

Metonymies are more literal than metaphors: evidence from ratings of German idioms*

DIANA MICHL

Department Linguistik, Universität Potsdam

(Received 23 February 2018 – Revised 04 April 2019 – Accepted 05 April 2019)

ABSTRACT

Metaphor and metonymy are likely the most common forms of non-literal language. As metaphor and metonymy differ conceptually and in how easy they are to comprehend, it seems likely that they also differ in their degree of non-literality. They frequently occur in idioms which are foremost non-literal, fixed expressions. Given that non-literality seems to be the defining criterion of what constitutes an idiom, it is striking that no study so far has focused specifically on differing non-literality in idioms. It is unclear whether and how metaphoric and metonymic structures and their properties are perceived in idioms, given that the comprehension of idioms is driven by a number of other properties that are connected. This study divides idioms according to their metonymic or metaphoric structure and lets participants rate their non-literality, familiarity, and transparency. It focuses on non-literality as key property, finds it strongly connected to transparency, and to be the one key factor in predicting idiom type. Specifically, it reveals that metonymies are generally perceived as rather or even extremely literal, while metaphors are generally perceived as highly non-literal.

1. Introduction

Metaphor and metonymy are fundamental thought patterns that structure our conceptual thinking and language use (Barcelona, 1997; Glucksberg & Keysar, 1990; Lakoff & Johnson, 1980; Radden, 2003, 2005). In a metaphor such as *to swim against the current*, what is said (*current* in a body of water)

[*] I thank all participants and test participants, everyone who helped distributing invitations to the surveys, and the reviewers of the paper. Address for correspondence: Department Linguistik, Universität Potsdam, Karl-Liebknecht-Str. 24-25, Haus 14, II.14.201, 14476 Potsdam. Email: dmichl@uni-potsdam.de

stands for something literally unrelated (a ‘general trend’). In a metonymy such as *to have an eye for detail*, *eye* refers to something that is literally or immanently related or part of the same concept (i.e., ‘ability to see details’). General consensus today is that metaphorically used words or phrases refer to a target in a distinct semantic concept, thus functioning in between two concepts or domains – whereas metonymically used words or phrases refer to a target within the same semantic concept (see also Kövecses & Radden, 1998; Lakoff & Johnson, 1980; Ruiz de Mendoza Ibáñez, 2003; Spieß & Köpcke, 2015; Turner & Fauconnier, 2003). Hence metonymy is based on a contiguity relationship between what is said and what is meant (Annaz, van Herwegen, Thomas, Fishman, Karmiloff-Smith, & Rundblad, 2009; Feyaerts, 2003; Bartsch, 2002; Croft, 1993; Dirven, 2002; Klepousniotou, 2002) while metaphor is based on a similarity or analogy relationship (see also Barnden, 2007; Bartsch, 2002; Bortfeld & McGlone, 2001; Bowdle & Gentner, 2005; Coulson & Matlock, 2001; Gentner, Bowdle, Wolf, & Boronat, 2001; Ortony, 1979). Metonymy is often suggested to be more basic to cognition and also easier to learn and comprehend than metaphor (Goossens, 1995; Taylor, 1995).

A vast number of studies have examined the processing of metaphors, significantly fewer have examined the processing of metonymies, and very few have compared them directly. In summary, experimental research suggests that metonymies are indeed easier to acquire and comprehend than metaphors: both adults (Rundblad & Annaz, 2010; see also partially Klepousniotou, 2002, and Weiland, Bambini, & Schumacher, 2014) and young children (Annaz et al., 2009) show faster comprehension and production with fewer errors in metonymies than metaphors.

Metaphors and metonymies can be conventionalized, which facilitates their comprehension. As such, they are found in many idioms (as implied by the term ‘frozen metaphors’ falsely used as a synonym for idioms in, for example, Handford & Koester, 2010; Ortony, Schallert, Reynolds, & Antos, 1978). Idioms are syntactically complex, more or less fixed expressions (Sailer, 2013) such as *to throw in the towel*, *to hit rock bottom*, *to drink somebody under the table*, or *to lose one’s heart to somebody*. They are a highly pervasive language phenomenon: speakers are estimated to use 7,000 idioms per week (Hoffman, 1984). Multiword expressions, including idioms, are believed to be stored in long-term semantic memory as complete units. This is mirrored by ample evidence of their processing advantage over non-idiomatic language (Gibbs & Gonzales, 1985), at least when they are presented in canonical chunks (Conklin & Schmitt, 2008; Schmitt & Underwood, 2004) and are equally familiar to a recipient (Libben & Titone, 2008; Schweigert, 1986). Faster processing indicates that idiomatic meanings are not entirely put together from their individual constituents upon being encountered (see also Keysar & Bly, 1999), but that comprehension is automatized to a certain extent.

Moreover, priming effects exist between idioms and words semantically or conceptually related to the idiomatic meaning (Cacciari & Tabossi, 1988; Sprenger, Levelt, & Kempen, 2006; Titone, Holzman, & Levy, 2002). This is evidence that, despite a certain degree of automatized comprehension, idiomatic meanings are semantically or conceptually accessed upon being encountered. Conceptual and semantic priming effects for conventional metaphors (which are often idiomatic) suggest that metaphoric structures affect processing (Coulson & Van Petten, 2002; Lai & Curran, 2013; Lai, Curran, & Menn, 2009). Ultimately, there is no study to our knowledge that examines the processing or perception of metaphoric or metonymic structures in idioms directly. It remains open whether metaphoric or metonymic structures are in fact processed in idioms or influence processing ease and, more fundamentally, whether metaphoric idioms are even differently perceived than metonymic idioms. As metaphor and metonymy differ conceptually and in how easy they are to comprehend, it seems likely that they also differ in their degree of non-literality. As non-literality is generally seen as a crucial property of idioms, exploring the properties of metaphoric compared to metonymic idioms should include exploring their non-literality. This study aimed to distinguish metonymic from metaphoric idioms by this very property. We performed four separate rating surveys to investigate whether native speakers perceive a difference in non-literality between metonymic versus metaphoric idioms, and how non-literality is linked to other idiom-typical properties.

1.1. NON-LITERALNESS IN METONYMIC AND METAPHORIC IDIOMS

Metaphors link distinct semantic concepts or domains (Barcelona, 2003; Lakoff & Turner, 1989; Sweetser, 1990) while metonymies work within one semantic concept or domain. Thus it seems that what is said is likely cognitively or semantically closer to what is meant in a metonymy than in a metaphor. This implies that there is a systematic difference in how literal metonymies are in comparison to metaphors. Consider someone expressing that a friend has a unique sense of fashion by saying *Sarah always swims against the current*. The ‘water current’ stands for the literally unrelated ‘sense of fashion’, and the swimming defines the action and adds meaning in the context of the water current. Hence the speaker does not intend this sentence to be understood literally at all. It can be concluded that this idiom – when used in its idiomatic sense – is highly non-literal. Consider the speaker wanting to express that a friend has a talent for being attentive to minor things: *Sarah has an eye for detail*. In this situation, the speaker uses *eye* to refer to a skill that is directly related to seeing or discovering by looking. The physical organ ‘eye’ – as in *eyeball*, *eyelid*, etc. – is not the strictly intended meaning in this sentence but it is the core ingredient to the mentioned skill. Hence this sentence is more literal than *Sarah always swims against the current*.

Comparing metaphoric and metonymic structures, it is expected that metaphoric idioms be perceived as more NON-LITERAL in this sense than metonymic idioms. Non-literality here is very similar to ‘figurativeness’, as metaphor and metonymy are generally described as kinds of figurative language. A plethora of linguistic expressions can be imagined as being positioned on a continuum of literalness with a literal and a figurative extreme, instead of within a clear-cut dichotomy of literal language on the one hand and figurative language on the other (for the difficulty of a clear distinction between literal and figurative, see also Gibbs & Colston, 2012, Chapter 2). ‘Figurative’ already implies a fairly high degree of ‘not literal’, and would not encompass language lower on the non-literality scale, such as any rather (but not completely) literal expressions. Metonymies (*money changes hands*) can in fact be close to the literal pole and resemble strictly literal language (*She washes her hands with soap*) more than very figurative language (*They are hand in glove with each other*). The term ‘non-literality’ seems more fitting than ‘figurativeness’ to capture the nature of many metonymic idioms especially.

1.2. PSYCHOLINGUISTIC CHARACTERISTICS OF IDIOMS

Many studies suggest that the processing advantage of idioms is driven by a number of properties, especially those related to semantic and cognitive processing (see also Nunberg, Sag, & Wasow, 1994). We chose five properties as key to the semantic and cognitive processing of idioms. While we may expect non-literality to be influential, familiarity, transparency (which we see as having the two aspects of comprehensibility and relation), and length are known to impact processing difficulty in idioms and other fixed expressions. In our study, metonymic and metaphoric idioms are rated on these properties, except for length, which is recorded in number of words.

1.2.1. *Familiarity*

Whether an idiom is familiar to a hearer strongly impacts processing ease (see Cronk & Schweigert, 1992; Gibbs, 1980; Libben & Titone, 2008; Nippold & Taylor, 1995, 2002; Schweigert, 1986; Tabossi, Fanari, & Wolf, 2009; Titone & Connine, 1994). Familiarity refers to how well known an idiom is to readers, which is driven by how frequently they encounter it. Processing models such as the Superlemma Hypothesis (Sprenger et al., 2006) or the Graded Salience Hypothesis (Giora, 1997, 2003) also assume familiarity to be one of the key properties determining processing difficulty. For example, highly conventional or familiar metaphors are processed as fast as literal sentences (Iakimova, Passerieux, Laurent, & Hardy-Bayle, 2005). Conventional or familiar metaphors also have a processing advantage

compared to less conventional and familiar ones (Blasko & Connine, 1993), and to novel ones (Bowdle & Gentner, 2005; Coulson & Van Petten, 2002; Lai & Curran, 2013). They also score higher in ratings on meaningfulness (Gildea & Glucksberg, 1983; Glucksberg, Newsome, & Goldvarg, 2001) and coherence (Giora, Fein, Kotler, & Shuval, 2015). As Citron, Cacciari, Kucharski, Beck, Conrad, and Jacobs (2016) point out, familiarity is a more adequate and reliable measure for idioms than frequency counts. It indicates how frequently an individual has been exposed to a particular idiom and is thus primarily a subjective and direct measure.

1.2.2. *Transparency*

‘Transparency’ is a broad term and has not been defined consistently in idiom research. Some contend transparency to be very closely related or identical to decomposability (Abel, 2003; Gibbs, Nayak, & Cutting, 1989; Gross, 1996; Zwitserlood, 1994), i.e., the degree to which an idiom’s figurative meaning can be ascertained from its component words (Libben & Titone, 2008; Nordmann & Jambazova, 2016). Others clearly demonstrate how idioms can be both non-decomposable and transparent (Nunberg et al., 1994) or show how transparency and decomposability correlate in ratings, but are different properties (Carrol, Littlemore, & Gillon-Dowens, 2018). On a second account, transparency is the ease with which an idiom can be comprehended (Boers & Demecheleer, 2001). It has also been shown that transparency is connected to the motivation of an idiomatic meaning. Motivation is often grounded in physical or other concrete experiences of language users. If a metaphor or metonymy in an idiom is recognized and combined with world knowledge, then an idiom can become relatively transparent (Boers & Webb, 2015). A third common account defines transparency as the relatedness between the literal and the non-literal meaning of an idiom (Cacciari & Glucksberg, 1995; Nippold & Taylor, 1995, 2002; Titone & Connine, 1999). Most of these authors have shown that transparency affects processing ease (for example, as found partly in adults and especially in children and adolescents by Cain, Towse, & Knight, 2009; Nippold & Duthie, 2003; Nippold & Rudzinsky, 1993; Nippold & Taylor, 2002). These transparency accounts may overlap yet address slightly different aspects of idioms.

1.2.2.1. *Comprehensibility*. We refer to the first aspect as ‘comprehensibility’. This is best captured by the following definition as “the ease with which the meaning of an idiomatic unit can be recovered” (Nunberg et al., 1994, p. 498; Boers & Demecheleer, 2001) while the motivation “needn’t be etymologically correct” (Nunberg et al., 1994, p. 498). It is equivalent to “comprehensibility” as in Katz, Paivio, Marschark, and Clark (1988) and to “meaningfulness” as in Titone and Connine (1994). An idiom is transparent “if [a reader or listener]

feels that there is a motivated relationship between the expression and its meaning” (Keysar & Bly, 1999, p. 1562). Individual cognition and experience differs, thus individuals tend to make sense of idioms in individual ways. Thus an individual may find an idiom quite easily comprehensible even though they might deem the relation between its literal and idiomatic meaning distant, which addresses a slightly different aspect of transparency.

1.2.2.2. Relation. The second aspect of transparency is the relatedness between the literal and the non-literal meaning of an idiom (Cacciari & Glucksberg, 1995; Nippold & Taylor, 2002; Titone & Connine, 1999). This notion defines the strength of the semantic or conceptual link between the literal and the idiomatic meaning of an idiom, which in our survey is referred to as ‘relation’. Literally, *to give the word* would thus be closely related to its meaning ‘to inform or notify somebody’, because *word* is immanently related to speaking and writing. In contrast, the literal *to wrap somebody around one’s finger* is very distantly or not at all related to its meaning of ‘purposefully seduce somebody’ as there is no obvious semantic link between the parts of the literal and the non-literal meaning.

Relation encourages a rather analytical perspective where each idiom is compared to its meaning and in terms of how close that relationship is. Judgment is thus based on a comparison of the idiomatic meaning with its literal meaning. High comprehensibility does not necessarily dictate a close relation, although a relationship between these two properties seems likely. Comprehensibility and relation thus capture the nature of transparency from two angles. Participants as young as ten years old are easily capable of rating relationships between idioms and meanings on a 3-step scale (Nippold & Taylor, 2002).

1.2.3. *Non-literality*

Non-literality is the key property researched here and is defined as the degree to which an idiom is non-literal. At the time of the design and execution of this study, it had not been considered as such in norming or processing studies of idioms. When ‘literality’ is a criterion in processing or rating studies, it is either defined as “how often an idiom phrase is used literally” (Cronk & Schweigert, 1992, p. 134), as literal plausibility (Bonin, Méot, & Bugaiska, 2013; Nordmann & Jambazova, 2016; Titone, Holzman, & Levy, 2002), or as an idiom’s potential literal interpretation (Libben & Titone, 2008; Titone & Connine, 1994), which all deviate from our definition. Katz et al.’s (1988) metaphor rating study is the only one to our knowledge that uses the term ‘metaphoricity’ equivalently to ‘non-literality’ as used here, and thus defined as the degree to which sentences are “literally or figuratively true” (Katz et al., 1988).

Citron et al. (2016) were the first to introduce ‘Metaphorizität’ (translated in English as ‘figurativeness’) in a German idiom norming study. This term

carries the assumption, however, that all idioms not strictly literal are metaphoric. Since we subdivide idioms according to metonymic versus metaphoric structure, this term is inapplicable to our study. Moreover, non-literality is not necessarily metaphoric: not only can it be metonymic, it can also be ironic, litotical (i.e., understated), hyperbolic, etc. For this reason, 'Metaphorizität' is too narrow to be synonymous to non-literality and it is possible that ratings of metaphoricality would differ from those of non-literality. The term 'non-literality' also has a practical advantage over 'figurativeness' for this study: 'literal/non-literal' are commonly used and well-known concepts in everyday language and life; speakers have clear intuitions about (non-)literally used language. From a young age on, children are necessarily exposed to it as non-literality permeates every sphere of language and everyday communication, and can be a topic of natural conversation itself. Accordingly, non-literality (or rather its opposite 'literality') has a second-nature status to speakers which sets it far apart from relation. 'Figurativeness', on the other hand, is a substantially less intuitive, natural, and well-known concept. As this study attempts to gather speakers' most natural and intuitive judgments, 'non-literality' seems to be the more adequate choice.

Similar reasons apply to the demarcation between relation and non-literality. They are related properties, and correlation is expected, but several differences make a division useful: once more, similarly to figurativeness, the relation between two meanings is a rather unfamiliar concept to a layperson and uncommon in life outside the language sciences. Second, relation focuses on a comparison. The strength of the link between two meanings requires considering both equally and analyzing their connection, while non-literality focuses on the idiom itself.

1.2.4. *Idiom length*

Length is usually a highly effective predictor in reading (as in Bonin et al., 2013; Just, Carpenter, & Woolley, 1982; Michl, unpublished observations) and in lexical decision (Ferrand et al., 2010). It is unclear whether it affects the perception of idiom properties, while so far the only property it correlates with in idioms seems to be predictability (Bonin et al., 2013; Tabossi, Arduino, & Fanari, 2011).

2. Method

There were several aims of this study. One was to collect idiom norms to be matched for use in semantic and cognitive processing studies. Another was to test whether processing differences in metonymies and metaphors are also mirrored in the perception of idioms. A third aim was to examine the relationships between these properties, as several other idiom rating studies

have done so far partly for other properties. In particular, it was assumed that familiarity ratings could influence ratings of the other properties, thus positively correlating with other properties, as has in fact been found for literality and decomposability (Nordmann, Cleland, & Bull, 2014), and transparency (Carrol et al., 2018). For these reasons, a large rating study was conducted in the form of four separate surveys, each containing instructions with examples to rate metonymic and metaphoric idioms on only one of four properties: familiarity, comprehensibility, relation, and non-literality.

At the time of planning and executing this study, no database with speaker ratings for German idioms existed. Today this is still the only idiom study that focuses directly on non-literality, which has not been a criterion in idiom norming studies before, as defined here, even though it is apparently seen as a key property of idioms, judging from idiom definitions. An additional, new feature of this database is the subdivision of idioms into metaphoric and metonymic idioms, thus a division by conceptual structure. It aims to shed light on the properties of metonymies and metaphors as perceived by native speakers with various backgrounds and to provide grounds for comparing the processing of metonymic and metaphoric idioms.

Finally, the present study is also more extensive than prior norming studies: it provides more ratings on each property than many other idiom studies (61–96 each on at least 122 idioms and 320 in the familiarity survey, as it contains 76 extra literal idioms outside the scope of this paper) and 397 participants with a wide range of age, educational and professional backgrounds, as compared to the other idiom rating studies.¹

2.1. MATERIAL

We selected all idioms from a modern German idiom dictionary (Schemann, 2011). The preliminary and explorative database comprised 1,800 idioms having a number of properties. An idiom was selected if it could be classified as metonymic or metaphoric, and if the metonymy or metaphor lay within the noun and not the verb alone. Furthermore, we selected idioms that consisted of 2 to 7 words (84% consisted of 3–5 words), had the syntactic structure VP+NP, VP+PP, or, rarely, VP+NP+PP, and that could be embedded in an indicative sentence and then end in a noun. This excludes stand-alone exclamations, proverbs, and other full sentences. We additionally excluded

[1] Five existing idiom rating studies are summarized in terms of quantity: Bonin et al. (2013): 27 ratings on 305 idioms, 187 participants in total; Nordmann and Jambazova (2016): 30 or more ratings per property on 90–100 idioms, 150 and 196 participants, respective of study; Titone and Connine (1994): 28–30 ratings from 171 to a final 40 idioms; Citron et al. (2016): at least 30 ratings on 619 idioms, 249 participants in total; Tabossi et al. (2011): 40–43 ratings per property on 245 idioms, 740 participants in total, all from school or university.

all idioms that were ironic, strongly hyperbolic, or strongly litotical, i.e., understated, very rare, old-fashioned, or very modern, dialectal, or regional, or very vulgar in nature. Idioms containing expletives, abusive, or onomatopoeic words, brand names and other neologisms were also excluded. Last, we excluded idioms if they required any specific background knowledge to be understood, if they were used only in specific jargons, or could only be made sense of in context or through knowledge of speaker-inherent intentions. We thus attempted to keep the selected idioms not specifically tied to a particular register, language style, or manner, comprehensible on their own, and familiar to the average adult German native speaker. Last, we excluded many idioms that used the same words, to avoid presenting participants with too many highly similar idioms (for example, 12 idioms containing the word 'eye'). This left 244 idioms, consisting of 87 metonymic and 157 metaphoric idioms. Please see Table 1 for examples.

Excluding neologisms, metonymic idioms mostly have a meronymic structure, specifically our idioms usually follow the PART FOR WHOLE paradigm (as opposed to the meronymic structure WHOLE FOR PART and the antonymic structures PRODUCER FOR PRODUCT or BRAND NAME FOR GENERIC). Examples are to *pay something out of one's own pocket* (*pocket* for 'belongings'), *have an evil tongue* (*evil tongue* for 'inclination to speak ill'). The classification into metonymic and metaphoric was done by the author and 17 independent raters, half of whom work in language-related fields. The 17 raters received working definitions of metonymy and metaphor referred to as idiom type 'A' or 'B', and classified their shares of idioms into either category 'A' or category 'B'. Every idiom received 5 classifications by 5 different people, namely the author plus 4 different raters. The classification was deemed successful if at least 80% of the classifications agreed on the type of idiom. Average agreement between each individual rater and the author was 86%. The amount of idioms reaching 80% classification agreement was 84% (89% in metonymic idioms, 80% in metaphoric idioms) while 57% of idioms reached 100% agreement (67% in metonymic, 48% in metaphoric idioms). In the surveys, participants were unaware of idiom type.

2.2. PROCEDURE

Each of the four surveys contained one variable to be rated from familiarity, comprehensibility, relation, or non-literalness. They were conducted separately to simplify the task and to avoid any bias in later ratings stemming from earlier ratings of another property. Moreover, letting participants rate all properties in one session not only increases each individual's impact on the results, but also cannot control whether they properly switch from rating one property to rating the next. Finally, explaining several unfamiliar properties after one another might make the task rather strenuous and confusing, which

TABLE 1. *Examples for the idiom types*

Idiom type	Example	Translation
Metonymic		
1)	<i>jmd. knurrt der Magen</i>	somebody growls the stomach
2)	<i>keine Papiere haben</i>	somebody is clearly hungry to have no papers
3)	<i>einen Luftsprung Machen</i>	to have no legal documents to make an air jump
4)	<i>die Nase rümpfen</i>	to jump for joy to wrinkle one's nose
5)	<i>etw. unter vier Augen besprechen</i>	to be contemptuous or disgusted to discuss something among four eyes
6)	<i>mit leeren Händen kommen</i>	to discuss something in private to come with empty hands to arrive empty-handed (without contribution)
7)	<i>sich die Kugel geben</i>	to give oneself the bullet to commit suicide by gun
8)	<i>ein offenes Wort sprechen</i>	to speak an open word
Metaphoric		
1)	<i>ans Licht kommen</i>	to come to light
2)	<i>das Eis brechen</i>	to break the ice to relax a socially stiff/uncomfortable situation
3)	<i>die Hand für jmd. ins Feuer legen</i>	to put one's hand in fire for somebody to vouchsafe for somebody's virtue
4)	<i>jmd. um den Finger wickeln</i>	to wrap somebody around one's finger
5)	<i>unter keinem guten Stern stehen</i>	to not stand under a good star to be ill-fated
6)	<i>im goldenen Käfig sitzen</i>	to sit in the golden cage to be wealthy or fortunate yet bound and unfree
7)	<i>noch feucht hinter den Ohren sein</i>	to still be moist behind the ears to be young and inexperienced
8)	<i>jmd. Dampf machen</i>	to make steam at somebody to impel somebody, usually at work

would decrease data quality. For these reasons, we used a between-subjects design, while the main goal was to recruit as many participants as feasible.²

[2] As very many participants were needed and it was not always possible to trace whether a particular participant had followed a particular invitation; some participants received two invitations to surveys. These were spaced several weeks up to three months apart to minimize any possible effects on rating behavior from completing their first survey. Furthermore, participants had to report partaking in another survey. Participants were usually not invited to partake in two consecutive surveys, and participation in both transparency surveys was not possible. In the end, 46 out of 397 participants completed two surveys. A comparison of a between- versus a within-subjects study showed no difference in property correlations of ratings (Nordmann & Jambazova, 2016).

The four surveys were released consecutively over a four-month period, with gaps of several weeks before and after each survey. We recruited participants via various channels, including e-mail, notices, and pamphlets. In each survey, the material was divided into two lists to decrease the load. Participants were randomly assigned to one list of each survey. Upon completion of that list, they could choose to continue with their second list or quit. Thus, completion of each questionnaire took 20 to 40 minutes, and each participant rated either 122 or 244 idioms on one variable, the number depending on their own choice. We designed and presented the surveys as questionnaires on the non-commercial social studies online platform *Sosci Survey*, Version 2.5.00-I (Leiner, 2014). The idioms were presented in individually randomized order on several consecutive web pages.

In each survey, participants had a five-point Likert scale. In all surveys except the familiarity survey, meanings of the 244 metonymic and metaphoric idioms were presented alongside the idioms to ensure that participants based their answers on the correct meanings, as they can have incorrect knowledge of a meaning, but be confident about it (see also Citron et al., 2016). In the familiarity survey, meanings were not given, to avoid biasing participants. They were asked to consider how often they encountered an idiom. They could answer on a five-point Likert scale reaching from encountering the idiom 'hardly ever' to 'very frequently', or instead choose the answer 'never encountered it before'. In the comprehensibility survey, participants rated how easily comprehensible they found each idiom on a scale from 'extremely difficult to understand' to 'extremely easy to understand'. In the relation survey, participants rated how closely related each literal meaning was to its idiomatic meaning on a scale from 'extremely distantly or not at all related' to 'extremely closely related'. In the non-literality survey, participants rated how literal each idiom was in comparison to its given meaning on a scale from 'extremely literal' to 'not at all literal'. Besides addressing different features of idioms, the relation study asked for an analytical perspective on an unfamiliar property, whereas the non-literality study allowed for a more intuitive answer on a familiar property.

2.3. PARTICIPANTS

410 early German monolingual participants participated in the study. They received either course credit or could partake in a raffle for €20 vouchers for online stores. 397 (39% male) were included in the analyses (96 for familiarity; 86 for comprehensibility; 111 for relation; 104 for non-literality). Two participants were excluded as they were only 10 and 13 years old; the rest was excluded due to 'fast-clicking', i.e., clicking through the questionnaires in 8 minutes and under, or almost consistently choosing

the same rating response ($SD < 0.6$ scale steps). The remaining participants were 18 to 94 years of age; 50% were 25 to 54 years old ($M = 38.5$, $SD = 16.1$). 47% were up to 30, 86% were up to 60 years old.

Participants were asked for their origin or their place of residence, depending on what they defined as their home. They came from every federal state of Germany, except for the Saarland. 39% came from Southern Germany, the most densely populated area; 38% came from the North Eastern federal states (the former DDR).

Of the 397 participants, 47% gave as their highest educational degree a university or college degree, 28% held a 12- or 13-year high school diploma, 18% held a secondary school 10-year diploma, a vocational baccalaureate diploma, or a completed apprenticeship, 6% held a PhD or higher, 2% were still in school, held a 9-year high school diploma (the lowest in Germany), or provided no answer.

2.4. ANALYSIS

We conducted most analyses with R Version 3.3.2 (R Core Team, 2016) and RStudio Version 1.0.136 (RStudio Team 2016) and the descriptive statistics with Microsoft Excel (Version 14.0.7229.5000; Microsoft Corporation, 2010). Means and standard deviations were obtained for all properties and each idiom. Length was calculated in number of words. Kendall's tau-b correlation tests were conducted to examine the relationship between the four different properties, using the R package 'Kendall' 2.2 (McLeod, 2015). This correlation test was chosen as it is rank-based on median ratings and works well with ties which can be common with medians.

Since the first key aim was to examine whether the properties could predict idiom type (of which participants were unaware), we performed a binomial logistic regression to predict metonymic or metaphoric idioms from ratings.

Mixed effect regressions were used to test participant effects on their ratings (Baayen, Davidson, & Bates, 2008; Field, Miles, & Field, 2014) because they allow accounting for inter-individual variances between participants as well as items within a single model. In this case, ordered cumulative logit models were fitted with the R package 'ordinal' 2018.8-25 (Christensen, 2015, 2018). Multicollinearity was tested using the R package 'fmsb' 0.6.3 (Nakazawa, 2018). Scripts, supplementary material, and concise idiom norms are available.³

[3] Files are available either at https://osf.io/uryfa/?view_only=ca27dfb5ff654ac9bc17523c2ccd8f1f or upon request to the author.

3. Results

3.1. DESCRIPTIVE STATISTICS

Familiarity, comprehensibility, relation, and non-literality all had average ratings of 3.1 to 3.8 for all idioms. The answer ‘never encountered the idiom before’ was chosen in less than 1% of cases, which were excluded from the analysis. Please see Table 2 for an overview of means and standard deviations. In terms of familiarity, the metonymic idioms ($M = 3.5$, $SD = 0.18$) had the same average ratings as the metaphoric idioms ($M = 3.5$, $SD = 0.21$). The comprehensibility ratings were higher for metonymic ($M = 4.1$, $SD = 0.22$) than metaphoric ($M = 3.6$, $SD = 0.21$) idioms. The relation ratings were distinctly higher for metonymic ($M = 3.8$, $SD = 0.18$) than metaphoric ($M = 2.8$, $SD = 0.13$) idioms. The non-literality ratings showed an even larger difference, with metonymic idioms ($M = 2.5$, $SD = 0.17$) being rated as much less non-literal than metaphoric idioms ($M = 4.0$, $SD = 0.19$). For a quick overview, Table 3 lists t -tests comparing mean values of the properties in metaphoric versus metonymic idioms. It can be seen that metonymic and metaphoric idioms do not differ in familiarity ($p = .53$), while they differ significantly with respect to comprehensibility, relation, and non-literality.

3.2. RELATIONSHIPS AMONG PROPERTIES

To keep the family-wise error rate low, we set the significance level for the correlation analysis to $\alpha = 0.001$. First, familiarity was compared to all other properties. Familiarity and comprehensibility showed a significant correlation ($r_b = 0.35$, $p < .001$). This connection with comprehensibility may express the phenomenon that more familiar phrases are likely more easily comprehensible. (Participants themselves were asked to gauge how strongly their familiarity with idioms influenced their ratings of other properties. Please see the ‘Appendix’ for a discussion.)

TABLE 2. *Descriptive statistics for each rated or calculated variable^a*

Measure	All idioms				Metonymic		Metaphoric	
	Variable	M	SD		M	SD	M	SD
Rating	familiarity	3.5	0.17		3.5	0.18	3.5	0.16
	comprehensibility	3.8	0.23		4.1	0.22	3.6	0.21
	relation	3.1	0.17		3.8	0.18	2.8	0.13
	non-literality	3.5	0.19		2.5	0.17	4.0	0.19
	length	4.2	1.11		4.3	1.14	4.2	1.1

NOTES: ^a Based on mean values for each idiom. The higher the rating, the more familiar, the more comprehensible, the closer the relation, the more non-literal.

TABLE 3. *Welch two sample t-tests on mean ratings of metaphoric and metonymic idioms*

Variable	t	Metonymic	Metaphoric	p
		M	M	
familiarity	-0.63	3.47	3.52	.53
comprehensibility	7.33	4.06	3.63	< .001***
relation	13.12	3.77	2.80	< .001***
non-literalness	-16.75	2.54	4.02	< .001***

Other correlations with comprehensibility were stronger, as values for comprehensibility and relation ($\tau_b = 0.63$, $p < .001$), and comprehensibility and non-literalness ($\tau_b = -0.53$, $p < .001$) indicated. The correlation of comprehensibility and relation showed that the more easily comprehensible an idiom was rated, the more closely related it was to its meaning. The moderate negative correlation between comprehensibility and non-literalness revealed that the more easily comprehensible an idiom, the more literal (the less non-literal) it tended to be.

The strongest correlation was obtained for relation and non-literalness ($\tau_b = -0.77$, $p < .001$), which indicated that the more distant the relationship between idiom and meaning is, the more likely it was perceived as highly non-literal. No significant or practically meaningful correlations were found between number of words (length) and the other properties. For a numeric and visual overview of correlations, please see Table 4 and Figure 1.

3.3. PREDICTING IDIOM TYPE

One key question was which of the properties could best predict idiom type. To answer this question, a binomial logistic regression was conducted with idiom type as dependent variable and familiarity, comprehensibility, relation, and non-literalness as numeric independent variables, based on medians.⁴ Idiom length, as recorded by number of words, was additionally used as an independent variable to make sure that the material was balanced according to idiom type. In testing for multicollinearity between these properties, all variance inflation factors were below 4.31, indicating no strong multicollinearity.

[4] In modelling ordinal data, the median is recommended as the more appropriate measure than the mean, because scalar intervals may not be evenly spaced and interpreted by raters in different ways, and because the median is not affected by outliers and skewness (Allen & Seaman, 2007; Lund & Lund, 2013), although preferences of median versus mean differ (for example, Sauro, 2016).

TABLE 4. *Correlations of properties*

	Familiarity	Comprehensibility	Relation	Non-literalness	Length in wds
a) Correlation of properties in all idioms					
familiarity	1	0.35***	0.11*	-0.12*	0.02
comprehensibility		1	0.63**	-0.53***	0.06
relation			1	-0.77***	0.04
non-literalness				1	0.03
length in words					1.00
b) Correlation of properties in metonymic idioms					
familiarity	1	0.33***	0.25**	-0.12	0.05
comprehensibility		1	0.69***	-0.55***	-0.00
relation			1	-0.79***	0.06
non-literalness				1	0.02
length in words					1.00
c) Correlation of properties in metaphoric idioms					
familiarity	1	0.41***	0.09	-0.12*	0.00
comprehensibility		1	0.56***	-0.43***	0.09
relation			1	-0.62***	0.04
non-literalness				1	0.05
length in words					1.00

Regression outputs can be found in Table 5. No effect was found for idiom length, which shows that the two kinds of idioms were balanced in terms of length. Thus length was excluded from further analysis. Familiarity, comprehensibility, and relation did not turn out to be significant predictors ($p > .05$). Non-literalness was the strongest predictor for metaphoric versus metonymic idioms, as indicated by its large effect size ($b = 2.25$) which was significant ($p < .001$). Indeed, metonymic idioms were often rated as highly literal and rarely as very non-literal, whereas the opposite was found for metaphoric idioms: a one-unit increase in non-literalness, i.e., an idiom being one unit ‘more non-literal’, means that this idiom is nine times more likely to be a metaphoric than a metonymic idiom. McFadden R^2 indicated 55% of variance to be covered by familiarity, comprehensibility, relation, and non-literalness. Please see Figure 2 for effect sizes.

3.3.1. *Predictive value*

The predictive value of the effects was tested, based on the full model. We tested whether idioms were correctly predicted to be metonymic or metaphoric. This was done by calculating a percentage of probability to which an idiom would be either metonymic or metaphoric. If the percentage was over 50% for the correct type, prediction was considered correct. Correctness was high: 85% of metonymic idioms were correctly predicted

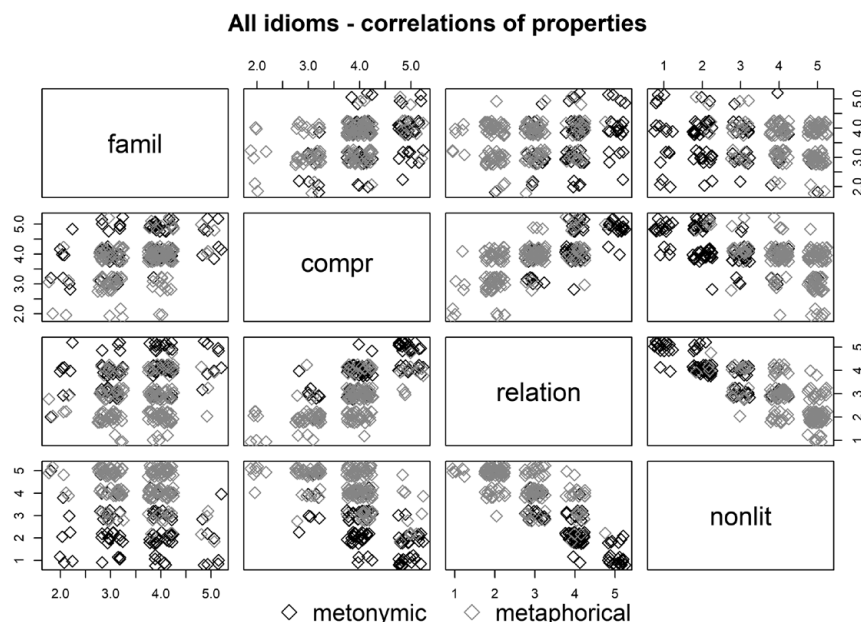


Fig. 1. Correlations are based on medians (ratings 1, 2, 3, 4, and 5). To make the individual items visible and avoid overplotting, the datapoints are jittered.

to be metonymic, while 87% of metaphoric idioms were correctly predicted to be metaphoric.

3.3.2. *Non-literality versus relation*

Two results suggest that relation would likely be a strong predictor of idiom type on its own if non-literality was excluded: one, the rather strong correlation between non-literality and relation; two, the absence of an effect of relation in predicting idiom type when non-literality was held constant. Thus, idiom type was once fit with only familiarity, comprehensibility, and relation, and compared to a model with familiarity, comprehensibility, and non-literality. As shown in Table 5b, relation then became a significant predictor ($p < .001$): if the relation rating decreased by one unit (i.e., if the relationship between idiom and meaning was rated as one unit more distant), this idiom was nine times more likely to be metaphoric than metonymical. Thus the relation effect was almost as strong as the non-literality effect in the full and reduced model (Table 5). However, the relation model had much lower fit than the non-literality model (as indicated by Akaike information criterion (AIC), Bayes information criterion (BIC), and residual deviance which all showed a difference of 58.4, which is large for these criteria).

TABLE 5. *Results from the top binomial logistic regression models 95% CI for odds ratio*

Predictor of idiom type	β (SE)	lower	odds ratio	upper	p
a) metaphoric versus metonymic	full model McFadden R ² : 0.55 Res. deviance = 142.8, AIC = 152.8, BIC = 170.3, Log-likelihood = -71.4				
familiarity	0.55 (0.36)	0.87	1.74	3.59	.12
comprehensibility	0.96 (0.51)	0.97	2.6	7.2	.059
relation	-0.41 (0.43)	0.29	0.66	1.56	.34
non-literalness	2.25 (0.37)	4.83	9.46	20.63	< .001***
b) metaphoric versus metonymic	model without non-literalness (relation model) McFadden R ² : 0.36 Res. deviance = 202.2, AIC = 210.2, BIC = 224.2, Log-likelihood = -101.1				
familiarity	0.38 (0.28)	0.84	1.46	2.55	.18
comprehensibility	0.48 (0.40)	0.73	1.62	3.58	.08
relation	-2.17 (0.31)	0.06	0.11	0.20	< .001***
c) metaphoric versus metonymic	model without relation (non-literalness model) McFadden R ² : 0.55 Res. deviance = 143.8, AIC = 151.8, BIC = 165.7, Log-likelihood = -71.9				
familiarity	0.38 (0.28)	0.95	1.88	3.84	.18
comprehensibility	0.48 (0.40)	0.86	2.07	5.07	.08
non-literalness	-2.17 (0.31)	6.39	11.35	22.67	< .001***

NOTES: Dependent variable: idiom type, reference level = metonymy; effect size indicates how much higher ratings are on average, comparing metaphoric to metonymic idioms.

Last, McFadden R² showed that the model without non-literalness (but with familiarity, comprehensibility, and relation) together covered 36% variance, whereas the model without relation (but with familiarity, comprehensibility, and non-literalness) together reached a value of 55% (see Table 5c). Given these large differences, we conclude that non-literalness and relation are not interchangeable: relation is a considerably weaker and less exact predictor for idiom type than non-literalness.

3.4. INDIVIDUAL DIFFERENCES

For each survey, ratings were fitted as a function of sex, education, and origin as categorical fixed effects, age as a numeric fixed effect, and random intercepts for participants and items. Models were fitted with all levels of each factor, as well as with collapsed levels of each factor, looking for large-scale effects of education (low, medium, and high), residential background (such as North, South, East, or West of the country, as well as effects of a North–South or East–West divide). No effects could be found. As age and education may

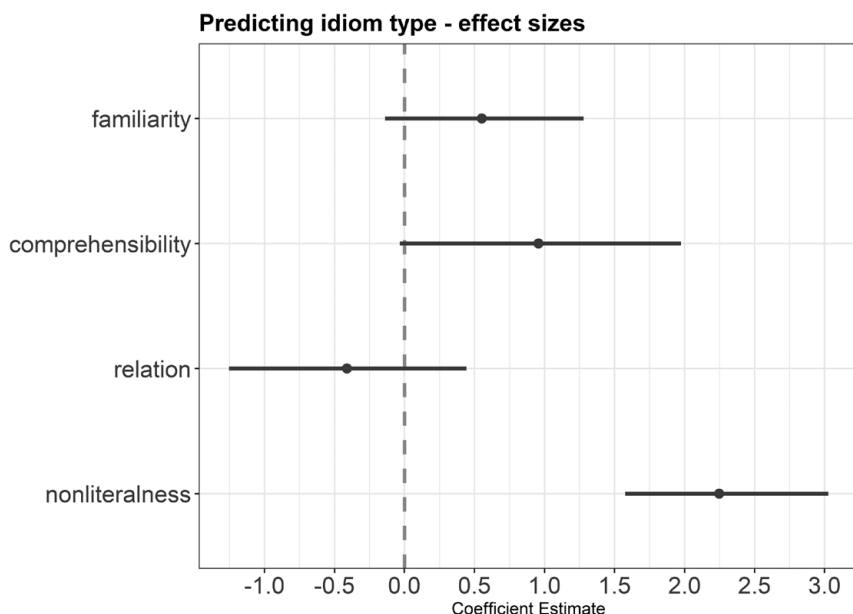


Fig. 2. Effect sizes of the four rated properties in the regression predicting idiom type.

correlate to a certain extent, an interaction term was compared to the individual effects in likelihood ratio tests for all four properties. No effects were found and the interaction term actually decreased model fits in three out of four surveys. With regards to the familiarity ratings, this indicates that the idioms presented were equally known across age and the country.

4. General discussion

This idiom rating study asked adult German native speakers to rate familiarity, comprehensibility, relation, and non-literality. We analyzed correlations between properties, property effects on predicting idiom type, non-literality ratings, and participant effects on ratings.

Our most interesting finding is that non-literality is the only strong and reliable predictor in predicting idiom type. Our results clearly suggest that the degree of non-literality predicts whether an idiom is metonymic or metaphoric. Metonymic idioms were rated as more literal than metaphoric idioms. Our findings indicate that degree of non-literality, and type of non-literality (metonymic or metaphoric), as well as relation are closely connected: an extremely non-literal idiom with very close relation to its meaning, or vice versa, is unlikely. Our data also suggest that we are rather unlikely to find an extremely non-literal metonymic idiom or an extremely

literal metaphoric idiom, a metonymic idiom very distantly related to its meaning or a metaphoric idiom extremely closely related to its meaning.

As discussed earlier, metonymies are easier and faster to process than metaphors. The rating study at hand indicates that the reason may be the higher literalness. This would make non-literalness an influential factor on processing difficulty. Non-literalness is connected to the other properties to varying degrees. Our result suggests that familiarity is roughly equally typical in rather literal and non-literal idioms. This result is not directly comparable to Katz et al. (1988), who found a somewhat strong correlation between metaphoricity and ‘felt familiarity’ (0.74) for literary metaphors: first, results are likely different for metaphors occurring in poetry and fiction compared to idioms which are highly lexicalized, fixed expressions largely present in everyday language. Second, felt familiarity measured the frequency with which ‘ideas’ in the sentences were ‘experienced’ (Katz, et al., 1988, p. 197), which is not a measure of how frequently a specific metaphor in its precise wording is encountered, as it was instructed in our study. In addition, the metaphors by Katz et al. had the form X IS A Y, which is not at all a common form in idioms and did not occur in our materials.

While non-literalness correlates only moderately with comprehensibility, its negative correlation with relation is strong. Hence the more non-literal an idiom, the more distant is its relation to its meaning, and vice versa. This is unsurprising, as transparency seems to be connected to the motivation of idiomatic meanings. Following Boers and Webb (2015), motivating an idiomatic meaning “involves an appreciation of the correspondence between a literal reading of the expression and the idiomatic, figurative meaning (...) If the literal reading is congruent with the idiomatic meaning (...), then the scene evoked by the literal reading can render the expression semantically transparent” (p. 370).

Citron et al. (2016) also found figurativeness and semantic transparency to be negatively correlated. Their finding is not strictly identical to ours for two reasons: one, Citron et al. allowed participants to rate more than one variable in one long session, which might produce different results; second, in contrast to the definition adopted here, they define transparency as decomposability, namely the degree to which an idiom’s meaning can be constructed from its parts. However, to the best of our knowledge, theirs is the only study that examines transparency in connection to figurativeness.

Nippold and Duthie (2003) examined transparency, defined here as relation, and its link to mental imagery in idioms. They found a significant difference in adults, who tended to produce more figurative mental images for transparent idioms and more literal ones for opaque idioms. While this does not directly relate to non-literalness, it shows a difference between transparent and opaque idioms with regard to their figurative interpretations. Other rating studies

that examined ‘literality’ define it as literal plausibility (Bonin et al., 2013; Nordmann & Jambazova, 2016; Tabossi et al., 2011; Titone et al., 2002). The meta-analysis by Nordmann and Jambazova (2016) demonstrated that there are no clear correlations between literality and other rated properties, which suggests that literality and non-literality are distinct properties functioning according to distinct principles.

Non-literality is connected to both comprehensibility and relation. Our findings indicate that metonymic idioms are more likely to be based on a close relationship between idiom and meaning, whereas the relationship between metaphoric idiom and meaning tends to be more distant. Relation, however, had no predictive value for idiom type as long as non-literality was controlled for. Without non-literality, relation was found to be a significant predictor with similar effect size for whether an idiom was metonymic or metaphoric, yet was significantly weaker and less exact than non-literality.

Comprehensibility is moderately correlated with relation, indicating that the more easily comprehensible an idiom, the more likely it is rather closely related to its meaning. The obvious implication is that a clear connection between idiom parts and their idiomatic meaning renders idioms more straightforward to native speakers (compare the examples translated from German: *to come with empty hands* and ‘to come without a gift or contribution’ versus *to make steam at somebody* and ‘to impel somebody’). The restriction to this finding is that we cannot examine whether these two properties are thus linked in a single rater. Yet there seems to be some tacit agreement across raters that high comprehensibility is moderately linked to close relation. This has also been concluded by Cacciari and Levorato (1998), Nippold and Duthie (2003), Nippold and Rudzinski (1993), and Nippold and Taylor (1995), who inferred that transparent idioms are generally easier to understand than opaque idioms. It seems that a clear connection between idiom parts and meaning, perceived as a close relation, tends to make idioms easily comprehensible. It should be borne in mind that the great majority of idioms in our study were rated as rather easily comprehensible, although there was a weak tendency to rate metaphoric idioms as more difficult.

As for a familiarity effect, our findings are partly in line with those of other studies. Nordmann and Jambazova (2016), for example, concluded from their own experiments that “the only thing that has any tangible effect on ratings is familiarity” (p. 205), which we did not find as such. Our participants’ own judgments of how their familiarity influenced their ratings of the other variables are mixed, but do not support a strong influence of familiarity (see ‘Appendix’). Nordmann and Jambazova’s meta-analysis of other idiom rating studies also found few strong correlations of familiarity with other subjective properties in both within- and between-subject designs. Our study cannot

examine directly whether participants' familiarity ratings influence their ratings of other properties, but our findings support an absence of correlations between familiarity and two other properties in idioms: a familiar idiom is not necessarily also highly literal, or closely related to its meaning. Comprehensibility is the only property with a moderate correlation with familiarity, as is also found to varying degrees in other idiom rating studies (Nippold & Taylor, 2002; Nordmann & Jambazova, 2016; for transparency, Carrol et al., 2018). This suggests that the more familiar an expression, the more often it has been encountered, thus the more easily comprehensible it may seem (see also Keysar & Bly, 1999). However, in our data, this is a tendency, not an absolute finding, and it might be present because a large number of idioms were found to be familiar. On the other hand, several idiom rating studies that have found a strong connection between familiarity and transparency also tested non-native speakers who were not familiar with all idioms. It is quite possible that L2 effects of familiarity on rating other properties are much stronger than L1 effects, especially because the degrees of familiarity differ much more strongly between L1 and L2 speakers than within L1 speakers alone.

For familiarity ratings, age has often been found to be a significant predictor in adolescents of different ages versus adults (for example, Chan & Marinellie, 2008), which is unsurprising. Nordmann and Jambazova (2016), however, found age in adults to be a marginally significant predictor in rating familiarity, meaning that older adults are more likely to rate idioms as familiar than younger adults. While this finding may seem plausible at first, we did not find such an effect, despite our large age range and our very similarly constructed material and analysis. The reason may be that older participants may have encountered all presented idioms more frequently during their lives and there may be fewer idioms unfamiliar to them. Still, it is very likely that they have encountered rare idioms much less frequently than highly frequent ones. That is, the absolute number of encountering all idioms may be much higher than for younger participants, but the relative differences remain similar. Our scale gave no absolute terms or numbers but asked for relative, not absolute comparisons; thus participants had to judge for themselves what 'extremely' or 'fairly familiar' meant to them.

Two other reasons for the diverging findings are possible. One: Nordmann and Jambazova (2016) used a 7-point Likert scale whereas our 5-point Likert scale might not have been sensitive enough to detect an age effect, although it has been shown that 5- and 7-point Likert scales render very similar or equivalent results and can easily translate into one another (Colman, Norris, & Preston, 1997; Dawes, 2008). Two: an age effect might be absent in this study because we randomized the order of idioms for every participant, whereas Nordmann and Jambazova (2016) did not. So it is theoretically possible that

older participants somehow responded differently to the specific order of idioms, although the authors argue that order effects are very unlikely.

Raters were not influenced in their judgments by how long an idiom was, demonstrated by the absence of significant correlations between length and all other properties. This is in slight contrast to Citron et al. (2016), who found a very weak, yet significant correlation between length and figurativeness. While their correlation coefficient is slightly larger than ours (0.13 versus 0.03), such small values should have no practical value.

Overall, the absence of demographic effects suggests that native speakers' judgment of all properties tested is not driven by their age, gender, home, or education, nor by the length of the idiom. Their judgments rather seem to be driven individually, as by a feeling for language, or other cognitive properties.

5. Conclusion

This idiom rating study examined the factor of non-literality. First, it structured non-literality in idioms by dividing them into metonymic and metaphoric. Second, it examined the influence of the degree of non-literality and found it to be the one crucial factor in predicting whether an idiom tends to be metonymic or metaphoric. Third, it found that non-literality is connected to different aspects of transparency. Our novel key finding is that metaphoric idioms are perceived as more non-literal than metonymic idioms.

Supplementary Material

To view supplementary material for this article, please visit https://osf.io/uryfa/?view_only=ca27dfb5ff654ac9bc17523c2ccd8f1f.

REFERENCES

- Abel, B. (2003). English idioms in the first language and second language lexicon: a dual representation approach. *Second Language Research* 19(4), 329–358.
- Ahmad, K., Brewster, C., Stevenson, M., & Wilks, Y. (Eds.) (2007). *Essays in honor of Yorick Wilks: words and intelligence: Vol. 2*. Dordrecht: Springer.
- Allen, I. E., & Seaman, C. A. (2007). Likert Scales and data analysis. *Quality Progress, Statistics Roundtable* 40(7), 64–65. Retrieved from <<http://rube.asq.org/quality-progress/2007/07/statistics/likert-scales-and-data-analyses.html>>.
- 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 and Communication Disorders* 44(6), 962–978.
- 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.
- Barcelona, A. (1997). Clarifying and applying the notions of metaphor and metonymy within cognitive linguistics. *Atlantis* 19, 21–48.
- Barcelona, A. (Ed.) (2003). *Metaphor and metonymy at the crossroads: a cognitive perspective*. Berlin/New York: Mouton de Gruyter.

- 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.
- Bartsch, R. (2002). Generating polysemy: metaphor and metonymy. In R. Dirven & R. Pörrings (Eds.), *Cognitive linguistics research: Vol. 20. Metaphor and metonymy in comparison and contrast* (pp. 49–74). Berlin/New York: Mouton de Gruyter.
- 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.
- Boers, F., & Demecheleer, M. (2001). Measuring the impact of cross-cultural differences on learners' comprehension of imageable idioms. *English Language Teaching Journal* **55**(3), 255–263.
- Boers, F., & Webb, S. (2015). Gauging the semantic transparency of idioms: Do natives and learners see eye to eye? In R. R. Heredia & A. B. Cieřlicka (Eds.), *Bilingual figurative language processing* (pp. 368–404). New York: Cambridge University Press.
- Bonin, P., Méot, A., & Bugajska, A. (2013). Norms and comprehension times for 305 French idiomatic expressions. *Behavior Research Methods* **45**(4), 1259–1271.
- Bortfeld, H., & McGlone, M. S. (2001). The continuum of metaphor processing. *Metaphor and Symbol* **16**(1/2), 75–86.
- Bowdle, B. F., & Gentner, D. (2005). The career of metaphor. *Psychological Review* **112**(1), 193–216.
- Brône, G., Feyaerts, K., & Veale, T. (Eds.) (2015). *Cognitive linguistics and humor research: applications of cognitive linguistics: Vol. 26*. Berlin/Boston, MA: De Gruyter Mouton.
- Cacciari, C., & Glucksberg, S. (1995). Understanding idioms: Do visual images reflect figurative meanings? *European Journal of Cognitive Psychology* **7**(3), 283–305.
- Cacciari, C., & Levorato, M. C. (1998). The effect of semantic analyzability of idioms in metalinguistic tasks. *Metaphor and Symbol* **13**(3), 159–177.
- Cacciari, C., & Tabossi, P. (1988). The comprehension of idioms. *Special Issue: Emerging Data Analysis* **27**(6), 668–683.
- Cacciari, C., & Tabossi, P. (Eds.). (1993). *Idioms: processing, structure, and interpretation*. New York/Hove: Psychology Press.
- Cain, K., Towse, A. S., & Knight, R. S. (2009). The development of idiom comprehension: an investigation of semantic and contextual processing skills. *Journal of Experimental Child Psychology* **102**(3), 280–298.
- Carrol, G., Littlemore, J., & Gillon-Dowens, M. (2018). Of false friends and familiar foes: comparing native and non-native understanding of figurative phrases. *Lingua* **204**, 21–44.
- Chan, Y.-L., & Marinellie, S. A. (2008). Definitions of idioms in preadolescents, adolescents, and adults. *Journal of Psycholinguistic Research* **37**(1), 1–20.
- Christensen, R. (2015, 2018). R package 'ordinal' [Computer software]. Retrieved from <<https://cran.r-project.org/>>.
- Citron, F. M. M., Cacciari, C., Kucharski, M., Beck, L., Conrad, M., & Jacobs, A. M. (2016). When emotions are expressed figuratively: psycholinguistic and affective norms of 619 idioms for German (PANIG). *Behavior Research Methods* **48**(1), 91–111.
- Colman, A. M., Norris, C. E., & Preston, C. C. (1997). Comparing rating scales of different lengths: equivalence of scores from 5-point and 7-point scales. *Psychological Reports* **80**(2), 355–362.
- 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.
- Coulson, S., & Matlock, T. (2001). Metaphor and the space structuring model. *Metaphor and Symbol* **16**(3/4), 295–316.
- Coulson, S., & Van Petten, C. (2002). Conceptual integration and metaphor: an event-related potential study. *Memory and Cognition* **30**(6), 958–968.
- Croft, W. (1993). The role of domains in the interpretation of metaphors and metonymies. *Cognitive Linguistics* **4**(4), 335–370.
- Cronk, B. C., & Schweigert, W. A. (1992). The comprehension of idioms: the effects of familiarity, literalness, and usage. *Applied Psycholinguistics* **13**(2), 131–146.

- Dawes, J. (2008). Do data characteristics change according to the number of scale points used? An experiment using 5-point, 7-point and 10-point scales. *International Journal of Market Research* 50(1), 61–77.
- Dirven, R. (2002). Metonymy and metaphor: conceptualisation strategies. In R. Dirven & R. Pörrings (Eds.), *Metaphor and metonymy in comparison and contrast* (Cognitive Linguistics Research 20) (pp. 75–111). Berlin/New York: Mouton de Gruyter.
- Dirven, R., & Pörrings, R. (Eds.) (2002). *Metaphor and metonymy in comparison and contrast* (Cognitive Linguistics Research 20). Berlin, New York: Mouton de Gruyter.
- Ferrand, L., New, B., Brysbaert, M., Keuleers, E., Bonin, P., Méot, A., & Pallier, C. (2010). The French Lexicon Project: lexical decision data for 38,840 French words and 38,840 pseudowords. *Behavior Research Methods* 42(2), 488–496.
- 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/New York: Mouton de Gruyter.
- Field, A. P., Miles, J., & Field, Z. (2014). *Discovering statistics using R*. London: Sage.
- Gentner, D., Bowdle, B. F., Wolff, P., & Boronat, C. (2001). *Metaphor is like analogy*: Cambridge, MA: MIT Press.
- Gibbs, R. W. (1980). Spilling the beans on understanding and memory for idioms in conversation. *Memory and Cognition* 8(2), 149–156.
- Gibbs, R. W., & Colston, H. L. (2012). *Interpreting figurative meaning*. Cambridge: Cambridge University Press.
- Gibbs, R. W., & Gonzales, G. P. (1985). Syntactic frozenness in processing and remembering idioms. *Cognition* 20(3), 243–259.
- 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.
- Gildea, P., & Glucksberg, S. (1983). On understanding metaphor: the role of context. *Journal of Verbal Learning and Verbal Behavior* 22(5), 577–590.
- Giora, R. (1997). Understanding figurative and literal language: the graded salience hypothesis. *Cognitive Linguistics* 8(3), 183–206.
- Giora, R. (2003). *On our mind: salience, context, and figurative language*. New York: Oxford University Press.
- Giora, R., Fein, O., Kotler, N., & Shuval, N. (2015). Know hope: metaphor optimal innovation, and pleasure. In G. Brône, K. Feyaerts, & T. Veale (Eds.), *Cognitive linguistics and humor research* (pp. 129–146). Berlin/Boston, MA: De Gruyter Mouton.
- Glucksberg, S., & Keysar, B. (1990). Understanding metaphorical comparisons: beyond similarity. *Psychological Review* 97(1), 3–18.
- 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.
- Goossens, L. (1995). 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/Philadelphia, PA: J. Benjamins.
- Goossens, L. (Ed.). (1995). *By word of mouth: metaphor, metonymy, and linguistic action in a cognitive perspective*. Amsterdam/Philadelphia, PA: J. Benjamins.
- Gross, G. (1996). *Les expressions figées en français: Noms composés et autres locutions*. Collection l'essentiel français. Paris: Ophrys.
- Handford, M., & Koester, A. (2010). 'It's not rocket science': metaphors and idioms in conflictual business meetings. *Text and Talk – An Interdisciplinary Journal of Language, Discourse and Communication Studies* 30(1), 27–51.
- Heredia, R. R., & Cieśllicka, A. B. (Eds.) (2015). *Bilingual figurative language processing*. New York: Cambridge University Press.
- Hoffman, R. R. (1984). Recent psycholinguistic research on figurative language. *Annals of the New York Academy of Sciences* 433(1), 137–166.
- Iakimova, G., Passerieux, C., Laurent, J.-P., & Hardy-Bayle, M.-C. (2005). ERPs of metaphoric, literal, and incongruous semantic processing in schizophrenia. *Psychophysiology* 42(4), 380–390.

- 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.
- Katz, A. N., Paivio, A., Marschark, M., & Clark, J. M. (1988). Norms for 204 literary and 260 non-literary metaphors on 10 psychological dimensions. *Metaphor and Symbolic Activity* **3**(4), 191–214.
- Keyser, B., & Bly, B. M. (1999). Swimming against the current: Do idioms reflect conceptual structure? *Journal of Pragmatics* **31**(12), 1559–1578.
- Klepousniotou, E. (2002). The processing of lexical ambiguity: homonymy and polysemy in the mental lexicon. *Brain and Language* **81**(1/3), 205–223.
- Kövecses, Z., & Radden, G. (1998). Metonymy: developing a cognitive linguistic view. *Cognitive Linguistics* **9**(1), 37–78.
- 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.
- Lai, V. T., Curran, T., & Menn, L. (2009). Comprehending conventional and novel metaphors: an ERP study. *Brain Research* **1284**, 145–155.
- Lakoff, G., & Johnson, M. (1980). The metaphorical structure of the human conceptual system. *Cognitive Science* **4**(2), 195–208.
- Lakoff, G., & Turner, M. (1989). *More than cool reason: a field guide to poetic metaphor*. Chicago, IL: University of Chicago Press.
- Leiner, D. (2014). Sosci Survey [Computer Software]. Retrieved from <<https://www.soscisurvey.de/>>.
- Libben, M. R., & Titone, D. A. (2008). The multidetermined nature of idiom processing. *Memory and Cognition* **36**(6), 1103–1121.
- Lund, A., & Lund, M. (2013). Measures of central tendency. Retrieved from <<https://statistics.laerd.com/statistical-guides/measures-central-tendency-mean-mode-median.php>>.
- McLeod, A. I. (2015). R package ‘Kendall’ 2.2 [Computer software]. Retrieved from <<https://cran.r-project.org/>>.
- Microsoft Corporation (2010). Excel: Microsoft Office Home and Student [Computer Software].
- Nakazawa (2018). R package ‘fmsb’ [Computer software]. Retrieved from <<https://cran.r-project.org/>>.
- Nippold, M. A., & Duthie, J. K. (2003). Mental imagery and idiom comprehension. *Journal of Speech, Language, and Hearing Research* **46**(4), 788–799.
- Nippold, M. A., & Rudzinski, M. (1993). Familiarity and transparency in idiom explanation. *Journal of Speech and Hearing Research* **36**(4), 728–737.
- Nippold, M. A., & Taylor, C. L. (1995). Idiom understanding in youth: further examination of familiarity and transparency. *Journal of Speech and Hearing Research* **38**(2), 426–433.
- Nippold, M. A., & Taylor, C. L. (2002). Judgments of idiom familiarity and transparency. *Journal of Speech, Language, and Hearing Research* **45**(2), 384–391.
- 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.
- Nordmann, E., & Jambazova, A. A. (2016). Normative data for idiomatic expressions. *Behavior Research Methods* **49**(1), 198–215.
- Nunberg, G., Sag, I. A., & Wasow, T. (1994). Idioms. *Language* **70**(3), 491–538.
- Ortony, A. (1979). The role of similarity in similes and metaphors. In A. Ortony (Ed.), *Metaphor and thought* (pp. 186–201). Cambridge: Cambridge University Press.
- Ortony, A. (Ed.) (1993). *Metaphor and thought*. Cambridge: Cambridge University Press.
- Ortony, A., Schallert, D. L., Reynolds, R. E., & Antos, S. J. (1978). Interpreting metaphors and idioms: some effects of context on comprehension. *Journal of Verbal Learning and Verbal Behavior* **17**(4), 465–477.
- Otal, J. L., Navarro i Ferrando, I., & Bellés Fortuño, B. (Eds.) (2005). *Cognitive and discourse approaches to metaphor and metonymy*. Castelló de la Plana: Publicacions de la Universitat Jaume I.

- R Core Team (2016). R [Computer software]. Retrieved from <<https://cran.r-project.org/>>.
- RStudio Team (2016). R Studio [Computer software]. Retrieved 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/New York: 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.
- Ruiz de Mendoza Ibáñez, F. J. (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/New York: Mouton de Gruyter.
- 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.
- Sailer, M. (2013). *Idiom and phraseology*. Oxford: Oxford University Press.
- Sauro, J. (2016). Can you take the mean of ordinal data? Retrieved from <<https://measuringu.com/mean-ordinal/>>.
- Schemann, H. (2011). *Deutsche Idiomatik*. Berlin/Boston, MA: De Gruyter.
- Schmitt, N. (Ed.) (2004). *Language learning and language teaching: Vol. 9. Formulaic sequences: 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. Formulaic sequences: Acquisition, processing and use* (pp. 173–190). Amsterdam: Benjamins.
- Schweigert, W. A. (1986). The comprehension of familiar and less familiar idioms. *Journal of Psycholinguistic Research* 15(1), 33–45.
- Spieß, C., & Köpcke, K.-M. (Eds.) (2015). *Metapher und Metonymie: Theoretische, methodische und empirische Zugänge*. Berlin/Boston, MA: De Gruyter.
- Sprenger, S., Levelt, W., & Kempen, G. (2006). Lexical access during the production of idiomatic phrases. *Journal of Memory and Language* 54(2), 161–184.
- Sweetser, E. (1990). *From etymology to pragmatics: metaphorical and cultural aspects of semantic structure* (Cambridge Studies in Linguistics 54). Cambridge/New York: Cambridge University Press.
- Tabossi, P., Arduino, L., & Fanari, R. (2011). Descriptive norms for 245 Italian idiomatic expressions. *Behavior Research Methods* 43(1), 110–123.
- Tabossi, P., Fanari, R., & Wolf, K. (2009). Why are idioms recognized fast? *Memory and Cognition* 37(4), 529–540.
- 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.
- Titone, D. A., & Connine, C. M. (1999). On the compositional and noncompositional nature of idiomatic expressions. *Journal of Pragmatics* 31(12), 1655–1674.
- Titone, D. A., Holzman, P. S., & Levy, D. L. (2002). Idiom processing in schizophrenia: literal implausibility saves the day for idiom priming. *Journal of Abnormal Psychology* 111(2), 313–320.
- 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/New York: 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>>.
- Zwitserlood, P. (1994). The role of semantic transparency in the processing and representation of Dutch compounds. *Language and Cognitive Processes* 9(3), 341–368.

Appendix

Effect of familiarity on other ratings by self-report

Our study also had participants self-report how strongly they felt their comprehensibility, relation, and non-literality ratings were influenced by how familiar the idiom was to them. The findings somewhat mirror the absence or weakness of correlations between familiarity and the other properties. In the comprehensibility survey, only 31% feel their rating was very strongly or strongly influenced by how familiar the idiom was to them, 17% report some influence, while 52% report no or very weak influence.

In rating the relation between idiom and meaning, 26% of participants felt (very) strongly influenced by how familiar the idiom was to them, 28% report some, while 46% report no or very weak influence.

In rating non-literality, 36% of participants felt (very) strongly influenced by how familiar the idiom was to them, 30% report some, while 34% report no or very weak influence.

It is unclear how difficult it is to block influence by familiarity, whether it occurs subconsciously, and whether its effect can be adequately gauged by participants themselves. These three surveys contained the help instruction to not let familiarity influence the ratings. So, on the one hand, participants likely tried their best to follow and could then gauge the influence on a scale of 1 to 5 ('none' to 'very strong influence'). Usage of the full scale in all surveys at least indicates large individual differences in how successful blocking the influence worked, or at least how participants perceived it. As the comparably large shares of ratings on the low end ('no/very weak influence') of the scale shows, it seemed possible for over one-third to over one-half of participants to largely block influence by familiarity. Although no rewards or sanctions were tied to succeeding in this respect, answers in favor of the instruction cannot be ruled out.