Psychonomic Bulletin & Review (2022) 29:253–261
https://doi.org/10.3758/s13423-021-01985-y

BRIEF REPORT

Can you mend a broken heart? Awakening conventional metaphors in the maze

Laura Pissani[1] · **Roberto G. de Almeida**[1]

Accepted: 13 July 2021 / Published online: 2 August 2021

Abstract

Conventional metaphors such as *broken heart* are interpreted rather fast and efficiently. This is because they might be stored as lexicalized, noncompositional expressions. If so, they require sense retrieval rather than sense creation. But can their literal meanings be recovered or "awakened"? We examined whether the literal meaning of a conventional metaphor could be triggered by a later cue. In a maze task, participants ($N = 40$) read sentences word by word (e.g., *John is an early bird so he can . . .*) and were presented with a two-word choice. Participants took longer and were less accurate when the correct word (*attend*) was paired with a literally-related distractor (*fly*) rather than an unrelated one (*cry*). This suggests that the literal meaning of a conventional metaphor is not circumvented, nor that metaphors simply involve sense retrieval. The metaphor awakening effect suggests that the mechanisms employed to process conventional metaphors are dynamic with both metaphorical sense and literal meaning being available.

Keywords Metaphor · Conventionality · Semantics-pragmatics interface · Language comprehension · The metaphor awakening effect · Maze task

In the song by the Bee Gees, "How Can You Mend a Broken Heart?," the lyrics invite us to attend to the literal sense of *broken heart*—or, instead, to *mend* it metaphorically. Either way the verse goes, it leads us to awaken the literal meaning of a conventional metaphor. Metaphorical expressions pepper our language use, and their literal interpretations appear to go unnoticed because they are usually linked to a conventional metaphorical—thus nonliteral—content. But how are these expressions represented and processed in real time? Are they represented in lexicalized form—thus linked to a concept that is noncompositional—that is, one that is not a function of the meanings of its constituent elements and how they combine? If they are represented as independent forms, can their literal meanings be recovered or "awakened" as the Bee Gees's verse invite us to do? If so, what does this process reveal about metaphorical processing? We were interested in understanding what conventional metaphor processing can tell us about the interface between arguably two different levels of comprehension: semantic (based on what is actually said) and pragmatic (what is intended by the speaker in a particular context).

Recent research has focused on how we calculate the content that a metaphor carries—whether *directly*, relying on what has been stored based on use, or *indirectly*, via a literal form of interpretation. While *direct* interpretation may rely primarily on the retrieval of stored representations (Gibbs, 1994; Gibbs & Colston, 2012), the *indirect* form relies on a proposition that is faithful to the literal interpretation—which,

Statement of relevance
Conventional metaphors such as *broken heart* are ubiquitous in language use and appear to be understood rather rapidly and efficiently. Metaphor theories have postulated that conventional metaphors are accessed directly, involving "sense retrieval" like idioms or other lexicalized expressions. In our study, we report a new phenomenon. The metaphorical content of a conventional expression is quickly accessed, but its literal meaning lingers and can be "awakened" by subsequent context. In a maze task, participants took longer and were less accurate selecting a target word (*attend*) to continue a sentence (*John is an early bird so he can . . .*) when the target was paired with a literally-related (*fly*) rather than an unrelated distractor (*cry*). We rely on these data to suggest that the literal meaning of a conventional metaphor is initially accessed and may cause the conventional metaphor to be reinterpreted in real time.

✉ Laura Pissani
laura.pissani@concordia.ca

Roberto G. de Almeida
roberto.dealmeida@concordia.ca

[1] Department of Psychology, Concordia University, 7141 Sherbrooke Street West, Montreal, QC H4B 1R6, Canada

when deemed false, triggers the search for an alternate meaning in line with context and as an *approximation* to the speaker's intentions (Grice, 1975; Searle, 1979).

A key issue in the dispute between direct and indirect theories is the conventionality of a metaphor. In the case of copular metaphors—those with the form *X is Y*—such as *My lawyer is a shark*, conventionality is taken to refer to the degree at which the vehicle *shark* is used to refer to a particular alternative sense (e.g., *aggressive*, *shrewd*), predicating this sense to the topic *lawyer*. The conventionality of *shark*, in this case, comes from its frequent use contributing that particular alternative sense to numerous expressions (*the banker is a shark*, *your friend is a shark*, etc.). Conventionality here is then distinct from but interacts with familiarity, which refers to how well-known a full expression such as *My lawyer is a shark* is (Bowdle & Gentner, 2005; Roncero & de Almeida, 2015).

Conversely, the conventionality of metaphors of the form *adjective–noun* can come from the modifier, the head, or both. For instance, in expressions such as *bright student* only the modifier *bright* is metaphorical, and it is conventionally used to refer to someone or something smart—such as in *bright idea*, *bright scientist*. In some expressions, however, only the head, but not the modifier, can be metaphorical: an *old flame* refers to an ex-lover, where *old* maintains its literal meaning as "former", while the head *flame* is used metaphorically to refer to a lover. In other expressions, both terms are used together to form a metaphor, as in the case of *broken heart*, where both *broken* and *heart* have productive, metaphorical uses. Here, we refer to all these *adjective–noun* types of metaphors as conventional, even when only one of the constituents is used metaphorically. We do so because metaphoricity comes from the combination of both terms, which are conventionally used together to refer to something figuratively.

Conventionality constitutes an important variable for understanding the nature of meaning processing because though conventional metaphors are blatantly false, like all other forms of metaphors, qua linguistic expressions, they are usually associated with a particular alternative content. Conventional metaphors, in this regard, are in stark contrast with novel metaphors, whose metaphorical contents are indeterminate and need to be calculated by the listener or reader. This contrast has been termed "stored" versus "fresh" (Morgan, 1979) or "lexicalized" versus "novel" (Blank, 1988), which Bowdle and Gentner (2005, p. 199) claim "invite sense retrieval" (conventional) in opposition to "sense creation" (novel). As Morgan (1979) put it, "recognizing the phrase [. . . *is a pig*], one knows immediately what is intended. It is an institutionalized metaphor, and knowledge that the phrase 'a pig' is used this way short circuits the process of figuring out the metaphor from literal meaning" (p. 129). Explicit in this contrast is the idea that there are two processes involved in understanding metaphors, depending on whether the metaphor is novel or conventional, with conventional ones requiring the retrieval of its associated content.

We examined whether conventional metaphor processing in fact bypasses or "short circuits" the literal meaning. In particular, we investigated whether or not the literal meaning could be recovered or "awakened" by a later cue. This shift in interpretation—from metaphorical to literal—to our knowledge has been investigated only by Goldstein et al. (2012), but with a radically different method. They induced thoughtful semantic processing of two-word metaphors by asking participants to make a semantic judgment and to interpret a subset of the materials. Event-related potential (ERPs) were recorded during a later task in which the subset of conventional metaphors that had been explained in terms of their literal meaning elicited a higher N400 component and a lower late positive component in contrast to the subset that had been left unexplained. According to Goldstein et al., this pattern suggests that explaining a conventional metaphor might create new structural relations between the two words. Although these results might indicate a shift in processing mode, this shift happened over time—from the exposure to the test phase—and was triggered by asking participants to explain a conventional metaphor rather than relying on moment-by-moment, real-time processes of interpretation.

We reasoned that if conventional metaphors indeed "bypass" the literal interpretation but can be "shifted" back to the literal meaning by subsequent cues, this literal meaning might in fact be available during initial processing. Thus, rather than engaging participants in a thinking task, we investigated the potential awakening of the literal meaning of *Adj+N* metaphors by employing a maze task (Forster, 2010; Forster et al., 2009). This task is taken to provide an accurate estimate of the processing cost of a target word in the presence of a given distractor during word-by-word self-paced reading. The maze task ensures that consecutive words in a sentence are integrated with its previous elements, as participants must continuously and rapidly understand unfolding segments of the sentence to be able to make an accurate lexical choice. Hence, the task yields a measure of word-by-word incremental interpretation (Gallant & Libben, 2020), with findings comparable with those obtained in eye-tracking studies (Boyce et al., 2020; Forster et al., 2009).

To our knowledge, no other study has investigated the potential metaphor-to-literal shift in real time. We hypothesized that if conventional metaphors are lexicalized, involving only meaning retrieval, meaning composition should be insulated from a cue that refers to the literal meaning of the metaphor. That is, the proposition that readers form should consist of the metaphorical content obtained by sense retrieval. However, if the literal meaning is available even in cases of conventional metaphors, a subsequent cue might cause an *awakening* effect by making the literal content of a conventional metaphor to be (re-)accessed in real time. Specifically, we predicted that the presence of a word such as *mend* after a metaphor such as *broken heart* would engender processing costs compatible with the hypothesis that the literal meaning of *broken heart* has been awakened.

Method

Participants

Participants were 40 native speakers of English between the ages of 18 and 47 years (M = 22.20, SD = 4.91; 39 females) with normal or corrected-to-normal vision. The sample size was based on similar studies that employed the same technique and investigated related phenomena (e.g., Gallant & Libben, 2020; Witzel & Forster, 2014). Participants were recruited via the Concordia University online participant pool and were compensated with course credit.

Materials

Materials consisted of 24 experimental sentences containing highly conventional two-word metaphorical expressions obtained from the combination of adjectives and nouns (e.g., *broken heart, early bird*) plus 48 filler sentences. The metaphorical expressions were chosen based on their familiarity ratings because familiarity is a good predictor of conventionality (Roncero & de Almeida, 2015). In a separate rating study, 10 participants rated 48 metaphors such as *early bird* and 48 other nonmetaphorical expressions (e.g., *red wine, handsome woman*) which were presented in upper case and embedded in simple carrier sentences. Their task was to rate how well they knew the expression on a scale of 1 to 7, with 7 being the most familiar. We then selected 24 metaphor combinations that were rated above 4 (M = 6.36, SD = 0.77). Based on those expressions, 24 sentences were created for the maze task (e.g., *John is an early bird so he can attend morning classes*; see Appendix). For each sentence, we selected two types of distractor words, which were to be entered as choices in the maze portion of the task: (1) a related distractor (e.g., *fly*), which was semantically associated with the literal meaning of the metaphor combination, and (2) an unrelated distractor (e.g., *cry*), which was not semantically associated with the literal meaning of the metaphor combination. Both types of distractors were matched to each other in overall frequency, length, and syntactic category based on data from the MRC database (Coltheart, 1981). Both distractors were grammatically equivalent, but none of them was semantically appropriate to continue the sentence.

Based on the two conditions, two mixed lists were created so that participants would see only one version of each sentence, but both distractor conditions. Thus, participants were presented with half of the items in the related-distractor condition, while the other half of the items was presented in the unrelated-distractor condition. In addition, for each sentence, the position of the distractor was counterbalanced for each list, so that half of the participants read the distractor on the right, and the other half read the same distractor on the left (e.g., *attend/fly* or *fly/attend*).

Procedure

The experiment was programmed using PsychoPy3 (Version 3.2.4; Peirce & MacAskill, 2018). Participants were tested via the Pavlovia online platform (pavlovia.org), which allows data collection remotely with high response time (RT) accuracy (Bridges et al., 2020; Grootswagers, 2020). Participants were randomly assigned to one of two lists when they registered for the experiment. After receiving a web link to access the experiment, participants were directed to a virtual consent form, which was followed by a demographics form. During the experiment, participants were presented with sentences word by word, in a self-paced manner (see Fig. 1). At any point during the sentence (one to three words after the metaphorical expression; M = 1.96, SD = 0.91), they were presented with two words to choose from so that the sentence could continue. Only one word was semantically appropriate to continue the sentence. Participants were required to select the appropriate word by pressing the left or right arrow according to the chosen word's position on the screen. In about 25% of the trials, a comprehension question was presented. Participants were instructed to answer the question by pressing the Y key for YES and the N key to for NO to answer the question. It was emphasized that they should move as quickly as possible through the sentence and make their decision as fast and as accurately as possible because their time was being recorded. Participants were presented with eight practice trials before the experimental phase (24 experimental trials + 48 fillers, randomly distributed). The experimental session lasted approximately 15 minutes.

Data preparation

Analyses of RT and accuracy were restricted to participants who achieved 75% or higher in both the lexical choice and comprehension question tasks. Overall accuracy scores were calculated for each participant. All participants achieved 89% or higher in the lexical choice task and 83% or higher when responding to the comprehension question. Based on these criteria, all 40 participants were included in the analysis. Incorrect trials (i.e., those where participants either selected the incorrect word or did not respond correctly to the probe question) were removed from the RT analyses only (12% of all data points).

All RTs above 3 seconds were trimmed prior to data analysis (3% of all data points). Further, outliers were defined as more than 2.5 standard deviations from the mean and calculated per participant to preserve individual variability. Outliers

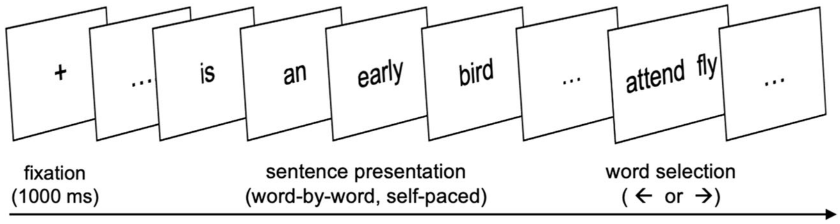

Fig. 1 Each trial in the maze task started with a 1,000-ms fixation cross, followed by the first word of the sentence. Participants were instructed to press the up arrow on the keyboard to indicate that they had read the word and move on to the next word in the sentence. The metaphor (*early bird*) was followed by one to three words (*so he can . . .*), followed by the word selection task—a maze juncture—consisting of the target (*attend*) presented together with a distractor that was semantically related (*fly*) to the literal meaning of the metaphor or an unrelated control (*cry*). Participants had to select which word would continue the sentence by pressing the corresponding left or right arrow

were replaced with 2.5 standard deviations above the mean for the upper tail (2% of all data points).

Results

Response times

We used R (R Core Team, 2012) and *lme4* (Bates et al., 2015) to perform a linear mixed-effects analysis of RTs to word selection. As fixed effects, we entered the type of distractor into our first model. As random effects, we had by-subject and by-item random intercepts. In a second model, we added the distance (in number of words) between the metaphor combination and the lexical choice as a fixed effect. We obtained *p* values by likelihood ratio tests of the full model against the null model including only random effects (subject and item). Our first model was a significantly better fit to the data than the null model, $\chi^2(1) = 17$, $p < .001$. The type of distractor affected the response time, increasing it by 117 ms, 95% CI [61.44, 171.77]. Moreover, the model with distance as a second predictor was not significantly better than the first model. There was no main effect, $\chi^2(1) = 1.90$, $p = .16$, or interaction, $\chi^2(1) = 0.39$, $p = .53$. Thus, the distance between the metaphorical expression and the maze juncture, which ranged between one and three words, did not affect RTs. Visual inspection of residual plots did not reveal any obvious deviations from homoscedasticity or normality.

As predicted, participants took longer to select the correct word (*attend*) when it was paired with the literally-related distractor (*fly*; M = 1,396 ms, SD = 188) rather than the literally-unrelated one (*cry*; M = 1,274 ms, SD = 161), as shown in Fig. 2a.

Additionally, due to the variation of position of the metaphorical word (i.e., first, second, or both) within the two-word combinations, we included position as a factor in our model. Results indicated no significant main effect, $\chi^2(2) = 0.59$, p = .74, or interaction, $\chi^2(2) = 1.75$, $p = .41$. Thus, the position of the metaphorical word, whether it was the first word (e.g., *warm welcome*), the second word (e.g., *early bird*), or both words (e.g., *red flag*), did not affect response times.

Accuracy

For the accuracy analyses, we used the *glm* function (R Core Team, 2012) to perform logistic regression by modeling the probability of observing a correct word selection as a function of the type of distractor. Results indicated a reliable effect of type of distractor (logit difference: +1.96, $SE = 0.29$, $z = 6.78$, $p < .001$). We then calculated the probabilities of the log odds, which predicted that the probability of observing a correct word selection was 81% when the correct answer (*attend*) was paired with a literally-related distractor (*fly*), while the probability of observing a correct word selection increased to 97% when the correct answer (*attend*) was paired with a literally-unrelated distractor (*cry*).

As predicted, participants were less accurate when selecting the correct word (*attend*) when it was paired with the literally-related distractor (*fly*; $M = 15.1$, $SD = 4.02$) rather than the unrelated one (*cry*; $M = 18.6$, $SD = 1.64$).

Meaningfulness ratings

After conducting our main experiment, we conducted a meaningfulness rating study to ensure that the correct answer (*attend*) had a higher chance of being selected than both distractors (*fly* and *cry*) and, further, to determine whether both distractors were equally anomalous to continue the sentence. For this task, we recruited 21 additional participants from the Concordia community (M = 23.86, SD = 3.8; 17 females), none of whom had participated in the main experiment. The study was conducted online, using the Pavlovia (pavlovia.org) platform, with all 24 experimental sentences and 24 fillers. Each trial consisted of an individual sentence segment presented in isolation (e.g., *John / is an / early / bird / so he can / attend / morning classes*). Participants were instructed to determine how meaningful each segment was as a continuation of the sentence by pressing a number between 1 (*very bad*) and 5 (*very good*) on the computer keyboard. Participants were encouraged to use the full scale with examples of what would constitute bad, good, and not-so-good/bad continuations. After each segment rating, participants were immediately presented with the next

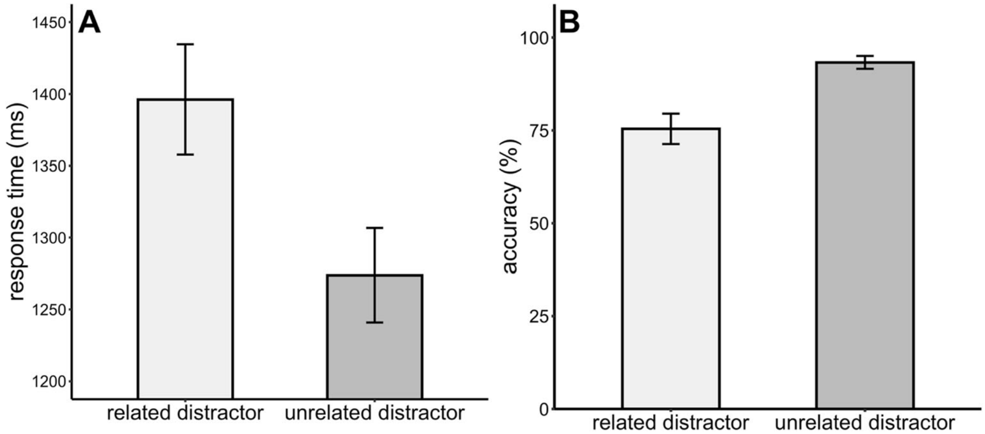

Fig. 2 **a** Response time to the correct word as a function of the related and unrelated distractors. **b** Accuracy selecting the correct word as a function of the related and unrelated distractors. Error bars are standard errors

segment, and so on until the end of the sentence. We avoided presenting syntactic units (e.g., *an early bird*) in order to obtain ratings for individual constituents of the metaphor expression and, more importantly, for the segment that contained either the correct maze choice (*attend*), the related distractor (*fly*), or the unrelated distractor (*cry*). Three lists were created, each one containing eight items of each kind (correct, related, or unrelated), so that participants only saw one of the versions of the same sentence.

We used R (R Core Team, 2012) and *lme4* (Bates et al., 2015) to perform a linear mixed-effects analysis of meaningfulness ratings, taking into account ratings provided at the key maze choice word. As fixed effects, we entered the type of answer into our full model. As random effects, we had by-subject and by-item random intercepts. We obtained *p* values by likelihood ratio tests of the full model against the null model including only random effects (subject and item). The full model was compared with a null model consisting of only random effects and was found to provide a statistically significant better fit to the data, $\chi 2(2) = 160.15$, $p < .001$, which also indicated an overall significant main effect of answer type. Planned comparisons between answer types showed that correct answers yielded significantly higher mean meaningfulness ratings than both related, $t(462) = 7.38$, $p < .001$, and unrelated distractors, $t(462) = 13.81$, $p < .001$. That is, regardless of whether the correct answer was paired with a related or an unrelated distractor, the correct answer was always the most meaningful (correct: $M = 4.36$, $SD = 0.93$; related: $M = 3.39$, $SD = 1.54$; unrelated: $M = 2.54$, $SD = 1.47$). In addition, the related distractor yielded significantly higher mean meaningfulness ratings than the unrelated distractor, $t(462) = 6.41$, $p < .001$. Therefore, to guarantee that both related and unrelated distractors were equally anomalous as maze choices, we removed seven items where the difference in mean meaningfulness ratings between the related and unrelated distractors was the highest. Afterwards, planned comparisons to the new data set including 17 items indicated that correct answers remained significantly higher than both related and unrelated distractors (correct: $M = 4.30$, $SD = 0.97$; related: $M = 2.99$, $SD = 1.58$; unrelated: $M = 2.96$, $SD = 1.49$), while meaningfulness ratings for both related and unrelated distractors were not significantly different, $t(322) = 2.09$, $p = .09$.

Based on the results of this meaningfulness rating task, we conducted further RT and accuracy analyses restricted to the 17 items for which the correct answer was more meaningful than both related and unrelated distractors, while related and unrelated distractors were rated equally anomalous. The rating scores for all materials as well as the items removed for the purposes of reanalyses are shown in the Appendix.

RT reanalyses

For response time re-analyses, we also used *lme*4 (Bates et al., 2015) to perform a linear mixed-effects analysis of RTs to word selection, using the same data-analytic procedures employed with the full data set. As in our main analyses, the full model was a significantly better fit to the data than the null model, $\chi^2(1) = 14.07$, $p < .001$. The type of distractor affected the response time, increasing it by 122 ms, 95% CI [58.53, 185.31]. Further, the model with distance as a second predictor was not significantly better than the first model. That is, there was no main effect. $\chi^2(1) = 2.28$, $p = .13$, or interaction, $\chi^2(1) = 1.09$, $p = .29$.

Similar to our results with the full data set, participants took longer to select the correct word (*attend*) when it was paired with the literally-related distractor (*fly*; $M = 1397$ ms, $SD = 186$) rather than the unrelated one (*cry*; $M = 1264$ ms, $SD = 168$). With regards to the variation in position of the metaphorical word within the two-word expression (first, second, or both), again results indicated no significant main effect, $\chi^2(2) = 2.15$, $p = .34$, or interaction, $\chi^2(2) = 0.35$, $p = .83$.

Accuracy reanalyses

For the accuracy reanalyses, we again used the *glm* function (R Core Team, 2012) to perform logistic regression by modeling the probability of observing a correct word selection as a function of the type of distractor. Results, again, indicated a reliable effect of type of distractor (logit difference: +1.73, $SE = 0.34$, $z = 5.06$, $p < .001$). We then calculated the probabilities of the log odds, which predicted that the probability of observing a correct word selection was 84% when the correct answer (*attend*) was paired with a related distractor (*fly*), while the probability of observing a correct word selection increased to 97% when the correct answer (*attend*) was paired with an unrelated distractor (*cry*). As obtained in the main analyses with the full data set, participants were less accurate when selecting the correct word (*attend*) when it was paired with the literally-related distractor (*fly*; $M = 15.9$, $SD = 4.09$) rather than the unrelated one (*cry;* $M = 18.6$, $SD = 1.66$).

General discussion

We employed a maze task to examine whether the literal meaning of a conventional metaphor could be triggered or “awakened” by a subsequent cue. This would indicate that the literal interpretation was available during conventional metaphor processing. Results support our awakening hypothesis, showing a significant increase in RTs to the correct alternative when it was paired with a literally-related distractor (*fly*) rather than a literally-unrelated one (*cry*). Furthermore, accuracy decreased significantly in the related condition in contrast to the unrelated one. In a maze task, an increase in response time and a decrease in accuracy are indicators of a higher processing cost to a target word as a function of the type of distractor. As Gallant and Libben (2020) suggested, the assumption is that at the maze juncture “it is impossible for the participant to suppress the activation of the distractor word or the consequences of its activation” (p. 7).

It is important to note that the maze task forces semantic composition, with each word’s meaning being integrated into the ongoing proposition. At the time the juncture is presented, the two-word metaphor (*early bird*) has already been interpreted according to its conventional content, which is triggered only by the individual content of the lexical items that are accessed. That is, in order to obtain the pragmatic content *EARLY RISER* from the conventional metaphor *early bird*, the two concepts *EARLY* and *RISER* need to be composed. Hence, the awakening effect cannot be accounted for by the “activation” of lexical items through simple association (e.g., *bird* → *fly*). Such associations are determined over the words’ conceptual representations—thus, they are established via the *content* that each word yields. Figure 3 depicts the model we propose for the awakening effect. Crucially, our proposal is that the concepts obtained by lexical

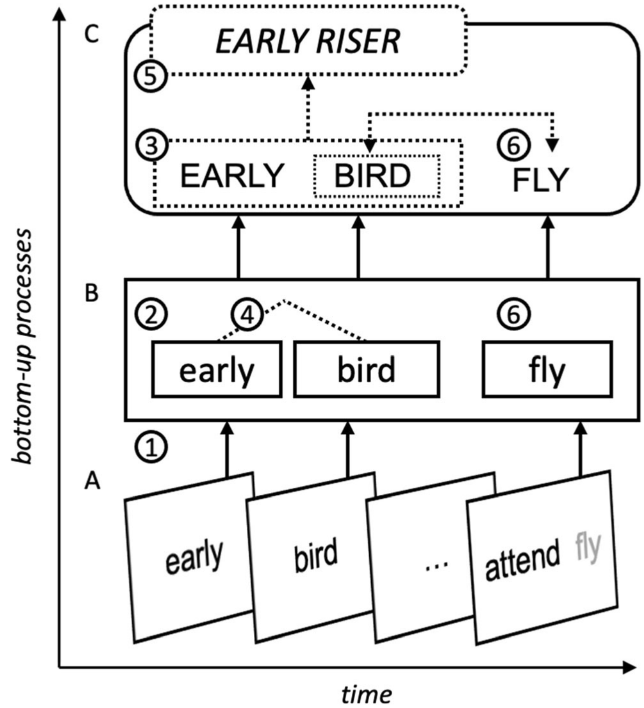

Fig. 3 The “awakening” effect. The model entails two main levels of representation and processing after (**a**) lexical input: one for (**b**) word recognition and linguistic-structural operations, and one for (**c**) conceptual processes. During incremental interpretation, (1) input tokens (2) are recognized (lowercase words) and their conceptual representations (uppercase) (3) are accessed and incrementally composed into propositions. Then, the next incoming token items (4) are combined via linguistic operations, such as phrase structuring. In turn, this triggers a (5) stored, conventional representation associated with the full phrase interpretation. At the same time, the concepts that are accessed during lexical input also trigger (6) associated concepts. Finally, the relation among the related concepts is enhanced by the later cue, thus “awakening” the concepts accessed during literal interpretation. We suggest that a proposition related to the “lower” literal interpretation is available together with one built with the “higher” pragmatic information

access may also form a proposition that is faithful to the input expression with the cue (e.g., *fly*) working to enhance the literal proposition built out of the lexical denotations of the token sentences.[1]

Several other studies have suggested that competing interpretations for sentences linger in working memory (see Christianson et al., 2001; Reyna et al., 2016). Data from Christianson et al. (2001), for instance, suggest that, for temporary ambiguous sentences such as *While the man hunted the deer ran into the woods*, subjects retain two propositions, a false one, compatible with the idea that [*the man hunted the deer*], which is derived from misparsing the sentence, and a true one, compatible with the idea that [*the deer ran into the*

[1] While we have assumed that lexical concepts engage in classical composition (see, e.g., de Almeida, 2018, for discussion), our results and model (Fig. 3) are also compatible with different forms of compositionality (see, e.g., Pelletier, 2017, for a wide range of views on the nature of compositionality).

woods]. However, if only *one* proposition about the sentence was constructed during incremental interpretation, then these two propositions would be mutually exclusive.[2] In the present case, we suggest that both literal and metaphorical interpretations linger, with the less salient literal content remaining a viable—if a dormant one—interpretation even for conventionalized metaphors. Although our study does not provide evidence for the nature of these interpretations—whether they are full propositions or only partial ones—our suggestion is that both literal and metaphorical interpretations are available simultaneously, compatible with theories of propositional memory representation that postulate "true" and "false" memory traces for sentences and events (see Reyna et al., 2016, for a review).

It could be argued, however, that the literally-related distractor *fly* may be working as an extension of the metaphor *early bird.* According to the conceptual metaphor theory (Lakoff & Johnson, 1980), conceptual metaphors (e.g., *LOVE IS A JOURNEY*) govern our thought and elicit multiple metaphorical expressions (e.g., *we are going too fast, it has been a bumpy road, this relationship is not going anywhere*). Thus, when arriving at the distractor *fly*, readers might assume that it is a continuation of the metaphor, which in turn might delay their response. Crucially, our experiment measures online responses, which do not allow the participant to reflect on the relationship between *early bird* and *fly*. Yet if the reader took *fly* as a metaphorical extension, that would entail that the literal meaning of *early bird* was being accessed in real time to link *fly* and *early bird*, which indeed supports our hypothesis. Conversely, if the reader took *fly* literally and still longer RTs were obtained, it would follow that the meaning of *fly* triggers the literal meaning of *early bird*, which ultimately supports our hypothesis that the literal meaning of conventional metaphors had been initially available.

Our results suggest that the literal meaning of a conventional metaphor is not necessarily "short circuited" (Morgan, 1979), nor that conventional metaphors simply involve the retrieval of an associated meaning. The results rather suggest that the conventional content may take priority but the literal meaning is available and may be awakened by a subsequent cue. This cue might yield a conflict between the metaphorical content and the literal meaning, which may force a reinterpretation of the conventional expression. If so, this reinterpretation may involve a process that is similar to that of a novel metaphor. In this regard, our results are partially in line with those of Goldstein et al. (2012), who found that after explaining conventional metaphors, participants' pattern of activation for these conventional metaphors resembled that of novel metaphors. Crucially, our study demonstrates that a shift from conventional to literal meaning may occur in real time, given a triggering cue. In our study, the distance between the conventional metaphor and the trigger was between one and three words, which is estimated to be in the 250 to 750-ms range (Forster et al., 2009), suggesting that the awakening effect occurs fast and automatically. Our accuracy data also supports this conclusion: although the literally-related distractor fits syntactically, it renders the sentence semantically anomalous. Yet in almost one-fourth of the trials, participants chose the literally-related distractor. These results were obtained with the full data set and with a subset of materials that showed greater contrast between correct choice and the two maze distractors. Overall, this pattern can be further interpreted as a conflict between the dominant metaphorical content and the availability of the literal meaning, which lingers briefly until it is awakened by the literal cue.

The classical dichotomy between direct and indirect metaphor interpretation rests on the assumption that metaphors are either accessed directly or via literal interpretation. The effect we obtained suggests that a different process might be at work. Word constituents quickly give rise to metaphorical content as they attempt to semantically compose, but their literal meanings remain available and can be enhanced by further information in the context. This suggests that literal meanings are always accessed, but their availability may be inhibited by conventional content.

The view we espouse is not far from what other theories propose. For instance, for Bowdle and Gentner (2005), conventional metaphors can access literal meanings but with associated figurative senses having primacy over literal meaning during comprehension. This model indeed proposes that conventional vehicles of copular metaphors (e.g., *shark* in *My lawyer is a shark*) are polysemous, for they "refer both to a literal concept and to an associated metaphorical category" (p. 199)—which meaning (or sense) wins depends on numerous factors "including the context of the metaphor and the relative salience of each meaning of the [vehicle] term" (p, 199).

In addition, our results—and model—are compatible with Giora's (2003) graded salience hypothesis. In her model, if metaphorical content is most salient, it is accessed during comprehension, with salience being determined by conventionality and other stimulus properties. Accordingly, the less salient meaning—possibly the literal meaning—"may not reach sufficient levels of activation to be visible" (p. 11). Beyond salience, the awakening effect might shed light on how literal meaning and pragmatic content might interface in the course of sentence interpretation. It suggests that pragmatic content is quickly computed, without however impeding on underlying processes of semantic (i.e., literal) access and composition. As words compose into propositions, there

[2] As suggested by an anonymous reviewer, our results are compatible with an incremental, constraint-satisfaction model of sentence interpretation, which takes into account several sources of information (e.g., lexical, syntactic, semantic, and even pragmatic) as every new lexical item is processed. However, it is also possible to assume that incremental interpretation is driven by structural principles (e.g., argument structure), largely shielded from pragmatics and world knowledge (see Ferreira & Nye, 2018, for discussion).

should not be a need to reinterpret the content of a conventional metaphor if all that is happening is simply *retrieval* of a conventional content. The awakening cue signals that information about the literal interpretation of a conventional metaphor lingers during interpretation processes. To wit, it suggests that broken hearts can be literally mended.

Appendix

Experimental materials employed in the maze task

Conventional metaphors are in boldface. The words presented in the maze juncture appear in italics, with the two distractors (related, unrelated) in parenthesis next to the target word. Familiarity ratings (scale of 1–7) of the metaphor appear in parenthesis at the end of each sentence. Meaningfulness ratings for the correct, related control, and unrelated control maze choices are presented, in order, within square brackets. Items removed from the reanalyses, based on the meaningfulness rating task, are preceded by an asterisk.

1. * Some people say that a **sharp tongue** can *ruin* (*cut/age*) a friendship. (4.1) [4.7;4.7;3.1]
2. * Most people agree that a **broken heart** can be *difficult* (*stitched/unfolded*) to overcome. (6.9) [4.7;4.4;2.4]
3. John is an **early bird** so he can *attend* (*fly*/*cry*) morning classes. (6.9) [4.4;2.4;2.1]
4. * Daniel got himself a **trophy wife** so he could *brag* (*display*/*stomach*) about her. (6.7) [4.6;4.7;2.1]
5. Mike is a **night owl** and he *hates* (*hunts*/*coins*) you if you don't go to bars. (7) [4.7;4;3.1]
6. Some people advise that when a **red flag** occurs, one should *fix* (*wave*/*smell*) it. (6.6) [4;3.1;2.1]
7. If a groom gets **cold feet**, he should *reflect* (*warm*/*camp*) and reconsider. (6.9) [4.7;4.1;3.3]
8. * Most guys think that dating an **old flame** can *wreck* (*burn*/*swim*) your life. (6) [4.7;4.1;1.9]
9. * Everyone agrees that a **hungry mind** can *learn* (*devour*/*harass*) to read faster. (5.4) [4.1;4;2]
10. * They say that a **blind date** can *be* (*see*/*war*) the best for you. (6.9) [4.6;3.7;1.3]
11. People know that a **foggy memory** can *affect* (*fade*/*ache*) younger and older people. (6.5) [4.6;3.9;2.7]
12. Moms always say that a **bright student** can *boost* (*illuminate*/*masquerade*) a whole class. (6.9) [4;4.1;3.7]
13. The old neighbour has a **bucket list** so he can *write* (*shovel*/*gallop*) his wishes in. (6.9) [4.6;2.7;1.3]
14. Jane said that the **warm welcome** better be *fun* (*hot*/*due*) if it is for her boyfriend. (6.9) [3.1;1.7;3.3]
15. Professionals agree that an **ice breaker** can *entertain* (*freeze*/*inject*) shy people. (6.8) [4.1;2.3;3.6]
16. People believe that a **melting pot** can *create* (*heat*/*ship*) a balanced society. (6.8) [4.3;3.7;2.4]
17. * People agree that a **black sheep** can become a *good* (*white*/*early*) person. (6.7) [4.1;4.7;3]
18. Everyone thinks that a **silver lining** should be *encouraging* (*gold*/*pale*) for victims. (6.6) [4.7;2.9;2.3]
19. Melissa has a **bubbly personality** that is *pleasing* (*sparkling*/*circling*) all her friends. (6.6) [4.4;3;1.7]

20.Mr. Harrison has a **smoking gun** that he will *show* (*shoot*/*paint*) to the attorney. (4.9) [4.7;3.4;3]

21 Erik's father has a **short fuse** that he should *manage* (*enlarge*/*testify*) soon. (6.6) [4.1;2.1;2.9]

22. Sara has **butter fingers** and she will *serve* (*melt*/*tame*) very important guests. (6.5) [4;2.9;3.4]
23. His family thought that it was a **bitter pill** to *swallow* (*sweeten*/*unscrew*) but still necessary. (5) [4.4;2.1;2.3]
24. The housewife has an **iron fist** that is really *intimidating* (*heavy*/*aware*) to everybody. (5.7) [4.1;2.3;1.9]

Acknowledgments This research was supported by a grant from the National Science and Engineering Research Council of Canada (NSERC).

Author note Research for this article was supported in part by a grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) to R. G. de Almeida.

References

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Blank, G. D. (1988). Metaphors in the lexicon. *Metaphor & Symbolic Activity, 3*(1), 21–36.

Bowdle, B. F., & Gentner, D. (2005). The career of metaphor. *Psychological Review*, *112*(1), 193–216.

Boyce, V., Futrell, R., & Levy, R. P. (2020). Maze Made Easy: Better and easier measurement of incremental processing difficulty. *Journal of Memory and Language, 111*, Article 104082. https://doi.org/10.1016/j.jml.2019.104082

Bridges, D., Pitiot, A., MacAskill, M. R., & Peirce, J. W. (2020). The timing mega-study: Comparing a range of experiment generators, both lab-based and online. *PeerJ*, *8*, Article e9414.

Christianson, K., Hollingworth, A., Halliwell, J. F., & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology*, 42(4), 368–407

Coltheart, M. (1981). The MRC psycholinguistic database. *The Quarterly Journal of Experimental Psychology Section A*, *33*(4), 497–505.

de Almeida, R. G. (2018). Composing meaning and thinking. In G. Preyer (Ed.), *Beyond semantics and pragmatics* (pp. 201–229). Oxford University Press.

Ferreira, F., & Nye, J. (2018). The modularity of sentence processing reconsidered. In R. G. de Almeida & L. R. Gleitman (Eds.), *On concepts, modules, and language* (pp. 63–86). Oxford University Press.

Forster, K. I., Guerrera, C., & Elliot, L. (2009). The maze task: Measuring forced incremental sentence processing sentence processing time. *Behavior Research Methods*, *41*(1), 163–171.

Forster, K. (2010). Using a maze task to track lexical and sentence processing. *The Mental Lexicon, 5,* 347–357.

Gallant, J., & Libben, G. (2020). Can the maze task be even more amazing? Adapting the maze task to advance psycholinguistic experimentation. *The Mental Lexicon, 15*(2), 366–38.

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

Gibbs, R. W., Jr., & Colston, H. L. (2012). *Interpreting figurative meaning.* Cambridge University Press. https://doi.org/10.1017/CBO9781139168779

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

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. Morgan (Eds.), *Syntax and semantics*. Academic Press.

Grootswagers, T. (2020). A primer on running human behavioural experiments online. *Behavior Research Methods, 52*, 2283–2286. https://doi.org/10.3758/s13428-020-01395-3

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

Morgan, J. L. (1979). Observations on the pragmatics of metaphor. In A. Ortony (Ed.), *Metaphor and thought* (pp. 136–147). Cambridge University Press.

Peirce, J. W., & MacAskill, M. R. (2018). *Building experiments in PsychoPy.* SAGE.

Pelletier, F. J. (2017). Compositionality and concepts—A perspective from formal semantics and philosophy of language. In J. A. Hampton & Y. Winter (Eds.), *Compositionality and concepts in linguistics and psychology* (pp. 31–94). Springer.

R Core Team. (2012). R: A language and environment for statistical computing [Computer software]. R Foundation for Statistical Computing. http://www.R-project.org/

Reyna, V. F., Corbin, J. C., Weldon, R. B., & Brainerd, C. J. (2016). How fuzzy-trace theory predicts true and false memories for words, sentences, and narratives. *Journal of Applied Research in Memory and Cognition, 5*(1), 1–9.

Roncero, C., & de Almeida, R. G. (2015). Semantic properties, aptness, familiarity, conventionality, and interpretive diversity scores for 84 metaphors and similes. *Behavior Research Methods, 47*(3), 800–812.

Searle, J. (1979). Metaphor. In A. Ortony (Ed.), *Metaphor and thought* (pp. 83–111). Cambridge University Press.

Witzel, J., & Forster, K. (2014). Lexical co-occurrence and ambiguity resolution. *Language, Cognition and Neuroscience, 29*(2), 158–185.