Boon or Burden? The Role of Compositional Meaning in Figurative Language Processing and Acquisition

Mila Vulchanova[1] · **Evelyn Milburn**[1] · **Valentin Vulchanov**[1] · **Giosuè Baggio**[1]

Abstract

We critically address current theories of figurative language, focusing on the role of literal or compositional meaning in the interpretation of non-literal expressions, including idioms and metaphors. Specifically, we formulate and discuss the processing hypothesis that compositional meaning may either facilitate or impede the recovery or construction of the intended figurative meaning depending on multiple factors, and in particular, on the expression's decomposability and on the "strength" of semantic relations between the compositional and figurative meanings. As a case study, we consider research on processing and acquisition of figurative expressions in highly verbal individuals with autism spectrum disorder (ASD) in comparison to neuro-typical individuals, and examine the factors that may account for the observed comprehension deficits in the ASD group. Based on this evidence, we discuss some of the strategies employed by language users in processing non-compositional or non-literal expressions, and we highlight implications for research on natural language comprehension and processing systems in the domain of figurative meaning.

Keywords Compositionality · Figurative language · Autism spectrum disorder (ASD) · Language processing · Language acquisition

1 Introduction

Figurative language is a pervasive phenomenon in everyday human communication. It covers a wide range of expressions or utterance types, such as idioms, metaphors, metonymy, jokes, irony and sarcasm, hyperbole, indirect requests, and stereotyped expressions, such as clichés. A study investigating the incidence of non-literal expressions in e-mails written by young people found that 94.30% of text messages included at least one non-literal statement. People used on average 2.90 non-literal

✉ Mila Vulchanova
  mila.vulchanova@ntnu.no

1   Language Acquisition and Language Processing Lab, Department of Language and Literature, Norwegian University of Science and Technology, 7491 Trondheim, Norway

expressions per e-mail (Whalen et al. 2009). Unlike literal language, where at least a tentative interpretation of a complex expression may be derived by composing the meaning of each of the expression's constituents, figurative language often seems to require *additional operations* in order to arrive at the intended meaning. Some types of literal expressions (e.g., aspectual and complement coercions) are also thought to engage additional semantic operations that may *not* be entirely lexically specified or mirrored by the syntax, as compositionality requires (for reviews of behavioral and M/EEG evidence, see Pylkkänen and McElree 2006; Baggio 2018; for representative research, see Traxler et al. 2002; Piñango et al. 2006; Pylkkänen and McElree 2007; Brennan and Pylkkänen 2008; Kuperberg et al. 2010; Baggio et al. 2010; Paczynski et al. 2014). However, it is the potential *specificity* of operations underlying figurative language that concerns us here.

An additional focus of the present paper is cognitive development. The competences and skills associated with mastery of figurative language seem to take much longer to develop than word knowledge (vocabulary) or core grammar. In typical language development, children first demonstrate appreciation of figurative expressions, such as idioms, at some point in their school years (Nippold 1998, 2006; Nippold and Duthie 2003; Cain et al. 2009; Levorato and Cacciari 1995). The development of this ability seems to pattern in ways similar to the emergence of so-called *dimensionality* in language competences or skills, as established in a recent large-scale cohort study covering pre-school to early school ages (LARRC 2015). In that study, the three core dimensions of language competence (i.e., vocabulary, grammar, and discourse) may be distinguished around third grade at school, but are not distinct earlier. One open question concerning the developmental trajectory of figurative language abilities is whether it displays a linear trend over time (Nippold 1998, 2006), or whether it is, instead, shaped by a quadratic trend peaking shortly before adolescence, with less change afterwards (Kempler et al. 1999; Laval and Bernicot 2002; Vulchanova et al. 2011). More intriguingly, figurative language skills, by taking longer to acquire, may manifest some vulnerability both in developmental deficits and across the life-span. Research in typical ageing suggests that older adults produce fewer idioms and may benefit more from cueing than younger speakers (Conner et al. 2011), and findings from acquired deficits, such as aphasia, have shown impaired idiom comprehension (Cacciari et al. 2006; Milburn et al. 2018). Problems with figurative language have been systematically documented in developmental deficits, for example autism spectrum disorder (ASD) (Volden and Phillips 2010; Ramberg et al. 1996). Recent research has established impairment or failure to understand pragmatic, non-literal aspects of language, such as metaphors, idioms, and other forms of figurative language, even when structural language appears to be largely intact (Gold and Faust 2010; see also Vulchanova et al. 2015 for a comprehensive and critical review of converging evidence from existing research). Similar generalised vulnerability of figurative language has been attested in schizophrenia, where metaphors, proverbs, idioms, and irony have been shown to present a definite challenge to these patients (Thoma and Daum 2006; Bambini et al. 2016b; Saban-Bezalel and Mashal 2017).

From a cognitive and neural perspective, non-literal uses of language are a natural part of the way in which language is represented in the minds of speakers and of

the processes that underlie semantic access and use. For instance, metaphor is typically based on *mappings across* concepts or conceptual domains (Lakoff and Johnson 1980; Tourangeau and Sternberg 1982; Lakoff 2008) and on the activation of semantic or conceptual information (e.g., some of the semantic features associated to sharks are applied to a human being in 'My lawyer is a shark'). This is a *relational computation*, reflecting semantic links (associative, distributional, categorial, logical etc.) between items in the mental lexicon (Baggio 2018). Verbal humour, jokes, and irony build on detecting incongruity between the literal interpretation of the verbal message and a (symbolic) representation of reality, which may often be resolved by factoring in the speaker's intention. Such processes are also involved in understanding ambiguity in literal expressions, where inferring the speaker's intentions in the given context are key to the correct interpretation. The human brain is adequately equipped with the machinery to process figurative expression, akin to its usual mode of operation.

If, from a semantic and cognitive perspective, humans are fully equipped to process figurative language—although this ability might take longer to develop compared to other language skills—the question then is what might make figurative language more challenging and open to vulnerability. A related question, given the relative complexity of figurative expressions, is what can explain the high prevalence of non-literal language in discourse. Below, we address critically current accounts of figurative language processing, focusing on (1) the role of compositional meaning in deriving figurative interpretations, and (2) some of the main factors that influence figurative language comprehension. The purpose of the present discussion is theory building, not a comprehensive review of research on figurative language processing (for surveys of the theoretical, developmental, and neurolinguistic literature, with a focus on idioms and metaphors, see, among others, Hattouti et al. 2016; Holyoak and Stamenković 2018; Vulchanova et al. 2015; Kalandadze et al. 2018; Baggio 2018).

2 Figurative Language Processing

Figurative language is commonly characterized as non-transparent in various ways: the comprehender has to go beyond the literal meanings of the constituent words in a figurative expression in order to recover the speaker's intended meaning. There is often a gap, and sometimes even a conflict, between the compositional meaning of a given expression (or "sentence meaning") and its intended interpretation in context (or "utterance meaning"). Figure 1 illustrates this idea for several kinds of figurative expressions. Predicative relations are shown of the forms 'S is P' (what the *sentence* means) and 'S is R' (what the *speaker* means by 'S is P' in context). For example, the speaker may say 'My lawyer is a shark' ('S is P') to mean that her lawyer is ruthless ('S is R'). For literal expressions, sentence meaning and utterance meaning coincide. However, the relationship varies for different types of figurative language. In simple metaphorical phrases, there is a *unique mapping* from the compositional (or literal) meaning (e.g., [S is a shark]) to the figurative meaning (e.g., [S is ruthless]). Here, the figurative meaning may be recovered *deterministically* from
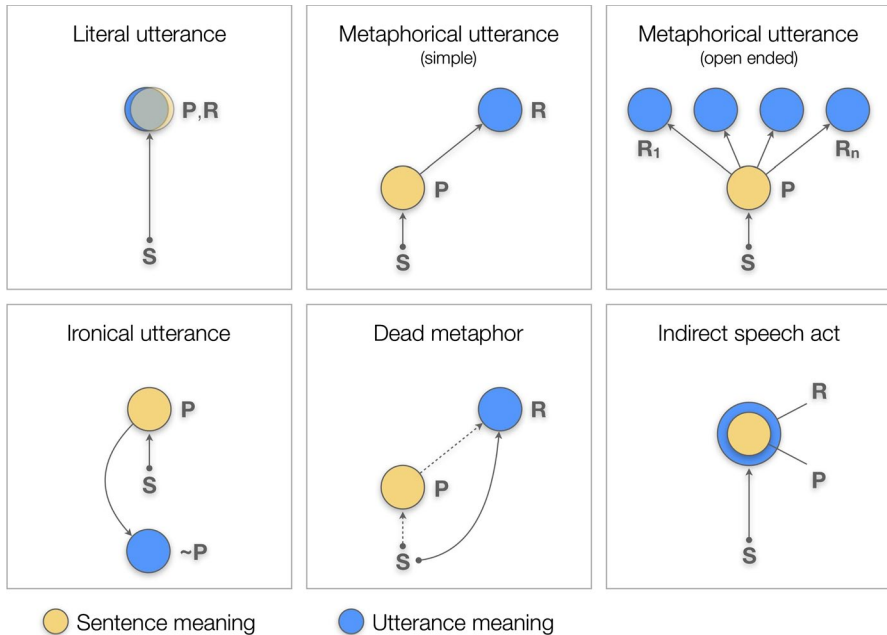
Fig. 1 Models of relations between a topic (S), sentence meaning (yellow), and utterance meaning (blue) in figurative expressions: metaphor, irony, and indirect speech acts. Adapted from Searle (1993, p. 110). See Bambini (2017) for further discussion. (Color figure online)

sentence meaning—that is, directly from the sentence's logico-syntactic form and from lexical meanings. In open-ended metaphorical utterances, multiple mappings link the given predicate (P) to a range of possible figurative meanings $(R_1,...,R_n)$; literary metaphors are often of this kind. In ironical utterances, P and R may be opposites semantically ('John is a rock', uttered ironically, means that he is weak), but the same kind of relation holds: the intended meaning can be recovered from sentence meaning by applying specific interpretation functions or by drawing certain (pragmatic) inferences.

The most interesting case for our purposes in this paper are non-literal expressions whose meanings *cannot* be recovered via the application of interpretive functions to compositional or literal meaning. To illustrate this point more concretely, consider the figurative meanings of expressions such as 'to hit the sack' (to go to sleep) or 'to kick the bucket' (to die). Here, meaning cannot be recovered based on the meanings of verbs and their complements. To derive the appropriate figurative interpretation, the comprehender must rely on *familiarity* with the expression—i.e., on knowledge of the form-meaning pair and its use in production and comprehension—as well as on the *context* in which the expression is embedded. One notable property of some of these expressions is that they have both literal *and* figurative meanings, which may be unrelated to one another. 'To kick the bucket' has the figurative meaning [to die] *and* the compositional meaning [to strike a bucket with one's foot]; crucially, these meanings are semantically *un*related. In some cases, the

relation between figurative and compositional or literal meanings may be more or less direct or transparent, as, e.g., in 'to pop the question' [to propose marriage]. This raises a series of questions concerning the specific role of compositional meaning in the derivation of figurative interpretations, in particular for expressions such as idioms (sometimes referred to as "dead metaphors") and proverbs, where the relation between the compositional and figurative meanings is indirect, opaque, or absent.

Most accounts of figurative language identify this tension between compositional or literal and figurative meanings as essential to understanding the exact cognitive and neural mechanisms underlying figurative language comprehension. In the past few decades, two accounts have dominated the theoretical landscape. First, *indirect access* theories (known as the "standard model") posit that non-literal or figurative expressions violate one or more conversational maxims (Grice 1975) and therefore that meaning may be recovered (largely inferentially) by exploiting the cooperative principle (Levinson 1983; Eco 1986). Under this view, the meaning of figurative expressions is constructed, or accessed indirectly via an extended interpretive, inferential process. Second, *direct access* theories assume that the context can, at least in some cases, support the *direct activation* or *construction* of figurative meaning (Gibbs 1990; but see Noveck et al. 2001 for a re-assessment of evidence). Recent pragmatic models put forth within the Relevance Theory framework have gravitated either towards a direct access account, construed as a "deflationary" model (Sperber and Wilson 2008) or towards the "lingering of the literal meaning" view (Carston 2010). Importantly, direct and indirect access models may not be completely opposed: the construction of figurative interpretations may or may not be mediated by access or computation of compositional meaning, and crucially this mediation may *depend on a number of factors*, such as the decomposability of the figurative expression (see below for details) and the properties of the lexical items in the expression (e.g., the frequency of words, constructions, or collocations). In order to derive adequate and testable hypotheses concerning how different types of non-literal expressions may be processed, such factors ought to be in focus, both theoretically and empirically.

In cases where some variant of the indirect access account does hold, the question becomes *under what circumstances does compositional (literal) meaning facilitate or obstruct the derivation of a figurative interpretation?* Historically, few models have been developed and tested by directly comparing the entirety of figurative language or just different types of figurative expressions (e.g., idioms vs. metaphors). Instead, several proposals have largely focused on one specific type of figurative expression, most often metaphor. This has resulted in a rich history of in-depth theoretical and empirical research on specific types of figurative expressions, but it has also meant that theories of processing of *one* type of expression are only rarely tested on *other* expressions. One exception is arguably the *graded salience hypothesis* (or GSH; Giora 1997, 2003), which has been tested and used to account for more expression types: idioms, metaphors, and irony. A successful theory of figurative language processing should account for interactions between literal and figurative meanings, regardless of the specific kinds of figurative expressions that the model is primarily designed to explain. The question of the role of compositional meaning in figurative language comprehension is particularly prominent in most current models

of idiom comprehension. In what follows, we thus focus on idioms as proxies for figurative language comprehension more generally.

2.1 Models of Idiom Processing

Idioms are figurative multi-word expressions, such as 'kick the bucket', 'sail close to the wind', or 'sing the blues'. Several older accounts of idiom comprehension and representation posit that idioms are stored as large "chunks", akin to single words, and are retrieved as whole units during processing (e.g. Swinney and Cutler's (1979) *Lexical Representation Hypothesis*, discussed below). Under this view, the semantics of individual words in the idiom should play no role in accessing or computing the figurative meaning, because they have no relation to the idiom's figurative meaning.

However, many idioms are at least partially analyzable. For example, one idiomatic component may function according to its typical collocational environment, and the other may require a metaphorical interpretation. This relationship is also known as *decomposability* and is frequently used in idiom research to investigate the tensions between literal lexical meaning and overall phrase meaning. For example, the idiom 'to bury the hatchet' contains a verb, 'to bury', which refers to a reconciliation event, and an NP, 'the hatchet', which refers to a disagreement (Jackendoff 1997). Because we can put the syntax *and* the meaning of 'bury the hatchet' into one-to-one correspondence, the idiom is said to be decomposable. In contrast, the syntactic and semantic components of 'kick the bucket' do not correspond, even metaphorically, to any part of the event [die] that it describes. These kinds of idioms are considered non-decomposable. Importantly, idioms vary widely along a spectrum of decomposability (Bulkes and Tanner 2017), and classifying idioms as decomposable or nondecomposable can additionally vary depending on the population being tested (Nordmann et al. 2014).

Interestingly, processing of even supposedly non-decomposable idioms appears to show effects of individual words. For example, head verbs in non-decomposable idioms can retain their aspectual features. This limits both the syntactic flexibility of these idioms and the contexts in which they may occur. For example, the sentence '*John lay kicking the bucket due to his chronic illness' is either unacceptable or costly in processing terms because it describes a temporally extended event (progressive aspect) that is incongruent with the punctual feature of 'kick' (Glucksberg 1991; Hamblin and Gibbs 1999). Therefore, idioms that would appear non-decomposable can still be affected by some semantic properties of their constituent words. This indicates that simply characterizing idioms as either decomposable or nondecomposable might obscure important nuances.

These properties of idioms have given rise to two main types of processing models: *non-compositional models*, in which idioms are stored and retrieved as multi-word chunks or constructions (form-meaning pairings), and *compositional models*, which focus instead on the possibility that individual lexical and syntactic

constituents can affect the interpretation of the idiom on-line. We critically assess these models and the predictions they make about idiom processing below.

2.1.1 Non-compositional Models

Non-compositional accounts of idiom processing assume that idioms are stored and processed as multi-word chunks, subject to the same processing mechanisms (e.g., access and selection) as most single lexical items. Non-compositional accounts build on the observation that figurative idiom meaning is not always a "function" of the literal meanings of its lexical constituents, in the sense of compositionality (Partee 1995; for a discussion in the context of language processing, see Baggio et al. 2012). In addition, in the linguistics tradition, several authors have highlighted the fact that many idioms are *syntactically frozen*: their structure often cannot be modified and exploited productively (see, e.g., Chomsky 1980; Cutler 1982; Nunberg et al. 1994; Jackendoff 2002), indicating that an idiom's syntactic form may be stored alongside its figurative meaning.

Non-compositional accounts largely assume that literal and figurative meanings are accessed *sequentially* during comprehension, but they may differ with respect to the order in which these meanings are accessed. For example, in the *direct access model* of idiom comprehension (Gibbs 1990, 1994), figurative meanings may be retrieved directly following relevant cues from the linguistic and other (e.g., communicative) context of the expression. Compositional analysis starts only if figurative meaning is found to be inappropriate after retrieval, given the context. This algorithm is highly reminiscent of the dual processing account of lexical access of irregular word-forms put forth by Pinker and Prince (1994). In contrast, the first step in processing according to the *standard pragmatic approach* involves activating the literal meaning of constituent words in the idiom. If compositional processing of literal meaning fails, as for other utterances, pragmatic inferencing is invoked to exit the impasse, possibly along the lines of the cooperative theory of communication inspired by Grice (1975), thereby achieving the intended figurative interpretation.

A third non-compositional account of idiom processing is Swinney and Cutler's (1979) *lexical representation hypothesis* (LRH). This proposal is critically different from the standard pragmatic approach and from the direct access model, because under the LRH compositional analysis and figurative meaning retrieval unfold simultaneously. The figurative meanings of idioms are assumed to be stored in the mental lexicon as multi-word units. Presentation of the first word of an idiom immediately triggers automatic retrieval of the idiom's figurative meaning *and* compositional analysis of its literal meaning. Retrieving one long word (the figurative meaning) is faster and easier than compositional analysis of the literal meaning, so the figurative meaning is processed first and has priority in comprehension. Under the LRH, figurative phrases should always be processed faster than literal phrases.

Swinney and Cutler (1979) found support for the LRH in a reading and acceptability judgment task. Participants read idioms and matched control phrases consisting of idioms with a single word replaced to create a literal, grammatical English phrase. Participants judged whether the phrase they had just read was an acceptable phrase in English. Participants judged idioms to be acceptable phrases faster

than matched control phrases, regardless of the idioms' "syntactic frozenness", of the transitional probabilities of words within the phrases, or how aware participants were that they were reading idioms. These results are compatible with the results of other studies documenting faster processing of idioms compared to literal expressions (Conklin and Schmitt 2008; Ortony et al. 1978).

Non-compositional models are among the very first accounts of idiom processing to have appeared historically. These models recognize that idiom meaning is often independent of compositional meaning, thus capturing an essential feature of idioms. In fact, more recent evidence supports a notion of language representation that allows for similar storage, access, and processing of single words *and* common strings of words, or lexical bundles. In general, comprehenders are sensitive to the *frequencies* of literal multi-word phrases, so that more frequent literal expressions would be processed faster (Arnon and Snider 2010; Siyanova-Chanturia et al. 2011) and may be remembered more accurately (Tremblay et al. 2011) than less frequent phrases. If the language comprehension system is able to process *literal* multi-word expressions in ways analogous to single words, then idioms—in essence, *figurative* multi-word phrases—may be treated similarly. Models that account for the tension between compositional and overall phrasal or sentential meanings might, therefore, help build processing models encompassing both literal and figurative language.

2.1.2 Compositional Models

Compositional accounts of idiom processing point to experiments showing effects of single word meanings on idiom interpretation (Caillies and Butcher 2007; Hamblin and Gibbs 1999; Nordmann et al. 2013) as evidence that idiom processing may involve (partial) analysis of individual words, regardless of the degree to which those words contribute to the idiom's figurative meaning. For example, Hamblin and Gibbs (1999) suggested that idiom interpretation depends on identifying the main constituents in the expression. Supporting this, they observed that the action denoted by an idiom's main verb affected how the whole idiom was interpreted, even for idioms that were otherwise non-decomposable: participants had consistent intuitions on the manner in which the events described by idioms took place, and they preferred replacement verbs that preserved this relationship rather than disrupting it.

The *Configuration Hypothesis* of idiom comprehension (Cacciari and Glucksberg 1991; Cacciari and Tabossi 1988) is a representative model wherein interpretation proceeds largely compositionally until the comprehender recognizes that the *configuration* of words that they are processing corresponds to an idiom, a point known as the *idiom key*. The figurative meaning of the idiom is then directly accessed and retrieved, and compositional analysis halts. Importantly, identification of the idiom is guided by co-occurrence frequencies of the words in the idiom, rather than by the semantic relationships between the idiom's constituents and its figurative meaning. The most important construct affecting comprehension is *familiarity of the phrase*, which may be analyzed as a function of high transitional probabilities between the words in the idiom and of the comprehender's sensitivity to these transitional probabilities. The more familiar a phrase is to the comprehender, the easier it is to recognize the given configuration.

Supporting the Configuration Hypothesis, Tabossi et al. (2009) found equally fast judgments of meaningfulness for decomposable and non-decomposable idioms as well as for compositional clichés, and concluded that familiar phrases are recognized faster than unfamiliar phrases, regardless of their idiomaticity.

However, the evidence is mixed regarding whether compositional analysis is *halted completely* when the comprehender recognizes the idiom, as the Configuration Hypothesis predicts. In an ERP study in Dutch, Rommers et al. (2013) presented participants with noun phrases following idiomatic and literal contexts. Noun phrases were either the normal continuation of the idiom, a semantically-related replacement, or contained unrelated words. Critically, all idioms were non-decomposable, meaning that their individual words should make no contribution to overall figurative meaning, and therefore that compositional analysis need not continue once the phrase has been recognized as an idiom. They used the amplitude of the N400 component in ERPs as a measure of lexical semantic processing difficulty: the N400 is reduced when context facilitates processing of the eliciting word (Kutas and Federmeier 2011; Baggio and Hagoort 2011). They found that semantically-related NPs elicited reduced N400s compared to unrelated NPs in literal contexts only, but that there were no effects of semantic relatedness in idiomatic contexts. Additionally, the amplitudes of the N400 components elicited by the unrelated and semantically-related NPs were similar, and both were larger than the N400 evoked by the expected idiomatic NP. The authors interpreted this result as indicating that compositional processing can be "switched off" when it is rendered unnecessary by the context, as for the key words of non-decomposable idioms. This is consistent with the Configuration Hypothesis: by the final word in the idiom, the idiom's key should have been reached and compositional analysis halted, making compositional processing of the final word unnecessary (Cacciari 2014).

In contrast to this result, Smolka et al. (2007) found activation of the literal meanings of German verbs even when they appeared at the end of figuratively-biased phrases. This may suggest that compositional and literal analysis can continue even *after* the figurative phrase has been identified, contrary to the predictions of the Configuration Hypothesis. Critically, although Smolka et al. did not use exclusively non-decomposable idioms, they embedded their idioms in strongly biasing contexts, which may have resulted in high predictability of the idiom's final words. According to the Configuration Hypothesis, however, the literal meanings of idiom-final words should not have been activated: by that point in the idiom, the figurative meaning should have been directly retrieved.

Compositional models of idiom comprehension identify individual lexical meanings as critical for idiom comprehension, even when idioms appear non-decomposable. Compositional accounts therefore recognize that the processing of idioms cannot be reduced to lexical access or lexical activation only (Cacciari and Tabossi 1988; Gibbs 1992; Vega-Moreno 2001), in contrast to non-compositional accounts. In particular, the Configuration Hypothesis represents a significant theoretical improvement over (strictly) non-compositional models of idiom processing, as it provides a *mechanism* by which the language system can recognize and react to the presence of an idiom.

2.1.3 The Hybrid Model

Both compositional and non-compositional models of idiom processing can explain important aspects of idiom processing. However, each model type explains findings that the other model type is *not* able to explain. For example, compositional models can explain findings of single-word influences on overall idiomatic meanings, while non-compositional models capture the nature of idioms as multi-word phrases and parallel recent work investigating processing of multi-word literal phrases.

To resolve this tension, and to account for evidence supporting both compositional and non-compositional processing of idioms, Titone and Connine (1999) proposed the *hybrid model of idiom comprehension*. Under this model, idiom comprehension follows two *simultaneous, parallel routes*, similar to Swinney and Cutler's LRH (1979): (a) *direct access* of the idiomatic meaning as soon as the idiom can be identified, and (b) *compositional analysis* based on the literal meanings of the idiom's constituents. In the hybrid model (HM), idioms can function simultaneously as arbitrary pairings of form and meaning (much like single words) and as compositional expressions.

Like earlier models of idiom comprehension, the HM identifies the tension between single-word and phrasal or sentential meaning as critical for explaining how idioms are processed. However, Titone and Connine (1999) specifically identify *decomposability* as the critical variable involved in resolving this tension, and therefore propose that idiom processing and representation can differ depending on the idiom's decomposability. This is markedly different from earlier models of idiom comprehension, which have assumed that processing and representation do *not* differ based on characteristics of the idiom. For example, under compositional *and* non-compositional models, 'to pop the question' and 'to kick the bucket' are processed in similar ways, despite the fact that one is decomposable and the other is not. However, several studies have reported processing advantages for decomposable over non-decomposable idioms. The figurative meanings of decomposable idioms are activated earlier than those of non-decomposable idioms (Caillies and Butcher 2007; Caillies and Declercq 2011), and activation of the literal *and* figurative meanings of decomposable idioms are facilitated compared to non-decomposable idioms (Titone and Connine 1999). In addition, decomposable idioms are read faster than non-decomposable idioms, and may be less disrupted by lexical changes (Gibbs et al. 1989).

To explain these results, Titone and Connine (1999) suggested that the literal and the figurative meanings of decomposable idioms are often (highly) semantically related, and that this relatedness might speed up comprehension of decomposable idioms. If the literal and idiomatic meanings of decomposable idioms are similar or related, concurrent compositional analysis of literal meaning may *facilitate* or *augment* the direct retrieval of the figurative meaning, resulting in faster processing. In contrast, slower processing for non-decomposable idioms is caused by *interference* between directly retrieved figurative meaning and the semantically dissimilar compositional meaning, which is activated concurrently during processing. Importantly, the HM is fully consistent with the overall view of semantic processing that has emerged from experimental research during the past

two decades, pointing to the existence of two *parallel, simultaneous, and inter-acting streams* for semantic processing in the brain: (a) a memory-based stream, which activates the meanings of words, constructions, and chunks, and tracks semantic relations (i.e., associative, distributional, categorial, logical) between them; and (b) a compositional stream, which binds together lexical meanings based on phrase and sentence-level constraints, including local syntactic relations (Kuperberg 2007; see Baggio 2018 for further details and a discussion of support-ing evidence).

In their first test of the HM, Titone and Connine (1999) examined reading times for decomposable and non-decomposable idioms. Idioms were presented accompa-nied by a context sentence, which appeared either before or after the idiom, and biased comprehension towards either the idiom's literal or its figurative meaning. Titone and Connine (1999) found that non-decomposable idioms were read more slowly when context preceded the idiom, irrespective of contextual bias. Decom-posable idioms were however read equally quickly, regardless of contextual bias and location of the context. They interpreted these results as suggesting that *both* literal and figurative meanings of idioms are activated during comprehension. This resulted in little or no processing cost for decomposable idioms because of a higher degree of relatedness between literal and idiomatic meanings. Integration of the contextually-appropriate meaning of a non-decomposable idiom, in contrast, was impaired because of on-line competition between the unrelated meanings. Addi-tional support for the HM comes from production studies. Individual constituent words of idioms primed retrieval of the entire idiom, indicating that idioms are accessed at least partly compositionally during production (Sprenger et al. 2006). In "tip-of-the-tongue" states, participants more frequently reported words related to the literal meanings of idioms that they could not produce, indicating that literal mean-ings of individual words in an idiom were available as speakers tried to access the idiomatic meaning (Nordmann et al. 2013). In consideration of the need to accom-modate compositional, literal analysis *and* direct access of idiom meaning, hybrid models of idiom comprehension may be the best account to date for the results of previous research.

3 Factors in Figurative Language Processing

In this section, we briefly discuss three factors that influence idiom comprehension: idiom decomposability, familiarity, and supportive context. However, several other potentially critical factors have been identified and studied. For example, Nunberg et al. (1994) proposed that all idioms may be described by the orthogonal factors of *compositionality*, *conventionality*, and *transparency*. Our intent in focusing on idiom decomposability, familiarity, and context is not to dismiss other characterizations of the factors that influence idiom comprehension. Instead, we have chosen to discuss these factors because they have a long history of investigation, using a wide variety of experimental measures and techniques. Examining these factors affords us a rich, detailed view of the mechanisms underlying idiom processing.

3.1 Idiom Decomposability

The processing and interpretation of figurative language depends on a number of critical properties. As illustrated by Titone and Connine's (1999) Hybrid Model (HM), the idiom's *decomposability* is a strong influence on idiom comprehension. Although operationalizations of decomposability vary from study to study, decomposability is often used either to measure how well the literal meanings of individual words in an idiom correspond (figuratively) to aspects of the idiom's figurative meaning (see the examples in Section 2; Gibbs et al. 1989; Nunberg et al. 1994), or less specifically, to indicate that constituent words in the idiom may contribute to the overall figurative meaning in some way (Caillies and Butcher 2007; Hamblin and Gibbs 1999; Titone and Connine 1999). These conceptual differences render it difficult to compare results between studies using different operationalizations of decomposability. Regardless of the exact definition adopted, the construct of decomposability is always used to characterize the semantic links between the idiom's literal and figurative meanings, and therefore captures a critical aspect of idiomatic language.

As previously mentioned, decomposability seems to facilitate idiom processing. One potential explanation for this result is that the semantic relatedness between literal and figurative meanings of a decomposable idiom speeds comprehension. Under the HM, this is because concurrent compositional analysis of literal meaning augments direct retrieval of the figurative meaning, resulting in faster processing. Moreover, this model presupposes that the two streams can interact continuously. In contrast, any variance or interference between the directly retrieved figurative meaning and the highly semantically dissimilar compositional meaning results in slower or more costly processing of non-decomposable idioms. Additional evidence of an advantage for decomposable idioms comes from studies of idiom processing in healthy aging. Westbury and Titone (2011) found that older adults were overall both slower and less accurate than younger adults when making decisions on whether the meaning of non-decomposable idioms was literal or not. They interpreted these data as showing that older adults have difficulty resolving semantic ambiguity, which is maximized by co-activation of unrelated literal and figurative meanings of a non-decomposable idiom.

However, in some situations higher decomposability can impair processing even for younger comprehenders. In a priming study, Titone and Libben (2014) observed that increased semantic decomposability actually *interferes* with idiom priming 1000 ms following idiom offset, in contrast to the predicted decomposability advantage. They argued that the primary advantage of higher decomposability may be in later stages of processing, when one specific interpretation of the idiom is being embedded into a larger context. Also, their priming paradigm only investigated meaning activation, which may not be as affected by idiom decomposability.

Additionally, processing of highly decomposable idioms may be slowed down when the figurative meaning of the idiom is dominant or relatively more frequent (Duffy et al. 1988) compared to the literal meaning (Milburn and Warren under review). In an eye-tracking study during reading, Milburn and Warren examined eye movement responses to idioms varying *both* in semantic relatedness between the

literal and figurative meanings *and* in dominance of the figurative meaning over the literal meaning. They found facilitated processing, as revealed by decreased go-past time, re-reading time, and total time measures, for idioms with highly related literal and figurative meanings, and with neither meaning being strongly dominant[1] over the other (e.g., 'deliver the goods'). In contrast, idioms with highly related literal and figurative meanings, and with strongly dominant figurative meanings (e.g., 'on the fence'), showed slower processing. They explained these results by suggesting that the facilitative effect of decomposability depends on the relative dominances of the literal and the figurative meanings: when the literal and the figurative meanings are more balanced, concurrent activation of literal and figurative semantics facilitates processing; however, when one meaning (e.g., figurative) is dominant, activation of the other non-dominant meaning (e.g., literal) interferes with processing.

To summarize, higher idiom decomposability appears to facilitate idiom processing, especially during later processing stages when idioms are integrated into a context. However, the status of some idioms as ambiguous units—with both comprehensible literal and figurative meanings—can result in interference from co-activated related meanings when one meaning is strongly dominant.

3.2 Idiom Familiarity

Another factor with a significant influence on idiom comprehension is the idiom's *familiarity*, or its subjective frequency for an individual comprehender. Familiar idioms are consistently easier for comprehenders compared to less-familiar idioms (Milburn et al. 2018; Qualls et al. 2003; Schweigert 1986; Titone and Libben 2014). Interestingly, this effect appears to be consistent regardless of whether the literal or figurative meaning is intended. For instance, Schweigert (1986) found that increased familiarity sped whole-sentence reading times irrespective of whether idioms were embedded in literally-biasing or in figuratively-biasing sentences. Congruent with this view, in an eye-tracking experiment, Milburn and Warren (under review) found that increased familiarity facilitated processing, as indexed by decreased go-past, re-reading, and total time measures, of the idiom itself, regardless of context bias.

Although subjective idiom familiarity is related to objective frequency, these factors are not identical. Additionally, the relationship between familiarity and other measures such as subjective frequency, meaningfulness, and contextual fit is as-yet unclear. For example, idiom familiarity interacts with both the frequency with which the idiom is used and the frequencies of individual words in the idiom. In a large-scale norming study, Bulkes and Tanner (2017) observed that familiarity ratings correlated positively with ratings of how well comprehenders knew the figurative meanings of idioms. Using principal component analysis (PCA), Bulkes and Tanner suggested that their measures of idiom familiarity and meaningfulness built on a single underlying construct, which was moreover separable from an idiom's given corpus frequency.

---

[1] For the present purposes, *dominance* can be defined as the relative frequency of one meaning over others (Duffy et al. 1988) or the meaning with which the comprehender is most familiar or is most strongly supported by the context (Giora 1997).

Further complicating the picture, an individual speaker's familiarity with an idiom might differ greatly from the idiom's overall frequency of use, consistent with other norming studies that found high variability in idiom familiarity across populations (e.g., Nordmann et al. 2014). These results are reminiscent of well-attested findings of high variability in lexical knowledge. Furthermore, idiom familiarity interacts with the frequencies of single constituent words. For example, the magnitude of the idiom familiarity effect seems to be diluted when the idiom contains low-frequency words as constituents (Cronk et al. 1993). This effect also seems to vary depending on whether the literal or figurative meaning is intended. An interesting question, in light of these findings, is how the frequencies of an idiom's lexical constituents, their collocational frequency, and the idiom's given frequency interact to drive semantic access, and whether this process is marked by competition or rather facilitation between literal and figurative meanings. We will address this question in more detail later on.

Critically for researchers interested in familiarity effects on idiom processing, there is currently no consensus on how to operationalize familiarity, much less which test could be used as an objective measure (see Thibodeau et al. 2017 for a discussion). Some research has used subjective measures, such as perceived experience with the figurative item (Blasko and Connine 1993), but others suggest that corpus frequency can be used as a viable objective measure due to its high correlation with familiarity (Thibodeau and Durgin 2011; but see Bulkes and Tanner 2017 for some evidence that subjective familiarity may be independent of corpus frequency). Thus, an important methodological issue in idiom research is how to operationalize all the factors that play a key role in idiom processing, and how to establish valid measures to be used in experimental design or modeling.

3.3 Context in Idiom Processing

An additional factor influencing idiom processing is the *context* in which the idiom is embedded. Although a figuratively-biasing context is not necessarily required to arrive at a figurative interpretation of an idiom, biasing contexts may facilitate retrieval of literal or figurative meaning. Qualls et al. (2003) observed that supportive contexts aided rural adolescents' comprehension of non-familiar idioms, although they used an offline definition selection task that did not directly allow the investigation of immediate context effects on idiom processing. However, Holsinger (2013) found an immediate effect of context on interpretation of idiomatic phrases: participants looked more at figurative probes when they heard idioms embedded in figurative contexts, and at literal probes when they heard idioms in literal contexts. This indicates that context can successfully drive idiom interpretation towards the literal or figurative.

However, other factors influencing idiom comprehension, such as decomposability and familiarity, may interact with context, resulting in instances where context bias *inhibits* successful comprehension. Ortony et al. (1978) showed that highly familiar idioms in a figuratively-biased context were understood more quickly than idioms in a literally-biased context. It is possible that, because their idioms were familiar to

participants, the figurative meanings were more accessible, thereby interfering with literal interpretation when the same idioms appeared in a literal context. Likewise, context can interact with idiom decomposability. Titone and Connine (1999) reported no effects of context bias, whether literal or figurative, for either decomposable or non-decomposable idioms. Instead, they found differences in processing depending on whether the context was located before the idiom, thereby biasing it, or after the idiom, thereby disambiguating it. Non-decomposable idioms were read more slowly when context preceded the idiom, regardless of contextual bias. Yet, decomposable idioms were read equally quickly regardless of contextual bias and placement of the context. They interpreted these data as showing that literal *and* figurative meanings of idioms were activated on-line, and that construction of contextually-appropriate meanings of non-decomposable idioms was impaired due to competition between the unrelated meanings, especially in the presence of preceding biasing context. The effects of context on idiom comprehension are therefore complex, and are qualified depending on other, inherent, characteristics of the idiom.

Familiarity and supportive context are important during processing of other types of figurative expressions. In particular, these two factors seem to affect how quickly and how accurately different populations of speakers can understand metaphors. In a cross-modal priming paradigm, highly familiar metaphors led to higher and faster activation of the figurative meaning compared to less familiar expressions (Blasko and Connine 1993). That study also found evidence that figurative activation was not caused by activation of individual words in the metaphorical expression, but rather by activation of the emergent metaphorical meaning of the phrase as a whole. The role of familiarity has been confirmed in other studies, after controlling for factors such as aptness (e.g., Damerall and Kellogg 2016; Holyoak and Stamenkovic Holyoak and Stamenkovic 2018). In a recent review of studies on metaphor, Holyoak and Stamenkovic (2018) highlight the importance of context, and identify this as an area for future research in the field, as few studies so far have used context in their designs (Gerring and Healy 1983; Gibbs and Gerrig 1989; Giora 2003; Nayak and Gibbs 1990; Ortony et al. 1978; Thibodeau and Durgin 2008). Thibodeau et al. (2017) however suggest that processing fluency and figurativeness are responsible for familiarity ratings and metaphor processing. This study also provides new evidence of the supportive role of context in understanding metaphors: target metaphorical sentences were processed more fluently when they were preceded by a context that included matching metaphoric language than when they were preceded by a context that included mixed metaphoric language or literal language. Moreover, expressions presented in matching figurative context received higher rating on comprehensibility and aptness by speakers.

4 Figurative Language in Atypical Development

The diversity and complexity of factors involved in figurative language processing may be especially challenging in developmental disorders, such as autism spectrum disorder (ASD). Difficulties in this domain are well attested (Tager-Flusberg 2006; Volden and Phillips 2010; Vulchanova et al. 2015), but their source remains largely unknown. Current debates center on whether the figurative language impairment in

autism mostly resides in difficulties in language competences and skills or is rather linked to aspects of autism symptomatology and of the autism phenotype (Norbury 2005; Gernsbacher and Pripas-Kapit 2012; Vulchanova and Vulchanov 2018).

In a series of specifically-designed experiments, we investigated performance on figurative language tasks (involving both idioms and metaphors) in highly verbal individuals with autism compared to IQ- and language ability-matched neuro-typical individuals (Chahboun et al. 2016; Chahboun et al. 2017). The participants in those studies came from two age groups, i.e., 10–12 years (children) and young adults in the age range 16–22 years in a cross-sectional design. The two age ranges and the cross-sectional design were included specifically to establish developmental trajectories in controls and in the experimental group. Also, the choice of highly verbal individuals with autism and the careful matching to controls allowed for excluding language problems per se as the cause for potential difficulty in figurative language comprehension in that group.

Our main findings may be summed up as follows. The main problems encountered by participants with autism were primarily reflected in greater reaction latencies in comparison to controls. The participants with autism performed at adequate levels of accuracy, although still displaying poorer responses in comparison to controls. Another significant finding is the different developmental trajectories between the experimental groups and controls: young adult participants with autism performed at the level of control children, but better than children with autism, as evidenced by main effects of Age and Group in our results. More importantly, we also find evidence of potentially different underlying strategies between individuals with autism and controls in processing of figurative language and in text comprehension.

One main finding in that research is that young adults with autism are less accurate than adults without autism. A valid question then is what types of errors are they making. The results in Chahboun et al. (2016) show that the responses they provide are more literal. In this study, a difference in degree of literalness was observed in response accuracy. The model revealed a main effect of Group (control/ASD) ((χ^2(1, 26) = 5.22, p = .022), with more literal responses by participants with autism and a smaller difference in accuracy between Age groups (children/young adults) ((χ^2(1, 26) = 3.51, p = .06). Furthermore, a two-way interaction between Age and Group was found (χ^2(1,26) = 4.89, p = .02). Additional multiple comparisons revealed that this interaction was likely due to a difference between control young adults and young adults with autism (p = .015), where young adults with autism converged on more literal responses than did their typically developing peers. Thus, the younger participants and participants with autism in our study interpreted the stimuli more often literally than did older participants and controls. These data provide support for findings in research on young children and individuals with autism documenting an overall tendency for literal interpretation (Mitchell et al. 1997). Data from the same study of figurative language processing suggest, in addition, that younger participants and participants with autism have specific difficulties with idioms with greater decomposability, but no such problems were observed with novel decomposable metaphors or literal expressions. This result was not expected and opens up for a number of possible accounts. Importantly, it suggests that idiom decomposability interferes negatively with the idiom processing and interpretation, and increases the

likelihood of non-figurative, literal interpretation in younger speakers and individuals with autism. In contrast, decomposability and transparency appear to provide an advantage in the processing of other decomposable expressions, such as e.g., novel metaphors. These data align with the studies reported above, where idiom decomposability has been documented to pose a problem when other factors were at play (e.g., lower familiarity). Furthermore, they are consistent with a recent ERP experiment of idiom comprehension in Chinese, where decomposable expressions, both idioms and free literal expressions, elicited greater ERP responses than non-decomposable idioms (Zhang et al. 2013), suggesting greater processing load. These data and observations suggest that decomposability is not a "one and all" factor, and whether it presents an advantage or not for processing may depend on a number of factors, such as the nature of the expression (free or figurative; idiom or metaphor), its lexical status, the likelihood that it is part of the speaker's lexicon (stored or not), the speakers' degree of exposure (familiarity with the expression) and the speaker's age. Furthermore, in certain contexts idiom decomposability increases the processing load specifically for participants with autism. A tentative account may be that in such contexts, the literal meanings of the idiom constituents are activated, thus causing greater competition between possible interpretations, and preventing access to the target, figurative interpretation. We address this issue in 5.2 below.

5 Processing Strategies in Figurative Language Comprehension

5.1 Neural Aspects

On the backdrop of the factors involved in processing non-literal language in typical individuals, and of the problems observed in highly verbal individuals with autism, a possible approach needs to look at what features of the expression would trigger literal (composition) strategies, procrastinating the target figurative interpretation. We aim to outline under what conditions this is more likely to happen, focusing on two main factors: the novelty and the decomposability of the figurative expression.

Recent studies using neuroscience methods point to different processing strategies for novel versus conventional figurative expressions. Early research (Winner and Gardner 1977; Rinaldi et al. 2004) showed that patients with right hemisphere (RH) lesions have difficulty interpreting figurative expressions, and prefer literal interpretations when these are available. These and similar results led to the hypothesis that the RH is primarily engaged during the construction of figurative meaning, whereas the LH subserves mainly processing of literal or compositional meaning (for a discussion of recent versions of this hypothesis, see Baggio 2018, Ch. 5; see also below). However, Luria (for discussion, see Bambini 2017) showed as early as the 1940s that LH-lesioned patients had problems both with compositional aspects of meaning *and* figurative language (e.g., metaphors and proverbs). These early results suggest that the division of processing labor between the LH and RH does not quite correspond to the distinction between literal and figurative meaning. The LH seems to be crucially involved in the construction of both literal and non-literal meaning, whereas the role of the RH remains to some extent elusive. Thus, most fMRI studies

on figurative language, in particular metaphor, have addressed two main questions: (1) the involvement of RH regions in the derivation of non-literal meanings, and (2) the engagement of areas known to be involved in mentalizing, perspective taking, or related social cognitive processes, such as the medial prefrontal cortex (mPFC). Two prominent theories of figurative language comprehension make definite predictions in this respect. The *coarse semantic coding* (CSC) theory (Jung-Beeman 2005) holds that the RH is more active when "distant semantic relations" are established, such as between the meanings of 'butcher' and 'surgeon' in the metaphoric expression 'That surgeon is a butcher'. The CSC theory posits that lexical meanings are represented asymmetrically in the two hemispheres, by a finer-grained code in the LH (reflecting the hierarchical structure of conceptual knowledge; Federmeier and Kutas 1999) and a coarser code in the RH. The two hemispheres would represent the same concepts; what differs are the types of semantic relationships *between concepts* coded in each hemisphere. Alternative to the CSC theory is the *Graded Salience Hypothesis* (or GSH; Giora 1997, 2003), which holds that the distinction between conventional and novel metaphors is key. The GSH also assumes that the literal meaning of novel metaphors becomes available first in the LH, and that figurative meaning is constructed by the RH, possibly at later stages. In contrast, the figurative meaning of conventional and known metaphors is immediately accessed via the LH.

The available evidence does not sit well either with the CSC theory or with GSH. For example, right temporal and superior frontal regions are involved in *early stages* of understanding novel metaphors (Arzouan et al. 2007a), which contradicts a view of GSH where the RH should be engaged *after* initial LH processing. Other research has shown that the left inferior frontal gyrus (LIFG) and the posterior superior temporal gyrus (pSTG) are rapidly activated by processing novel metaphors (Schneider et al. 2014), which seems difficult to reconcile with the CSC theory. Here, one might argue that there is almost *parallel access* to the literal meaning of a figurative phrase in the LH and generation of the non-literal meaning in the RH. However, that is unlikely to result in an empirically adequate model. The reason is evidence for the involvement of LH regions, including classical perisylvian language regions, in the construction of figurative meaning. For example, a meta-analysis of imaging experiments by Bohrn et al. (2012) showed that known metaphors engaged primarily regions such as LIFG and left STG, whereas novel metaphors produced stronger responses also in the left middle frontal gyrus and left mPFC. A direct comparison between conventional and novel metaphors confirmed that the former class of expressions is processed by the core LH language network (LIFG and STG), whereas novel metaphors also activated the right IFG and cingulate regions. Moreover, a contemporaneous meta-analysis by Rapp et al. (2012) indicated that conventional metaphors and idioms are processed by the LH with activation foci in the LIFG and MTG/STG, while novel metaphors also activated middle and medial frontal cortices, RIFG, and the parahippocampal region. A novel set of findings concerns the engagement of left inferior parietal lobe regions in figurative language processing (Bambini et al. 2011; Benedek et al. 2014; Obert et al. 2014). The role of parietal cortex in semantic and pragmatic processing is still debated (Catani and Bambini 2014; Baggio 2018). These results show that processing figurative language relies heavily on LH systems, and that additional regions (mPFC, IPL, the right hemisphere etc.) are only engaged by novel figurative expressions.

EEG studies on metaphor have reported modulations of all known ERP components directly or indirectly associated with semantic processing: the N400, the P600, and post-N400 sustained frontal negativities (for discussion, see Baggio 2018, Chs. 2 and 5). For example, Pynte et al. (1996) showed that the same word (e.g., 'lions') produces a larger N400 effect when it functions as a metaphor vehicle (e.g., 'Those fighters are lions') relative to its occurrence in a literal sentence (e.g., 'Those animals are lions'). Modulations of the N400 amplitude in studies such as this one suggest that meaning is activated with a similar time course in figurative and literal contexts. But does the N400 reflect the computation of non-literal meanings or just differences in strength of semantic relations between words in metaphoric and literal sentences? The latter hypothesis seems the most plausible. Coulson and van Petten (2002) compared literal sentences ('He knows that whiskey is a strong intoxicant'), conventional metaphors ('He knows that power is a strong intoxicant'), and literal cross-domain mappings ('He has used cough syrup as an intoxicant'). In the latter case, the relevant mapping would link the conceptual domain of actual intoxicants (alcoholic beverages) to the broader domain of substances that may be used for similar purposes, such as cough syrup. For conventional metaphors, the mapping between intoxicants and power is structurally similar but in addition yields a *figurative effect*. In this study, metaphors produced the largest N400, followed by literal mappings and by literal expressions, in that order. Only metaphors also triggered a post-N400 positivity, resembling the P600. These data indicate that the N400 reflects brain processes that track semantic relations in the input and in memory. 'Power' and 'intoxicant' are less semantically related than 'cough syrup' and 'intoxicant', thus the N400 effect will be larger in the former case. The process of deriving a figurative interpretation of the sentence will however be reflected in the post-N400 window, either by the P600 or by a sustained anterior negativity (SAN) in about the same time range (~500–800 ms). P600 effects have been found with metaphor (Bambini et al. 2016a), metonymy (Nieuwland and Van Berkum 2005; Schumacher 2011, 2014), and idioms (Canal et al. 2017). While the standard interpretation of the P600 as just a marker of syntactic processing has been abandoned, it is quite possible that syntax still plays an *indirect role* here. The P600 would reflect a *conflict* between the output of the two streams described early on in this paper: (a) a memory-based stream, that derives meanings based on stored or contextually-available semantic relations between words (and can directly access the meaning of conventional and otherwise familiar figurative expressions), and (b) a syntax-driven stream, that generates meanings via compositional analysis of input strings (for further details, see Baggio 2018, Chs. 2 and 5; see also Bambini 2017, Ch. 3, for an attempt at unifying current P600 findings).

The ERP correlates of processing *novel metaphors* are however different from N400 and P600 effects. Experiments that compared conventional metaphors to novel ones found sustained negative ERPs in response to novel metaphors, such as 'Brain waves are stethoscopes' (Arzouan et al. 2007b; Lai et al. 2009; Goldstein et al. 2012; Bambini et al. 2019). These sustained negative shifts in ERPs likely reflect the construction of metaphoric meaning, not the novelty of the conceptual mapping as such (Davenport and Coulson 2011). In addition, they are similar to the SAN effects evoked by jokes and humor, which require the joint presence of novelty (one should not have heard that

joke before) and a non-strictly-compositional meaning (Coulson and Kutas 2001). SAN effects are also found in response to sentence that do *not* involve figurative interpretation (see Baggio et al. 2008; Wittenberg et al. 2014; Paczynski et al. 2014 for some examples). These data suggest that literal and figurative language processing alike occur in two successive stages or time frames: (1) activation of lexical meanings and tracking of semantic relations between words (N400), followed by (2) construction of a literal or a figurative interpretation (SAN) and comparison of the resulting discourse model with the output of syntax-driven composition (P600). None of these stages is specific to either literal or figurative language processing: what differs is the *content* of the interpretation that is constructed in each case, and whether content is mostly provided in the first stage (lexical activation and relational processing) or in the second (construction of a discourse model, or interpretation). Early relational semantic processes will generally suffice for comprehension of most literal expressions and of many *conventional* figurative expressions. In addition, this processing strategy will be computationally efficient for highly frequent expressions (e.g., common idioms and metaphors, collocations etc). When the relations between the constituent expressions in discourse cannot be matched, partially or completely, to stored semantic relations, additional interpretive operations will be engaged. The bottom line is that much of figurative language processing may be conceptualized as a two-step algorithm, where an early search for meaning in memory is followed by a creative process of interpretation, less constrained by stored knowledge.

5.2 Meaning Activation and Competition

As discussed above, one key factor that needs to be considered in the case of idioms is their decomposability, and to what extent properties of the constituent words can trigger competition between literal and figurative meaning. Some studies have used the notion of *semantic plausibility*, suggesting that, in the absence of biasing context, both a literal and a figurative interpretation may be equally plausible. For instance, the idiom 'to pull someone's leg' seems equally plausible figuratively and on a direct literal analysis. But other expressions may not easily yield such interpretations. For example, 'I am a bit under the weather today' cannot possibly make sense on a strict literal interpretation. It is to be expected that only the semantically plausible idioms would trigger competition between the literal and the figurative meanings, whereas the less plausible ones would directly cue the target figurative meaning, which is the only plausible interpretation to be accessed. Another approach may be to assess the semantic relatedness, similarity, or "closeness" of the two available interpretations, the literal and the idiomatic (Milburn 2018). This approach may hold some promise in circumventing issues arising from the need to categorize expressions according to their decomposability, conventionality, transparency, and other features. Furthermore, semantic similarity can be measured fairly easily. On this approach, frequency could be easily added in the equation to estimate the collocational probability of one part of the expression co-occurring with another part in comparison to its "collocational alternatives", e.g., the same word co-occurring with other lexical items.

   To give a concrete example of an idiom like 'kick the bucket', one can estimate the probability of the NP 'the bucket' co-occurring with the head verb 'kick' against

the probability of the same head verb co-occurring with other fillers of the complement position. The relevant measure is the *cloze probability* of the dependent constituent. Cloze probability can be defined as the probability that a given word will be produced in a given context on a sentence completion task (Coulson 2007). In other words, this is the probability that a specific word will complete a specific syntactic frame, e.g., the (missing) object noun phrase in a verb phrase. Typically, cloze probability is measured as the percentage of native speaker individuals completing the phrase with the same word/phrase. Thus, cloze probability is an appropriate measure in that it reflects native speaker expectations of a given word occurring in a certain context. Concerning idioms and free expressions based on the same verb, we may assume that the degree of activation of possible candidates for the verb complement position occurring after the verb will depend on the *ratio of cloze probabilities of the respective filler phrases*. In the case of Verb-NP idioms, this is the ratio of the most frequent literal filler of the argument position to the NP filler in the idiom, as expressed in the formula:

$$\text{ClProb NP(MaxFreq) / ClProb NP(idiom)}$$

Cloze probabilities for NP fillers following head verbs may be estimated reliably either in norming studies with native speakers or by comparing frequencies of lemma occurrences in large-scale corpora. Furthermore, reliable correlations have been observed between cloze probabilities measured in sentence completion tasks with native speakers and in on-line corpora (Hammerås 2017). For instance, counts concerning 'kick the bucket' show that the most frequently occurring filler in the context of 'kick' is the NP 'the ball': native speaker cloze probability is 54%, which, on most accounts, is considered high cloze probability. Additionally, corpora list 'the ball' as the top most frequent complement filler after 'kick' (cf. iWeb Word Web Corpus). Regarding the collocational frequency of 'the bucket' as the argument filler in the idiom, counts vary depending on the corpus. A search in the Corpus of Contemporary American English (COCA; Davies 2008) yields a value, estimated by using the above formula, of 5.4 for 'ball'/'bucket' ratio in the context of 'kick', indicating that 'ball' is by far more frequent after 'kick'. The value is based on comparing the collocational frequencies of the two expressions, the literal and the figurative one. In contrast, consider the idiom 'to wear the pants'. A search of COCA shows that the most common filler following 'wear' is 'uniform', and by using the above formula, the value estimated for the 'uniform'/'pants' ratio is 2. Even though 'uniform' is twice more frequent after 'wear' than 'pants', the ratio is smaller than that for 'ball'/'bucket', suggesting that both 'uniform' and 'pants' as equally plausible completions of the phrase. This, according to our line of thinking, is what might lead to equal activation of both candidates ('uniform' and 'pants'), and thus, to greater competition for the complement slot at initial stages of processing the idiom, whereas no similar situation is expected for 'ball'/'bucket' in the context of 'kick'.

We assume that the likelihood that literal interpretations are activated is measured as the ratio between the cloze probabilities of the two argument filler candidates. We further hypothesize that values around 1 (according to the formula above) will lead to greater competition between the literal and figurative meaning, as a result of equal likelihood of literal and figurative activation of the Verb-NP collocation, as, e.g.,

illustrated by the ratio in 'wear the pants'. It can be further stipulated that values bigger than 1, and greater values in general, would augment processing and suppress the competition, as a result of clear collocational frequency distinctions between the literal and the figurative collocations. This is, for example, similar to the values we observe for 'kick the bucket', which has very low statistics of occurrence and presents no real competition with literal interpretations of 'kick'. The cloze probability value may even prove to be sufficient, as it appears to largely deal with the plausibility of both figurative and literal interpretation (i.e. semantic similarity measure) as well. This type of approach can be tested experimentally in a controlled design with carefully selected stimuli (cf. Milburn et al. in progress). Given known correlations between N400 amplitudes and cloze or transition probabilities of the eliciting word, predictions for ERP experiments may also be derived and implemented from this model. We should note, however, that the N400 is sensitive to a wide range of other variables—lexical, contextual, semantic etc. (reviewed by Kutas and Federmeier 2011; see Baggio and Hagoort 2011; Baggio 2018 for a unifying model)—which again poses the problem of generating sets of stimuli where cloze probabilities are manipulated while controlling for all other modulating factors. A general cautionary note is that a multi-factorial model is needed to fully account for the prevalence or the likelihood of accessing literal and figurative meanings of idioms, in particular to the extent that the decomposability, conventionality, transparency, or other features of idioms are not reflected in cloze or transitional probabilities, as analyzed above. The construct around which the current proposal builds is cloze probability. This does not preclude, however, possible alternative ways in which the occurrence of competing constituents is estimated, e.g., entropy measures, which is an approach we do not pursue in the current paper.

The issues discussed in this paper have a broad range of potential applications. We have argued that the factors that impact on figurative language processing and the cognitive and neural mechanisms that underlie non-literal language comprehension are largely similar to those employed for literal processing, and depend on the properties of the constituent words (frequencies, probability of occurrence) and the processes of access and activation. Many of the debates on figurative language that have been critically addressed here carry over to debates in the domain of lexical storage and access and processing of morphologically complex words (for a detailed computational model in the area of morphology, see O'Donnell 2015; for a discussion of related theoretical issues across areas of linguistics, see the contributions in Nooteboom et al. 2002; Pirelli et al. 2019). In those fields as well, hybrid models have been adopted as most plausible. Yet, there remain some underexplored areas where the analogy with idioms may prove illuminating. For example, the formal distinction between endocentric compounds (e.g., 'blackboard', where the head noun 'board' determines the syntactic category, N, and the overall meaning of the compound) and exocentric compounds (e.g., 'turnout', an N resulting from compounding a V and a P) may be similar to the distinction between free literal expressions and figurative expressions, because exocentric compounds are largely idiomatic in nature and interpretation. The same factors identified above for idioms (decomposability, conventionality, transparency etc.) could play a similar role in processing

exocentric compounds. This point serves to illustrate the potential generalizability of some of our observations.

6 Conclusion

Figurative language is highly prevalent in everyday discourse. However, seemingly paradoxically, its processing places additional load on language users, partly due to ambiguity between the intended figurative meaning and the available compositional interpretations, modulo presence of a biasing context. It has even been argued that this is the communicative value of such expressions: by introducing ambiguity, the expression becomes communicatively more efficient (Piantadosi et al. 2012). In this paper, we have addressed issues arising from some of the factors that impact on the processing of figurative language against common assumptions and accounts. Based on evidence from behavioural studies on idiom and metaphor processing in autism and in neurotypical individuals, we have proposed an approach which may capture the common problems encountered by special populations, and often by children, in the processing of non-literal language, at the same time offering a solution to how to operationalize key aspects of idiom processing in measurable and meaningful ways, consistent with how the human brain may be handling the task. The approach we envisage is based on the construct of cloze probability, which already has a long tradition in the neuroscience of language with results corroborating its relevance to native speakers' expectations in language processing. Such an approach is testable, and as such, can be useful for future experimentation and computational modeling of language.

References

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language, 62*(1), 67–82.

Arzouan, Y., Goldstein, A., & Faust, M. (2007a). Dynamics of hemispheric activity during metaphor comprehension: Electrophysiological measures. *NeuroImage, 36*(1), 222–231.

Arzouan, Y., Goldstein, A., & Faust, M. (2007b). Brainwaves are stethoscopes: ERP correlates of novel metaphor comprehension. *Brain Research, 1160,* 69–81.

Baggio, G. (2018). *Meaning in the brain*. Cambridge, MA: MIT Press.

Baggio, G., Choma, T., Van Lambalgen, M., & Hagoort, P. (2010). Coercion and compositionality. *Journal of Cognitive Neuroscience, 22*(9), 2131–2140.

Baggio, G., & Hagoort, P. (2011). The balance between memory and unification in semantics: A dynamic account of the N400. *Language and Cognitive Processes, 26*(9), 1338–1367.

Baggio, G., Van Lambalgen, M., & Hagoort, P. (2008). Computing and recomputing discourse models: An ERP study. *Journal of Memory and Language, 59*(1), 36–53.

Baggio, G., Van Lambalgen, M., & Hagoort, P. (2012). *The processing consequences of compositionality. The Oxford handbook of compositionality* (pp. 655–672). Oxford: Oxford University Press.

Bambini, V. (2017). *Il cervello pragmatico*. Rome: Carocci.

Bambini, V., Arcara, G., Bechi, M., Buonocore, M., Cavallaro, R., & Bosia, M. (2016a). The communicative impairment as a core feature of schizophrenia: Frequency of pragmatic deficit, cognitive substrates, and relation with quality of life. *Comprehensive Psychiatry, 71,* 106–120.

Bambini, V., Bertini, C., Schaeken, W., Stella, A., & Di Russo, F. (2016b). Disentangling metaphor from context: An ERP study. *Frontiers in Psychology, 7,* 559.

Bambini, V., Canal, P., Resta, D., & Grimaldi, M. (2019). Time course and neurophysiological underpinnings of metaphor in literary context. *Discourse Processes, 56*(1), 77–97.

Bambini, V., Gentili, C., Ricciardi, E., Bertinetto, P. M., & Pietrini, P. (2011). Decomposing metaphor processing at the cognitive and neural level through functional magnetic resonance imaging. *Brain Research Bulletin, 86*(3–4), 203–216.

Benedek, M., Beaty, R., Jauk, E., Koschutnig, K., Fink, A., Silvia, P. J., et al. (2014). Creating metaphors: The neural basis of figurative language production. *NeuroImage, 90,* 99–106.

Blasko, D. G., & Connine, C. M. (1993). Effects of familiarity and aptness on metaphor processing. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 19*(2), 295.

Bohrn, I. C., Altmann, U., & Jacobs, A. M. (2012). Looking at the brains behind figurative language—A quantitative meta-analysis of neuroimaging studies on metaphor, idiom, and irony processing. *Neuropsychologia, 50*(11), 2669–2683.

Brennan, J., & Pylkkänen, L. (2008). Processing events: Behavioral and neuromagnetic correlates of aspectual coercion. *Brain and Language, 106*(2), 132–143.

Bulkes, N. Z., & Tanner, D. (2017). "Going to town": Large-scale norming and statistical analysis of 870 American English idioms. *Behavior Research Methods, 49*(2), 772–783.

Cacciari, C. (2014). Processing multiword idiomatic strings: Many words in one? *The Mental Lexicon, 9*(2), 267–293.

Cacciari, C., & Glucksberg, S. (1991). *Understanding idiomatic expressions: The contribution of word meanings. Advances in psychology* (Vol. 77, pp. 217–240). Amsterdam: North-Holland.

Cacciari, C., Reati, F., Colombo, M. R., Padovani, R., Rizzo, S., & Papagno, C. (2006). The comprehension of ambiguous idioms in aphasic patients. *Neuropsychologia, 44*(8), 1305–1314.

Cacciari, C., & Tabossi, P. (1988). The comprehension of idioms. *Journal of Memory and Language, 27*(6), 668–683.

Caillies, S., & Butcher, K. (2007). Processing of idiomatic expressions: Evidence for a new hybrid view. *Metaphor and Symbol, 22*(1), 79–108.

Caillies, S., & Declercq, C. (2011). Kill the song—Steal the show: What does distinguish predicative metaphors from decomposable idioms? *Journal of Psycholinguistic Research, 40*(3), 205–223.

Cain, K., Towse, A. S., & Knight, R. S. (2009). The development of idiom comprehension: An investigation of semantic and contextual processing skills. *Journal of Experimental Child Psychology, 102*(3), 280–298.

Canal, P., Pesciarelli, F., Vespignani, F., Molinaro, N., & Cacciari, C. (2017). Basic composition and enriched integration in idiom processing: An EEG study. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 43*(6), 928.

Carston, R. (2010). XIII—Metaphor: Ad hoc concepts, literal meaning and mental images. In *Proceedings of the Aristotelian Society* (Vol. 110, No. 3_pt_3, pp. 295–321). Oxford: Oxford University Press.

Catani, M., & Bambini, V. (2014). A model for social communication and language evolution and development (SCALED). *Current Opinion in Neurobiology*, *28*, 165–171.

Chahboun, S., Vulchanov, V., Saldaña, D., Eshuis, H., & Vulchanova, M. (2016). Can you play with fire and not hurt yourself? A comparative study in figurative language comprehension between individuals with and without autism spectrum disorder. *PLoS ONE, 11*(12), e0168571.

Chahboun, S., Vulchanov, V., Saldaña, D., Eshuis, H., & Vulchanova, M. (2017). Can you tell it by the prime? A study of metaphorical priming in high-functioning autism in comparison with matched controls. *International Journal of Language & Communication Disorders, 52,* 766–785. https://doi.org/10.1111/1460-6984.12314.

Chomsky, N. (1980). *Rules and representations*. New York: Columbia University Press.

Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied Linguistics, 29*(1), 72–89.

Conner, P. S., Hyun, J., O'Connor Wells, B., Anema, I., Goral, M., Monereau-Merry, M. M., et al. (2011). Age-related differences in idiom production in adulthood. *Clinical Linguistics & Phonetics, 25*(10), 899–912.

Coulson, S. (2007). Electrifying results: ERP data and cognitive linguistics. In M. Gonzales-Marques, I. Mittelberg, S. Coulson, & M. Spivey (Eds.), *Methods in cognitive linguistics*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Coulson, S., & Kutas, M. (2001). Getting it: Human event-related brain response to jokes in good and poor comprehenders. *Neuroscience Letters, 316*(2), 71–74.

Coulson, S., & Van Petten, C. (2002). Conceptual integration and metaphor: An event-related potential study. *Memory & Cognition, 30*(6), 958–968.

Cronk, B. C., Lima, S. D., & Schweigert, W. A. (1993). Idioms in sentences: Effects of frequency, literalness, and familiarity. *Journal of Psycholinguistic Research, 22*(1), 59–82.

Cutler, A. (1982). Idioms: The colder the older. *Linguistic Inquiry, 13*(2), 317–320.

Damerall, A., & Kellogg, R. (2016). Familiarity and aptness in metaphor comprehension. *The American Journal of Psychology, 129*(1), 49–64.

Davenport, T., & Coulson, S. (2011). Predictability and novelty in literal language comprehension: An ERP study. *Brain Research, 1418,* 70–82.

Davies, M. (2008). *The corpus of contemporary American English (COCA): 560 million words, 1990-present.* https://corpus.byu.edu/coca/.

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language, 27,* 429–446.

Eco, U. (1986). *Semiotics and the philosophy of language*. Bloomington: Indiana University Press.

Federmeier, K. D., & Kutas, M. (1999). Right words and left words: Electrophysiological evidence for hemispheric differences in meaning processing. *Cognitive Brain Research, 8*(3), 373–392.

Gernsbacher, M., & Pripas-Kapit, S. (2012). Who's missing the point? A commentary on claims that autistic persons have a specific deficit in figurative language comprehension. *Metaphor and Symbol, 27*(1), 93–105.

Gerring, R. J., & Healy, A. F. (1983). Dual processes in metaphor understanding: Comprehension and appreciation. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 9*(4), 667.

Gibbs, R. (1994). *The poetics of mind: Figurative thought, language and understanding*. Cambridge: Cambridge University Press.

Gibbs, R. W., Jr. (1990). Comprehending figurative referential descriptions. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 16*(1), 56.

Gibbs, R. W. (1992). Categorization and metaphor understanding. *Psychological Review, 99,* 572–577.

Gibbs, R. W., Jr., & Gerrig, R. J. (1989). How context makes metaphor comprehension seem 'special'. *Metaphor and Symbol, 4*(3), 145–158.

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*(5), 576–593.

Giora, R. (1997). Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics, 8*(3), 183–206.

Giora, R. (2003). *On our mind: Salience, context and figurative language*. Oxford: Oxford University Press.

Glucksberg, S. (1991). Beyond literal meanings: The psychology of allusion. *Psychological Science, 2*(3), 146–152.

Gold, R., & Faust, M. (2010). Right hemisphere dysfunction and metaphor comprehension in young adults with Asperger syndrome. *Journal of Autism and Developmental Disorders, 40*(7), 800–811.

Goldstein, A., Arzouan, Y., & Faust, M. (2012). Killing a novel metaphor and reviving a dead one: ERP correlates of metaphor conventionalization. *Brain and Language, 123*(2), 137–142.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics 3: Speech acts* (pp. 41–58). Cambridge, MA: Academic Press.

Hamblin, J. L., & Gibbs, R. W. (1999). Why you can't kick the bucket as you slowly die: Verbs in idiom comprehension. *Journal of Psycholinguistic Research, 28*(1), 25–39.

Hammerås, M. L. (2017). Probing sensitivity to argument structure in two proficiency level groups: An exploratory study with Norwegian learners of English. MA thesis, NTNU.

Hattouti, J., Gil, S., & Laval, V. (2016). The development of idiom comprehension: A review. *L'Année psychologique, 116*(1), 105–136.

Holsinger, E. (2013). Representing idioms: Syntactic and contextual effects on idiom processing. *Language and Speech, 56*(3), 373–394.

Holyoak, K. J., & Stamenković, D. (2018). Metaphor comprehension: A critical review of theories and evidence. *Psychological Bulletin, 144*(6), 641.

Jackendoff, R. (1997). *The architecture of the language faculty* (Vol. 28). Cambridge: MIT Press.

Jackendoff, R. (2002). What's in the lexicon? In S. Nooteboom, F. Weerman, & F. Wijnen (Eds.), *Storage and computation in the language faculty* (pp. 23–58). Dordrecht: Kluwer.

Jung-Beeman, M. (2005). Bilateral brain processes for comprehending natural language. *Trends in cognitive sciences, 9*(11), 512–518.

Kalandadze, T., Norbury, C., Nærland, T., & Næss, K. A. B. (2018). Figurative language comprehension in individuals with autism spectrum disorder: A meta-analytic review. *Autism, 22*(2), 99–117.

Kempler, D., VanLancker, D., Marchman, V., & Bates, E. (1999). Idiom comprehension in children and adults with unilateral brain damage. *Developmental Neuropsychology, 15*(3), 327–349.

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research, 1146,* 23–49.

Kuperberg, G. R., Choi, A., Cohn, N., Paczynski, M., & Jackendoff, R. (2010). Electrophysiological correlates of complement coercion. *Journal of Cognitive Neuroscience, 22*(12), 2685–2701.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, 62,* 621–647.

Lai, V. T., Curran, T., & Menn, L. (2009). Comprehending conventional and novel metaphors: An ERP study. *Brain Research, 1284,* 145–155.

Lakoff, G. (2008). The neural theory of metaphor. In R. W. Gibbs (Ed.), *The cambridge handbook of metaphor and thought* (pp. 17–38). Cambridge: Cambridge University Press.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

LARRC. (2015). The dimensionality of language ability in young children. *Child Development, 86,* 1948–1965.

Laval, V., & Bernicot, J. (2002). Tu es dans la lune: Understanding idioms in French speaking children and adults. *Pragmatics, 12,* 399–413.

Levinson, S. (1983). *Pragmatics*. Cambridge: Cambridge University Press.

Levorato, M. C., & Cacciari, C. (1995). The effects of different tasks on the comprehension and production of idioms in children. *Journal of Experimental Child Psychology, 60*(2), 261–283.

Milburn, E. (2018). *The effects of meaning dominance and meaning relatedness on ambiguity resolution: Idioms and ambiguous words*. Doctoral Dissertation, University of Pittsburgh.

Milburn, E., & Warren, T. (under review). Idioms show effects of meaning relatedness and dominance similar to those seen for ambiguous words. *Psychonomic Bulletin & Review*.

Milburn, E., Warren, T., & Dickey, M. W. (2018). Idiom comprehension in aphasia: Literal interference and abstract representation. *Journal of Neurolinguistics, 47,* 16–36.

Mitchell, P., Saltmarsh, R., & Russell, H. (1997). Overly literal interpretations of speech in autism: Understanding that messages arise from minds. *Journal of Child Psychology and Psychiatry, 38*(6), 685–691.

Nayak, N. P., & Gibbs, R. W. (1990). Conceptual knowledge in the interpretation of idioms. *Journal of Experimental Psychology: General, 119*(3), 315.

Nieuwland, M. S., & Van Berkum, J. J. (2005). Testing the limits of the semantic illusion phenomenon: ERPs reveal temporary semantic change deafness in discourse comprehension. *Cognitive Brain Research*, *24*(3), 691–701.

Nippold, M. (1998). *Later language development: The school-age and adolescent years* (2nd ed.). Austin, TX: Pro-Ed.

Nippold, M. A. (2006). Language development in school-age children, adolescents and adults. In K. Brown (Ed.), *Encyclopedia of language and linguistics* (2nd ed., Vol. 6, pp. 368–372). Oxford: Elsevier Publishing.

Nippold, M. A., & Duthie, J. K. (2003). Mental imagery and idiom comprehension: A comparison of school-age children and adults. *Journal of Speech, Language, and Hearing Research, 46*(4), 788–799.

Nooteboom, S. G., Weerman, F., & Wijnen, F. N. K. (Eds.). (2002). *Storage and computation in the language faculty*. Berlin: Springer.

Norbury, C. (2005). The relationship between theory of mind and metaphor: Evidence from children with language impairment and autistic spectrum disorders. *British Journal of Developmental Psychology, 23,* 383–399. https://doi.org/10.1348/026151005x26732.

Nordmann, E., Cleland, A. A., & Bull, R. (2013). Cat got your tongue? Using the tip-of-the-tongue state to investigate fixed expressions. *Cognitive Science, 37*(8), 1553–1564.

Nordmann, E., Cleland, A. A., & Bull, R. (2014). Familiarity breeds dissent: Reliability analyses for British–English idioms on measures of familiarity, meaning, literality, and decomposability. *Acta Psychologica, 149,* 87–95.

Noveck, I. A., Bianco, M., & Castry, A. (2001). The costs and benefits of metaphor. *Metaphor and Symbol, 16*(1–2), 109–121.

Nunberg, G., Sag, I. A., & Wasow, T. (1994). Idioms. *Language, 70*(3), 491–538.

Obert, A., Gierski, F., Calmus, A., Portefaix, C., Declercq, C., Pierot, L., et al. (2014). Differential bilateral involvement of the parietal gyrus during predicative metaphor processing: An auditory fMRI study. *Brain and Language, 137,* 112–119.

O'Donnell, T. J. (2015). *Productivity and reuse in language: A theory of linguistic computation and storage*. Cambridge: MIT Press.

Ortony, A., Schallert, D. L., Reynolds, R. E., & Antos, S. J. (1978). Interpreting metaphors and idioms: Some effects of context on comprehension. *Journal of Verbal Learning and Verbal Behavior, 17*(4), 465–477.

Paczynski, M., Jackendoff, R., & Kuperberg, G. (2014). When events change their nature: The neurocognitive mechanisms underlying aspectual coercion. *Journal of Cognitive Neuroscience, 26*(9), 1905–1917.

Partee, B. (1995). Lexical semantics and compositionality. In Lila R. Gleitman & Mark Liberman (Eds.), *Invitation to cognitive science* (Vol. 1, pp. 311–360). Cambridge, MA: MIT Press.

Piantadosi, S. T., Tily, H., & Gibson, E. (2012). The communicative function of ambiguity in language. *Cognition, 122*(3), 280–291.

Piñango, M. M., Winnick, A., Ullah, R., & Zurif, E. (2006). Time-course of semantic composition: The case of aspectual coercion. *Journal of Psycholinguistic Research, 35*(3), 233–244.

Pinker, S., & Prince, A. (1994). Regular and irregular morphology and the psychological status of rules of grammar. *The Reality of Linguistic Rules*, *321*, 51.

Pirelli, V., Plag, I., & Dressler, W. U. (Eds.). (2019). *Word knowledge and word usage: A cross-disciplinary guide to the mental lexicon*. Berlin, New York: De Gruyter.

Pylkkänen, L., & McElree, B. (2006). *The syntax-semantics interface: On-line composition of sentence meaning. Handbook of psycholinguistics* (2nd ed., pp. 539–579). Amsterdam: Elseiver.

Pylkkänen, L., & McElree, B. (2007). An MEG study of silent meaning. *Journal of Cognitive Neuroscience, 19*(11), 1905–1921.

Pynte, J., Besson, M., Robichon, F. H., & Poli, J. (1996). The time-course of metaphor comprehension: An event-related potential study. *Brain and Language, 55*(3), 293–316.

Qualls, C. D., O'Brien, R. M., Blood, G. W., & Hammer, C. S. (2003). Contextual variation, familiarity, academic literacy, and rural adolescents' idiom knowledge. *Language, Speech, and Hearing Services in Schools, 34*(1), 69–79.

Ramberg, C., Ehlers, S., Nydén, A., Johansson, M., & Gillberg, C. (1996). Language and pragmatic functions in school-age children on the autism spectrum. *European Journal of Disorders of Communication, 31*(4), 387–413.

Rapp, A. M., Mutschler, D. E., & Erb, M. (2012). Where in the brain is nonliteral language? A coordinate-based meta-analysis of functional magnetic resonance imaging studies. *Neuroimage, 63*(1), 600–610.

Rinaldi, M. C., Marangolo, P., & Baldassarri, F. (2004). Metaphor comprehension in right brain-damaged patients with visuo-verbal and verbal material: A dissociation (re) considered. *Cortex, 40*(3), 479–490.

Rommers, J., Dijkstra, T., & Bastiaansen, M. (2013). Context-dependent semantic processing in the human brain: Evidence from idiom comprehension. *Journal of Cognitive Neuroscience, 25*(5), 762–776.

Saban-Bezalel, R., & Mashal, N. (2017). Comprehension and hemispheric processing of irony in schizophrenia. *Frontiers in Psychology, 8,* 943.

Schneider, S., Rapp, A. M., Haeußinger, F. B., Ernst, L. H., Hamm, F., Fallgatter, A. J., et al. (2014). Beyond the N400: Complementary access to early neural correlates of novel metaphor comprehension using combined electrophysiological and haemodynamic measurements. *Cortex, 53,* 45–59.

Schumacher, P. B. (2011). The hepatitis called electrophysiological evidence. *Experimental Pragmatics/Semantics, 175,* 199.

Schumacher, P. B. (2014). Content and context in incremental processing: "The ham sandwich" revisited. *Philosophical Studies, 168*(1), 151–165.

Schweigert, W. A. (1986). The comprehension of familiar and less familiar idioms. *Journal of Psycholinguistic Research, 15*(1), 33–45.

Searle, J. (1993). Metaphor. In A. Ortony (Ed.), *Metaphor and thought* (pp. 83–111). Cambridge: Cambridge University Press.

Siyanova-Chanturia, A., Conklin, K., & Van Heuven, W. J. (2011). Seeing a phrase "time and again" matters: The role of phrasal frequency in the processing of multiword sequences. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 37*(3), 776.

Smolka, E., Rabanus, S., & Rösler, F. (2007). Processing verbs in German idioms: Evidence against the configuration hypothesis. *Metaphor and Symbol, 22*(3), 213–231.

Sperber, D., & Wilson, D. (2008). A deflationary account of metaphors. *The Cambridge handbook of metaphor and thought, 18,* 84–105.

Sprenger, S. A., Levelt, W. J., & Kempen, G. (2006). Lexical access during the production of idiomatic phrases. *Journal of Memory and Language, 54*(2), 161–184.

Swinney, D. A., & Cutler, A. (1979). The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behavior, 18*(5), 523–534.

Tabossi, P., Fanari, R., & Wolf, K. (2009). Why are idioms recognized fast? *Memory & Cognition, 37*(4), 529–540.

Tager-Flusberg, H. (2006). Defining language phenotypes in autism. *Clinical Neuroscience Research, 6*(3–4), 219–224.

Thibodeau, P., & Durgin, F. H. (2008). Productive figurative communication: Conventional metaphors facilitate the comprehension of related novel metaphors. *Journal of Memory and Language, 58*(2), 521–540.

Thibodeau, P. H., & Durgin, F. H. (2011). Metaphor aptness and conventionality: A processing fluency account. *Metaphor and Symbol, 26*(3), 206–226.

Thibodeau, P. H., Sikos, L., & Durgin, F. H. (2017). Are subjective ratings of metaphors a red herring? The big two dimensions of metaphoric sentences. *Behavior Research Methods, 50*(2), 759–772.

Thoma, P., & Daum, I. (2006). Neurocognitive mechanisms of figurative language processing—Evidence from clinical dysfunctions. *Neuroscience and Biobehavioral Reviews, 30*(8), 1182–1205.

Titone, D. A., & Connine, C. M. (1999). On the compositional and non-compositional nature of idiomatic expressions. *Journal of Pragmatics, 31*(12), 1655–1674.

Titone, D., & Libben, M. (2014). Time-dependent effects of decomposability, familiarity and literal plausibility on idiom priming: A cross-modal priming investigation. *The Mental Lexicon, 9*(3), 473–496.

Tourangeau, R., & Sternberg, R. J. (1982). Understanding and appreciating metaphors. *Cognition*, *11*(3), 203–244.

Traxler, M. J., Pickering, M. J., & McElree, B. (2002). Coercion in sentence processing: Evidence from eye-movements and self-paced reading. *Journal of Memory and Language, 47*(4), 530–547.

Tremblay, A., Derwing, B., Libben, G., & Westbury, C. (2011). Processing advantages of lexical bundles: Evidence from self-paced reading and sentence recall tasks. *Language Learning, 61*(2), 569–613.

Vega-Moreno, R. E. (2001). Representing and processing idioms. *UCL Working Papers in Linguistics, 13,* 73–107.

Volden, J., & Phillips, L. (2010). Measuring pragmatic language in speakers with autism spectrum disorders: Comparing the children's communication checklist—2 and the test of pragmatic language. *American Journal of Speech-Language Pathology, 19*(3), 204–212.

Vulchanova, M., Saldaña, D., Chahboun, S., & Vulchanov, V. (2015). Figurative language processing in atypical populations: The ASD perspective. *Frontiers in Human Neuroscience, 9,* 24.

Vulchanova, M., & Vulchanov, V. (2018). *Figurative language processing in autism. The interactive mind: Language, vision and attention* (pp. 67–76). Basingstoke: Macmillan.

Vulchanova, M., Vulchanov, V., & Stankova, M. (2011). Idiom comprehension in the first language: A developmental study. *Vigo International Journal of Applied Linguistics, 8,* 206–234.

Westbury, C., & Titone, D. (2011). Idiom literality judgments in younger and older adults: Age-related effects in resolving semantic interference. *Psychology and Aging, 26*(2), 467.

Whalen, J. M., Pexman, P. M., & Gill, A. J. (2009). "Should be fun—Not!" Incidence and marking of nonliteral language in e-mail. *Journal of Language and Social Psychology, 28*(3), 263–280.

Winner, E., & Gardner, H. (1977). The comprehension of metaphor in brain-damaged patients. *Brain, 100*(4), 717–729.

Wittenberg, E., Paczynski, M., Wiese, H., Jackendoff, R., & Kuperberg, G. (2014). The difference between "giving a rose" and "giving a kiss": Sustained neural activity to the light verb construction. *Journal of Memory and Language, 73,* 31–42.

Milburn, E., Vulchanov, V., Baggio, G., Eigsti, I-M., Magnuson, J & Vulchanova, M. (in progress). Idioms in the L1 and the L2: Speaker sensitivities and expectations.

Zhang, H., Yang, Y., Gu, J., & Ji, F. (2013). ERP correlates of compositionality in Chinese idiom comprehension. *Journal of Neurolinguistics, 26*(1), 89–112.