When emotions are expressed figuratively: Psycholinguistic and Affective Norms of 619 Idioms for German (PANIG)

Francesca M. M. Citron[1,2] · Cristina Cacciari[3] · Michael Kucharski[1] · Luna Beck[1,4] ·
Markus Conrad[1,5] · Arthur M. Jacobs[1,6]

Abstract Despite flourishing research on the relationship between emotion and literal language, and despite the pervasiveness of figurative expressions in communication, the role of figurative language in conveying affect has been underinvestigated. This study provides affective and psycholinguistic norms for 619 German idiomatic expressions and explores the relationships between affective and psycholinguistic idiom properties. German native speakers rated each idiom for emotional valence, arousal, familiarity, semantic transparency, figurativeness, and concreteness. They also described the figurative meaning of each idiom and rated how confident they were about the attributed meaning. The results showed that idioms rated high in valence were also rated high in arousal. Negative idioms were rated as more arousing than positive ones, in line with results from single words. Furthermore, arousal correlated positively with figurativeness (supporting the idea that figurative expressions are more emotionally engaging than literal expressions) and with concreteness and semantic transparency. This suggests that idioms may convey a more direct reference to sensory representations, mediated by the meanings of their constituting words. Arousal correlated positively with familiarity. In addition, positive idioms were rated as more familiar than negative idioms. Finally, idioms without a literal counterpart were rated as more emotionally valenced and arousing than idioms with a literal counterpart. Although the meanings of ambiguous idioms were less correctly defined than those of unambiguous idioms, ambiguous idioms were rated as more concrete than unambiguous ones. We also discuss the relationships between the various psycholinguistic variables characterizing idioms, with reference to the literature on idiom structure and processing.

Keywords Idioms · Affect · Valence · Arousal · Emotion · Concreteness · Familiarity · Semantic transparency · Figurativeness

✉ Cristina Cacciari
cristina.cacciari@unimore.it

[1] Cluster of Excellence "Languages of Emotion", Free University of Berlin, Berlin, Germany

[2] Humanities Council, Princeton University, Princeton, NJ, USA

[3] Department of Biomedical, Metabolic and Neurological Sciences, University of Modena and Reggio-Emilia, Modena, Italy

[4] Institute for Quality Development and Education, Humboldt University of Berlin, Berlin, Germany

[5] Department of Cognitive, Social and Organisational Psychology, University of La Laguna, La Laguna, Spain

[6] Dahlem Institute for the Neuroimaging of Emotion (D.I.N.E.), Free University of Berlin, Berlin, Germany

Introduction

In the last decade, a growing body of research on the relationships between language and affect has shown that the emotional content of words affects comprehension processes, challenging semantic models of word recognition and text comprehension that typically have not considered this important aspect (Jacobs, 2011; Jacobs et al., 2015). More specifically, emotionally laden words are recognized faster and have processing priority when compared to neutral words. Emotionally laden words also show larger amplitudes of the event-related potential (ERP) components associated with emotional-stimulus processing; furthermore, their processing is subserved by a network of brain regions functionally

associated with affective states, as has been revealed by brain-imaging studies (for reviews, see Citron, 2012; Kissler, Assadollahi, & Herbert, 2006). Texts containing emotional information specifically activated the bilateral ventromedial prefrontal cortex (vmPFC) and the left amygdala, both regions associated with emotion processing, whereas texts containing chronological or spatial information activated different and distinct networks (Ferstl, Rinck, & von Cramon, 2005; Ferstl & von Cramon, 2007). In addition, emotionally negative texts activated brain regions associated with theory of mind as well as the vmPFC more strongly than did texts containing emotionally neutral information (Altmann, Bohrn, Lubrich, Menninghaus, & Jacobs, 2012).

According to dimensional models of affect, *valence* describes the extent to which a stimulus is positive or negative, and *arousal* refers to its degree of physiological activation (i.e., how calming or exciting/agitating a stimulus is; Lang, Bradley, & Cuthbert, 1997; Reisenzein, 1994; Russell, 2003). These dimensions typically show a quadratic relationship, whereby highly positive and negative stimuli are also highly arousing, whereas emotionally neutral stimuli tend to be low in arousal (e.g., Bradley & Lang, 1999; Võ et al., 2009). Several single-word corpora have also suggested that negative word meanings are more arousing than positive ones (e.g., Citron, Weekes, & Ferstl, 2014b; Schmidtke, Schröder, Jacobs, & Conrad, 2014; Võ et al., 2009). Although correlated, these two dimensions also show partial distinction, as evidenced by rating as well as neuroimaging studies (cf. Citron et al., 2014b; Lewis, Critchley, Rotshtein, & Dolan, 2007).

Emotion and figurative language

To date, most of the psychological and neuroscientific research on the relationship between language and emotion has been centered on literal language, despite the pervasiveness of nonliteral expressions in everyday communication. In fact, estimates based on simple frequency counts showed that people use approximately six nonliteral expressions per minute of discourse (Pollio, Barlow, Fine, & Pollio, 1977). Furthermore, at least in American English, there seem to be as many fixed expressions as there are words (Jackendoff, 1995). Nevertheless, the role of figurative language in conveying affect is still poorly understood (Citron & Goldberg, 2014; Fainsilber & Ortony, 1987; Fussell & Moss, 1998; Schrott & Jacobs, 2011).

Generally speaking, a word's or sentence's meaning is considered figurative if the conveyed meaning differs from the literal meaning assigned to the word(s)—that is, when the speaker/hearer has to go beyond the conventional word meaning and construct the intended sentential meaning by also using knowledge stored in semantic memory: For example, the interpretation of *She had a rough day* requires the reader to assign to *rough* the nonliteral meaning *difficult* or *straining*, rather than

the literal reference to texture. Figurative language is formed by a variety of different types of expressions (e.g., metaphors, proverb, idioms, and oxymora). The present study tested the affective and psycholinguistic characteristics of idioms (e.g., *to spill the beans*), which are in many languages the most frequent instance of figurative language. Idioms are strings of words whose global meaning cannot generally be inferred solely on the basis of the meaning of the constituent words, and therefore has to be retrieved from semantic memory. The relationship between lexical items and phrasal meaning is to a large extent arbitrary and learned, although this does not imply that individual lexical items do not constrain the semantic and syntactic operations that an idiom can undergo while still retaining a figurative interpretation (for an overview, see Cacciari, 2014). The idiomatic meaning and the default idiom structure are stored in long-term semantic memory together with word meanings, concepts and many other types of multiword strings. Idioms are different from metaphors (although some idioms can diachronically come from metaphors) since metaphors (even the most frozen ones) do not possess a unique standardized meaning, and can convey more than one meaning depending on context; This occurs also in highly conventionalized metaphors such as *Bob is an elephant*, which can mean that he is clumsy, extremely big, a blunderer, and so forth. Idioms indeed have a unique meaning that can be specialized but not radically modified by context. As Konopka and Bock (2009) pointed out, speakers cannot retrieve and productively combine words online to create an idiomatic expression. Some idioms allow for some forms of variation—for example, *She did not spill a single bean*, *The beans have been spilled*, and *I'm going to spill the beans*, none of which change the idiomatic meaning of "to reveal a secret." However, this should not to be confounded with true metaphorical language. In contrast, we can create a metaphor on the fly, although not necessarily an apt one. Basically, metaphors concern categorization processes, whereas idioms require meaning retrieval from semantic memory (Cacciari & Glucksberg, 1994; Cacciari & Papagno, 2012; Glucksberg, 2001). Idioms, then, differ from proverbs—for example, *You can't get blood from a stone* or *Two wrongs don't make a right*—since the proverbs are temporarily undefined full sentences, signaled by specific grammatical, phonetic, and/or rhetorical patterns, or by a binary structure (theme/comment); in general, proverbs are literally and figuratively true statements (Ferretti, Schwint, & Katz, 2007; Turner & Katz, 1997).

Why do we use figurative rather than—or together with—literal language to speak about affect? This seemingly easy question has received few answers. Pioneering work by Fainsilber and Ortony (1987) has shown that figurative language is preferred to literal language in oral descriptions of autobiographical emotional experiences. In particular, participants used significantly more metaphors for describing *how they felt* during a specific event than *what they did* in the same

circumstances. Furthermore, participants used more figurative expressions when asked to describe emotionally intense events than mildly intense ones (Fainsilber & Ortony, 1987; Fussell & Moss, 1998). Discourse analysis has shown that idiomatic expressions are preferred when speakers express complaints (Drew & Holt, 1988, 1998), presumably to elicit empathy in the addressee and thus become more convincing. Specifically, speakers were more likely to use idiomatic than literal expressions when summarizing their complaints in the presence of nonempathic interlocutors (Drew & Holt, 1988) and in topic transitions (Drew & Holt, 1998). Recent brain-imaging evidence has shown that nonliteral sentences evoked stronger implicit emotional responses than literal sentences (Bohrn, Altmann, & Jacobs, 2012). Similarly, a study on taste metaphors showed that metaphorical sentences elicited enhanced activation of the amygdala compared to their literal counterparts, which were matched for valence and arousal (Citron & Goldberg, 2014).

A few neurocognitive studies to date have controlled the affective characteristics of figurative as compared to literal language, with a predominant interest in metaphors (Bohrn, Altmann, Lubrich, Menninghaus, & Jacobs, 2012; Citron & Goldberg, 2014; Forgács et al., 2012; Forgács, Lukács, & Pléh, 2014). However, to the best of our knowledge, no study has yet provided descriptive norms of affective variables for the most common among figurative expressions—namely, idioms. To start filling this gap, the present descriptive study offers norms of affective and psycholinguistic properties for a set of 619 German idiomatic expressions (see also Fellbaum & Geyken, 2005, for a linguistically descriptive database). These data provide a structured tool for selecting experimental stimuli for future studies investigating the role of affect in nonliteral language. We chose idioms rather than other types of figurative expressions because idioms are frequent, highly conventionalized nonliteral strings of words with shared meanings. Thus, native speakers of a target language can easily rate many of their properties. Furthermore, these expressions are often semantically and syntactically flexible and can be embedded in different contexts without losing or changing their core meaning, and this allows idioms to be employed in many different experimental designs.

Descriptive norms of idiomatic expressions already exist for a few languages, including American English (Abel, 2003; Cronk, Lima, & Schweigert, 1993; Libben & Titone, 2008; Titone & Connine, 1994b), Italian (Tabossi, Arduino, & Fanari, 2011), and French (Bonin, Méot, & Bugaiska, 2013; Caillies, 2009). These databases provide ratings for the psycholinguistic properties of idioms (e.g., ambiguity, familiarity, and knowledge of the idiomatic meaning), and some studies also include reaction time data from different tasks—namely, online reading comprehension (Bonin et al., 2013), offline and online meaningfulness judgment (Libben & Titone, 2008), and self-paced reading (Cronk et al., 1993). However, these

studies have not considered whether idiomatic meanings conveyed an affective content (differentiated for valence and arousal), or whether their meaning was concrete or abstract. Hence, the present study not only offers norms for the main psycholinguistic variables affecting idiom processing in German, but also provides new data on variables ignored in prior normative studies. In the present study, we also aimed to explore the relationships within affective variables and between these variables and the psycholinguistic norms.

The psycholinguistic characteristics of idioms

Several models have been proposed to account for idiom comprehension (for overviews, see Cacciari, 2014; Cacciari, Padovani, & Corradini, 2007; Libben & Titone, 2008; Titone & Connine, 1994a). According to *lexical look-up models*, idioms are fixed expressions listed in the mental lexicon, either together with other lexical units (Swinney & Cutler, 1979) or in a separate list (Bobrow & Bell, 1973). In this model, linguistic processing of the string and retrieval of the idiomatic meaning proceed in parallel, with the retrieval of the idiomatic meaning being faster than the computation of its literal meaning. According to the *configuration hypothesis* (Cacciari & Tabossi, 1988; Vespignani, Canal, Molinaro, Fonda, & Cacciari, 2009), idioms are processed word by word, like any other piece of language, until enough information has accumulated to render the sequence of words identifiable as—or highly expected to be—an idiom. At this point, the idiomatic meaning is retrieved.

The recent studies on idiom processing that led to nonlexical models highlighted that idioms differ in many respects, and that studies on idiom comprehension must take this variability into account in order to satisfactorily account for their comprehension. In fact, many psycholinguistic properties have been shown to affect idiom processing (for overviews, see Cacciari, 2014; Cacciari & Glucksberg, 1994; Libben & Titone, 2008). In this study, we collected descriptive norms of the most important of these variables, together with norms for concreteness, valence and arousal of the idiomatic meaning. Below we provide a definition for each of the psycholinguistic and affective characteristics investigated (familiarity, knowledge, confidence, ambiguity, semantic transparency, figurativeness, concreteness, valence, and arousal), accompanied by a summary of previous results. The details concerning data collection are presented in the "Method" section.

(a) *Familiarity* refers to the subjective frequency of exposure to idioms (Titone & Connine, 1994b)—namely, how often one has read or heard an idiom. Familiarity may differ from objective frequency estimates, which are based on written and spoken databases. Subjective frequency estimates may provide information that helps the choice of the experimental materials of idiom studies for several

reasons: (1) Idiomatic expressions are not necessarily processed word by word; (2) very common idioms may contain nonfrequent or old-fashioned words but nonetheless be understood very quickly and easily; (3) an idiom's meaning is not necessarily associated with the meaning of its constituent words, and as such, the frequency of its constituent words may play less of a role than in the case of literal sentences. Highly familiar idioms have been shown to be comprehended faster and more accurately than less familiar ones (Cronk et al., 1993; Cronk & Schweigert, 1992; Libben & Titone, 2008). Previous norming studies reported highly positive correlations between familiarity, intended as subjective frequency of exposure, *other*-based familiarity [i.e., an estimate of how well others know the idiom; Tabossi et al. (2011) and Bonin et al. (2013) operationally defined this notion as "how well you think the expression is known by people like you, independently of whether you know it" (Tabossi et al., 2011, p. 115)], confidence about one's own knowledge, and knowledge of the idiomatic meaning (Bonin et al., 2013; Libben & Titone, 2008; Tabossi et al., 2011; Titone & Connine, 1994b). In general, the vast majority of idioms are estimated to have been acquired at approximately 9 years of age (Libben & Titone, 2008; Tabossi et al., 2011). Familiar and well-known idioms are also estimated to have been acquired earlier. However, Bonin et al. (2013) reported that the estimated age of acquisition was a better predictor of comprehension times than familiarity, with faster reaction times to idioms that were acquired earlier.

The subjective estimate of the frequency of exposure to idiomatic expressions is a better predictor of idiom processing than is a measure of frequency obtained by combining the single frequencies of an idiom's constituent words (Bonin et al., 2013; Libben & Titone, 2008). This may reflect the fact that idiomatic meanings are often arbitrarily related to the meanings of their constituent words and that sometimes familiar idioms contain words that are no longer used out of the idiomatic context. For example, *Flausen* is only used in *Flausen im Kopf haben*, meaning to have nonsense/weird ideas in mind in German. Finally, the reliability of estimates of *other*-based familiarity (Bonin et al., 2013; Tabossi et al., 2011) can be problematic, since it is more likely that participants can reliably estimate their own frequency of exposure to an idiomatic expression than how well other people know such expressions (Cronk et al., 1993; Libben & Titone, 2008; Titone & Connine, 1994b).

(b) *Knowledge of idiomatic meaning* refers to whether or not the correct idiomatic meaning is known. Some studies (Tabossi et al., 2011) tested this variable by asking participants to provide a written explanation of the idiomatic meaning, whereas others measured the extent to which participants were confident about their own knowledge of the idiomatic meaning (Bonin et al., 2013; Libben & Titone, 2008; Titone & Connine, 1994b). However, the latter procedure does not necessarily provide a reliable measure of the actual idiom knowledge. In fact, speakers may be very confident about their incorrect knowledge of an idiomatic meaning, particularly for less-familiar idioms. Since *knowledge of the idiomatic meaning* and *confidence* do differ, we tested these two variables separately. Confidence about the knowledge of the idiomatic meaning was rated before participants wrote down a definition of the idiom meaning. Confidence about one's own knowledge as well as *other*-based familiarity have been shown to speed up online reading comprehension times (Bonin et al., 2013).

(c) *Ambiguity* (also referred to as *literality*) refers to whether an idiom string also has a semantically plausible literal meaning (Cronk et al., 1993). In fact, some idioms are ambiguous, since they have both a literally plausible and an idiomatic meaning (e.g., *kick the bucket* can describe a literally plausible action, beyond the idiomatic meaning "to die"). In some normative studies (e.g., Bonin et al., 2013; Tabossi et al., 2011), participants have been asked to rate how often they came across an idiom used in a literal sense. To avoid any bias due to the preponderant figurative use of idiom strings, we did not ask participants to provide ratings of ambiguity. Rather, the experimenters divided the German idioms into ambiguous and unambiguous idioms on the basis of the presence versus absence of a semantically plausible literal meaning. Idiom ambiguity generally shows a less consistent pattern of correlations with other variables across studies: Ambiguity correlated negatively with other-based familiarity (Tabossi et al., 2011) and with confidence (Libben & Titone, 2008), and positively with subjective frequency (Bonin et al., 2013; Cronk et al., 1993). This suggests that participants rated the literal meaning as being more plausible when others were supposed to be familiar with the idiomatic meaning, and also when idioms had a high subjective frequency. Idiom ambiguity significantly predicted accuracy (of participants' meaningfulness judgments), in that responses were more accurate to idioms with literally plausible meanings than to idioms without a literal counterpart (Libben & Titone, 2008). However, it still is unclear whether literally plausible and literally implausible idioms are comprehended with similar ease and through the same processes (Cacciari, 2014). Reaction time studies have suggested that the figurative meaning of literally plausible idioms is accessed faster than that of implausible idioms (Cronk & Schweigert, 1992; Libben & Titone, 2008; Mueller & Gibbs, 1987). This has been supported by a case study on semantic dementia

(Papagno & Cacciari, 2010) and a study on aphasic patients (Papagno & Caporali, 2007), but has been contradicted by a different study on aphasic patients (Cacciari et al., 2006) that reported impaired comprehension of literally plausible than implausible idioms, possibly due to difficulty in inhibiting the literal meaning.

In the present study, we specifically tested whether the psycholinguistic and affective properties of literally plausible and implausible idioms differed and/or showed different patterns of correlations.

(d) *Semantic transparency* refers to the ease with which the figurative meaning of an idiom can be inferred from the constituent words' literal meanings. For example, *keep in touch* is a relatively transparent idiom, since its figurative meaning "maintain social contact with someone" can be easily inferred from the constituent words. In contrast, the idiomatic meaning of *kick the bucket* is opaque, and the figurative meaning "to die" must be learned. This variable is quite problematic and unstable, because transparency estimates are based on readers'/listeners' intuitions that derive from the knowledge of the idiomatic meaning. In fact, familiar idioms tend to be perceived as more transparent than unfamiliar ones, because they are *repeatedly used* with that stipulated meaning (Keysar & Bly, 1995). Studies reporting faster responses to semantic transparent, or *decomposable*, idioms than to nontransparent ones (Gibbs, Nayak, & Cutting, 1989) have predominantly used offline paradigms, which measure late interpretative phases rather than real-time comprehension processes (Tabossi, Fanari, & Wolf, 2008). Some studies have reported that the more semantically transparent an idiom is, the more familiar (Abel, 2003), the less ambiguous (Libben & Titone, 2008), and the better known (Tabossi et al., 2011) it seems to be. However, no correlation between semantic transparency and familiarity was found by Tabossi et al. (2011). This lack of consensus may also be due to the high variability in individual participants' ratings of semantic transparency (Cacciari & Glucksberg, 1995; Levorato & Cacciari, 1999; Tabossi et al., 2011).

(e) *Figurativeness* refers to the extent to which an idiomatic expression is perceived as expressing a nonliteral meaning. For instance, the Italian idiom *dormire come un ghiro* ("to sleep as a dormouse," in English *to sleep like a log*) denotes a way of sleeping—an action literally expressed by the verb—and may be perceived as less figurative than idioms in which the verb changes its basic meaning, as for instance in *to get cold feet* (i.e., to become afraid and to refrain from doing something). Although this variable has not been considered in previous normative studies on idioms, we decided to include it in order to investigate whether the perceived degree of idiomaticity correlates with other variables.

(f) *Length* is measured by either the number of composing *words* or the number of *letters*. Evidence has shown that, all other things being equal, the meaning of short idioms (i.e., with few words) is not yet available at the string offset, unless the prior context creates a bias toward an idiomatic interpretation. In contrast, the idiomatic meaning of long idioms is available at the string offset, regardless of the preceding context (Fanari, Cacciari, & Tabossi, 2010). Idiom length (in letters) significantly explained most of the variance in reading times (Bonin et al., 2013).

(g) *Concreteness* refers to the extent to which an idiomatic meaning refers to a state or event that one can experience in one or more sensory modalities (cf. Paivio, 2007; Paivio, Yuille, & Madigan, 1968). This characteristic has not been assessed in previous normative studies of idioms, despite the fact that a vast literature has shown that concrete words are more easily accessed and processed than abstract words (e.g., Adorni & Proverbio, 2012; Zhang, Guo, Ding, & Wang, 2006). Concreteness is sometimes confused with *imageability*, which instead refers to the ability to create a mental image of a word (Paivio, 2007). Imageability also facilitates word processing (e.g., Sabsevitz, Medler, Seidenberg, & Binder, 2005). Emotionally valenced abstract words are rated as more imageable than neutral abstract words (Altarriba & Bauer, 2004). Imageability differs from concreteness in that even abstract concepts may be imageable (e.g., *joy*), whereas some concrete concepts (e.g., *sloth*) may be less so. Imageability and concreteness are usually positively correlated, and most of the variance they explain tends to overlap. We decided to measure only concreteness, since idiomatic meanings may be rather difficult to imagine in their nonliteral sense, due to the interference of the literal meaning of the constituent words (Cacciari & Glucksberg, 1995).

(h) *Emotional valence* describes the extent to which a stimulus is positive or negative (Russell, 1980). Since normative studies on idiomatic expressions have not rated this variable (and arousal), we briefly review the literature on single words. Once a range of psycholinguistic variables have been controlled, emotionally valenced words have processing priority as compared to neutral words, leading to faster reaction times and higher accuracy in a variety of tasks (e.g., Citron, Weekes, & Ferstl, 2014a; Kousta, Vinson, & Vigliocco, 2009; Larsen, Mercer, & Balota, 2006). Furthermore, emotionally valenced words elicit a larger amplitude of ERP components associated with the processing of emotional stimuli (i.e., the early posterior negativity (EPN) and the late positive component (LPC)). Such words also elicit enhanced activation of brain regions associated with emotion processing (for an overview, see Citron, 2012). The results concerning the polarity of valence (positive vs. negative)

have been mixed. Some studies reported processing facilitation and enhanced brain activity in response to positive but not to negative words, once the level of arousal had been matched (Citron, Gray, Critchley, Weekes, & Ferstl, 2014; Herbert et al., 2009; Kuchinke et al., 2005; Recio, Conrad, Hansen, & Jacobs, 2014). In contrast, other studies reported no difference between positive and negative words (Citron et al., 2014a; Kousta et al., 2009; Larsen et al., 2006), unless a block design was used (Algom, Chajut, & Lev, 2004; Nasrallah, Carmel, & Lavie, 2009).

(i) *Emotional arousal* describes the excitation potential of a stimulus, independently of whether it is positive or negative (Barrett & Russell, 1998). Arousal ratings of words typically show a quadratic, U-shaped relationship with valence ratings (e.g., Bradley & Lang, 1999; Võ et al., 2009): The more emotionally valenced a word is, the more arousing it typically is. However, note that this overall U-shaped distribution involves a particularly strong negative linear correlation of arousal with valence within the domain of negative words, which sometimes leads to an overall negative linear correlation (e.g., Citron et al., 2014b; Schmidtke et al., 2014; Võ et al., 2009). Highly arousing words are processed faster and more accurately and elicit stronger neural responses than nonarousing words, when valence is kept constant (Bayer, Sommer, & Schacht, 2012; Hofmann, Kuchinke, Tamm, Võ, & Jacobs, 2009; Recio et al., 2014). Nevertheless, emotional valence seems to be a stronger predictor of response speed and accuracy than arousal (Estes & Adelman, 2008; Kousta et al., 2009).

The present study

The aims of this study were (1) to provide descriptive norms for psycholinguistic and affective properties of a large set of German idioms, and (2) to explore the relationships between psycholinguistic and affective properties of idioms. Toward this aim, 624 idioms (see Table 1 for examples) were rated,

using Likert scales, for emotional valence, arousal, familiarity, semantic transparency, figurativeness, and concreteness. Knowledge of the idiomatic meaning was assessed by asking participants to write down an explanation of each idiom's meaning and then to rate their confidence. Ambiguity was categorically determined by the experimenters.

In terms of the psycholinguistic variables tested in previous normative studies, we expected to replicate the positive correlations between familiarity, knowledge, and confidence (Bonin et al., 2013; Libben & Titone, 2008; Tabossi et al., 2011). Since concreteness, figurativeness, length in words and in letters, and valence and arousal have not yet been tested for idioms, we did not have a priori predictions, bur rather we explored their possible correlations. We also tested for the first time whether the properties of ambiguous and unambiguous idioms differ, and how they are correlated. This might help clarify which psycholinguistic properties underlie their differences, if any. Finally, as for the relationships between affective variables, we expected to replicate the results based on single words (e.g., Bradley & Lang, 1999; Citron et al., 2014b; Schmidtke et al., 2014)— namely, a quadratic relationship between valence and arousal (i.e., the more highly valenced an idiom, the more arousing it is) and a negative linear relationship (i.e., negative idioms are rated as more arousing than positive idioms).

Method

Materials

Idiom selection A total of 624 idiomatic expressions were selected from different Web sources (http://german.about. com, www.spruecheportal.de, and www.staff.uni-marburg. de/~naeser/idiom-ak.htm; Udem, 2001) and from a list of figurative expressions collected by Verena Simon during the "Bilingualism and Affectivity in Reading" project at the Cluster of Excellence "Languages of Emotion." The criteria for identifying a figurative expression as an idiom were as follows: It consists of a verb phrase (VP) with one or more arguments—for example, *to spill*[(VP)] *the beans*[(dir obj)], *to*

Table 1 Examples of idioms from our database

Code	German Idioms	Literal Translation	Meaning
005	sich jeden Bissen vom Munde absparen	to save oneself every bite from the mouth	to be very poor and have to save money to be able to afford anything
018	Pfeffer im Hintern haben	to have pepper in the ass	to be lively; to be full of energy; not to be able to sit still
053	Böses im Schilde führen	to carry evilness in the shield	to have evil intentions
076	jemanden auf die falsche Fährte locken	to bring someone on the wrong trail	to entice someone to believe the wrong/false lead
174	einen Geistesblitz haben	to have a genius lightning	to suddenly some up with a very good idea
418	jemanden sein Herz ausschütten	to spill one's heart to someone	to openly talk about one's problems with someone; to have an honest and confessing conversation

give(VP) *someone*(indir obj) *a hard time*(dir obj); the verb can be inflected for person and time; its meaning is conventionalized (to distinguish it from a metaphor); and it should not be formed by an entire sentence that cannot be altered, as in proverbs—for example, *A man's home is his castle*. Because we expected that the variables to be rated (i.e., emotional valence, arousal, familiarity, concreteness, figurativeness, semantic transparency, confidence, and knowledge) would generalize to inflected forms, we only presented idioms in the standard, infinitival form. Lengths in letters and in words were calculated with Excel. The lengths of the 619 idioms that were retained (please refer to the Data Analysis section) ranged from two to nine words, and from nine to 43 letters.

Variables determined by the experimenters Ambiguity— that is, whether an idiom had or did not have a semantically plausible literal meaning—was first established separately by the first, third, and fourth authors, who examined each of the 624 idioms and classified them as ambiguous or unambiguous, depending on whether a literal interpretation of the idiom is plausible or not. For instance, "to bite into the sour apple" (idiom no. 001) describes an event that can actually happen, and therefore we considered this an ambiguous idiom. On the other hand, "to have someone in the pear" (idiom no. 003) does not describe a plausible event, and therefore we classified it as an unambiguous idiom. Then, the individual decisions were compared; possible differences were discussed until an agreement on the categorization was reached; and a categorical variable (ambiguous vs. nonambiguous) was created.

Procedure

Instructions for the rating task Instructions were presented in written form and contained a definition of each of the variables to be rated, some examples of sentences rated with extreme values, an explanation of the Likert scale, and an explanation of the labels of the extreme and middle points. The original German instructions, an English translation, and a screenshot of one of the questionnaires may be found in Appendix A. *Familiarity* referred to the frequency with which the participant had heard or read the idiom. The rating scale ranged from 1 (*never heard/read*) to 7 (*often heard/read*). *Semantic transparency* referred to the extent to which the figurative meaning of an idiom could be inferred from the meaning of its constituent words. The scale ranged from 1 (*semantically transparent*) to 7 (*semantically opaque*). *Figurativeness* was actually labeled *Metaphoricity* in the instructions. This was because the average German participant is more familiar with the concept of metaphor or metaphorical meaning than with figurative expression or idiom. In fact, in current German, "metaphoric" is used as a synonym of

"figurative" (Bibliographisches Institut GmbH, 2013), despite the different and more specific uses of these terms in linguistics (e.g., Cacciari & Glucksberg, 1994; but please see the "introduction"). In this way, we aimed to increase participants' understanding of their task. The Figurativeness/Metaphoricity scale referred to how much the meaning of an idiom was perceived as nonliteral, on a scale ranging from 1 (*not at all figurative/metaphorical*) to 7 (*very figurative/metaphorical*). *Concreteness* referred to the extent to which the figurative meaning could be experienced with one or more sensory modalities. The scale ranged from 1 (*totally abstract*) to 7 (*totally concrete*). *Confidence* was measured by asking participants to rate their knowledge of the idiomatic meaning on a scale from 1 ("*I don't know the meaning at all*") to 7 ("*I know the meaning very well*"). *Knowledge of the idiomatic meaning* was assessed right after confidence by asking participants to write down the idiomatic meaning. *Emotional valence* referred to the extent to which the idiomatic meaning was positive or negative. The scale ranged from –3 (*very negative*) to +3 (*very positive*), through 0 (*neutral*). *Arousal* referred to the extent to which the idiomatic meaning was stimulating, on a scale from *not stimulating at all* to *very exciting or agitating*, independently of whether it was positive or negative. The rating scale ranged from 1 (*not at all arousing*) to 7 (*very arousing*). At the end of each scale, the option *unknown* was given.

Questionnaires Online questionnaires were created using SurveyMonkey. Six separate questionnaires were used to measure the emotional valence, arousal, familiarity, concreteness, figurativeness, and semantic transparency of the entire set of 624 idioms. Each of these questionnaires contained the full set of idioms. Another questionnaire measured participants' confidence about their own knowledge of the idiomatic meaning (through ratings), as well as their actual knowledge (through written definitions). Since this task required more time to be completed, we split the total number of idioms into two parts. Hence, each confidence/knowledge questionnaire contained only half of the idioms.

Eight randomized orders of the 624 idioms were first created. Then, each variable to be rated (except confidence and knowledge) was randomly assigned to four different randomizations, therefore creating four versions of each questionnaire for each variable. The confidence/knowledge variables were randomly assigned to three different randomizations, and each version was split in two halves, therefore creating six distinct questionnaires, each containing half of the stimuli.

Each variable was rated by at least 30 participants, who were each randomly assigned to a specific variable. Each participant was allowed to complete more than one questionnaire (each one on a different variable) and was rewarded accordingly. The online survey lasted approximately one and a half hours.

Participants

A total of 249 native speakers of German from the Berlin area (131 women, 118 men), between 19 and 67 years of age (*Median* = 30), completed the online survey. Participants were recruited through an online newsletter from the "Languages of Emotion" research cluster, to which our participant pool subscribed themselves. Our group included students from different faculties studying in Berlin and Potsdam, as well as unemployed or self-employed people. As a reward, participants were either paid €8 or given a lottery ticket for a raffle through which 20 Amazon vouchers worth €20 were awarded. One of the questionnaires was randomly assigned to each participant and accessed by him/her upon receipt of a unique URL via email. We report here only data from participants who completed at least 9/10 of the questions on each questionnaire and completed them accurately.

We compared gender proportion and mean age across the different subgroups of participants who rated the different variables, to make sure that these demographic characteristics were balanced across variables. The analyses showed no significant differences in either gender proportion [$\chi^2(6) = 4.25$, n.s.] or mean age [$F(6, 248) = 1.39$, n.s.] (see Table 2 for descriptive statistics).

Data analysis

We calculated the mean rating and standard deviation of each idiom for emotional valence, arousal, familiarity, semantic transparency, figurativeness, concreteness, and confidence. Knowledge of the idiomatic meaning was calculated by counting the number of correct definitions for each idiom. To determine the correctness of these definitions, the first author, together with a native German-speaking linguist, went through all definitions and compared them with the definitions given in our database (by a professional translator), as well as with the German definitions given at the "Redensarten-Index" webpage (Udem, 2001). Definitions that correctly matched

Table 2 Gender proportion and age of the participants across all rated variables

Variables	Gender			Age in Years		
	Women	Men	Total	Min.	Max.	Median
Em. valence	13	19	32	19	60	29
Arousal	15	16	31	21	60	30
Familiarity	15	15	30	21	67	30
Concreteness	20	11	31	20	67	29
Figurativeness	18	14	32	20	66	31
Sem. transparency	18	14	32	20	59	40
Knowledge/confidence	32	29	61	19	65	31

the idiomatic meaning were considered correct. For example, the idiom *seine Worte abwägen* (i.e., "to weigh one's words," idiom no. 289) means "to think about something carefully before speaking"; in this case, we categorized as correct definitions such as "to think before speaking" or "to think carefully before speaking."

In addition, we calculated the total number of valid responses relative to each variable for each idiom (i.e., either ratings or definitions), the percentage of "unknown" responses, and the number of omissions. An additional variable was created by calculating the square of emotional valence, which was called *valence²*. This variable represents the degree of absolute emotionality of a stimulus, independent of its polarity—namely, independently of whether its linear valence is positive or negative (Udem, 2001). Valence² enabled us to explore quadratic relationships between emotional valence and other variables.

Upon inspection of the idioms and rated variables, we found that for one idiom all participants reported not knowing its meaning in the confidence/knowledge questionnaire. We therefore excluded it from our database. We also found that four idioms appeared twice: One pair had identical items, two pairs had slightly different forms (i.e., the presence of an additional word that was not essential to constitute the idiomatic meaning), and one pair had a right and a wrong version. Thus, we eliminated one of the identical items in the first pair (chosen randomly), the least frequent items in the other two pairs, and the wrong item in the last pair, therefore retaining a total of 619 idiomatic expressions.

Distribution of variables and statistical methods For each variable obtained, we calculated mean value, standard deviation, median value, minimum, maximum, mean percentage of valid responses, of "unknown" responses and of omissions. For the length measures, we also calculated the first five values. Most of the variables were not normally distributed. Length in letters, words, emotional arousal, and concreteness were slightly positively skewed, and confidence was negatively skewed; these variables were successfully logarithmically transformed. However, other variables could not be transformed successfully: Familiarity and figurativeness were slightly negatively skewed; knowledge of the idiomatic meaning was strongly negatively skewed (given that we sampled relatively common idiomatic expressions that are almost all well known by native speakers); semantic transparency had a platykurtic distribution; emotional valence had a (natural) binomial distribution; and valence² had a quadratic distribution (and was hence strongly skewed). In order to make up for the lack of normality, we applied a bootstrapping technique to all our parametric statistical analyses (1,000 samples, 95% percentile confidence interval); this procedure allows for the estimation of the sampling distribution of almost any statistic through resampling the observed data, and is therefore

distribution-independent (cf. Bradley & Lang, 1999; Citron et al., 2014b).

Relationships among variables We calculated Pearson partial correlations between each linear affective variable and each other variable, as well as between each pair of nonaffective variables, by partialing out the effects of all remaining ones. On the basis of the literature (Efron & Tibshirani, 1993), emotional valence and arousal ratings were plotted against each other and showed a quadratic relationship (see Fig. 1). Therefore, we also computed a quadratic regression predicting arousal ratings from valence: In a first step, we entered all variables of no interest in order to partial out their effects, and in a second step we entered valence and valence^2 as predictors (i.e., the quadratic regression equation). Finally, we further conducted quadratic regressions for any nonaffective variable that correlated significantly with valence, to explore whether a quadratic function would better explain their relationship. Significant partial correlations up to $\pm .1$ are referred to as "small correlations," between $\pm .1$ and $\pm .3$ as "moderate correlations," and between $\pm .3$ and $\pm .5$ as "large correlations."

Ambiguous versus unambiguous idioms Since we were interested in possible differences between these two types of idioms, rather than in the relationship between ambiguity and other variables, we calculated descriptive statistics for each continuous variable (broken down by idioms' ambiguity) and ran t tests to compare the two conditions. We also recalculated the correlations just described, separately for ambiguous and unambiguous idioms.

Reliability analysis We conducted a reliability analysis based on internal consistency (Cronbach's alpha), also referred to as *intraclass correlations*. We chose this analysis because it
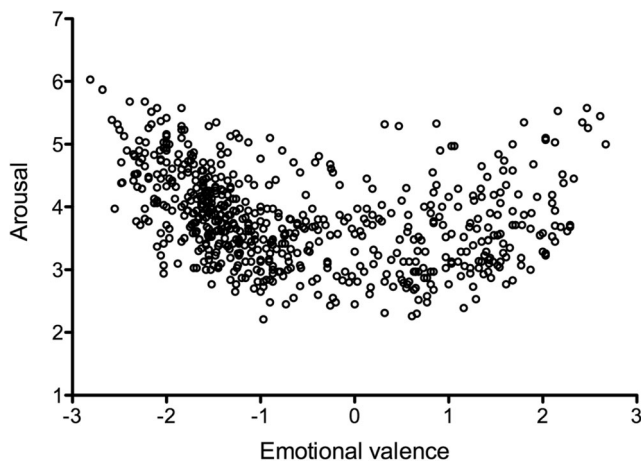

Fig. 1 Distribution of our idioms' affective properties. Emotional valence ratings (–3 = *very negative*, +3 = *very positive*, 0 = *neutral*) plotted against arousal ratings (1 = *not at all arousing*, 7 = *very highly arousing*)

represents a more reliable measure than the split-half procedure (also referred to as *product-moment correlations*), as outlined, for example, in Cicchetti (1994). We used the raw ratings from each participant as a different variable and the single idioms as cases and obtained different Cronbach's alpha values for emotional valence, arousal, familiarity, concreteness, figurativeness, and semantic transparency. However, the confidence and knowledge questionnaires were organized differently: Three different randomizations of all idioms were first applied, and then the randomized idioms were split into two halves. Thus, the six resulting questionnaires had to be analyzed separately (with sample sizes of nine to 12 participants only), since each of them contained a unique combination of idioms.

Results and discussion

The full list of 619 idiomatic expressions with their literal translation in English, their idiomatic meanings, and the means and standard deviations for all variables are reported in the Supplementary Material.

Descriptive statistics for each variable are reported in Table 3. The mean emotional valence and arousal values across idioms varied from very negative to very positive values and from very low to very high arousal values (please refer to Fig. 1). The median valence value was negative, suggesting a higher proportion of negatively than positively valenced items. Overall, the idioms were rated as being familiar and having predominantly abstract meanings. Furthermore, they were evaluated as moderately figurative and semantically transparent. Knowledge of the idiomatic meaning was high (*Median* = 94%), as was confidence (*Median* = 6.23), suggesting that the selected idioms were known and frequently used (see the Familiarity paragraph above). The mean percentage of omissions (1.84%) and of "I don't know" responses (1.80%) were very low. However, the percentages slightly increased for the variables *confidence* and *idiom knowledge*, suggesting that, when explicitly asked about their knowledge of the idiomatic meanings, participants admitted that they did not know some of them.

In the following subsections, we report the partial correlations between the psycholinguistic and affective variables. A matrix of simple linear correlations is reported in Appendix B, Table C1.

Relationships between affective variables

The idiom list contained more than twice as many idioms with negative valence ($N = 422$) as with positive valence ($N = 194$; three idioms had valence = 0). This difference is unlikely to be due to the specific sample selected. Rather, it may reflect the fact that, since idiomatic expressions typically provide an

Table 3 Descriptive statistics for each rated or calculated variable

Measure	Variable	Mean	SD	Median	Minimum	Maximum	Mean Valid Responses	Mean % Unknown	Mean % Omissions
Rating	Em. valence	−0.59	1.32	−1.04	−2.81	2.67	30.65	1.62	2.51
	Valence^2	2.09	1.61	1.88	0.00	7.90	–	–	–
	Arousal	3.82	0.74	3.74	2.21	6.03	30.52	0.98	0.50
	Familiarity	4.76	0.68	4.83	2.67	6.24	28.98	1.71	1.50
	Concreteness	3.34	0.80	3.23	1.66	5.81	30.17	0.51	2.03
	Figurativeness	4.70	0.65	4.75	2.39	6.09	31.47	1.10	0.49
	Sem. transparency	4.13	0.97	4.16	1.83	6.22	31.08	0.80	1.97
	Confidence	6.10	0.55	6.23	2.54	6.90	28.21	3.94	3.49
% correct definitions	Knowledge	90.82	10.33	94.00	40.00	100.00	27.66	3.77	2.25
Value	Length in letters	23.00	6.00	22.00	9.00	43.00	–	–	–
	Length in words	4.00	1.00	4.00	2.00	9.00	–	–	–

Mean valid responses consist of all responses given (i.e., ratings or definitions) except "unknown" responses.

indirect form of communication (Cacciari, 1998; Drew & Holt, 1988), they are preferred over literal expressions for negative statements.

In the quadratic regression predicting arousal from valence ratings, the first model, including all psycholinguistic variables, accounted for 12% of the variance ($R^2 = .12$, $r = .35$), $F(8, 610) = 10.70$, $p < .001$, whereas the second model, also including valence and valence squared, accounted for an additional 26% ($R^2 = .38$, $r = .62$), $F(2, 608) = 128.29$, $p < .001$, with both valence and valence^2 as significant predictors. [The regression line was as follows: estimated arousal = 0.15 × familiarity + 0.13 × concreteness + 0.29 × figurativeness + 0.15 semantic transparency − 0.13 × valence + 0.48 × valence^2.] Thus, the more emotionally valenced an idiomatic meaning was, the higher its level of arousal; the quadratic relationship between valence and arousal can be seen in Fig. 1. This result is in line with the typical U-shaped relationship between emotional dimensions that has repeatedly been found for single words (e.g., Bradley & Lang, 1999; Citron et al., 2014b; Schmidtke et al., 2014; Võ et al., 2009). In addition, higher level of arousal were attributed to negative than to positive idioms [$t(614) = 5.89$, $p < .0001$], a result often reported for single words (Citron et al., 2014b; Schmidtke et al., 2014; Võ et al., 2009).

Correlations between affective and psycholinguistic variables

In what follows, we present statistically significant partial correlations between variables ($\alpha = .05$) (see the specific tables for Pearson's r and p values).

Arousal and figurativeness, concreteness, semantic transparency, familiarity A first interesting result was a moderate positive partial correlation between arousal and figurativeness

(see Table 4 for correlations) that suggests an association between the nonliterality of the string and the attributed emotional–physiological intensity.

A second interesting result was a small positive partial correlation between arousal and concreteness: The more concrete an idiom's meaning was, the more emotionally arousing it was rated. The fact that concrete concepts have direct reference to one or more sensory modalities may have led to higher ratings of physiological arousal. A positive correlation between arousal and imageability was reported by Citron et al. (2014b) for single words. Our study and Citron et al.'s (2014b) study showed positive correlations within samples that contained mostly abstract items. However, larger word corpora with more balanced distributions of concrete and abstract words showed a negative quadratic correlation between arousal and concreteness (Montefinese, Ambrosini, Fairfield, & Mammarella, 2014) and negative

Table 4 Linear partial correlations between affective and psycholinguistic variables

	Em. Valence	Arousal (log 10)
Familiarity	.14***	.12**
Concreteness (log 10)	−.08	.08*
Figurativeness	−.01	.30***
Sem. transparency	.07	.25***
Confidence (log 10)	−.01	−.05
Knowledge	.03	.04
Length in letters (log 10)	−.03	.00
Length in words (log 10)	.09*	−.01

In each correlation, the effects of all other nonaffective variables, as well as those of the affective variables, were partialed out. The numbers in the columns represent Pearson's r values; p values are expressed as specified.
* $p < .05$; ** $p < .01$; *** $p < .001$

linear and quadratic correlations between arousal and imageability (Schmidtke et al., 2014). Therefore, it seems that more research on the relationship between arousal and concreteness in idioms is needed to confirm and generalize our result.

Arousal also showed a moderate positive partial correlation with semantic transparency, in that the more transparent was the meaning of an idiom, the more arousing it was rated. Perhaps it may be easier to attribute high arousal values to idioms in which the literal meaning of the constituent words clearly contributes to the idiomatic interpretation.

Finally, arousal also had a moderate positive partial correlation with familiarity, in line with the results from single words (Montefinese et al., 2014).

Valence and familiarity Emotional valence had a moderate positive partial linear correlation and a significant quadratic relationship with familiarity. The first regression model, which included all psycholinguistic variables and arousal, accounted for 19% of the variance ($R^2 = .19$, $r = .44$), $F(8, 610) = 17.85$, $p < .0001$, whereas the second model, which also included valence and valence squared, accounted for an additional 3% ($R^2 = .22$, $r = .47$), $F(2, 608) = 11.17$, $p < .001$. [The regression line was as follows: estimated familiarity = 0.19 × arousal (log 10) + 0.38 × knowledge – 0.23 × length in letters (log 10) + 0.11 × valence – 0.14 × valence^2.]

In sum, the more positive an idiom, the more familiar it was rated, and the more highly valenced an idiom, the less familiar it was rated. This linear relationship is in line with previous findings on emotional words (Citron et al., 2014b), and the quadratic relationship confirms the first result: In fact, in the present corpus the most highly valenced idioms were mostly negative. This may reflect the bias hypothesized by Citron et al. (2014b) in a normative study, according to which participants may be more prone to declare that they are familiar with positive than with negative concepts (Citron et al., 2014b). However, for idioms, there is a much smaller number of items that convey positive than negative meanings, possibly leading to a higher frequency of use of positive than of negative idioms.

Emotional valence also showed a small positive partial linear correlation with length in words, but no significant quadratic relationship (R^2 change = .004), $F(2, 608) = 2.71$, n.s.

Correlations between nonaffective variables

Familiarity, confidence, and knowledge Familiarity had a large partial positive correlation with knowledge of the idiomatic meaning, suggesting that the more familiar an idiom is, the better it is known; however, we found no significant correlation with confidence (please refer to Table 5). Furthermore, knowledge and confidence were not correlated. These results suggest that measuring participants' knowledge of

idiomatic meanings, without controlling whether or not their knowledge was correct (Bonin et al., 2013; Libben & Titone, 2008; Titone & Connine, 1994a), may be problematic. Only Tabossi et al. (2011) have measured idiom knowledge by asking participants to explain the idiomatic meaning. They also found a large positive correlation between knowledge and familiarity ($r = .49$). However, it should be noted that Tabossi et al. did not test the subjective frequency of idioms, but rather only other-based familiarity (see the Familiarity paragraph in the introduction), therefore obtaining potentially different estimates.

Figurativeness, concreteness, and semantic transparency Figurativeness had a large partial negative correlation with semantic transparency and a moderate negative partial correlation with concreteness. These results suggest that the more figurative a meaning was, the less semantically transparent and the less concrete it was rated. *Semantic transparency* refers to how easily the idiomatic meaning can be inferred from the literal meanings of the constituent words. Therefore, a transparent idiom may be perceived as less figurative than a semantically opaque one. The negative partial correlation between figurativeness and concreteness suggests that the less abstract an idiom was, the less figurative it was considered. It should be noted that we had more abstract than concrete idioms overall, and that this reflects the fact that typically idioms refer to abstract events. Finally, semantic transparency showed a moderate positive partial correlation with concreteness.

Length, familiarity, semantic transparency, and figurativeness Finally, length in letters had a moderate partial negative correlation with familiarity. The tendency of short idioms to be more frequently encountered or produced is in line with what is typically found for single words (Bird, Franklin, & Howard, 2001; Citron et al., 2014b; Stadthagen-Gonzalez & Davis, 2006). Furthermore, idiom length in words showed positive moderate and small partial correlations with figurativeness and semantic transparency, respectively. Perhaps the more semantic information participants had, the easier it was to rate the degree of figurativeness and semantic transparency of the idiom. In contrast, short idioms convey less semantic information, and their interpretation can be less dependent on the literal meaning of the word or word string.

Ambiguous versus unambiguous idioms

Ambiguous idioms had slightly lower valence^2 [$t(602.63) = 3.42$, $p < .01$] (i.e., were less emotionally valenced, independently of whether positively or negatively) and lower arousal mean values [$t(616.41) = 3.35$, $p < .01$] than unambiguous idioms (please refer to Table 6 for descriptive statistics).

Table 5 Linear partial correlations between nonaffective variables

	Familiarity	Concreteness (log 10)	Figurativeness	Sem. Transparency	Confidence (log 10)	Knowledge	Length in Letters (log 10)	Length in Words (log 10)
Familiarity	1							
Concreteness (log 10)	.01	1						
Figurativeness	−.00	−.19***	1					
Sem. transparency	−.03	.14**	−.68***	1				
Confidence (log 10)	−.03	−.03	.04	.00	1			
Knowledge	.38***	−.04	−.03	.02	−.02	1		
Length in letters (log 10)	−.16***	.06	.06	.08	.01	.06	1	
Length in words (log 10)	.05	.00	.13**	.10*	.06	−.05	.71***	1

In each correlation, the effects of all other nonaffective variables as well as the effects of affective variables were partialed out. The numbers in the columns represent Pearson's r values, and p values are expressed as specified below the table. * $p < .05$; ** $p < .01$; *** $p < .001$.

Furthermore, ambiguous idioms were rated as significantly more concrete [$t(617) = 8.85$, $p < .001$] than unambiguous ones, and were less correctly defined [$t(617) = 2.79$, $p < .01$]. How can we interpret these differences? One possibility is that, since ambiguous idioms also have a literal plausible meaning, they may more easily evoke a concrete (literal) meaning as well as a more abstract, idiomatic meaning. This would be consistent with the results of a study by Cacciari and Glucksberg (1995), in which participants were asked to produce and describe a mental image for a set of semantically ambiguous idioms. The images they obtained overwhelmingly reflected the literal meanings of the idiomatic strings, rather than the idiomatic meanings. In addition, these literally oriented mental images interfered with idiomatic paraphrase verification times.

Since unambiguous idioms only possess a figurative interpretation, the fact that they are rated as more arousing and emotionally valenced than ambiguous idioms fits nicely with recent neuroimaging data showing that figurative formulations are more emotionally engaging than their literal counterparts (Citron & Goldberg, 2014).

We did not observe significant differences in the correlations when they were calculated separately for ambiguous and unambiguous idioms.

Reliability analysis

The analyses showed high reliability of the measures of the variables for which we had 30 or more raters (see Table 7). However, we could not obtain high alpha values for the confidence and knowledge variables, due to the small sample sizes (please see Appendix B, Table C2, for more details).

General discussion

The aims of this study were to provide norms for psycholinguistic and affective properties of a large number of German idioms and to explore, for the first time, the relationships between affective and psycholinguistic properties. In what follows, we summarize the main results of this descriptive study, starting from the affective characteristics of the 619 German idioms, since their investigation represents the major contribution of this study.

We found that the more emotionally valenced an idiomatic meaning was, the higher its level of arousal, with negative idioms being evaluated as leading to a higher level of arousal than positive idioms. Although interesting, this result may have been partly influenced by the composition of our idiom list, wherein more than two thirds of the items had negative valence (422 out of 619). However, these results may also reflect the fact that nonliteral language tends to be preferred over literal language when speakers make negative statements (cf. Cacciari, 1998; Drew & Holt, 1988). This is also indirectly supported by the fact that the more figurative an idiom was, the more "arousing" it was rated. Additionally, the concreteness of the idiomatic meaning was positively correlated with emotional arousal, presumably reflecting the fact that concrete concepts with a direct reference to sensory modalities may be seen as more linked to physiological states.

However, since studies on single words have already consistently shown that a high proportion of negative words tend to elicit higher arousal ratings (e.g., Citron et al., 2014b; Võ et al., 2009), our results may not be idiom-specific, but rather reflect a general feature of language.

The observed positive correlations of semantic transparency and arousal may reflect the fact that it may be easier to attribute high arousal values to idioms in which the literal meaning of the constituent words clearly contributes to the idiomatic interpretation. If so, it would be necessary to separate the arousal value of the single words from that of the entire idiom.

Emotional valence had a positive linear correlation and a negative quadratic correlation with familiarity. Since the number of idiomatic expressions used to convey positive concepts is much smaller than the number of negative idiomatic

Table 6 Descriptive statistics of each rated or calculated variable, broken down my meaning type: Ambiguous versus unambiguous

Measure	Variable	Ambiguous Idioms (N = 293)					Unambiguous Idioms (N = 326)				
		Mean	SD	Median	Minimum	Maximum	Mean	SD	Median	Minimum	Maximum
Rating	Em. valence	−0.63	1.21	−1.06	−2.47	2.33	−0.55	1.41	−1.02	−2.81	2.67
	Valence^2	1.86	1.37	1.66	0.00	6.10	2.29	1.78	2.04	0.00	7.90
	Arousal	3.71	0.65	3.65	2.26	5.35	3.92	0.79	3.84	2.21	6.03
	Familiarity	4.75	0.67	4.83	2.68	6.24	4.78	0.69	4.85	2.67	6.21
	Concreteness	3.63	0.81	3.60	1.69	5.81	3.08	0.70	3.00	1.66	5.81
	Figurativeness	4.67	0.69	4.69	2.39	6.09	4.72	0.60	4.78	2.97	6.00
	Sem. transparency	4.06	1.01	4.06	1.87	6.19	4.20	0.93	4.22	1.83	6.22
	Confidence	6.12	0.53	6.23	3.87	6.86	6.09	0.57	6.22	2.54	6.90
% correct definitions	Knowledge	89.61	10.84	93.00	40.00	100.00	91.91	9.74	96.00	42.00	100.00
Value	Length in letters	23.00	5.00	23.00	9.00	41.00	23.00	6.00	22.00	10.00	43.00
	Length in words	5.00	1.00	4.00	2.00	8.00	4.00	1.00	4.00	2.00	9.00

Table 7 Measures of internal consistency (Cronbach's α) for variables with a minimum sample size of $N = 30$

Variables	Cronbach's α	αs If Each Item Deleted	Participant No.
Emotional valence	.98	all > .98	32
Arousal	.89	all > .88	31
Familiarity	.80	all > .78	30
Concreteness	.83	all > .81	31
Figurativeness	.86	all > .84	32
Sem. transparency	.93	all > .92	32

expressions, positive idioms may be more frequently used than negative ones.

We now turn to the relationships between psycholinguistic variables. Familiarity (i.e., subjective frequency) was positively correlated with knowledge of the idiomatic meaning, confirming previous findings (Bonin et al., 2013; Libben & Titone, 2008; Tabossi et al., 2011; Titone & Connine, 1994b). However, unlike in previous studies, we did not observe any significant correlation between familiarity and confidence or between knowledge and confidence. These discrepancies may reflect, at least in part, differences in the ways in which these variables were measured. For example, Tabossi et al. (2011) tested idiom knowledge by asking participants to write down the idiomatic meaning (as in our study) and measured other-based familiarity. In contrast, other studies reported an estimate of participants' own knowledge (i.e., their confidence) and measured familiarity in terms of subjective frequency (Bonin et al., 2013; Libben & Titone, 2008). Our results suggest that confidence may not necessarily be a reliable measure of the actual knowledge of an idiomatic meaning, which is better captured by asking participants to write the meaning down. Furthermore, idiom familiarity, conceptualized in terms of subjective frequency of exposure to an idiom, can provide a more reliable measure than other-based familiarity.

The perceived level of figurativeness of an idiom was negatively correlated with concreteness and semantic transparency; specifically, the more idiomatic a meaning was, the less semantically transparent and concrete it was rated. In sum, the meanings of most idiomatic strings were unrelated to the literal meaning of the constituent words, and predominantly conveyed abstract contents. Shorter idioms were perceived as being more familiar, but longer idioms provided more semantic information than shorter ones, facilitating the evaluation of idioms' figurativeness as well as semantic transparency.

Semantic transparency was not correlated with familiarity (in line with Tabossi et al., 2011; but see Abel, 2003) or with idiom knowledge (unlike Tabossi et al., 2011). This seems to reflect the fact that the idiomatic meaning of known, familiar idioms is stored in semantic memory and retrieved regardless of the fact that we detect a clear relationship between the component word meanings and the global figurative interpretation of the string (Bonin et al., 2013; Libben & Titone, 2008; Titone & Connine, 1994b). It should also be mentioned that the notion of semantic transparency is hard to capture, varies across participants, and often reflects a post-hoc attribution of a link between the idiomatic meaning that we have already apprehended and the individual words (Cacciari, 2014).

Ambiguous idioms were overall less emotionally salient (i.e., were rated as less valenced and arousing), less correctly defined, and more linked to concrete, sensory-based information than were unambiguous idioms. Since people cannot bypass the meanings of the constituent words en route to accessing (or generating) the idiom's figurative meaning, this might represent a source of possible interference in ambiguous idioms, leading to wrong meaning definitions. However, when the correct meaning of ambiguous idioms is known, this may evoke a more intense emotional response with a more direct link to sensory domains.

Finally, this descriptive study provides a useful tool for researchers interested in exploring the relationships between figurative language and affect using German figurative expressions with empirically determined variables. To our knowledge, this is the first descriptive study on idioms that provides ratings for affective variables, beyond other psycholinguistic variables. It also shows high reliability—that is, internal consistency. In addition, variables such as concreteness and figurativeness were not tested in previous idiom norms, although they have been shown to correlate with affective and psycholinguistic properties of idioms. Thus, all of these variables should be taken into account when designing experiments on idiom processing.

Appendix A: Instructions for the online rating of German idioms

A screenshot of one of the questionnaires is also included. *Important note*: We acknowledge that the labeling of arousal in terms of "emotional intensity" in the instructions is not accurate; in fact, *arousal* refers to the degree of physiological activation evoked by an emotional stimulus. However, we

found the immediately following definition—that is, "to what extent an event is stimulating on a scale from not at all stimulating to very exciting or agitating"—to be quite accurate with respect to physiological activation.

ALLGEMEINE INSTRUKTIONEN

Lieber Teilnehmer,

Willkommen zu unserer Umfrage.

Im Folgenden werden dir metaphorische Sätze präsentiert, d.h. Sätze deren Gesamtbedeutung nicht aus der Bedeutung der einzelnen Wörter hergeleitet werden kann. Ein Beispiel: Den Satz „Sie ist sauer auf ihn" versteht man, wenn man weiß, dass „sauer" hier metaphorisch „wütend" bedeutet und nicht sauer im wortgetreuen Sinne des Geschmacks. Im Satz „Er musste eine schwere Entscheidung treffen" bedeutet „schwer" metaphorisch „schwierig bzw. kompliziert" und nicht schwer im wortgetreuen Sinne des Gewichts.

Wir bitten Dich, jeden Satz hinsichtlich einer Eigenschaft zu bewerten. Die genauen Instruktionen zu der Eigenschaft erhältst Du vor der Satz-Liste.

Damit wir Deine Daten auswerten können, ist es wichtig, dass Du keine Bewertung auslässt und den Fragebogen komplett ausfüllst.

Vielen Dank!

SPEZIFISCHE INSTRUKTIONEN

KENNTNISSE ÜBER DIE METAPHORISCHE BEDEUTUNG

Wir bitten dich deine Kenntnisse über die metaphorische Bedeutung unterschiedlicher Sätze zu bewerten, auf einer Skala von 1 bis 7. Hier bedeutet 1 „Ich kenne die Bedeutung der Metapher gar nicht", 4 bedeutet „Ich kenne die Bedeutung der Metapher ungefähr" und 7 bedeutet „Ich kenne die Bedeutung der Metapher sehr gut".

Nach der Bewertung jedes Satzes, schreibe bitte, was die Metapher deiner Meinung nach bedeutet.

GEBRÄUCHLICHKEIT

Wir bitten dich zu bewerten, wie gebräuchlich ein metaphorischer Satz für dich ist, d.h. wie oft du ihn gehört oder gelesen hast. Zum Beispiel wird der Satz „Ich bin tot" im Sinne von „erschöpft" sehr häufig benutzt, wohingegen der Satz "vom Hundertsten ins Tausendste kommen" fast nie benutzt wird.

Die Bewertungsskala erstreckt sich von 1 bis 7. Hier bedeutet 1 „Ich habe die Metapher nie gehört/gelesen", 4 bedeutet „Ich habe die Metapher manchmal gehört/gelesen" und 7 bedeutet „Ich habe die Metapher sehr oft gehört/gelesen".

KONKRETHEIT

Ein Satz überträgt eine konkrete Bedeutung, wenn er auf eine Erfahrung in einem oder mehreren der fünf Sinne hindeutet (z.B. Tasten, Sehen). Ein Beispiel: die Bedeutung des Satzes „Der Hund wird gestreichelt" ist konkret, weil sie sich auf die Erfahrung vom Tasten bezieht. Im Gegensatz

hierzu ist die Bedeutung des Satzes „Sie hatte einen traurigen Gedanken" abstrakt, weil sie mit keinem der fünf Sinne assoziiert werden kann.

Wir bitten dich zu bewerten, in welchem Ausmaß die metaphorische Bedeutung jedes Satzes auf eine Erfahrung, einen Zustand oder ein Ereignis in einem oder mehreren der fünf Sinne hindeutet. Die Bewertungsskala erstreckt sich von 1 bis 7, wobei 1 bedeutet „die metaphorische Bedeutung ist sehr abstrakt, d.h. kann in KEINEM der fünf Sinne erfahren werden" und 7 bedeutet „die metaphorische Bedeutung ist sehr konkret, d.h. kann sehr einfach in (zumindest) EINEM der fünf Sinne erfahren werden."

SEMANTISCHE DURCHSICHTIGKEIT

Semantische Durchsichtigkeit bezieht sich auf das Ausmaß, in welchem man die metaphorische Bedeutung eines Satzes aus der Bedeutung der einzelnen Wörter herleiten kann. Beispiel: „Dieses Kind ist eine Klette" ist semantisch durchsichtig, weil man leicht erahnen kann, dass das Kind lästig und anhänglich ist, auch wenn man diesen metaphorischen Satz noch nie im Leben gehört oder gelesen hat. Im Gegensatz dazu ist „Ich habe einen Kater" semantisch undurchsichtig, da man aus den einzelnen Wörtern nicht die metaphorische Bedeutung „Ich fühle mich schlecht von zu viel Alkohol" ableiten kann.

Wir bitten dich für jeden Satz zu bewerten, wie einfach es ist, aus der wortgetreuen Bedeutung der einzelnen Wörter die metaphorische Bedeutung abzuleiten auf einer Skala von 1 „überhaupt nicht ableitbar, d.h. sehr undurchsichtig" bis 7 „sehr einfach ableitbar, d.h. sehr durchsichtig".

METAPHORIZITÄT

Wir bitten dich zu bewerten, in welchem Ausmaß die Bedeutung jedes Satzes metaphorisch ist. Beispiel: In dem Satz „Sie sieht so alt aus wie meine Oma" lässt sich „Oma" leicht mit der Eigenschaft „sehr alt" assoziieren, weshalb der Satz wenig metaphorisch ist und sehr ähnlich mit seiner wortgetreuen Entsprechung „sie sieht sehr alt aus". Im Gegensatz dazu hat der Satz „Ich habe einen Kater" nichts mit dem Besitzen einer männlichen Katze zu tun und ist deshalb sehr metaphorisch. Wir bitten dich, die folgenden Sätze auf einer Skala von 1 „überhaupt nicht metaphorisch" bis 7 „sehr metaphorisch" zu bewerten. Bitte beachte, dass alle Sätze, die du bewertest, mindestens ein bisschen metaphorisch sind; deshalb wirst du den Skalenpunkt 1 fast nie benutzen.

EMOTIONALE VALENZ

Emotionale Valenz beschreibt, in welchem Ausmaß ein Ereignis positiv oder negativ ist. Zum Beispiel beschreibt der Satz „Die dunkle Vorhersage enttäuschte sie" ein negatives Ereignis, wohingegen der Satz „Er ist vor Freude an die Decke gesprungen" etwas Positives beschreibt.

Wir bitten dich für jeden Satz auf einer Skala von -3 bis +3 zu bewerten, wie positiv oder negativ seine metaphorische Bedeutung ist. -3 bedeutet „sehr negativ", +3 bedeutet „sehr positiv" und 0 bedeutet „weder positiv noch negativ", wie z.B. „Der Stuhl steht im Zimmer".

EMOTIONALE INTENSITÄT (AROUSAL)

Emotionale Intensität bezeichnet, in welchem Ausmaß ein Ereignis stimulierend ist auf einer Skala von gar nicht stimulierend bis sehr aufregend oder rührend. Zum Beispiel wird der Satz „Er schrie vor Schmerz" als sehr intensiv empfunden, ebenso „Sie tanzte voller Freude". Im Gegensatz dazu werden Sätze wie z.B. "Wolken bildeten sich am Himmel" oder „Sie las ein gutes Buch" als nicht intensiv empfunden.

Wir bitten dich die Intensität der metaphorischen Bedeutung jedes Satzes zu bewerten auf einer Skala von 1 bis 7, wobei 1 „überhaupt nicht intensiv" bedeutet, 4 „mäßig intensiv" und 7 „sehr intensiv" bedeutet.

BEMERKUNG: DER PARAGRAPH HIER UNTEN WIRD NACH JEDER SPEZIFISCHEN INSTRUKTION ERSCHEINEN

Es gibt keine Zeitbegrenzung. Wir möchten aber, dass du nicht zu lange über jeden Satz nachdenkst sondern eher spontan entscheidest. Falls du die Bedeutung eines Satzes nicht kennst, kreuze bitte die Option „unbekannt" an.

Bitte klicke jetzt auf START

Instructions for the online rating of German idioms (English translation)

Important note: This is a translation of the German instructions into English for the sake of transferring information to non-German speakers. Therefore, most sentence examples would not make any sense to English or non-German speakers: They are just translations of the original German figurative expressions.

GENERAL INSTRUCTIONS

Dear participant, welcome to our survey.

You will be presented with nonliteral sentences, i.e., sentences whose overall meaning cannot be extracted by the meaning of its single words. For example, the sentence "she is sour at him" can only be understood if one knows that here "sour" metaphorically means "angry" and not sour in the literal sense of taste. In the sentence "he had to take a heavy decision," "heavy" metaphorically means "difficult, complicated" and not heavy in the literal sense of weight.

We would like you to judge specific characteristics of each of the sentences that form the lists below. Follow the instructions that are specified at the beginning of each list.

In order to be able to use your data, it is important that you rate each sentence and that you complete the whole questionnaire.

Thanks a lot!

SPECIFIC INSTRUCTIONS

KNOWLEDGE OF THE METAPHORICAL MEANING

Please rate your knowledge of the metaphorical meaning of each of the following sentences. The rating scale goes from 1 to 7 where 1 means "I don't know the meaning of the metaphor at all," and 7 means "I know the meaning of the metaphor very well." A rating of 4 means "I know the meaning of the metaphor only approximately."

Just after each rating, please write down what you think the metaphor means.

SUBJECTIVE FREQUENCY (the term used in German is kind of "USAGE")

Please rate the frequency with which you have heard or read each of the following metaphorical expressions. For instance, the sentence "I am dead" in the sense of "exhausted" is used very frequently, whereas the sentence "to come from the hundredth into the thousandth" is almost never used. The rating scale goes from 1 to 7 where 1 means "I never heard/read this expression," and 7 "I often heard/read this expression." A rating of 4 means "I sometimes heard/read this expression."

CONCRETENESS

A sentence conveys a concrete meaning when it denotes an experience that refers to one or more sensory modalities (e.g., touch, vision). For instance, the meaning of "the dog is/gets stroked" refers to the experience of touch. In contrast, the meaning of "she had a sad thought" is abstract since it cannot be associated with a sensory modality.

Please rate the extent to which the metaphorical meaning of the sentence refers to an experience, state or event that can be experienced with the senses. The rating scale goes from 1 to 7 where 1 means "the metaphorical meaning is totally abstract," and 7 means "the metaphorical meaning is totally concrete."

SEMANTIC TRANSPARENCY

Semantic transparency refers to the extent to which you can infer the metaphorical meaning of a sentence from the meaning of the individual words that compose the expression. For instance, "the child is a limpet" can be considered semantically transparent since you can easily infer that the child is annoying and dependent (metaphorical meaning) from the literal meaning of the words forming the expression, even if you have never heard it before. In contrast, "I have a cat" is

semantically opaque since the meaning of the words forming the expression has nothing to do with its metaphorical meaning "I feel sick because of too much alcohol."

Please rate the semantic transparency of each metaphorical expression using the rating scale that goes from 1 to 7 where 1 means "the metaphorical meaning is semantically transparent," and 7 means "the metaphorical meaning is semantically opaque."

FIGURATIVENESS (in German METAPHORICITY is more easily understood)

Please rate the extent to which the meaning of each sentence is metaphorical. For instance, in the expression "she looks like my grandma" one can easily associate "grandma" with the feature "very old" and therefore this sentence is (a) little metaphorical and very similar to its literal counterpart "she looks very old." On the contrary, the expression "I have a cat" has nothing to do with the possession of a male cat and therefore it is very metaphorical.

Please rate the following sentences on a scale from 1 "not at all metaphorical" to 7 "very metaphorical." Please bear in mind that all sentences you rate are somehow metaphorical; therefore you will almost never use the extreme point 1.

EMOTIONAL VALENCE

Valence describes the extent to which an event is positive or negative. For example, the sentence "the dark forecast disappointed her" describes a negative event, whereas the sentence "he jumped on the blanket for joy" describes something positive.

Please rate how positive or negative the metaphorical meaning of the following sentences is, on a scale from –3 to +3. –3 means "very negative," +3 means "very positive," and 0 means "neither positive nor negative," e.g., "the chair is in the room."

AROUSAL (in German EMOTIONAL INTENSITY)

Emotional intensity describes to what extent an event is stimulating on a scale from not at all stimulating to very exciting or agitating. For example, the sentence "he screamed because of pain" is perceived as very intense as well as the sentence "she danced of full joy." On the contrary, sentences such as "clouds were appearing in the sky" or "she read a good book" are perceived as not intense.

Please rate the intensity of the metaphorical meaning of the following sentences on a scale from 1 to 7, whereby 1 means "not intense at all," 4 means "quite intense" and 7 means "very intense."

NOTE: THE PARAGRAPH BELOW WILL APPEAR AFTER EACH FEATURE'S INSTRUCTIONS

There is no time limit for the response. Nevertheless, we would like you to not think about each sentence for too long but to decide spontaneously instead. In case you don't know the meaning of a sentence, please tick the button "unknown."

Please click now on START

KENNTNISSE ÜBER DIE METAPHORISCHE BEDEUTUNG

Wir bitten dich deine Kenntnisse über die metaphorische Bedeutung unterschiedlicher Sätze zu bewerten, auf einer Skala von 1 bis 7. Hier bedeutet 1 „Ich kenne die Bedeutung der Metapher gar nicht", 4 bedeutet „Ich kenne die Bedeutung der Metapher ungefähr" und 7 bedeutet „Ich kenne die Bedeutung der Metapher sehr gut".

Nach der Bewertung jedes Satzes, schreibe bitte, was die Metapher deiner Meinung nach bedeutet. Falls du die Metapher nicht kennst, trage bitte "unbekannt" in das Feld ein.

* 7. aus der Haut fahren

1 überhaupt nicht	2	3	4 mäßig	5	6	7 sehr	unbekannt
○	○	○	○	○	○	○	○

Bedeutung

* 8. aufs Ganze gehen

1 überhaupt nicht	2	3	4 mäßig	5	6	7 sehr	unbekannt
○	○	○	○	○	○	○	○

Bedeutung

Appendix B

Table 8 Pearson's correlations between each pair of psycholinguistic and affective variables (without partialing out the effects of the other variables)

	Em. Valence	Valence Squared	Arousal (log 10)	Familiarity	Concrete-ness (log 10)	Figura-tiveness	Sem. Trans-parency	Confidence (log 10)	Knowl-edge	Length in Letters (log 10)	Length in Words (log 10)
Em. valence	1										
Valence squared	−.27***	1									
Arousal (log 10)	−.28***	.54***	1								
Familiarity	.13**	−.01	.10*	1							
Concreteness (log 10)	−.06	−.04	.07	−.03	1						
Figurativeness	−.10*	.03	.18***	.00	−.36***	1					
Sem. transparency	.07	.16***	.07	−.02	.37***	−.68***	1				
Confidence (log 10)	−.00	−.02	−.04	−.06	−.05	.08	−.05	1			
Knowledge	.07	.13**	.06	.39***	−.02	−.05	.05	−.05	1		
Length in letters (log 10)	.01	−.08	.05	−.18***	.10*	.06	.14**	.08*	−.04	1	
Length in words (log 10)	.07	−.05	.04	−.10*	.06	.10**	.11**	.10**	−.05	.73***	1

Table 9 Measures of internal consistency (Cronbach's α) for variables with small sample sizes

Questionnaires	Confidence		Knowledge		Participant No.	Idiom No.
	Cronbach's α	αs If Each Item Deleted	Cronbach's α	αs If Each Item Deleted		
List 1, 1st half	.65	all > .60	.37	all > .31	12	310
List 1, 2nd half	.44	all > .34	.51	all > .45	10	309
List 2, 1st half	.30	all > .17	.29	all > .17	10	310
List 2, 2nd half	.33	all > .15	.34	all > .26	9	309
List 3, 1st half	.51	all > .40	.49	all > .37	10	309
List 3, 2nd half	.38	all > .25	.50	all > .37	9	310

References

Abel, B. (2003). English idioms in the first language and second language lexicon: A dual representation approach. *Second Language Research, 19,* 329–358.

Adorni, R., & Proverbio, A. M. (2012). The neural manifestation of the word concreteness effect: An electrical neuroimaging study. *Neuropsychologia, 50,* 880–891. doi:10.1016/j.neuropsychologia.2012.01.028

Algom, D., Chajut, E., & Lev, S. (2004). A rational look at the emotional Stroop phenomenon: A generic slowdown, not a Stroop effect. *Journal of Experimental Psychology, 133,* 323–338. doi:10.1037/0096-3445.133.3.323

Altarriba, J., & Bauer, L. M. (2004). The distinctiveness of emotion concepts: A comparison between emotion, abstract and concrete words. *American Journal of Psychology, 117,* 389–410.

Altmann, U., Bohrn, I., Lubrich, O., Menninghaus, W., & Jacobs, A. M. (2012). The power of emotional valence—From cognitive to affective processes in reading. *Frontiers in Human Neuroscience, 6,* 192. doi:10.3389/fnhum.2012.00192

Barrett, L. F., & Russell, J. A. (1998). Independence and bipolarity in the structure of current affect. *Journal of Personality and Social Psychology, 74,* 967–984.

Bayer, M., Sommer, W., & Schacht, A. (2012). P1 and beyond: Functional separation of multiple emotion effects in word recognition. *Psychophysiology, 49,* 959–969. doi:10.1111/j.1469-8986.2012.01381.x

Bibliographisches Institut GmbH. (2013). *Duden* [Online database]. Retrieved February, 2015, from www.duden.de/rechtschreibung/figurativ

Bird, H., Franklin, S., & Howard, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior Research Methods, Instruments, & Computers, 33,* 73–79. doi:10.3758/BF03195349

Bobrow, S. A., & Bell, S. M. (1973). On catching on to idiomatic expressions. *Memory & Cognition, 1,* 343–346. doi:10.3758/BF03198118

Bohrn, I., Altmann, U., & Jacobs, A. M. (2012a). Looking at the brains behind figurative language—A quantitative meta-analysis of neuroimaging studies on metaphor, idiom, and irony processing. *Neuropsychologia, 50,* 2669–2683. doi:10.1016/j.neuropsychologia.2012.07.021

Bohrn, I., Altmann, U., Lubrich, O., Menninghaus, W., & Jacobs, A. M. (2012b). Old proverbs in new skins—an fMRI study on defamiliarization. *Frontiers in Psychology, 3,* 204. doi:10.3389/fpsyg.2012.00204

Bonin, P., Méot, A., & Bugaiska, A. (2013). Norms and comprehension times for 305 French idiomatic expressions. *Behavior Research Methods, 45,* 1259–1271. doi:10.3758/s13428-013-0331-4

Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Simuli, instruction manual and affective ratings*. Gainesville: NIMH Center for Research in Psychophysiology, University of Florida.

Cacciari, C. (1998). Why do we speak metaphorically? Reflections on the functions of metaphor in discourse and reasoning. In A. N. Katz, C. Cacciari, R. W. Gibbs, & M. Turner (Eds.), *Figurative language and thought* (pp. 119–158). Oxford: Oxford University Press.

Cacciari, C. (2014). Processing multiword idiomatic strings: Many words in one? *Mental Lexicon, 9,* 267–293.

Cacciari, C., & Glucksberg, S. (1994). Understanding figurative language. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 447–477). San Diego: Academic Press.

Cacciari, C., & Glucksberg, S. (1995). Understanding idioms: Do visual images reflect figurative meaning? *European Journal of Cognitive Psychology, 7,* 283–305.

Cacciari, C., & Papagno, C. (2012). Neuropsychological and neurophysiological correlates of idiom understanding: How many hemispheres are involved? In M. Faust (Ed.), *The handbook of the neuropsychology of language* (pp. 368–385). New York: Wiley Blackwell.

Cacciari, C., & Tabossi, P. (1988). The comprehension of idioms. *Journal of Memory and Language, 27,* 668–683.

Cacciari, C., Reati, F., Colombo, M. R., Padovani, R., Rizzo, S., & Papagno, C. (2006). The comprehension of ambiguous idioms in aphasic patients. *Neuropsychologia, 44,* 1305–1314. doi:10.1016/j.neuropsychologia.2006.01.012

Cacciari, C., Padovani, R., & Corradini, P. (2007). Exploring the relationship between individuals' speed of processing and their comprehension of spoken idioms. *European Journal of Cognitive Psychology, 19,* 417–445. doi:10.1080/09541440600763705

Caillies, S. (2009). Descriptions de 300 expressions idiomatiques: Familiarité, connaissance de leur signification, plausibilité littérale, "décomposabilité" et "prédictabilité.". *L'Année Psychologique, 109,* 463–508.

Cicchetti, D. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6,* 284–290.

Citron, F. M. M. (2012). Neural correlates of written emotion word processing: A review of recent electrophysiological and hemodynamic neuroimaging studies. *Brain and Language, 122,* 211–226. doi:10.1016/j.bandl.2011.12.007

Citron, F. M. M., & Goldberg, A. E. (2014). Metaphorical sentences are more emotionally engaging than their literal counterparts. *Journal of Cognitive Neuroscience, 26,* 2585–2595. doi:10.1162/jocn_a_00654

Citron, F. M. M., Gray, M., Critchley, H. D., Weekes, B. S., & Ferstl, E. C. (2014). Emotional valence and arousal affect reading in an interactive way: Neuroimaging evidence for an approach-withdrawal framework. *Neuropsychologia, 56,* 79–89. doi:10.1016/j.neuropsychologia.2014.01.002

Citron, F. M. M., Weekes, B. S., & Ferstl, E. C. (2014a). Arousal and emotional valence interact in written word recognition. *Language, Cognition and Neuroscience, 29,* 1257–1267. doi:10.1080/23273798.2014.897734

Citron, F. M. M., Weekes, B. S., & Ferstl, E. C. (2014b). How are affective word ratings related to lexico-semantic properties? Evidence from the Sussex Affective Word List (SAWL). *Applied Psycholinguistics, 35,* 313–331. doi:10.1017/S0142716412000409

Cronk, B. C., & Schweigert, W. A. (1992). The comprehension of idioms: The effects of familiarity, literalness, and usage. *Applied Psycholinguistics, 13,* 131–146.

Cronk, B. C., Lima, S. D., & Schweigert, W. A. (1993). Idioms in sentences: Effects of frequency, literalness, and familiarity. *Journal of Psycholinguistic Research, 22,* 59–82.

Drew, P., & Holt, E. (1988). Complainable matters: The use of idiomatic expressions in making complaints. *Social Problems, 35,* 398–417.

Drew, P., & Holt, E. (1998). Figures of speech: Figurative expressions and the management of topic transition in conversation. *Language in Society, 27,* 495–522.

Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap*. Boca Raton: Chapman & Hall/CRC.

Estes, Z., & Adelman, J. S. (2008). Automatic vigilance for negative words is categorical and general: Comment reply. *Emotion, 8,* 453–457. doi:10.1037/a0012887

Fainsilber, L., & Ortony, A. (1987). Metaphorical uses of language in the expression of emotions. *Metaphor and Symbolic Activity, 2,* 239–250.

Fanari, R., Cacciari, C., & Tabossi, P. (2010). The role of idiom length and context in spoken idiom comprehension. *European Journal of Cognitive Psychology, 22,* 321–334.

Fellbaum, C., & Geyken, A. (2005). Transforming a corpus into a lexical resource: The Berlin idiom project. *Revue Française de Linguistique Appliquée, X,* 49–62.

Ferretti, T. R., Schwint, C. A., & Katz, A. N. (2007). Electrophysiological and behavioral measures of the influence of literal and figurative contextual constraints on proverb comprehension. *Brain and Language, 101,* 38–49. doi:10.1016/j.bandl.2006.07.002

Ferstl, E. C., & von Cramon, D. Y. (2007). Time, space and emotion: fMRI reveals content-specific activation during text comprehension. *Neuroscience Letters, 427,* 159–164. doi:10.1016/j.neulet.2007.09.046

Ferstl, E. C., Rinck, M., & von Cramon, D. Y. (2005). Emotional and temporal aspects of situation model processing during text comprehension: An event-related fMRI study. *Journal of Cognitive Neuroscience, 17,* 724–739. doi:10.1162/0898929053747658

Forgács, B., Bohrn, I., Baudewig, J., Hofmann, M. J., Pléh, C., & Jacobs, A. M. (2012). Neural correlates of combinatorial semantic processing of literal and figurative noun noun compound words. *NeuroImage, 63,* 1432–1442. doi:10.1016/j.neuroimage.2012.07.029

Forgács, B., Lukács, Á., & Pléh, C. (2014). Lateralized processing of novel metaphors: Disentangling figurativeness and novelty. *Neuropsychologia, 56,* 101–109. doi:10.1016/j.neuropsychologia.2014.01.003

Fussell, S. R., & Moss, M. M. (1998). *Figurative language in emotional communication [Working paper].* Pittsburgh: Carnegie Mellon University, Human–Computer Interaction Institute. Retrieved from http://repository.cmu.edu/cgi/viewcontent.cgi?article=1081&context=hcii

Gibbs, R. W., Jr., Nayak, N. P., & Cutting, C. (1989). How to kick the bucket and not decompose: Analyzability and idiom processing. *Journal of Memory and Language, 28,* 576–593. doi:10.1016/0749-596X(89)90014-4

Glucksberg, S. (2001). *Understanding figurative language: From metaphors to idioms.* New York: Oxford University Press.

Herbert, C., Ethofer, T., Anders, S., Junghofer, M., Wildgruber, D., Grodd, W., & Kissler, J. (2009). Amygdala activation during reading of emotional adjectives—An advantage for pleasant content. *Social Cognitive and Affective Neuroscience, 4,* 35–49. doi:10.1093/scan/nsn027

Hofmann, M. J., Kuchinke, L., Tamm, S., Võ, M. L.-H., & Jacobs, A. M. (2009). Affective processing within 1/10th of a second: High arousal is necessary for early facilitative processing of negative but not positive words. *Cognitive, Affective, & Behavioural Neuroscience, 9,* 389–397. doi:10.3758/9.4.389

Jackendoff, R. (1995). *Languages of the mind: Essays on mental representation.* Cambridge: MIT Press.

Jacobs, A. M. (2011). Neurokognitive Poetik: Elemente eines Modells des literarischen Lesens. In R. Schrott & A. M. Jacobs (Eds.), *Gehirn und Gedicht: Wie wir unsere Wirklichkeiten konstruieren* (pp. 492–520). Munich: Hanser.

Jacobs, A. M., Võ, M. L.-H., Briesemeister, B. B., Conrad, M., Hofmann, M. J., Kuchinke, L., & Braun, M. (2015). Ten years of BAWLing into affective and aesthetic processes in reading: What are the echoes? *Manuscript under review at Frontiers in Psychology.*.

Keysar, B., & Bly, B. (1995). Intuitions of the transparency of idioms: Can one keep a secret by spilling the beans? *Journal of Memory and Language, 34,* 89–109. doi:10.1006/jmla.1995.1005

Kissler, J., Assadollahi, R., & Herbert, C. (2006). Emotional and semantic networks in visual word processing: insights from ERP studies. *Progress in Brain Research, 156,* 147–183. doi:10.1016/S0079-6123(06)56008-X

Konopka, A. E., & Bock, K. (2009). Lexical or syntactic control of sentence formulation? Structural generalizations from idiom production. *Cognitive Psychology, 58,* 68–101. doi:10.1016/j.cogpsych.2008.05.002

Kousta, S.-T., Vinson, D. P., & Vigliocco, G. (2009). Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition, 112,* 473–481. doi:10.1016/j.cognition.2009.06.007

Kuchinke, L., Jacobs, A. M., Gubrich, C., Võ, M. L.-H., Conrad, M., & Herrmann, M. (2005). Incidental effects of emotional valence in single word processing: An fMRI study. *NeuroImage, 28,* 1022–1032. doi:10.1016/j.neuroimage.2005.06.050

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). Motivated attention: Affect, activation, and action. In P. J. Lang, R. F. Simons, & M. T. Balaban (Eds.), *Attention and orienting: Sensory and motivational processes* (pp. 97–135). Hillsdale: Erlbaum.

Larsen, R. J., Mercer, K. A., & Balota, D. A. (2006). Lexical characteristics of words used in emotional Stroop experiments. *Emotion, 6,* 62–72. doi:10.1037/1528-3542.6.1.62

Levorato, C., & Cacciari, C. (1999). Idiom comprehension in children: Are the effects of semantic analysability and context separable? *European Journal of Cognitive Psychology, 11,* 51–66.

Lewis, P. A., Critchley, H. D., Rotshtein, P., & Dolan, R. J. (2007). Neural correlates of processing valence and arousal in affective words. *Cerebral Cortex, 17,* 742–748. doi:10.1093/cercor/bhk024

Libben, M. R., & Titone, D. A. (2008). The multidetermined nature of idiom processing. *Memory & Cognition, 36,* 1103–1121. doi:10.3758/MC.36.6.1103

Montefinese, M., Ambrosini, E., Fairfield, B., & Mammarella, N. (2014). The adaptation of the Affective Norms for English Words (ANEW) for Italian. *Behavior Research Methods, 46,* 887–903. doi:10.3758/s13428-013-0405-3

Mueller, R. A. G., & Gibbs, R. W. (1987). Processing idioms with multiple meanings. *Journal of Psycholinguistic Research, 16,* 63–81.

Nasrallah, M., Carmel, D., & Lavie, N. (2009). Murder, she wrote: Enhanced sensitivity to negative word valence. *Emotion, 9,* 609–618. doi:10.1037/a0016350

Paivio, A. (2007). *Mind and its evolution: A dual coding theoretical approach.* Mahwah: Erlbaum.

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology, 76*(1, Pt. 2), 1–25. doi:10.1037/h0025327

Papagno, C., & Cacciari, C. (2010). The role of ambiguity in idiom comprehension: The case of a patient iwth a reversed concreteness effect. *Journal of Neurolinguistics, 23,* 631–643. doi:10.1016/j.neuroling.2010.06.002

Papagno, C., & Caporali, A. (2007). Testing idiom comprehension in aphasic patients: The effects of task and idiom type. *Brain and Language, 100,* 208–220. doi:10.1016/j.bandl.2006.01.002

Pollio, H. R., Barlow, J. M., Fine, H. J., & Pollio, M. R. (1977). *Psychology and the poetics of growth: Figurative language in psychology, psychotherapy, and education.* Hillsdale: Erlbaum.

Recio, G., Conrad, M., Hansen, L. B., & Jacobs, A. M. (2014). On pleasure and thrill: The interplay between arousal and valence during visual word recognition. *Brain and Language, 134,* 34–43. doi:10.1016/j.bandl.2014.03.009

Reisenzein, R. (1994). Pleasure–arousal theory and the intensity of emotions. *Journal of Personality and Social Psychology, 3,* 525–539.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39,* 1161–1178. doi:10.1037/h0077714

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110,* 145–172. doi:10.1037/0033-295X.110.1.145

Sabsevitz, D. S., Medler, D. A., Seidenberg, M. S., & Binder, J. R. (2005). Modulation of the semantic system by word imageability. *NeuroImage, 27,* 188–200. doi:10.1016/j.neuroimage.2005.04.012

Schmidtke, D. S., Schröder, T., Jacobs, A. M., & Conrad, M. (2014). ANGST: Affective norms for German sentiment terms, derived from the affective norms for English words. *Behavior Research Methods, 46,* 1108–1118. doi:10.3758/s13428-013-0426-y

Schrott, R., & Jacobs, A. M. (2011). *Gehirn und Gedicht: Wie wir unsere Wirklichkeiten konstruieren*. Munich: Hanser.

Stadthagen-Gonzalez, H., & Davis, C. J. (2006). The Bristol norms for age of acquisition, imageability, and familiarity. *Behavior Research Methods, 38,* 598–605. doi:10.3758/BF03193891

Swinney, D. A., & Cutler, A. (1979). The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behavior, 18,* 523–534. doi: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,* 313–327. doi:10.1037/0278-7393.34.2.313

Tabossi, P., Arduino, L., & Fanari, R. (2011). Descriptive norms for 245 Italian idiomatic expressions. *Behavior Research Methods, 43,* 110–123. doi:10.3758/s13428-010-0018-z

Titone, D. A., & Connine, C. M. (1994a). Comprehension of idiomatic expressions: Effects of predictability and literality. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 1126–1138. doi:10.1037/0278-7393.20.5.1126

Titone, D. A., & Connine, C. M. (1994b). Descriptive norms for 171 idiomatic expressions: Familiarity, compositionality, predictability and literality. *Metaphor and Symbolic Activity, 9,* 247–270.

Turner, N., & Katz, A. (1997). The availability of conventional and of literal meaning during the comprehension of proverbs. *Pragmatics and Cognition, 5,* 199–233.

Udem, P. (2001). *Redensarten-Index: Wörterbuch für Redensarten, Redewendungen, idiomatische Ausdrücke, feste Wortverbindungen.* Retrieved December, 2013, from www.redensarten-index.de

Vespignani, F., Canal, P., Molinaro, N., Fonda, S., & Cacciari, C. (2009). Predictive mechanisms in idiom comprehension. *Journal of Cognitive Neuroscience, 22,* 1682–1700. doi:10.1162/jocn.2009.21293

Võ, M. L.-H., Conrad, M., Kuchinke, L., Urton, K., Hofmann, M. J., & Jacobs, A. M. (2009). The Berlin Affective Word List Reloaded (BAWL-R). *Behavior Research Methods, 41,* 534–538. doi:10.3758/BRM.41.2.534

Zhang, Q., Guo, C., Ding, J., & Wang, Z. (2006). Concreteness effects in the processing of Chinese words. *Brain and Language, 96,* 59–68. doi:10.1016/j.bandl.2005.04.004