Processing of literal and metaphorical meanings in polysemous verbs: An experiment and its methodological implications

Processing of literal and metaphorical meanings in polysemous verbs: An experiment and its methodological implications

Abstract

This paper discusses methodological issues related to the implementation of experimental studies that investigate the processing of literal and metaphorical expressions. The first and central aim of the paper is to discuss methodological questions arising from current experimental research on figurative language, such as use of heterogeneous stimuli and their suitability for investigating research questions related to the processing of figurative language. At the center of the methodological discussion is the contextual and cognitive dominance of literal and metaphorical meanings in polysemous words as an important variable in experimental research on figurative language. In this paper, we discuss its importance and propose a possible way of controlling for this variable. The second aim of this paper is to present experimental results that complement the methodological discussion and show that metaphor processing in polysemous words (i.e. very conventionalized figurative meanings) is not necessarily more difficult or special in comparison to processing of literal language if the stimuli are properly balanced. In this paper, we suggest a way to design stimuli that are balanced both in terms of traditional lexical measures as well as dominance of the literal and metaphorical meanings retrieved from corpus data.

Keywords: figurative language, literal language, language processing, polysemy, meaning dominance

1. Introduction

Conventional metaphorical expressions such as *open an account* or *support a friend* have a complex status in everyday speech. On the one hand, they are undoubtedly metaphorical[1] and more abstract than expressions in which the same verb is used literally, e.g. *open a door* or *support a roof*. On the other hand, they are so conventional that most of the time they are not considered metaphorical or special by native speakers (see Gibbs, 2017 for a recent review). Moreover, some verbs used in expressions of this kind are, in fact, more prominent in their metaphorical meaning, like *support*. Some studies (e.g. Gibbs, 1984; Glucksberg and Keysar, 1990; Giora, 1997; Glucksberg, 2001; Bowdle and Gentner, 2005) argue that when metaphors are as conventional as this, they are processed as a type of polysemous words. The meaning, then, is not processed by comparing two different domains, as in the more classic comparison view of metaphor comprehension (e.g. Lakoff and Johnson, 1980; Gentner, 1983; Gentner et al., 2001) but rather by semantic categorization or lexical disambiguation, much like ordinary polysemous words. Moreover, some theorists argue that metaphorical meaning is accessed later and that the literal meaning has primacy over the metaphorical meaning (Grice, 1975; Searle, 1979). Within the last 40 years, many studies have explored this issue. However, recently more sophisticated experimental methods have shown that more complex measures may also yield more complex results, and the debate on whether there is something special about processing of figurative meaning is therefore still very much alive.

[1] An anonymous reviewer noted that the metaphorical status of these examples is questionable since these can be seen as extensions of their core meaning. In line with the cognitive linguistics literature, we assume that these are, in fact, semantic extensions of the literal meaning based on metaphors (e.g., Sweetser, 1990).

Most of the studies mentioned above focus on metaphors expressed in their direct form (e.g. *Socrates is a midwife*), where both terms are explicitly expressed. In fact, metaphors are mostly expressed in indirect forms (Steen et al., 2010), where a polysemous term that can have both a literal and a metaphorical meaning is used metaphorically in a given context (e.g. *support a friend*). In direct metaphorical expressions, there is a cross-domain mapping which is usually expressed as some form of comparison (Steen et al., 2010: 774), while indirect forms of metaphor arise when a word is used in a specific context in such a way that it constructs a contrast between the contextual meaning of a lexical unit and its more basic meaning (Pragglejaz Group, 2007).[2] The basic meaning (usually more concrete) is therefore absent from the context, but it is observable in other contexts. For instance, when the verb *support* is used in the context of psychological help, its contextual meaning is substantially different from the basic meaning of supporting e.g. a roof with beams. Therefore, *support* is used metaphorically in that context.

This paper has two main objectives. The main objective is to discuss methodological questions that arise from current research on figurative language. The concern is specifically with studies that use a variety of stimuli construction methods which then result in heterogeneous stimuli that may not be suitable for investigating the research questions. We claim that for a theory to correctly predict differences between literal and figurative processing, the tools and stimuli used must be homogenous and controlled, especially regarding the contextual and cognitive dominance of different meanings in polysemous words, for instance, literal and metaphorical meanings. We discuss this issue and suggest a feasible method of constructing stimuli that are balanced both in terms of traditional lexical measures and in terms of dominance of literal and metaphorical meanings, retrieved from corpus data. By controlling for meaning dominance, we provide a proof of concept that this variable plays a critical role in the debate on the direct vs. indirect access to metaphor. The second objective is to present experimental results that complement the methodological discussion and show that metaphorical processing is not necessarily more effortful than literal processing in the case of highly conventionalized metaphorical meanings. We show the results of an experiment that uses such controlled stimuli to investigate processing of indirect metaphorical expressions in comparison to literal meanings of the same verbs.

2. Current state of research and open questions

2.1. The debate about figurative language processing

The status of literal vs. metaphorical meaning in language processing has been widely discussed in the literature with two opposing models. The ‘indirect access’ model is traditionally associated with Grice’s theory (1975) that suggests that we access the literal meaning first and then process the figurative meaning (see also Searle, 1979). Grice assumes that figurative language flouts the maxim of quality: the primary meaning of an expression is a literal one, and to assign a metaphorical meaning, the speaker first needs to recognize the primary meaning to assign a new contextually appropriate meaning – in this case a figurative one. Some of the early experimental

[2] The basic meaning is defined as more concrete, related to bodily action, more precise, historically older (Pragglejaz Group, 2007).

evidence supporting this theory comes from Clark and Lucy (1975). Their study supports the claim that literal language is easier to process than figurative language and therefore confirms the indirect access model. Janus and Beaver (1985) also found differences between literal and figurative processing but do not reject the hypothesis that the cognitive mechanisms behind the two types of processing are the same.

Shortly after the emergence of the indirect access theory, Harris (1976), Gildea and Glucksberg (1983), Gibbs (1984), and others set out the case against it. Their 'direct access' model suggests that figurative meaning and literal meaning are processed using the same cognitive mechanisms and that there is no primacy of the literal. In other words, metaphors are processed quickly and non-optionally (Gildea and Glucksberg, 1983). As Harris (1976: 314) points out, metaphor "is not a highly-specialized form of language that becomes comprehensible only after the use of inferential processes operating on some literal or more basic meaning". Other early psycholinguistic studies such as Ortony et al. (1978), Inhoff et al. (1984), and McElree and Nordlie (1999) also provided support for the direct access model by showing that when the context supports it, there is no difference between processing figurative and literal language. Eventually, there emerged "a consensus in the field that literal meaning does not have unconditional priority" (Glucksberg, 2003: 92).

Since the emergence of these theories, other theoretical accounts have developed to account for different aspects of figurative language comprehension, chiefly the Graded Salience Hypothesis (Giora, 1997) and the Career of Metaphor Theory (Bowdle and Gentner, 2005), but also constraint-based approaches and hybrid models. The former two theories will be discussed in more detail below in Section 2.2.1. Constraint-based approaches, such as the one by Katz and Ferretti (2001), assume that "comprehension involves utilizing all the sources of information that a person has at his or her command at any one instant" (Katz and Ferretti, 2001: 214), i.e. contextual constraints, literality, saliency, familiarity. Moreover, McRae and colleagues (1998) suggest that different types of information may be accessible during language comprehension and activated in parallel to explore different alternative possibilities. In the case of figurative language such as proverbs, this would mean exploring whether the context makes sense for a literal or figurative interpretation. This information is then taken as probabilistic evidence needed to resolve the ambiguity in interpretation. In the idiom literature, another prominent proposal comes from Titone and colleagues (Titone and Connine, 1999; Libben and Titone, 2008) who suggest that idiom comprehension highly relies on the time-dependent availability of different linguistic constraints. Among these more complex models, some hybrid models combine various theories, such as Coulson and Matlock's Space Structuring Model (Coulson and Matlock, 2001) which posits that metaphor is more than a set of mappings between two domains, and thus understanding a metaphor includes parallel accessing a variety of information from the "integration network" (Coulson and Matlock, 2001: 300). This means that metaphor comprehension comprises of integrating or *blending* different conceptual domains (this terminology led Fauconnier and Turner to develop Blending Theory later on, in 2002). The theory also assumes that there is no discrete metaphorical meaning, but similarly to other models, proposes that the same process of blending conceptual domains is included in processing of literal meaning as well, therefore suggesting that comprehension of figurative and non-figurative language is similar.

In addition to these, relatively recent EEG studies have cast some doubt on this picture, with EEG evidence showing that there is, in fact, a difference between processing literal and metaphorical meaning (Pynte et al., 1996; Lai et al., 2009; De

Grauwe et al., 2010; Bambini et al., 2016). Bambini et al. (2016) argue for an account that emphasizes the importance of context, claiming that the effects found in metaphorical sentences in their study are compatible with lexical access that is governed by contextual expectation and that some effects can be attributed to a pragmatic mechanism of figurative language interpretation. Lai et al. (2009) also found significant differences between processing metaphorical and literal meanings, and while metaphorical and literal meanings are rated similarly in an offline judgment task by participants, conventional metaphorical expressions still require more processing effort. Nevertheless, they claim that the requirement for this increased effort may have been caused by the need to select from multiple meanings, and in this case, the literal meaning was available first. In conclusion, the EEG evidence shows that there is a difference between processing literal and metaphorical meaning. However, they emphasize that these differences may be due to context or the nature of the stimuli so that these results cannot be assumed as a matter of course to support the indirect access model.

Other studies suggest different interpretations. Studies such as Bonnaud et al. (2002) and Weiland et al. (2014) interpret the results of their studies as supporting the primacy of the literal meaning. Bonnaud et al. (2002) tested semantic links in metaphorical and literal structures, finding differences in processing between the metaphorical and the literal meaning. They interpret their results as indicating that metaphorical expressions have a special status in the semantic memory and that they are processed in a manner that differs from the way literal expressions are processed. Similarly, Weiland et al. (2014) suggest that the results of experiments they conducted support the indirect access view since the processing times accord with effects that signal difficulties in processing metaphorical expressions. Another recent eye-tracking study suggests that metaphor processing might be difficult, at least in the case of *A is B* metaphors. Ashby et. al (2018) compared how participants read *A is B* metaphors and similes. Their results showed that metaphors are more difficult to process, which suggests that one primary interpretation of an expression is initially assumed, and that reanalysis is necessary. The authors dub this the 'metaphor effect': even though the only difference between metaphors and similes was one word (*is* or *like*) the participants had more difficulties with metaphors since their surface form (*A is B*) could also be the form of a literal statement, while a simile is already a literal statement. In other words, the comparison is explicit in similes, while in the case of *A is B* metaphors an initial, literal interpretation might be assumed since the comparison is not explicit.

To sum up, while the question of how metaphorical language is processed has been a central topic in theoretical and experimental research for over 40 years now, no clear consensus has so far emerged. In this paper, we suggest that this is partly motivated by methodological issues. A wide variety of stimuli and methods have been deployed to test this research question. Yet it seems that while the behavioral experimental methods provide more straightforward results, more complex methods that often involve measurement of brain activity are more sensitive to the variety of stimuli and methods. Consequently, many of these new studies have difficulty explaining the results, especially with respect to the theoretical models available. The aim of this paper is to address some of these challenges, mainly with respect to one variable, i.e. meaning dominance. In this paper, we focus on showing that meaning dominance is one of the most important measures to consider when constructing experimental stimuli, and suggest a possible way of controlling for this variable.

2.2. Methodological challenges

2.2.1. Type of stimuli

Studies dealing with the question of figurative processing use a variety of stimuli. Many studies use direct metaphors (*A is B* forms), such as Glucksberg et al. (1982), Gildea and Glucksberg (1983), Blasko and Connine (1993), Pynte et al. (1996), Bowdle and Gentner (1999), McElree and Nordlie (1999), Coulson and van Petten (2002), De Grauwe et al. (2010), Weiland et al. (2014), etc. Some other studies use sentential metaphorical expressions (described below and exemplified in Fig. 1), like Inhoff et al. (1984) and Janus and Beaver (1985). Meanwhile, Lai and colleagues (Lai et al., 2009, Lai and Curran, 2013) use conventional metaphorical expressions (see definition in Introduction), and Mashal et al. (2007), who look at novel metaphor comprehension, use word pairs. Interestingly, apart from exhibiting a great variety in the type of stimuli, figurative language studies also differ in the type of task they use, which is also relevant from a methodological perspective and might shed a different light on the interpretation of results (see Kalandadze et al. (2019) for an in-depth overview).

Even though these studies look at different figurative structures, they are mostly taken as evidence for the same effect, that is, even though they use different kinds of stimuli, they explore a similar question: what is the difference between processing of literal and figurative language? However, this does not take account of the fact that metaphor appears in language in different forms, and these are potentially processed differently. In this section, we discuss the various stimuli used and highlight the main differences. We also discuss the potential implications of the use of different stimuli on processing.

Author and year	Type of stimuli	Example used in the study
Ortony et al., 1978	metaphorical expressions idioms	*Regardless of the danger, the troops marched on.* *to let the cat out of the bag*
Glucksberg et al., 1982	*A is B* expressions	*Some jobs are jails.*
Gildea and Glucksberg, 1983	*A is B* expressions	*All criminals are germs.*
Inhoff et al., 1984	sentential metaphorical expression	*Regardless of the danger, the troops marched on.*
Janus and Beaver, 1985	sentential metaphorical expression	*[…] They had once been very happy, but after several years of marriage, they had become discontented […] The fabric had begun to fray.*
Blasko and Connine, 1993	*A is B* expressions	*The belief that hard work is a ladder is common to this generation.*
Pynte et al., 1996	*A is B* expressions	*Those fighters are lions.*
Bowdle and Gentner, 1999	*A is B* expressions, similes	*A ballerina is (like) a butterfly.*
McElree and Nordlie, 1999	*A is B* expressions	*Some hearts are stone.*
Bonnaud et al., 2002	metaphorical expressions	*will of iron*
Coulson and van Petten, 2002	*A is B* expressions	*The actor says interviews are always a headache.*
Mashal et al., 2007	word pairs	*bright student, pearl tears*
Lai et al., 2009	conventional metaphors	*every point in my argument was attacked*
De Grauwe et al., 2010	*A is B* expressions	*Unemployment is a plague.*
Lai and Curran, 2013	conventional metaphors	*Life can sometimes be bumpy.*

Weiland et al., 2014	*A is B* expressions	*Those lobbyists are hyenas.*
Bambini et al., 2016	non-lexicalized metaphorical expressions	*shark* (‘squalo in Italian)
Bambini et al., 2018	*A* of *B* expressions	*eyes of steel* (‘gli occhi d’acciaio’ in Italian)

Figure 1. Overview of stimuli used in previous experimental studies that investigate processing of metaphorical meaning

There is a considerable difference between an idiom, an *A is B* metaphor, a sentential metaphorical expression, and an indirect metaphorical expression. One of the features shared by these structures is their conventionality; however, they are inherently different in terms of their semantic and syntactic characteristics.

Many studies that focus on figurative language investigate idiom processing, even though not all idioms are necessarily metaphorical in the same sense as A is B metaphors or conventionalized indirect expressions. Idioms, e.g. *to let the cat out of the bag*, are usually semantically non-transparent, non-compositional, and fixed in their structure. Wray and Perkins (2000: 1) state that one of the characteristics of formulaic sequences (e.g. idioms) is that they are stored and retrieved as units, and are not susceptible to “generation or analysis by the language grammar”. In addition to this, it has repeatedly been shown that there is a processing advantage for idioms (Gibbs, 1980; Swinney and Cutler, 1979), which has often been attributed to their fixed formulaic nature.[3] In other words, idioms represent a special form of figurative language which differs significantly from other types such as *A is B* metaphors or indirect metaphorical expressions.

A is B metaphors, such as *fighters are lions*, are semantically transparent. Their form signals a mapping that tells the speaker that something is being described in terms of something else, i.e. it signals a similarity relationship or an analogy (Goatly, 2011).[4] Holme (2004) points out that *A is B* metaphorical expressions most clearly show the paradoxical nature of saying that something is what it is not, which is why the contrastive nature of figurative language is most easily identified in this type of metaphor. Furthermore, these constructions are syntactically special since they include a copula and are in that respect distinct from verbal figurative expressions such as *the criticism stung him* (Stowe and Palmer, 2018: 18). Even though this type of metaphor has been at the center of attention in cognitive linguistics ever since Lakoff and Johnson’s (1980) seminal work, and it has been claimed that copular constructions are the form of figurative language that is most easily identifiable and understood (Low, 2008; Cameron, 2003; Gentner and Toupin, 1986), research indicates that copular metaphorical expressions are not very common in everyday language. Cameron (1999) shows that, in a spoken corpus, verb metaphors are more frequent than noun metaphors, of which the most explicit form is the *A is B* construction. Cameron (2012: 344) points out that figurative speech mainly consists of vehicle terms

[3] There is recent research that shows that the idiom advantage might partly be caused by the nature of the stimulus and its formulaicity/convention (Kyriacou et al., 2018), that is, once predictability is balanced through context, the advantage disappears. This has also been shown in Carrol and Conklin (2019) for other types of formulaic phrases by statistically controlling for frequency and cloze predictability in the analysis.

[4] However, note that even though the surface form of *A is B* metaphors is transparent and signalling, similes are in fact literally true, while metaphors are not (Ashby et al., 2018).

that are in some way embedded in the "flow of topic talk", rather than copular constructions. In other words, even though copular metaphors have long dominated research on figurative language, research suggests that they are not the most prototypical type of figurative language. This is also supported by extensive corpus analyses showing that such copular metaphors (or direct metaphors) are significantly less frequent than indirect metaphors across various genres and registers (Steen et al., 2010). Holme (2004: 59) identifies three major difficulties with treating *A is B* metaphor as the central type of metaphorical language. First, they are rare; second, they may be the cause of erroneous conclusions concerning the nature of metaphorical processes since they are atypical; and finally, many languages do not have copular structures, so the conclusions based on these constructions may be limited to specific languages rather than being generally or universally applicable, compounding the potential for erroneous conclusions concerning the nature of metaphorical processes.

Sentential metaphorical expressions are metaphorical structures in which the entire sentence carries metaphorical meaning, rather than just certain parts of it:

A: What kind of mood did you find the boss in?

B: The lion roared. (Gibbs, 1994: 213)

These are compositional and non-formulaic. Gibbs (1994) points out that they are understood like comparison statements (*A is B* metaphors), and the irrelevance of their literal meaning to the context of the utterance signifies to the listener that the sentence needs to be interpreted figuratively. The relationship between the sentence and its context enables the conversational participants to recognize the sentence as metaphorical even though there are no explicit markers.

Indirect metaphorical expressions such as those used in Lai et al. (2009) and Lai and Curran (2013)[5] form most of our everyday uses of metaphorical language. These are expressions that are often characterized as polysemous words, with one or more literal and one or more metaphorical meanings. For instance, when words like *attack* or *defend* are used to talk about debates, they are used in a metaphorical sense on the basis of the conceptual metaphor ARGUMENT is WAR (Lakoff and Johnson, 1980). Britton (1978: 2) notes that 93 of 100 most frequently used English words (different word classes) are polysemous. Different meanings of a word often differ with respect to their degree of concreteness, and an abstract meaning of a word may often be considered to be a metaphorical meaning created by extension of a literal meaning. However, as Gibbs (1994) indicates, it cannot be assumed that a more concrete or more literal meaning is necessarily 'primary' in semantic or etymological terms.

Goatly (2011: 34–35) characterizes metaphorical meanings of such polysemous words as 'inactive' metaphors, which he contrasts with 'active' and 'dead' metaphors. According to Goatly, they are in the middle of the conventionality "cline", with active metaphors at one end and dead metaphors on the other. Dead metaphors, such as those referring to a student as a pupil, are considered homonyms by Goatly and exhibit no metaphorical connections; for inactive metaphors, by contrast, the metaphorical connections can be evoked and he considers such words to be polysemous. In such words, the metaphorical meaning is in some way motivated by or connected to the literal meaning. Active metaphors such as *his tractor of blood stopped thumping* (Goatly, 2011: 34-35) meaning *his heart stopped* exhibit no lexical relationship with a literal counterpart and are strongly dependent on the context.

[5] Note that Lai calls them 'conventional metaphors', while we here use the term indirect metaphorical expressions in line with work by Steen et al. (2010).

In terms of processing, direct access models (Gibbs, 1984; Glucksberg and Keysar, 1990; Giora, 1997; Glucksberg, 2001; Bowdle and Gentner, 2005) assume that indirect metaphorical expressions are processed with ease and effortlessly, without comparison of the two domains.

To summarize, even though all of these different expressions convey figurative meaning, research has shown that they differ in fundamental ways: they differ in their level of formulaicity, in their type of structure, and the transparency of their figurativeness. In some studies, such differences have been predicted to cause corresponding processing differences. This topic is at the center of the next section.

2.2.2. Effect of type of stimulus on cognitive processing

In the previous section, we discussed the different types of stimuli used in studies on figurative language processing. Factors such as the structure of the expression, as well as their figurative transparency, can influence how a figurative expression is processed and perceived. In this section, we discuss theoretical accounts that discuss how differences in the type of metaphorical expression can influence processing. In addition to this, one of the central issues discussed in this section is the dominance of literal and figurative meaning and the effect that such dominance can have on processing. Meaning dominance refers to the feature of ambiguous items in which one of the meanings of a lexical item or a phrase is more frequently used than the others (Simpson, 1981; Foraker and Murphy, 2012), and therefore, is assumed more easily accessible in language processing.[6] While the literature discusses both dominance and salience of meaning, we will be using the term *dominance* when referring to the prevalence of one of the possible meanings that occur within a single lexical item. A polysemous lexical item can carry many different meanings of which some can be metaphorical or literal, and some can be more dominant in use than others. We assume that salience is a term that is related but not equivalent to dominance. Salience of a certain meaning is determined by the frequency and prototypicality in the context. To illustrate, while the literal meaning of a certain word can be its most dominant meaning overall, it does not have to be the most salient one in a context that does not support it.

Issues concerning the types of metaphorical expressions we encounter in everyday communication are central to the Career of Metaphor Theory (Bowdle and Gentner, 2005). Bowdle and Gentner propose that when an *A is B* metaphor first appears (e.g. *an obsession is a tumor*), this establishes the connection/similarity between the two concepts. In the course of being used more frequently, the relation between the topic and vehicle term becomes more salient. With use, this relation is extended to other topics in different contexts (e.g. *a doubt is a tumor, a grudge is a tumor*). At this point, the term with its new meaning becomes conventional and polysemous. Bowdle and Gentner (2005: 198–199) also propose that differences in conventionality can be expected to generate differences in processing: the more

[6] Meaning dominance, as defined here, is different from metaphor dominance, which is typically defined as the frequency of the use of polysemous word's metaphorical meaning, divided by the frequency of its literal meaning (e.g., Dulcinati et al., 2014). Metaphoricity is a specific type of polysemy, which characterizes lexical entries that have a figurative, typically more abstract, meaning, and a literal, typically more concrete, meaning (e.g., see Reijnierse et al., 2019 for a recent discussion about polysemy and metaphoricity). The dominance of the metaphorical meaning of a polysemous word is measured as its occurrence in corpora, relative to the occurrence of the literal meaning of the same word form.

conventional metaphorical expressions become, the more their meaning is likely to be interpreted by categorization rather than comparison. Comprehension by categorization works indirectly: instead of directly comparing the two metaphorical terms in order to understand the mappings on which the metaphor is constructed, the categorization view for conventional metaphors suggests that the topic of the metaphor is treated as a member of a superordinate category, which is exemplified by the vehicle of the metaphor. The vehicle, therefore, provides direct access to a superordinate category that can be used to characterize the topic, whereas the topic constrains the range of dimensions by which it could be characterized. For example, in comprehending the metaphor *Socrates is a midwife*, the topic *Socrates* is associated with the ad hoc category *someone who helps and supports other people in difficult situations that involve generating new entities* to which the vehicle *midwife* typically belongs. In evoking the ad hoc category, the term *Socrates* facilitates the attribution of features related to generating ideas and developing intuitions, while blocking out irrelevant features such as those related to the physical process of delivering a baby (Bowdle, 1998; Bowdle and Gentner, 1995; Gentner and Bowdle, 2001; Gentner and Wolff, 1997).

We focus our research on only one particular type of conventional metaphorical expression: those expressed indirectly, which are not signaled explicitly and are the most frequent type of metaphor found in natural language. We are interested in seeing whether there are any differences in processing literal meanings in comparison to metaphorical meanings of polysemous words. The novelty of our approach lies in the polysemous items we chose as our stimuli, by carefully balancing them for the dominance of their meaning and other psycholinguistic factors. As with any polysemous word, it is predictable that a more frequent or more dominant meaning will be accessed faster than a less frequent or less dominant one. This factor is frequently not fully taken into account in previous studies since they do not test the corpus frequency of metaphorical vs. literal meanings of a single polysemous term. Goatly (2011: 107) points out that some words are more frequent in their metaphorical meaning, such as the noun *crane*, for which only 5 % uses in CoBuild corpus are literal uses. Similarly, consider the verb *support.* Oxford dictionary (OED 2018) lists its literal meaning as the primary meaning: *bear all or part of the weight of; hold up,* and its metaphorical meaning as secondary: *give assistance to*. However, corpus evidence from the English Web 2015 corpus which was accessed through SketchEngine (Kilgarriff et al., 2014) shows that *support* primarily appears in its metaphorical meaning, with the most frequent objects of *support* appearing in combination with the metaphorical meaning, e.g. *to support + development, project, effort, work, family, program, research,* etc. Conversely, if we consider the word *digest*, the more salient meaning is the literal one, according to SketchEngine: one can *digest + lactose, cellulose, sludge, plasmid, carbohydrate, food*. Finally, there are verbs like *build*, which are dominant in both of their meanings, so one can build, in order of frequency in SketchEngine, *relationship, capacity, house, community, environment, network, bridge, etc*.

The question of salience of meaning is treated as a core issue in Giora's work (1997). Giora (1997) proposes the Gradient Salience Hypothesis which posits that the difference between metaphorical and literal meaning is not a question of the primacy of one over the other in any given context, but rather a question of salience, that is, the relevant divide is not the figurative/non-figurative divide but rather salient/non-salient. Salience of meaning depends on its conventionality, familiarity, frequency, and

its status in a certain context. Regarding language processing, the salient meaning is always accessed first, regardless of whether it is metaphorical or literal. For instance, the conventional literal meaning can be processed slowly and with difficulty if it is not salient in a certain context. This means that if a lexical item is ambiguous between two meanings, the one that is more prototypical or more frequently used in that context will be the more salient one, and therefore easier to process. Giora emphasizes that the salient meaning is always accessed, and has "unconditional priority" (Giora, 1997: 186) over less salient novel interpretations. In the case of non-salient meanings, the processing predictions resemble the indirect access model: salient meaning is accessed first, then rejected, and reinterpreted.

The issue of differing degrees of dominance of different meanings of polysemous words has also been previously discussed in literature dealing with polysemy in general. Much of the literature on polysemous words emphasizes that meaning dominance affects how ambiguous meaning is processed (Simpson, 1981; Tabossi et al., 1987; Wiliams, 1992; Foraker and Murphy, 2012; Milburn, 2018; Milburn and Warren, 2019). For instance, Simpson (1981) shows that ambiguous words are positioned in semantic memory with respect to relative frequency. This means that when the context is neutral, or in the absence of context, the dominance will determine which of the meanings is accessed. Duffy et al. (1988) also demonstrate that a meaning which is more frequently used is more easily and quickly activated than a less frequent one. More recent studies such as Foraker and Murphy (2012) offer similar results. They show that when there is a neutral context, the dominant meaning is easier to access than the non-dominant one. In the case of idioms, Milburn and Warren (2019) investigated whether meaning relatedness and dominance affect idiom processing. The results show that idioms with more related figurative and literal meanings are processed faster, like ambiguous words, and that meaning relatedness and meaning dominance interact in the same way they do in polysemous ambiguous words. Some of these studies (Duffy et al., 1988; Foraker and Murphy, 2012; Milburn and Warren, 2019) rely on participants' judgments to determine meaning dominance of phrases and lexical items, while some rely on researchers' judgments during stimuli construction (Simpson, 1981; Tabossi et al., 1987)

Meaning dominance has tended to be neglected in figurative language studies, though some studies highlight its significance. For instance, Ortony et al. (1978) and studies such as Inhoff et al. (1984) and Janus and Beaver (1985), which used the stimuli of Ortony et al., control for how congruent the context is with respect to literal and metaphorical meaning, but they do not control for dominance of either meaning. Bonnaud et al. (2002) constructed a list of stimuli based on semantic links but do not specify either the strength of those links or whether the strength is controlled across the conditions. By contrast, a number of studies highlight the role of meaning dominance in figurative expressions. Lai et al. (2009) and Lai and Curran (2012) use stimuli that have been checked for their 'metaphorical sense frequency' using a crowd-sourcing method (Bethard et al., 2009). The model developed by Bethard et al. (2009) predicts how often a certain word is used metaphorically. However, it does not go far enough, as is evident from an example such as *attack a point* vs. *attack a soldier*. While Bethard et al.'s method measures how frequently *attack* appears in its metaphorical sense in a given phrase, it fails to capture whether the frequency of *attack a point* and *attack a soldier* is the same, which might influence the results. Bethard et al. also identify that their model does not take into account the probability of a word appearing in its metaphorical meaning in a certain context, and identify this as a future goal for their work.

To summarize, when it comes to conventionalized metaphorical expressions, their conventionality plays an important role in their processing and is tightly connected to both dominance and salience of meaning. In this paper, we focus on meaning dominance, arguing that the salience of a certain meaning of a polysemous word is determined by frequency and prototypicality in context. Even though research dealing with both figurative language as well as with polysemy has shown that meaning dominance influences how we access meaning in words that might have more than one meaning, the issue of meaning dominance has not been given much attention in research that looked into conventional metaphorical expressions.

2.2.3. Experimental evidence: The current study

The first and central aim of this paper is to highlight and address key methodological challenges that arise when investigating the processing of literal and metaphorical meanings of polysemous words. The second aim is to propose a possible experimental methodology that addresses the problem of literal or metaphorical dominance in polysemous words. The experiment supports the methodological claims addressed in this paper.

In line with the findings of the empirical literature reviewed in Sections 2.2.1 and 2.2.2., in this paper, we argue firstly, that any claims concerning the processing of metaphorical meanings need to be situated within a certain level of conventionality of metaphors, and secondly, that it is difficult if not impossible to make convincing hypotheses about metaphorical processing if dominance of literal and metaphorical meaning is not well-balanced.

We also contribute to the theoretical debate on metaphor processing with original experimental evidence that meets the methodological requirements for a coherent and well-controlled analysis in which our variable (i.e. metaphoricity) is isolated. We describe the process that was adopted for designing stimuli and present experimental data based on the created stimuli. The study is intended to test whether indirect metaphorical expressions are processed with ease, in a similar way as literal meanings. The hypothesis is that metaphorical meanings are processed with the same ease as literal meanings in highly conventionalized figurative expressions when the different meanings' dominance is balanced across conditions.

3. The experiment

This experiment investigates whether the meaning of indirect metaphorical expressions is as easily accessed as the literal meaning. The experiment was a cross-modal semantic priming study combined with a lexical decision task in which reaction times of the answers were recorded.

3.1. Stimuli

3.1.1. General notes

The experiment that we report here is a cross-modal semantic priming task combined with a lexical decision task (described in detail in Section 3.3. Procedure). This type of task exploits a well-established psycholinguistic effect called (semantic) priming: making a lexical decision on whether a sequence of letters (i.e. a target) is a genuine word or not is faster when the sequence of letters is preceded by a semantically related

word (i.e. a prime) than when the sequence of letters is unrelated to the prime. The cross-modality is due to the fact that the prime is presented as an audio prompt through the participant's headset, while the target is written on the screen.

The critical experimental stimuli consisted of 48 transitive verbs which were polysemous and could convey literal or metaphorical meaning. The verbs served as targets and 144 nouns (across three conditions: literal, metaphorical, unrelated) which served as primes. The primes were divided across three conditions: first, nouns that prime the literal meaning of the verb, second, nouns that prime the metaphorical meaning of the verb, and third, nouns that are semantically unrelated to the target verb. Altogether, the stimuli included 48 critical target words (verbs), 16 fillers, and 64 pseudowords. The pseudowords were constructed with the help of The ARC Non-word Database (Rastle et al., 2002). Some pseudowords were items from the database, while some were made up by the authors who adopted the principles used in the ARC Non-word database. Pseudowords included items such as *scolve*, *gleach*, *thrutt*, *splear*, *thwool*, *droothe*, *deliebe*, *glief*, etc.

3.1.2. The triplets

3.1.2.1. Lexical measures

The stimuli consisted of triplets which included a transitive verb that can have both a literal and a figurative meaning, such as *digest*, e.g. one can 'digest food' (literal) but one can also 'digest ideas' (metaphorical). The two meanings of *digest* point to the underlying conceptual metaphor IDEAS are FOOD. For examples of stimuli, see Figure 2. The relationship between the target and the prime was verb – object. For instance, *digest milk* or *digest news* both form a V + O phrase.

Target	Literal prime	Metaphorical prime	Unrelated prime
spend	fortune	holiday	radio
invest	cash	effort	garden
digest	milk	news	seat
devour	pizza	novel	bishop
climb	fence	rank	storm
touch	ground	heart	camp
cover	floor	story	church
release	prisoner	survey	cottage
expose	skin	truth	winter

Figure 2. Examples of stimuli

The triplets were chosen based on several criteria and were carefully balanced for several different measures, i.e. word length, single word frequency, collocation frequency and collocation association score. To control for word frequency and length we conducted single-factor ANOVAs to test whether there was any difference between literal, metaphorical and unrelated primes. The values for frequency and phoneme/character length (depending on whether the stimulus is presented auditory or visually) were extracted from N-Watch (Davis, 2005). There were no significant differences between CELEX frequencies ($F(2,141) = 0.73$, $p = .485$), Kučera & Francis frequencies ($F(2,141) = 1.97$, $p = .142$) and length ($F(2,141) = 2.11$, $p = .125$) of primes across the conditions (see Fig. 3 for more information). The target words were all verbs with a mean length of 5.16 (max = 9, min = 3, SD = 1.31).

To control for dominance of meaning, we collected additional measures on the frequency of the verb-object collocation and their relative association scores. These

were extracted from SketchEngine (Kilgarriff et al., 2014), using the Word Sketch function and the related scores.[7] A t-test was conducted to ensure there was no significant difference between collocation frequency (t = -0.58, p = .565) and association score (t = 1.46, p = .15) in the literal and the metaphorical condition.

	CELEX freq		KF freq		Length		Collocation freq		Association score	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
lit	60.21	55.91	59.10	54.3	4	1.03	8437.79	11316.33	6.72	1.18
met	74.86	68.75	86.18	77.27	4.31	1.05	9208.04	12404.05	6.55	1.16
unr	64.62	57.79	66.81	72.64	3.9	1.01	-		-	

Figure 3. Frequency in CELEX and Kučera & Francis frequency corpora, length measures of primes, and collocation frequency and association scores for literal, metaphorical and unrelated condition

Association scores and collocation frequency were the primary criteria in stimuli construction. This means that the most prototypical objects of the ambiguous verbs were not always chosen but rather the ones that match the scores of their metaphorical or literal counterpart. For instance, in the case of *digest*, we selected *digest milk* (Coll. Freq. = 658, Assoc. Score = 6.29), and matched it with *digest news* (Coll. Freq. = 645, Assoc. Score = 4.44), instead of the more prototypical or expected *food* because it was impossible to find a metaphorical pair that is as frequent as *digest food.* These types of issues are especially prominent in verbs where one of the meanings is highly dominant, such as the literal meaning of *digest* or metaphorical meaning of *support*.

It is important to note that it is not assumed that the frequency of collocation or the association score are actual measures of dominance or salience of meaning for the lexical item. We aim to show that they are, nevertheless, useful since they allow balancing one lexical item that might otherwise be more dominant in either the literal or metaphorical sense. Since it is difficult to create a list of stimuli where all polysemous words are not more dominant in either of the meanings, and it is questionable that such lexical items exist, we aimed to create a list of stimuli where at least the frequency of corpus appearance of the two meanings is comparable. Furthermore, we acknowledge that additional lexical measures/criteria could also be included and might be relevant for experimental results; however, introducing these might make the stimuli construction overly complicated to the point where control of the basic lexical measures could not be preserved. In this paper, we focus on meaning dominance as an important semantic feature of polysemous words that has frequently been neglected in figurative language literature even though it is one of the central semantic features of any linguistic expression with more than one possible meaning.

Another possible issue connected to the structure of stimuli is the effect to which the experiment is sensitive. Since the stimuli consist of verb + object combinations, it might be problematic to claim that the experiment measures meaning activation, rather than collocation activation. We claim that by balancing out the collocation frequency and association scores, this issue is resolved as well, at least in the case of literal vs. metaphorical meaning. Even though this experiment might, in fact, be sensitive to

[7] The association score "indicates the amount of statistical association between two words" (Rychlý, 2008: 6). It is calculated using logDice: a statistical measure for identifying collocation candidates which is based on the frequency of words and the bigram, and is not affected by the size of the corpus (Kilgarriff et al., 2014).

collocational effects, they should be minimal, since the frequency and association scores are balanced across conditions. However, the meaning effect should still occur, since, to perform a lexical decision task, the participant needs to access the right meaning of the polysemous verb, which will be either the literal or the metaphorical meaning, depending on what was activated by the prime.

3.1.2.2. Norming study

Once the triplet lists had been constructed, we normed them with native English speakers using Qualtrics. We recruited 30 native English speakers who saw a list of 96 phrases provided in a randomized order. They were asked to give a judgment about the perceived naturalness of the phrase, on a Likert scale. We used the following instructions: *For each phrase mark how natural it sounds from 1 to 7: 1 being the least natural, 7 being the most natural.* This was carried out due to the fact that the populations of participants in psycholinguistic experiments generally consist mainly of university undergraduates, while the texts from which the corpus draws its data may not be compatible with the type of discourse that undergraduates would be familiar with. For instance, they may not be familiar with some traditional expressions/collocations that have decreased in frequency in spoken discourse over time.

The norming study also partially targets the issue of speaker-specific dominance. Since not all meanings and all pairs are equally dominant for all speakers, by collecting not only corpus measures but also norming values, we aimed to construct a set of stimuli that is representative and possibly reduces the speaker variation.

Groups	Mean	P(T<=t) two-tail	t Stat
Lit vs. met	lit = 6.09 met = 5.92	p = .087	t = 1.75
Lit vs. unr	lit = 6.09 unr = 2.08	p < .001 ***	t = 38.83
Met vs. unr	met = 5.92 unr = 2.08	p < .001 ***	t= 35.42

Figure 4. Results of the norming study

As Fig. 4 shows, a two-tail paired t-test demonstrated that the difference in norming values between the literal (M = 6.09) and metaphorical (M = 5.92) phrases was not significant (t =1.75, p = .087). The difference between the literal and unrelated (M = 2.08) phrases was significant (t = 38.83, p < .001), as was the difference between metaphorical and unrelated (t = 35.42, p < .001) phrases. The general rule was for the metaphorical and literal phrases to have mean norming values above 5 while unrelated phrases had mean norming values below 3. A few phrases diverged from this rule but owing to the borderline values and other factors that influenced the choice of the stimuli, they were nevertheless included in the final stimuli. In the literal condition the phrases were *digest milk* (M = 4.97) and *scrap a ship* (M = 4.53), while in the metaphorical condition the phrases were *spend a holiday* (M = 4.8), *devour a novel* (M = 4.6), *weave a plot* (M = 4.83) and *hold a show* (M = 4.2). In the unrelated condition these were *bring an empire* (M = 4.3), *carry a devil* (M = 3.13), *open desire* (M = 3.63) and *possess sand* (M = 3.23). In relation to the methodological aims of our study set out in the Introduction, it should be mentioned that the norming study was conducted twice. We obtained some significant differences between the literal and the

metaphorical collocations after the first round of norming. Therefore, we revised our stimuli and replaced the outliers with new words that were fully balanced with respect to the above-mentioned lexical measures, and ran the norming study a second time.[8]

3.1.3. Audio stimuli

The audio primes were recorded in Audacity (44.1 kHz) by a male native English speaker in his thirties with previous experience in recording stimuli in an experimental setting. The stimuli were recorded in a sound-proof room using a professional quality microphone (Rode NT-USB). The words were cut manually by the experimenter and the volume was equalized using PRAAT (Boersma & Weenink, 2018) and Audacity (Audacity Team, 2010).

3.2. Participants

We recruited 48 native speakers of British English, aged between 18 – 50, with normal vision or corrected-to-normal vision, and no known language/neurological/hearing disorder. The participants were mainly Oxford students/residents. Their participation was rewarded by £5 or course credit.

3.3. Procedure

The experiment was conducted in the Language and Brain Lab at the University of Oxford using experimental software by Reetz and Kleinmann (2003). The auditory primes were played on individual headphones (Sennheiser PX200-II), and the visual targets were shown on a 16-inch computer screen. Participants made their lexical decisions using a custom-made two-button control box.

The experiment lasted about 6 minutes. Latin Square design was used to create three lists of 48 targets so each participant only saw each target in each condition once. The participants were given the following instructions: *As soon as the word appears on the screen as fast as you can, press YES if this is a word in English, and NO if this is not a word in English.* The participants first heard a beep, 300 ms before each prime. The visual target, which was either a verb or a pseudoword, followed straight after the prime and was displayed for 300 ms. The participants had 1500 ms to press YES/NO. They were able to press YES/NO as soon as the word appeared on the screen, but had to do so before the next beep was heard. For instance, in the literal condition, the participant would hear *cash*, and see *INVEST*; in the metaphorical condition, they would hear *effort* and see *INVEST*; and in the unrelated condition, participants would hear *garden* and see *INVEST* (Fig. 6). The participants had to decide if the written form on the screen (e.g. *INVEST* (word) or *STREAR* (pseudoword)) was an existing English word and they were given instructions before the experiment.

Before the experimental trials, participants did a practice test which consisted of 2 blocks of 12 trials. The experimental trials consisted of 8 blocks of 16 trials. In between blocks, the participants heard three beeps before the next block started.

[8] Note that the results reported in Fig 2, Fig 3a and 3b are based on the revised version of our stimuli, which are fully endorsed by the results of the second round of norming.

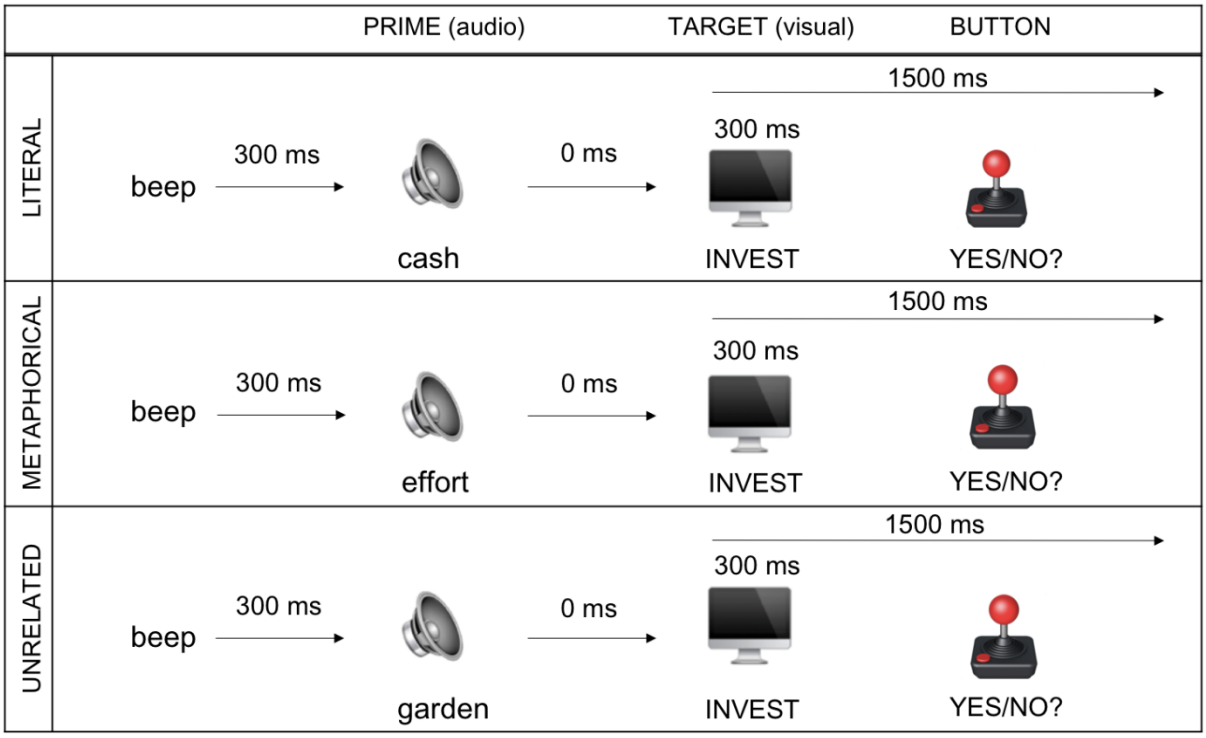

Figure 6. Experimental design

3.4. Analysis and Results

The data were analyzed using a linear mixed-effects model with the lme4 (Bates et al., 2014) R package (R Development Core Team, 2011). The fixed effect was condition (literal, metaphorical or unrelated) and random effects were participant and item. Both random slopes and intercepts were used for the random effect of participant, and only random intercepts were used for the random effect of item (Winter 2019). Comparisons between conditions were run using the lsmeans R package. No participants were excluded on the grounds of having made a high percentage of errors; however, two targets (*bury* and *seize*) were excluded due to a high percentage of errors (>20%). Reaction times outside ±2 standard deviations from the mean were excluded as outliers. As a result, 11.24 % of the data was excluded from the analysis.

There was a significant main effect of condition, $F(2, 91.59) = 5.36$, $p = .006$ (R^2: .458). A comparison test showed a significant priming effect for the literal condition ($t = -3.12$, $p = .002$) and for the metaphorical condition ($t = -1.85$, $p = .015$) compared to the unrelated condition (see Fig 7b. and 8a). In other words, both literal and metaphorical primes led to significant facilitation. However, not surprisingly, when compared directly, there was no significant difference between the literal and metaphorical condition ($t = -0.36$, $p = .719$) (see Fig 7a. and 8). Note that the direct comparison between literal and metaphorical condition is not a viable measure because many factors differ between the conditions. Our results show precisely this difference: a direct comparison leads to a null result, while when you compare with unrelated stimuli, we get a priming effect. These results demonstrate that both the literal and the metaphorical meanings of the polysemous words were activated and that they were activated with the same ease.[9] Error rates were analyzed using a

[9] The raw results file is available at this link:
https://osf.io/3p6a4/?view_only=69e0be53f07c4b7abb2dc63814b5cc3b

generalized linear model with a binomial distribution. The analysis (Table 7b) showed no effect of Condition (p = .287).

Main effect: Condition	**F (2, 91.59)**	**= 5.36**	***p* =.006***
Comparison: literal vs. unrelated	t (108.28)	= -3.12	p = .002*
Comparison: metaphorical vs. unrelated	t (45.89)	= -1.85	p = .015*
Comparison: literal vs. metaphorical	t (217.23)	= -0.36	p = .719

Figure 7a. Results of the mixed-effects model analysis: main effect and comparisons

	Estimate	Std. Error	z value
(Intercept)	3.16	0.18	17.28
conditionmet	-0.29	0.24	-1.20
conditionunr	0.07	0.26	0.79

Figure 7b. Generalised linear model of error data

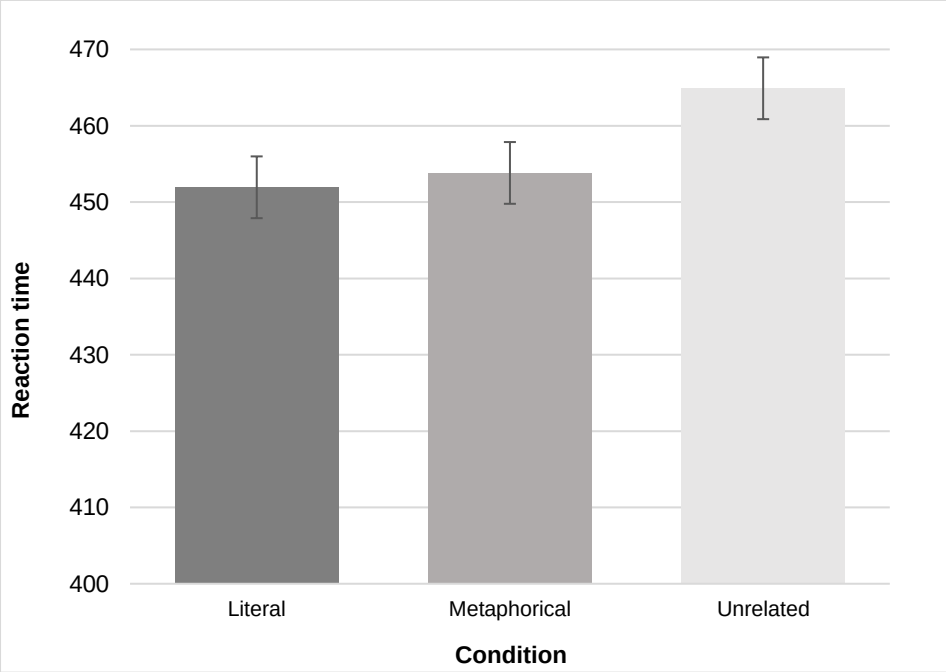

Figure 8. Reaction times by condition

4. Discussion

4.1. Methodological implications of the experimental results

The central experimental hypothesis that we entertained was that metaphorical meanings of polysemous words have both a literal and a metaphorical meaning which are both activated with the same ease. Our results confirm the hypothesis. We find no significant difference in reaction times when directly comparing literal and metaphorical conditions, but when compared to the unrelated condition, robust and significant priming effects are observed in both conditions.

The results could be interpreted as speaking in favor of the Graded Salience Model (Giora, 1997). Since the stimuli's metaphorical and literal dominance has been balanced across the conditions, the Graded Salience hypothesis predicts that there should be no differences in lexical access of the two different meanings, which is what our results show. The results are also in line with the Career of Metaphor Theory (Bowdle and Gentner, 2005), showing that very conventionalized figurative meanings, such as the ones used here, indeed do not require any extra processing effort, and therefore suggesting that in the case of conventionalized metaphorical expressions processing happens by categorization rather than by comparison. In brief, the results are in line with models that assume that conventionality and dominance of meaning influence the processing of figurative expressions. However, there is little support for models such as Coulson and Matlock's hybrid model, since it assumes that processing metaphorical expressions requires a variety of information and that there is no discrete metaphorical meaning. The fast pace of our study and no presence of a context might suggest that the activation of figurative meaning is simple and straightforward, with no need for blending the conceptual domains in the case of conventionalized expressions.

In light of these results, we will now focus on two methodological questions. First, we will show that there is substantial methodological variability across studies that consider how figurative language is processed. Second, we will argue that dominance of different kinds of meanings needs to be one of the central lexical measures in stimuli construction in studies on figurative language.

The question of methodological variability is concerned with a wide variety of metaphorical expressions found in the type of metaphorical expressions that are being tested. We have argued that certain types of figurative expressions, for instance, *A is B* metaphorical expressions, are inherently different from indirect metaphorical expressions. This poses a problem for the interpretation of experimental evidence reported in the literature. A key concern is that while most studies use *A is B* metaphorical expressions as stimuli, *A is B* expressions are not, in fact, the most frequently used type of metaphorical expression (Holme, 2004; Steen et al., 2010; Cameron, 2012). It is therefore difficult to claim that the processing of *A is B* expressions is informative about metaphor processing in general. While *A is B* expressions can be as conventional as indirect metaphoric expressions, their form is much more transparent because both terms are explicitly expressed. In other words, while all these studies are investigating metaphorical meaning, they are looking at different types of it. This needs to be acknowledged more fully and explicitly when interpreting the results. In the current study, we have used a certain type of indirect metaphorical expression, i.e. transitive verbs, and primed them with nouns. The role of the noun is the activation of a certain meaning in the verb, i.e. metaphorical, literal or unrelated. Using this type of uniform stimulus has resulted in clear and interpretable results showing for indirect metaphorical expressions that the ease of access of the metaphorical meaning does not differ from that of the literal meaning. In addition, there was a clear priming effect in both conditions in comparison to the unrelated condition. To sum up, our results can be interpreted as showing that when it comes to the most frequent type of metaphorical language, i.e. indirect metaphorical expressions, figurative language does not seem to be more difficult to process. However, it is important to note that these results do not imply that the same holds for other types of metaphorical language, such as *A is B* metaphors or sentential metaphors.

The other methodological question concerns the psycholinguistic measures taken into account in the empirical literature for designing balanced stimuli. While most

studies control for measures related to individual words, which include familiarity, frequency, length, or even comprehensibility of the expression, we contend that *meaning dominance* should be a key measure, i.e. it needs to be established whether the literal and metaphorical meanings are comparable in terms of dominance. While previous psycholinguistic literature has shown that this is a crucial measure that influences processing of polysemous words, this measure is rarely taken into account in studies investigating the processing of metaphor.

It is useful in this context to consider what role the measure of meaning dominance plays in those studies that do take account of it. Lai et al. (2009) and Lai and Curran (2013) consider how dominant the metaphorical meaning is for certain items using a model developed by Bethard et al. (2009), which predicts how often a certain word is used metaphorically. As mentioned earlier, this model does not consider the probability of a word appearing in its metaphorical meaning in a certain context. We claim that it is important to ensure that the dominance of different meanings is balanced across conditions with respect to the phrases used in the actual experiment, rather than, for example, just the metaphorical word. We suggest that existing corpus-based tools such as SketchEngine can be suitable for helping to achieve this goal.

The results contribute to the discussion on processing of literal vs. metaphorical meanings. The results show that when metaphorical meaning is conventional and inactive, both literal and metaphorical meanings are primed in the same way, with the same ease. However, we contend that such a carefully balanced set of stimuli does not in itself allow us to generalize about metaphor processing. Our results only tell us one part of the story. They help to predict how indirect metaphorical expressions are processed. Our results do not tell us anything about how *A is B* direct metaphors are processed. We suspect that this latter type of metaphor is more likely to be processed by comparison because the copula signals a comparison. On the one hand, the results of a controlled study such as the experiment reported here do not permit generalizations concerning the nature of metaphor processing as a general phenomenon. On the other hand, it is difficult to interpret empirical results appropriately and identify the source of the identified effects when the stimuli involve a range of variables. This permits the following conclusion: while limiting a study to a certain type of stimulus will correspondingly limit the possible theoretical predictions, such limitation coupled with full and explicit presentation of data, stimuli, and procedures is likely to make the study more informative and useful for subsequent research, and ultimately more productive for an understanding of the complex phenomenon of figurative language.

The current study suggests that the creation of stimuli for such studies should take several different factors into account. Apart from controlling for traditional lexical measures such as length and frequency, frequency and association scores need to be balanced also for the phrase, not only for the individual words. Dominance of meaning in context plays a crucial role in studies that investigate metaphorical processing (Giora, 1997). We suggest using the corpus frequency of phrases since their association scores can be a very efficient way of controlling for meaning dominance. In this paper, we focus on the measure of dominance to show why it is important and to introduce one of the possible ways of controlling for it. However, it must be emphasized that other lexical measures that we have not controlled for here can also be important. For instance, in addition to traditional lexical measures (length, frequency), figurative language studies also focused on comprehensibility, cloze probability, aptness, imagibility, and familiarity of expression. In this study, we limited

our focus on dominance to show how careful consideration of one variable can play an important role in the results. Introducing this many different variables could make the stimuli construction extremely complex. A possible solution might be controlling for traditional lexical measures through stimuli construction, and introducing other measures as variables in the statistical analysis.

In addition to the method proposed in this paper, there are other ways of controlling for dominance. For instance, Milburn and Warren (2019) collected dominance values from participants. This type of approach relies on participants' subjective judgments on how dominant one meaning is over another. On the one hand, this approach is sensitive to possible individual differences within the participant pool, while on the other hand, the judgments are completely subjective, and perhaps sometimes not very reliable since participants might not understand what they are being asked to judge. In that regard, relying on corpus data might be a more objective measure that is easily acquired. Nevertheless, the approach we suggest is certainly not the only way we can control for dominance using the corpus. For instance, another possible approach is to look at, for example, ten of the most frequent collocations, and determining whether the literal or the metaphorical meaning is the more dominant one.[10] The strength of this approach lies in the fact that this allows us to consider the dominance of the actual lexical item, that is, whether it more frequently appears in its literal or figurative meaning, and in this way, it is similar to the approach suggested by Bethard et al. (2009). However, this process would include getting subjective judgments on several collocations to determine whether the meaning of the verb is figurative or literal, which is not always an easy task, even with tools such as MIPVU (Steen et al., 2010). Both approaches, especially the latter, focus on a slightly different type of dominance, i.e. overall verb dominance, rather than the expression-specific meaning dominance. This might also be an important factor in experimental design. Again, introducing another variable into stimuli construction could prove to be a hugely complex task in practice, which might be avoided by controlling for this additional measure in the statistical analysis rather than through the stimuli construction. A possible problem that arises here is deciding which variables should be included during the stimuli construction stage. A choice needs to be made on a case-to-case basis: while meaning dominance might be important when testing indirect metaphorical expressions, the familiarity of expression could play a prominent role in *A is B* expressions.

5. Conclusion

The main aim of this paper is to show that the theoretical debate about processing literal and figurative meanings and the supporting experimental evidence are often hampered by methodological shortcomings. This was undertaken through an experimental study in which we have shown how in the case of indirect metaphorical expressions, the ease of access for metaphorical meaning does not differ from the literal meaning. The results of the study speak in favor of the Graded Salience Hypothesis and Career of Metaphor Theory.

We claim that studies investigating the processing of figurative language need to use a set of stimuli that is carefully controlled for the dominance of different types of meaning to obtain robust and straightforward results. If the meaning dominance is not accounted for, any processing differences that may be observed can be ascribed

[10] This approach was suggested by an anonymous reviewer.

to the differences between literal and metaphorical processing, or to the differing dominance of different meanings of a polysemous word. This paper takes meaning dominance to be one of the central semantic features of polysemous expressions that plays a crucial role in processing of figurative language, especially indirect metaphorical expressions. The central contribution of this paper is both theoretical and practical: we discuss the importance of meaning dominance in the construction of experimental stimuli and show how such stimuli might be constructed using corpus tools.

Finally, while our study uses stimuli that are balanced in terms of dominance, one of the possible future research directions certainly includes investigating cases in which stimuli dominance is not balanced (using the methodology we propose) or where the context is manipulated to further explore how these differences affect processing.

Appendix. Full stimuli list

Target	Literal prime	Metaphorical prime	Unrelated prime
spend	fortune	holiday	radio
invest	cash	effort	garden
digest	milk	news	seat
devour	pizza	novel	bishop
climb	fence	rank	storm
touch	ground	heart	camp
cover	floor	story	church
release	prisoner	survey	cottage
grow	tomato	revenue	cinema
expose	skin	truth	winter
feed	pig	ego	dome
hide	scar	flaw	haste
build	bridge	trust	noise
weave	wool	plot	pine
break	window	promise	uncle
hurt	ankle	pride	cave
grasp	handle	essence	tunnel
catch	trout	train	farm
cut	grass	budget	beach
wipe	stain	smirk	hawk
run	mile	test	bush
save	dollar	planet	alarm
wrestle	lion	demon	gossip
seize	arm	chance	age
approach	vehicle	topic	sugar
bring	camera	relief	empire
carry	weapon	burden	devil
lose	key	inch	rain

drop	bomb	price	birth
share	meal	belief	neck
clear	table	mind	size
spread	butter	fear	hotel
construct	roof	theory	breath
close	mouth	case	foot
fight	battle	cancer	mirror
open	bottle	account	desire
possess	wealth	strength	sand
bury	corpse	emotion	golf
defeat	enemy	purpose	pattern
navigate	terrain	menu	kettle
give	gift	life	mood
support	weight	goal	van
land	plane	client	card
sell	gold	soul	pain
flood	road	brain	voice
conquer	foe	death	fox
scrap	ship	tax	cloud
hold	baby	show	street

References

Audacity Team, 2010. Audacity (Version 1.3.12). Retrieved from http://audacity.sourceforge.net/.

Ashby, Jane, Roncero, Carlos, de Almeida, Roberto G., Agauas, Stephen J., 2017. The early processing of metaphors and similes: Evidence from eye movements. The Quarterly Journal of Experimental Psychology, 1-9.

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

Bambini, Valentina, Canal, Paolo, Resta, Donatella, Grimaldi, Mirko, 2018. Time Course and Neurophysiological Underpinnings of Metaphor in Literary Context. Discourse Processes 56 (1), 77-97.

Bates, Douglas, Maechler, Martin, Bolker, Ben, Walker, Steven, Christensen, Rune Haubo Bojesen, Singmann, Henrik, ... Bolker, Ben, 2014. Package 'lme4' (p. 12). Vienna: R foundation for statistical computing.

Blasko, Dawn G., Connine, Cynthia M., 1993. Effects of familiarity and aptness on metaphor processing. Journal of experimental psychology: Learning, memory, and cognition 19 (2), 295-308.

Bonnaud, Veronique, Gil, Roger, Ingrand, Pierre, 2002. Metaphorical and non-metaphorical links: a behavioral and ERP study in young and elderly adults. Neurophysiologie Clinique/Clinical Neurophysiology 32 (4), 258-268.

Boersma, Paul, Weenink, Don, 2018. Praat: doing phonetics by computer [Computer program]. Version 6.0.39. retrieved 1 May 2018 from http://www.praat.org/.

Bowdle, Brian F., 1998. Conventionality, polysemy, and metaphor comprehension. Unpublished doctoral dissertation, Northwestern University, Evanston, IL.

Bowdle, Brian F., Gentner, Dedre, 1995. The Career of Metaphor. Poster session presented at the Thirty-Sixth Annual Meeting of the Psychonomic Society, Los Angeles, CA.

Bowdle, Brian F., Gentner, Dedre, 2005. The Career of Metaphor. Psychological review 112 (1), 193-216.

Cameron, Lynne, 1999. Identifying and describing metaphor in spoken discourse data. In: Cameron, L., Low, G. (Eds.) Researching and Applying Metaphor. Cambridge: Cambridge University Press, pp. 105-132.

Cameron, Lynne, 2003. Metaphor in educational discourse. Advances in Applied Linguistics, Bloombury Publishing.

Cameron, Lynne, 2011. Metaphor in spoken discourse. In: Gee, J. P., Handford, M. (Eds.) The Routledge Handbook of Discourse Analysis. London: Routledge, pp. 342-355.

Carrol, Garreth, Conklin, Kathy, 2020. Is all formulaic language created equal? Unpacking the processing advantage for different types of formulaic sequences. Language and speech, 63, 1, 95-122.

Clark, Herbert, Lucy, Peter, 1975. Understanding what is meant from what is said: A study in conversationally conveyed requests. Journal of Verbal Learning and Verbal Behavior 12, pp. 335-359.

Coulson, Seana, Matlock, Teenie, 2001. Metaphor and the space structuring model. Metaphor and Symbol 16 (3-4), 295-316.

Davis, Colin J, 2005. N-Watch: A program for deriving neighborhood size and other psycholinguistic statistics. Behavior Research Methods 37 (1), 65-70.

Dulcinati, Giulio, Mazzarella, Diana, Pouscoulous, Nausicaa, & Rodd, Jennifer, 2014. Processing metaphor: The role of conventionality, familiarity and dominance. UCL Working Papers in Linguistics.

Duffy, Susan A., Morris, Robin K., Rayner, Keith, 1988. Lexical ambiguity and fixation times in reading. Journal of Memory and Language 27, 429-446.

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

Foraker, Stephani, Murphy, Gregory L., 2012. Polysemy in sentence comprehension: Effects of meaning dominance. Journal of memory and language 67 (4), 407-425.

Gentner, Dedre, 1983. Structure mapping: A theoretical framework for analogy. Cognitive Science, 7, 155-170.

Gentner, Dedre, 1989. The mechanisms of analogical learning. In: Vosniadou, S., Ortony, A. (Eds.), Similarity and analogical reasoning. Cambridge: Cambridge University Press, pp. 199-241.

Gentner, Dedre, Bowdle, Brian, F., Wolff, Phillip, Boronat, Consuelo., 2001. Metaphor is like analogy. In: Gentner, D., Holyoak, K., Kokinov, B. (Eds.), Analogical mind: Perspectives from cognitive science. Cambridge, MA: MIT Press, pp. 199-253.

Gentner, Dedre, Bowdle, Brian F., 2001. Convention, form, and figurative language processing. Metaphor and Symbol, 16, 223-247.

Gentner, Dedre, Wolff, Phillip, 1997. Alignment in the processing of metaphor. Journal of Memory and Language 37, 331-355.

Gentner, Dedre, Toupin, Cecile, 1986. Systematicity and surface similarity in the development of analogy. Cognitive science 10 (3), 277–300.

Gibbs, Raymond W., 1981. Your wish is my command: Convention and context in interpreting indirect requests. Journal of Verbal Learning and Verbal Behavior 20, 435-444.
Gibbs, Raymond W., 1984. Literal meaning and psychological theory. Cognitive science 8 (3), 275-304.
Gibbs, Raymond W., 1994. The poetics of mind: Figurative thought, language, and understanding. Cambridge: Cambridge University Press.
Gibbs, Raymond W., 2017. Metaphor Wars. Cambridge: Cambridge University Press.
Gildea, Patricia, Glucksberg, Sam, 1983. On understanding metaphor. The role of context. Journal of Verbal Learning and Verbal Behavior 22, 577-590.
Giora, Rachel, 1997. Understanding figurative and literal language: The graded salience hypothesis. Cognitive Linguistics 8 (3), 183-206.
Glucksberg, Sam, 2001. Understanding figurative language: From metaphors to idioms. New York: Oxford University Press.
Glucksberg, Sam, 2003. The psycholinguistics of metaphor. Trends in cognitive sciences 7 (2), 92–96.
Glucksberg, Sam, Gildea, Patricia, Bookin, Howard, 1982. On understanding non-literal speech: Can people ignore metaphors? Journal of Verbal Learning and Verbal Behavior 1, 85-96.
Glucksberg, Sam, Keysar, Boaz, 1990. Understanding metaphorical comparisons: Beyond similarity. Psychological Review 97, 3-18.
De Grauwe, Sophie, Swain, Abigail, Holcomb, Phillip J., Ditman, Tali, Kuperberg, Gina. R., 2010. Electrophysiological insights into the processing of nominal metaphors. Neuropsychologia 48 (7), 1965–1984.
Grice, Herbert Paul, 1975. Logic and Conversation. In: Cole, P., Morgan, J. (Eds.), Syntax and Semantics, 3: Speech Acts, New York: Academic Press, pp. 41-58.
Harris, Richard Jackson, 1976. Comprehension of metaphors: A test of the two-stage processing model. Bulletin of the Psychonomic Society 8 (4), 312-314.
Holme, Randal, 2004. Mind, metaphor and language teaching. Palgrave Macmillan UK.
Inhoff, Albrecht W., Lima, Susan D., Carroll, Patrick J., 1984. Contextual effects on metaphor comprehension in reading. Memory and Cognition 12 (6), 558-567.
Janus, Raizi A., Bever, Thomas G., 1985. Processing of metaphoric language: An investigation of the three-stage model of metaphor comprehension. Journal of Psycholinguistic Research 14, 473-487.
Kalandadze, Tamara, Bambini, Valentina, Næss, Kari-Anne B., 2019. A systematic review and meta-analysis of studies on metaphor comprehension in individuals with autism spectrum disorder: Do task properties matter? Applied Psycholinguistics 40 (6), 1421-1454.
Katz, Albert N., Ferretti, Todd R., 2001. Moment-by-moment reading of proverbs in literal and nonliteral contexts. Metaphor and Symbol 16 (3-4), 193-221.
Kilgarriff, Adam, Baisa, Vít, Bušta, Jan, Jakubíček, Miloš, Kovář, Vojtěch, Michelfeit, Jan, Rychlý, Pavel, Suchomel., Vít, 2014. The Sketch Engine: ten years on. Lexicography 1, 7-36.
Kyriacou, Marianna, Conklin, Kathy, Thompson, Dominic, 2018. Fathoming the depths of the idiomatic advantage: an eye-tracking investigation. Talk given at FTL 4 in Braga, Portugal in October 2018.
Lai, Vicky Tzuyin, Curran, Tim, Menn, Lise, 2009. Comprehending conventional and novel metaphors: An ERP study. Brain Research 1284, 145-155.

Lai, Vicky Tzuyin, Curran, Tim, 2013. ERP evidence for conceptual mappings and comparison processes during the comprehension of conventional and novel metaphors. Brain and language 127 (3), 484-496.
Lakoff, George, Johnson, Mark, 1980. Metaphors we live by. Chicago, IL: University of Chicago.
Lemhöfer, Kristin, Broersma, Mirjam, 2012. Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. Behavior Research Methods 44, 325-343.
Libben, Maya R., Titone, Debra A., 2008. The multidetermined nature of idiom processing. Memory & Cognition 36 (6), 1103-1121.
Mashal, Nira, Faust, Miriam, Hendler, Talma, & Jung-Beeman, Mark, 2007. An fMRI investigation of the neural correlates underlying the processing of novel metaphoric expressions. Brain and language 100 (2), 115-126.
McElree, Brian, Nordlie, Johanna, 1999. Literal and figurative interpretations are computed in parallel. Psychonomic Bulletin 6, 486-494.
McRae, Ken, Spivey-Knowlton, Michael J. S., Tanenhaus, Michael K., 1998. Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. Journal of Memory and Language 38, 283-312.
Milburn, Evelyn, 2018. The effects of meaning dominance and meaning relatedness on ambiguity resolution: Idioms and ambiguous words (Doctoral dissertation). University of Pittsburgh, Pittsburgh, USA.
Milburn, Evelyn, Warren, Tessa, 2019. Idioms show effects of meaning relatedness and dominance similar to those seen for ambiguous words. Psychonomic bulletin & review 26 (2), 591-598.
OED Online, Oxford University Press, December 2018, https://en.oxforddictionaries.com/definition/support. Accessed 15 February 2019.
Ortony, Andrew, Schauert, Diane L., Reynolds, Ralph E., Antos, Stephen J., 1978. Interpreting metaphors and idioms: Some effects of context on comprehension. Journal of Verbal Learning and Verbal Behavior 17, 465-477.
Pragglejaz Group, 2007. MIP: A method for identifying metaphorically used words in discourse. Metaphor and Symbol 22 (1), 1-39.
Pynte, Joël, Besson, Mireille, Robichon, Fabrice-Henri, Poli, Jézabel, 1996. The time-course of metaphor comprehension: An event-related potential study. Brain & Language 55, 293-316.
R Development Core Team, 2011. R: A language and environment for statistical computing. Austria: R Foundation for Statistical Computing Vienna. ISBN:3-900051-07-0.
Rastle, Kathleen, Harrington, Jonathan, Coltheart, Max, 2002. 358,534 nonwords: The ARC Nonword Database. Quarterly Journal of Experimental Psychology 55A, 1339-1362.
Reetz, Henning, & Kleinmann, Achim, 2003. Multi-subject hardware for experiment control and precise reaction time measurement. In: Solé, M.J., Recasens, D., J. Romero, J. (Eds.), Proceedings of the 15th International Congress of Phonetic Sciences. Barcelona, Spain, pp. 1489-1492.
Reijnierse, Gudrun, Burgers, Christian, Bolognesi, Marianna, Krennmayr, Tina, 2019. How Polysemy Affects Concreteness Ratings: The Case of Metaphor. Cognitive Science 43, 8, e12779.

Rychlý, Pavel, 2008. A Lexicographer-Friendly Association Score. In: Sojka, P., Horák, A. (Eds.) Proceedings of Recent Advances in Slavonic Natural Language Processing, RASLAN 2008. pp. 6-9.
Searle, John, 1979. Expression and meaning. Cambridge: Cambridge University Press.
Simpson, Greg B., 1981. Meaning dominance and semantic context in the processing of lexical ambiguity. Journal of Memory and Language 20 (1), 120-136.
Steen, Gerard, Dorst, Lettie, Herrmann, Berenike, Kaal, Anna, Krennmayr, Tina, 2010. Metaphor in usage. Cognitive Linguistics 21 (4), 765-796.
Steen, Gerard, Dorst, Lettie, Herrmann, Berenike, Kaal, Anna, Krennmayr, Tina, 2010. A method for linguistic metaphor identification. From MIP to MIPVU. Amsterdam: John Benjamins.
Stowe, Kevin, Palmer, Martha, 2018. Leveraging Syntactic Constructions for Metaphor Identification. In: Proceedings of the Workshop on Figurative Language Processing, New Orleans, pp. 17-26.
Sweetser, Eve, 1990. From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure. Cambridge: Cambridge University Press.
Tabossi, Patrizia, Colombo, Lucia, Job, Remo, 1987. Accessing lexical ambiguity: Effects of context and dominance. Psychological research 49 (2-3), 161-167.
Titone, Debra A., Connine, Cynthia M., 1999. On the compositional and noncompositional nature of idiomatic expressions. Journal of Pragmatics 31 (12), 1655-1674.
Weiland, Hanna, Bambini, Valentina, Schumacher, Petra B., 2014. The role of literal meaning in figurative language comprehension: Evidence from masked priming ERP. Frontiers in Human Neuroscience, 8, 583.
Williams, John N., 1992. Processing polysemous words in context: Evidence for interrelated meanings. Journal of Psycholinguistic Research 21 (3), 193-218.
Winter, Bodo, 2019. Statistics for Linguists: An Introduction Using R. Routledge.
Wray, Alison, & Perkins, Mick, 2000. The functions of formulaic language: An integrated model. Language and Communication 20, 1-28.