms undergo both compositional analysis and direct retrieval of the idiomatic meaning. If conventionality (as driven by familiarity and likely comprehensibility) of idioms is constant and compositionality has no effect, increasing transparency is the only facilitator. This would once again predict that literal idioms will be processed fastest, followed by metonymic, and metaphoric idioms, due to the increasing disparity between literal and idiomatic meaning. Within a neutral context, this prediction is partly met in the comparably large processing advantage for metonymic compared to metaphoric idioms. Within a biasing context, this prediction is met in the tenuous tendency for a hierarchical processing advantage from literal over metonymic to metaphoric idioms as the least advantageous. The slightness of these differences in experiment 2, this has to be regarded with caution. The bootstraps unanimously suggest that the processing advantage of metaphoric idioms is overestimated and may not exist at all, which would strengthen support for the Hybrid Model.

The different effects for metonymic and metaphoric idioms in the neutral-context condition suggest that it is not purely idiomaticity (i.e. being an idiom) that causes processing advantages, indicating that idiomatic meanings are not equally quickly retrieved, accessed or composed as soon as an idiom is recognized. If that were the case, all idioms should be processed equally fast. Instead, it seems that metonymic and metaphoric structures are indeed processed even in idioms with a neutral context: Our results provide evidence that metonymies are easier to process and more basic to cognition than metaphors, even in idioms, to a lesser degree also with a biasing context.

In our studies, literal idioms only differ from the other idiom types in their nonliteralness and show a clear processing advantage in both context conditions, and the largest one in a biasing context. However, they are not processed faster than metonymic idioms in a neutral context. Thus, the advantage of higher literalness seems to be limited, at least in a neutral context. It is possible that literal idioms have lower idiomaticity, precisely because they are literal, which might make them less salient as idioms. This might lead to later idiom recognition—or in terms of the Configuration Hypothesis, the idiomatic key would

occur later (Cacciari and Tabossi 1988). This would slow down reading times because an idiom is treated as a non-idiomatic sequence for longer than idioms with higher idiomaticity, and upon recognition possibly also needs to be verified as an idiomatic construction. After this step, time is saved in the process as reinterpretation is unnecessary because the idiom can be understood literally. Thus we suggest there is a trade-off between the properties of idiomaticity and nonliteralness in a neutral context. It seems that both literalness and idiomaticity expedite processing. An idiom in which both properties are as high as the other allows, their accumulated advantages would lead to the maximal processing advantage. This is found for the metonymic type. After all, metonymic idioms are still fairly literal expressions with a semantically *non*-distinct meaning, but are apparently highly idiomatic as long as they are familiar. Consequently, meaning composition or retrieval is fast in metonymic idioms. Familiar metaphoric idioms, then, also seem to have rather high idiomaticity; otherwise, a processing disadvantage would be expected. Yet their high nonliteralness, as seen in their semantically distinct idiomatic meaning, seems to create higher processing load, possibly resulting from re-interpretation or meaning retrieval costs, and thus lead to no or no tangible processing advantage.

In sum, our results lend support to accounts suggesting processing differences depending on contextual information, such as the Graded Salience Hypothesis (Giora 1997, 2003). Our findings especially demonstrate the importance of the context when familiarity and comprehensibility are constant. A facilitating context leads at least to a reduction in processing differences between idiom types. It can strongly decrease the effect of nonliteralness because idiomatic meanings become more salient when their semantic content is expected. The results also stress that matching on familiarity and comprehensibility alone is insufficient because then other idiomatic properties determine processing ease. As per the Graded Salience Hypothesis, it is possible that literal idioms have no processing advantage over metonymic ones because they are less salient as idioms due to their lower idiomaticity.

Yet as the unreliability of the effect in metaphoric idioms suggests, nonliteralness still seems to lead to a slight processing advantage of metonymic (and a larger one for literal) over metaphoric idioms. Together, the findings are unsuitable to support any theory predicting that all idiomatic meanings are automatically retrieved upon recognition, such as the Superlemma Hypothesis (Sprenger et al. 2006).

Lastly, no effect of compositionality could be found in either experiment. This is in line with other studies that found only a very limited effect on processing speed (Libben and Titone 2008; Sprenger et al. 2006) or no compositionality effect at all (Cutting and Bock 1997, in production errors). As Sprenger et al. (2006) state: “Once the level of (lexical) concepts has been passed, processing decompositional and nondecompositional idioms does not differ anymore. Decomposition of idiomatic expressions is thus relegated to the conceptual [...] domain” (p. 164). Our findings are also congruent with Tabossi et al. (2008), who found no compositionality effect in a speeded idiom reading and recognition task, at least in relatively unbiased contexts. Results thus suggest that compositionality is irrelevant in self-paced reading, even when idioms are presented as complete chunks. An alternative explanation might be that compositionality ratings are flawed because the task is too challenging for participants. In fact, 47% reported at least some uncertainty in rating compositionality. This uncertainty may be due to the novelty of the concept for most participants, or to the mixture of analytical and intuitive abilities needed for the task which may itself lead to an uncertainty of whether one’s rating is “correct”. Conversely, the level of uncertainty portrayed here could be overestimated as a result of the rather fine-grained 7-step scale which participants used. After all, a

participant might be certain an idiom is quite decomposable (a “2” or “3” on the scale), but unsure of whether to settle for “2” or “3”. The self-reports may mirror some uncertainty, but cannot tell the experimenter that the uncertainty may have been only on a small scale, such as between “2” and “3” (and not between e.g. “1” and “7”), or whether the uncertainty was present for some idioms only.

There was some variance in the ratings within idioms. In light of this finding, it might be possible that “comprehenders’ intuitions about compositionality are often unreliable” (Libben and Titone 2008: 1104, see also Titone and Connine 1994; Nordmann et al. 2014). However, there was still a clear unifying trend to literal and metonymic idioms as more decomposable than metaphors in both experiments. In addition, there is no objective measures of how decomposable an idiom is. Thus, instead of referring to compositionality ratings as unreliable, we rather infer that native speakers make sense of idioms in subjective ways. We also stress there are great similarities in how compositionally different idiom types are perceived. These are not necessarily reflected in processing difficulty and mostly likely not in self-paced reading, but can rather be captured in offline analytical tasks such as ratings. Consequently, our findings cannot support any models distinguishing processing according to decomposable and nondecomposable idioms.

Conclusion

Two self-paced reading experiments showed that idiom processing in a moving-window chunk-by-chunk design is determined by the sentential context. If the context does not foster expectation of the semantic content of an impending idiom, then nonliteralness along with idiomaticity affects processing ease. More literal idioms are processed faster than non-idiomatic language. Very nonliteral idioms have little to no processing advantage. Nonliteralness seems to cause higher processing load and can override idiomaticity when familiarity, understandability, and length are controlled. With a facilitating context, this effect is weakened. Yet, there still seems to be the implication that the more literal an idiom, the more easily it is processed. To confirm or refute this potential effect, further experimental research with several hundred participants would be needed. It can be concluded, however, that idioms differ in their processing ease depending on their nonliteralness and on whether their semantic content is expected. Overall, these findings can support hybrid accounts of idiom storage and processing.

Acknowledgements This study was funded by the author’s personal finances which are provided through a scholarship by the Friedrich-Naumann-Stiftung für die Freiheit. I thank an anonymous reviewer for their insightful and helpful comments. Furthermore, I am very thankful to Francesca Citron (University of Lancaster) for the very valuable discussion on the design of the first experiment, to Nina Julich (Universität Leipzig) for her very helpful input on an earlier version of this paper, to Gisbert Fanselow (Universität Potsdam) for reading and commenting an earlier draft, to Ruth Keßler (Universität Tübingen) for her great help in recruiting participants, and to Anatol Stefanowitsch (FU Berlin) and Titus von der Malsburg (Universität Potsdam) for their continuous expert support in making this research possible.

Compliance with Ethical Standards

Conflict of interest The author declares that they have no conflict of interest.

Informed Consent Informed consent was obtained from all individual participants included in the study.

Appendix 1

Theoretical Alternative for Building Matching Control Sentences

In building control sentences, constraints well-known to idiom experimenters arise: idioms require a fixedness of word choice such that synonyms or semantic neighbors can rarely replace words without rendering the idiom unrecognizable, marked, or changing the meaning. Some idioms require certain contexts to be understandable or unambiguous, some contain words rather rare in absolute terms yet very common within the idiomatic use, while the idioms themselves are actually very common or well-known themselves. However, exact overall match of length and frequency was done wherever possible. While it is most common to match control items as closely as possible to the test items by using as many of the same words in the same order as possible, this kind of “formal matching” causes an unavoidable difference in the meanings of sentences, which cannot be quantified. Size and effect of divergence or even fundamental difference in meanings of items remain unobservable and unmeasurable in participants’ minds. The problem of this “black box of meanings” could be dissolved by matching sentences for meaning. This would be especially profitable in the case of idioms as they can be translated into a fairly exact non-idiomatic meaning. Consequently, it could be argued that for idioms, it makes sense to form control sentences that are matched by meaning only. For example, the item *For two hours, Marie’s stomach has been growling* would be matched by *For two hours, Marie has been very hungry*. The advantage of matching by meaning is that it enables a direct comparison between the processing difficulty of an idiom and its meaning which in theory provides a more precise and pointed answer as to whether a fixed expression is indeed easier to process than its non-idiomatic sentence expressing the identical meaning. This matching, however, poses many challenges and leads to serious unavoidable differences in construction and word forms between sentence pairs: syntax, word categories, word number, and other grammatical differences can diverge to a degree where a control sentence cannot actually serve as true control anymore. Length and frequency matching become even more difficult as words from the idiom can hardly be repeated in its control. At the same time, the choice of potential words is automatically very constrained when the control sentence is meant to express a very particular meaning. Occasionally, an idiom’s meaning may be too ambiguous or complex to be fully captured in a control sentence with a fairly fixed required number of words. In sum, practical emerging difficulties and resulting differences between idiom and control sentences make matching by meaning very problematic and offer too many potential confounds, despite the undeniable advantage of a matched and controlled meaning and the theoretical strength of a very pointed answer to the question whether idioms are cognitively and semantically easier to process than non-idiomatic sentences. We conclude that “formal matching” is the sounder choice because it controls for more possible confounds, but matching by meaning could be valuable for a complementary study serving as a direct comparison for the current study.

Appendix 2

Prestudy on Ratings of Nonliteralness in Literal Idioms

Nonliteralness ratings on literal, metonymic, and metaphoric idioms were collected in two separate studies for two reasons: one, to test whether results from one study could be

repeated for literal idioms, thus to decrease the chance of chance findings; second, it was attempted to make the studies as simple and outcomes as clear as possible by only demarcating two different types of idioms each time.

To check for significance of the effect of idiom type and to account for random individual differences of items and participants, an ordinal mixed effects regression was performed. Literalness ratings were fitted as a function of idiom type as a categorical fixed effect, random intercepts for items and random slopes for type by participant. Both studies revealed idiom type to have a significant effect. In study 1 containing the metaphoric idioms, the effect was stronger ($b = -2.11$, $z = -14.43$, $p < 0.001$ as opposed to $b = -0.53$, $z = -4.56$, $p < 0.001$ in study 2), which is expected, given that metaphoric idioms were rated as much more nonliteral than metonymic idioms. This indicates that literal idioms are indeed perceived as substantially more literal than both metonymic and metaphoric idioms. Furthermore, the difference in nonliteralness between these two groups confirm the results of the initial rating study (Michl 2019) on the nonliteralness of metonymic compared to metaphoric idioms and show that the effect can also be found on a less detailed rating scale.

Appendix 3

Effect of Transparency Ratings on Reading Times

Given that transparency can influence processing ease depending on the task, metonymic and metaphoric idioms were also rated on it by adult German native speakers (see, Michl 2019). Transparency is defined as the closeness of relation between what is said and what is meant (or the literal and the idiomatic meaning) in an idiom (see also Nippold and Taylor 2002; Titone and Connine 1999). 111 participants rated it on a 5-point Likert scale on which 5 indicated “completely transparent”. Metonymic idioms received a mean rating of 3.8 ($sd = 0.18$), metaphoric idioms received a mean rating of 2.8 ($sd = 0.13$). To control for potential transparency effects in the present reading experiments, median ratings for each idiom were once included in the final models. Results remained largely the same. Transparency itself was completely irrelevant in a non-biasing context (Table 7a, experiment 1), yet turned out to be a significant predictor when idioms were presented in a biasing context (Table 7b, experiment 2). One caveat needs to be borne in mind: literal idioms were not rated on transparency by multiple participants. Instead, literal idioms received ratings of “completely transparent” by the author. This was decided because literal idioms have only one meaning, so their literal and the idiomatic meaning are the same, so they should be completely transparent by definition. Due to this difference in data collection and for the sake of simplicity, transparency ratings are not considered in the final data analysis.

Table 7 Log-transformed reading times from self-paced reading (a) experiment 1, (b) experiment 2, lmer coefficients and standard error, t-value, confidence intervals, and p-value

	β (SE)	t-value	CI-lower	CI-upper	Pr(> t)
(a)					
Controls (intercept)	7.26 (0.06)	112.05	7.131	7.388	< 0.001***
Literal	−0.05 (0.02)	−3.07	−0.08	−0.018	0.002***
Metonymic	−0.06 (0.02)	−4.3	−0.094	−0.035	< 0.001***
Metaphoric	−0.02 (0.01)	−1.33	−0.04	0.008	0.183
Length	0.03 (0.01)	3.94	0.014	0.042	< 0.001***
Transparency rating	0 (0.01)	0.02	−0.017	0.018	0.983
(b)					
Controls (intercept)	7.08 (0.05)	151.57	6.983	7.167	< 0.001***
Literal	−0.05 (0.01)	−3.72	−0.074	−0.023	< 0.001***
Metonymic	−0.03 (0.01)	−2.4	−0.052	−0.005	0.017*
Metaphoric	−0.03 (0.01)	−2.79	−0.045	−0.008	0.005**
Length	0.07 (0.01)	9.76	0.059	0.089	< 0.001***
Transparency rating	−0.02 (0.01)	−2.57	−0.038	−0.005	0.012*

***p < 0.001

**p < 0.01

*p < 0.05

References

- Aaronson, D., & Scarborough, H. S. (1976). Performance theories for sentence coding: Some quantitative evidence. *Journal of Experimental Psychology: Human Perception and Performance*, 2(1), 56–70. <https://doi.org/10.1037/0096-1523.2.1.56>.
- Annaz, D., van Herwegen, J., Thomas, M., Fishman, R., Karmiloff-Smith, A., & Rundblad, G. (2009). Comprehension of metaphor and metonymy in children with Williams syndrome. *International Journal of Language & Communication Disorders*, 44(6), 962–978. <https://doi.org/10.3109/13682820802525005>.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Special Issue: Emerging Data Analysis*, 59(4), 390–412. <https://doi.org/10.1016/j.jml.2007.12.005>.
- Barcelona, A. (1997). Clarifying and applying the notions of metaphor and metonymy within cognitive linguistics. *Atlantis*, 19, 21–48.
- Barnden, J. A. (2007). Metaphor, semantic preferences and context-sensitivity. In K. Ahmad, C. Brewster, M. Stevenson, & Y. Wilks (Eds.), *Words and intelligence: Vol. 2. Essays in honor of Yorick Wilks* (pp. 39–62). Dordrecht: Springer. https://doi.org/10.1007/1-4020-5833-0_2.
- Bartsch, R. (2002). Generating polysemy: Metaphor and metonymy. In R. Dirven & R. Pörrings (Eds.), *Cognitive linguistics research* (Vol. 20, pp. 49–74). Metaphor and metonymy in comparison and contrast Berlin: Mouton de Gruyter.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). R package “lme4”. [Computer software]. Retrieved July 20, 2019 from <https://cran.r-project.org/>.
- Blasko, D., & Connine, C. (1993). Effects of familiarity and aptness on metaphor processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(2), 295–308.
- Bobrow, S. A., & Bell, S. M. (1973). On catching on to idiomatic expressions. *Memory & Cognition*, 1(3), 343–346. <https://doi.org/10.3758/BF03198118>.
- Bogdanovich, J. M., Sykes, J. R., & Barr, D. J. (1997). Metaphor idiom comprehension. *Special Issue: Emerging Data Analysis*, 37(2), 141–154. <https://doi.org/10.1006/jmla.1996.2506>.
- Bortfeld, H., & McGlone, M. S. (2001). The continuum of metaphor processing. *Metaphor and Symbol*, 16(1&2), 75–86.

- Cacciari, C., & Glucksberg, S. (1991). Chapter 9: Understanding idiomatic expressions: The contribution of word meanings. In G. B. Simpson (Ed.), *Advances in psychology. Understanding word and sentence* (Vol. 77, pp. 217–240). Amsterdam: Elsevier. [https://doi.org/10.1016/S0166-4115\(08\)61535-6](https://doi.org/10.1016/S0166-4115(08)61535-6).
- Cacciari, C., & Tabossi, P. (1988). The comprehension of idioms. *Journal of Memory and Language*, 27, 668–683.
- Caillies, S., & Butcher, K. (2007). Processing of idiomatic expressions: Evidence for a new hybrid view. *Metaphor and Symbol*, 22(1), 79–108. <https://doi.org/10.1080/10926480709336754>.
- Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied Linguistics*, 29(1), 72–89. <https://doi.org/10.1093/applin/amm022>.
- Conklin, K., & Schmitt, N. (2012). The processing of formulaic language. *Annual Review of Applied Linguistics*, 32, 45–61. <https://doi.org/10.1017/S0267190512000074>.
- Coulson, S., & Matlock, T. (2001). Metaphor and the space structuring model. *Metaphor and Symbol*, 16(3–4), 295–316. <https://doi.org/10.1080/10926488.2001.9678899>.
- Coulson, S., & van Petten, C. (2002). Conceptual integration and metaphor: An event-related potential study. *Memory & Cognition*, 30(6), 958–968. <https://doi.org/10.3758/BF03195780>.
- Croft, W. (1993). The role of domains in the interpretation of metaphors and metonymies. *Cognitive Linguistics*, 4(4), 335–370. <https://doi.org/10.1515/cogl.1993.4.4.335>.
- Cronk, B. C., & Schweigert, W. A. (1992). The comprehension of idioms: The effects of familiarity, literalness, and usage. *Applied Psycholinguistics*, 13(2), 131–146.
- Cutting, J. C., & Bock, K. (1997). That's the way the cookie bounces: Syntactic and semantic components of experimentally elicited idiom blends. *Memory & Cognition*, 25(1), 57–71.
- de Mendoza Ibáñez, R., & José, F. (2003). The role of mappings and domains in understanding metonymy. In A. Barcelona (Ed.), *Metaphor and metonymy at the crossroads: A cognitive perspective* (pp. 109–132). Berlin: Mouton de Gruyter.
- Dirven, R. (2002). Metonymy and metaphor: Conceptualisation strategies. In R. Dirven & R. Pörrings (Eds.), *Cognitive linguistics research* (Vol. 20, pp. 75–111). Metaphor and metonymy in comparison and contrast Berlin: Mouton de Gruyter.
- Dobrovol'skij, D. (1995). Schiß und Espenlaub: Idiome der Angst. *Folia Linguistica*. <https://doi.org/10.1515/flin.1995.29.3-4.317>.
- Fauconnier, G., & Turner, M. (1996). Blending as a central process of grammar. *Conceptual Structure, Discourse, and Language*, 113, 130–153.
- Feyaerts, K. (2003). Refining the inheritance hypothesis: Interaction between metaphoric and metonymic hierarchies. In A. Barcelona (Ed.), *Metaphor and metonymy at the crossroads: A cognitive perspective* (pp. 59–78). Berlin: Mouton de Gruyter.
- Field, A. P., Miles, J., & Field, Z. (2014). *Discovering statistics using R (Repr)*. London: Sage.
- Forster, J., & Forster, K. (2003). DMDX. DMASTR software. [Computer Software].
- Gentner, D., Bowdle, B. F., Wolff, P., & Boronat, C. (2001). *Metaphor is like analogy*. Cambridge: The MIT Press.
- Gibbs, R. W. (1980). Spilling the beans on understanding and memory for idioms in conversation. *Memory & Cognition*, 8(2), 149–156. <https://doi.org/10.3758/BF03213418>.
- Gibbs, R. W., Nayak, N. P., & Cutting, C. (1989). How to kick the bucket and not decompose: Analyzability and idiom processing. *Journal of Memory and Language*, 28(5), 576–593. [https://doi.org/10.1016/0749-596X\(89\)90014-4](https://doi.org/10.1016/0749-596X(89)90014-4).
- Giora, R. (1997). Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics*, 8(3), 183–206. <https://doi.org/10.1515/cogl.1997.8.3.183>.
- Giora, R. (2003). *On our mind: Salience, context, and figurative language*. New York: Oxford University Press.
- Glucksberg, S., & Keysar, B. (1990). Understanding metaphoric comparisons: Beyond similarity. *Psychological Review*, 97(1), 3–18. <https://doi.org/10.1037/0033-295X.97.1.3>.
- Glucksberg, S., Newsome, M. R., & Goldvarg, Y. (2001). Inhibition of the literal: Filtering metaphor-irrelevant information during metaphor comprehension. *Metaphor and Symbol*, 16(3–4), 277–298. <https://doi.org/10.1080/10926488.2001.9678898>.
- Goossens, L. (Ed.). (1995a). *By word of mouth: Metaphor, metonymy, and linguistic action in a cognitive perspective*. Amsterdam: John Benjamins Publishing Company.
- Goossens, L. (1995b). Metaphonymy: The interaction of metaphor and metonymy in expressions for linguistic action. In L. Goossens (Ed.), *By word of mouth: Metaphor, metonymy, and linguistic action in a cognitive perspective* (pp. 159–174). Amsterdam: John Benjamins Publishing Company.
- Hamblin, J., & Gibbs, R. (1999). Why you can't kick the bucket as you slowly die: Verbs in idiom comprehension. *Journal of Psycholinguistic Research*, 28(1), 25–39.

- Holsinger, E. (2013). Representing idioms: Syntactic and contextual effects on idiom processing. *Language and Speech*, 56(3), 373–394. <https://doi.org/10.1177/0023830913484899>.
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87(4), 329–354. <https://doi.org/10.1037/0033-295X.87.4.329>.
- Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, 111(2), 228–238. <https://doi.org/10.1037/0096-3445.111.2.228>.
- Katz, A. N., Paivio, A., Marschark, M., & Clark, J. M. (1988). Norms for 204 literary and 260 nonliterary metaphors on 10 psychological dimensions. *Metaphor and Symbolic Activity*, 3(4), 191–214. https://doi.org/10.1207/s15327868ms0304_1.
- Keating, G. D., & Jegerski, J. (2015). Experimental designs in sentence processing research. *Studies in Second Language Acquisition*, 37(1), 1–32. <https://doi.org/10.1017/S0272263114000187>.
- Keysar, B., & Bly, B. M. (1999). Swimming against the current: Do idioms reflect conceptual structure? *Journal of Pragmatics*, 31(12), 1559–1578. [https://doi.org/10.1016/S0378-2166\(99\)00004-1](https://doi.org/10.1016/S0378-2166(99)00004-1).
- Klepousiotou, E. (2002). The processing of lexical ambiguity: Homonymy and Polysemy in the mental lexicon. *Brain and Language*, 81(1–3), 205–223. <https://doi.org/10.1006/brln.2001.2518>.
- Kövecses, Z., & Radden, G. (1998). Metonymy: Developing a cognitive linguistic view. *Cognitive Linguistics*, 9(1), 37–78. <https://doi.org/10.1515/cogl.1998.9.1.37>.
- Kuznetsova, A., Brockhoff, P. B., & Haubo, R. (2016). (R package “lmerTest”. [Computer software]. Retrieved July 20, 2019 from <https://cran.r-project.org/>.
- Lai, V. T., & Curran, T. (2013). ERP evidence for conceptual mappings and comparison processes during the comprehension of conventional and novel metaphors. *Brain and Language*, 127(3), 484–496. <https://doi.org/10.1016/j.b&l.2013.09.010>.
- Lai, V. T., Curran, T., & Menn, L. (2009). Comprehending conventional and novel metaphors: An ERP study. *Brain Research*, 1284, 145–155. <https://doi.org/10.1016/j.brainres.2009.05.088>.
- Lakoff, G., & Johnson, M. (1980). The metaphoric structure of the human conceptual system. *Cognitive Science*, 4(2), 195–208. https://doi.org/10.1207/s15516709cog0402_4.
- Laurent, J.-P., Denhières, G., Passerieux, C., Iakimova, G., & Hardy-Baylé, M.-C. (2006). On understanding idiomatic language: The salience hypothesis assessed by ERPs. *Brain Research*, 1068(1), 151–160. <https://doi.org/10.1016/j.brainres.2005.10.076>.
- Leiner, D. (2014). Sosci Survey. [Computer Software]. Retrieved April 20, 2018 from <https://www.sosci-survey.de/>.
- Libben, M. R., & Titone, D. A. (2008). The multidetermined nature of idiom processing. *Memory & Cognition*, 36(6), 1103–1121. <https://doi.org/10.3758/MC.36.6.1103>.
- Michl, D. (2019). Metonymies are more literal than metaphors: Evidence from ratings of German idioms. *Language and Cognition*, 11(1), 98–124.
- Nippold, M. A., & Taylor, C. L. (2002). Judgments of idiom familiarity and transparency. *Journal of Speech Language and Hearing Research*, 45(2), 384–391. [https://doi.org/10.1044/1092-4388\(2002/030\)](https://doi.org/10.1044/1092-4388(2002/030)).
- Nordmann, E., Cleland, A. A., & Bull, R. (2014). Familiarity breeds dissent: Reliability analyses for British-English idioms on measures of familiarity, meaning, literality, and decomposability. *Acta Psychologica*, 149, 87–95. <https://doi.org/10.1016/j.actpsy.2014.03.009>.
- Nunberg, G. (1978). *The pragmatics of reference (Dissertation)*. Bloomington: Indiana University Linguistics Club.
- Nunberg, G., Sag, I. A., & Wasow, T. (1994). Idioms. *Language*, 70(3), 491–538. <https://doi.org/10.2307/416483>.
- Ortony, A. (Ed.). (1979). The Role of Similarity in Similes and Metaphors. In *Metaphor and thought* (pp. 186–201). Cambridge: Cambridge University Press.
- Ortony, A. (Ed.). (1993). *Metaphor and Thought*. Cambridge: Cambridge University Press.
- Pynte, J., Besson, M., Robichon, F. H., & Poli, J. (1996). The time-course of metaphor comprehension: An event-related potential study. *Brain and Language*, 55(3), 293–316. <https://doi.org/10.1006/brln.1996.0107>.
- R Core Team. (2019a). R. [Computer software]. Retrieved July 21, 2019 from <https://cran.r-project.org/>.
- R Core Team. (2019b). R Studio. [Computer software]. Retrieved July 21, 2019 from <https://cran.r-project.org/>.
- Radden, G. (2003). How metonymic are metaphors? In A. Barcelona (Ed.), *Metaphor and metonymy at the crossroads: A cognitive perspective* (pp. 93–108). Berlin: Mouton de Gruyter.
- Radden, G. (2005). The ubiquity of metonymy. In J. L. Otal, I. Navarro i Ferrando, & B. Bellés Fortuño (Eds.), *Cognitive and discourse approaches to metaphor and metonymy* (pp. 11–28). Castelló de la Plana: Publicacions de la Universitat Jaume I.

- Rapp, A. M., Erb, M., Grodd, W., Bartels, M., & Markert, K. (2011). Neural correlates of metonymy resolution. *Brain and Language*, 119(3), 196–205. <https://doi.org/10.1016/j.b&l.2011.07.004>.
- Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R., & Clifton, C. (1989). Eye movements and on-line language comprehension processes. *Language and Cognitive Processes*, 4(3–4), SI21–SI49. <https://doi.org/10.1080/01690968908406362>.
- Rundblad, G., & Annaz, D. (2010). Development of metaphor and metonymy comprehension: Receptive vocabulary and conceptual knowledge. *British Journal of Developmental Psychology*, 28(3), 547–563. <https://doi.org/10.1348/026151009X454373>.
- Sailer, M. (2013). Idiom and phraseology. Oxford Bibliographies Online Datasets.
- Schemann, H. (2011). *Deutsche idiomatik*. Berlin: De Gruyter.
- Schmitt, N. (Ed.). (2004). *Language learning and language teaching: Formulaic sequences: Vol. 9. Acquisition, processing and use*. Amsterdam: Benjamins.
- Schmitt, N., & Underwood, G. (2004). Exploring the processing of formulaic sequences through a self-paced reading task. In N. Schmitt (Ed.), *Language learning and language teaching* (Vol. 9, pp. 173–190). Formulaic sequences: Acquisition, processing and use Amsterdam: Benjamins.
- Schweigert, W. A. (1986). The comprehension of familiar and less familiar idioms. *Journal of Psycholinguistic Research*, 15(1), 33–45. <https://doi.org/10.1007/BF01067390>.
- Shibata, M., Toyomura, A., Motoyama, H., Itoh, H., Kawabata, Y., & Abe, J.-I. (2012). Does simile comprehension differ from metaphor comprehension? A functional MRI study. *Brain and Language*, 121(3), 254–260. <https://doi.org/10.1016/j.b&l.2012.03.006>.
- Spieß, C., Köpcke, K. M., (Eds.). (2015). *Metapher und Metonymie: Theoretische, methodische und empirische Zugänge*: De Gruyter. Retrieved January 10, 2019 from <https://books.google.de/books?id=kQJfCAAAQBAJ>.
- Sprenger, S., Levelt, W., & Kempen, G. (2006). Lexical access during the production of idiomatic phrases. *Journal of Memory and Language*, 54(2), 161–184. <https://doi.org/10.1016/j.jml.2005.11.001>.
- Stowe, L. A. & Kaan, E. (2006). *Developing an experiment. Techniques and design*. [unpublished textbook].
- Swinney, D. A., & Cutler, A. (1979). The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behavior*, 18(5), 523–534. [https://doi.org/10.1016/S0022-5371\(79\)90284-6](https://doi.org/10.1016/S0022-5371(79)90284-6).
- Tabossi, P., Fanari, R., & Wolf, K. (2008). Processing idiomatic expressions: Effects of semantic compositionality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(2), 313–327. <https://doi.org/10.1037/0278-7393.34.2.313>.
- Tabossi, P., Fanari, R., & Wolf, K. (2009). Why are idioms recognized fast? *Memory & Cognition*, 37(4), 529–540. <https://doi.org/10.3758/MC.37.4.529>.
- Tabossi, P., & Zardon, F. (1993). The activation of idiomatic meaning in spoken language comprehension. In C. Cacciari & P. Tabossi (Eds.), *Idioms: Processing, structure, and interpretation*. New York: Psychology Press.
- Taylor, J. (1995). *Linguistic categorization*. Oxford: Clarendon.
- Titone, D. A., & Connine, C. M. (1994). Descriptive norms for 171 idiomatic expressions: Familiarity, compositionality, predictability, and literality. *Metaphor and Symbolic Activity*, 9(4), 247–270. https://doi.org/10.1207/s15327868ms0904_1.
- Titone, D. A., & Connine, C. M. (1999). On the compositional and noncompositional nature of idiomatic expressions. *Journal of Pragmatics*, 31(12), 1655–1674. [https://doi.org/10.1016/S0378-2166\(99\)00008-9](https://doi.org/10.1016/S0378-2166(99)00008-9).
- Turner, M., & Fauconnier, G. (2003). Metaphor, metonymy, and binding. In A. Barcelona (Ed.), *Metaphor and metonymy at the crossroads: A cognitive perspective* (pp. 133–145). Berlin: Mouton de Gruyter.
- Weiland, H., Bambini, V., & Schumacher, P. B. (2014). The role of literal meaning in figurative language comprehension: Evidence from masked priming ERP. *Frontiers in Human Neuroscience*, 8, 583. <https://doi.org/10.3389/fnhum.2014.00583>.