UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Episodic memory demands modulate novel metaphor use during event narration

Permalink

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 43(43)

ISSN

1069-7977

Authors

G. Djokic, Vesna
Shutova, Ekaterina
Dankers, Verna

Publication Date

2021

Peer reviewed

Episodic memory demands modulate novel metaphor use during event narration

Vesna G. Djokic (vesna@imsquared.eu)
Department of Computing, Goldsmiths University of London
ILLC, University of Amsterdam

Verna Dankers (vernadankers@gmail.com)
ILCC, University of Edinburgh

Ekaterina Shutova (e.shutova@uva.nl)
ILLC, University of Amsterdam

Abstract

Metaphor is an important part of everyday thought and language. Although we are often not aware of metaphor in everyday speech, on occasion, a particularly creative or novel use of metaphor will make us pay attention. It has been hypothesized that one of the driving cognitive factors behind the use of novel metaphor is a need to describe a new reality (as opposed to a preexisting reality) that would otherwise be difficult to convey using conventionalized metaphor. To this extent, novel metaphor use in everyday language may be more associated with episodic memory demands in contrast to conventional metaphor that is associated with semantic memory. To test this idea we analyzed novel metaphor use in the Hippocorpus—a corpus of more than 5000 recalled and imagined stories about memorable life events in the first person perspective. In this dataset, recalled events have been shown to rely on episodic memory to a greater extent than descriptions of imagined events (i.e., narrating an event as if it happened to you but not describing an event that actually happened to you), which largely draw on semantic memory. We hypothesized that novel metaphor use during event narration should be modulated by the extent to which language users are able to draw on primary experience to describe events. We found that novel metaphor counts in recalled events were significantly higher than imagined events. Importantly, we found that factors that influence the extent to which language users are able to draw on primary experience during event narration (i.e., openness to experience, similarity to one's own experience, and how memorable or important an event was) modulated novel metaphor use in different ways in imagined compared to recalled events. The work paves the way for using large scale corpora to analyze underlying cognitive processes that modulate metaphorical language use.

Keywords: Conventional Metaphor; Novel Metaphor; Semantic and Episodic Memory; Emotion and Metaphor

Introduction

Metaphor is an important part of everyday thought and communication serving a broad range of functions. However, little is known about the underlying cognitive factors that may modulate everyday metaphorical language use. According to conceptual metaphor theory, conventional metaphors reflect systematic conceptual mappings that have been established over the course of experience and are reflected in everyday language use (Lakoff & Johnson, 1980). While novel metaphor use may also reflect underlying conceptual mappings, it typically needs to involve a creative association of ideas (or use of language) to convey these mappings in a new way. Additionally, novel metaphor use may also involve making entirely new associations between disparate concepts. Critically, it has been hypothesized that language users may draw on novel metaphor to describe a new reality (as opposed to a preexisting reality) when more conventional metaphors are inadequate (Lakoff & Johnson, 1980).

While a great deal of research has been devoted to understanding differences in the comprehension of conventional versus novel metaphor (Gentner & Bowdle, 2005), less attention has been placed on identifying the cognitive factors that drive the production of conventional versus novel metaphor use in everyday contexts. One possibility, in line with the idea that novel metaphor is used primarily to convey a 'new reality' as opposed to 'preexisting reality', is that one cognitive factor associated with novel metaphor use may be the ability to draw to a greater degree on episodic memory in contrast to conventional metaphor that relies mainly on semantic memory. That is, a language user might be more likely to draw on novel metaphor to describe an event from primary experience than when describing an event that they have little to no primary experience to draw upon. Tulving (1972) first proposed a distinction between episodic and semantic memory. Episodic memory is associated with 'temporally dated episodes or events' with an autobiographical reference. In contrast, semantic memory is associated primarily with commonsense knowledge necessary for the use of language.

In order to test this idea we leveraged NLP techniques to investigate the use of novel metaphor in the Hippocorpus. The Hippocorpus consists of a large corpus of recalled and imagined stories about life events (Sap, Horvitz, Choi, Smith, & Pennebaker, 2020). In this dataset, NLP techniques were used to show that recalled events rely to a greater extent on markers of episodic memory in language than descriptions of imagined events (i.e., narrating an event as if it happened to you but not describing an event that actually happened to you), which draw largely on semantic memory (Sap et al., 2020). We hypothesized that novel metaphor use during event narration should be modulated by the extent to which language users are able to draw on primary experience to describe events. To this extent, we further hypothesized that cognitive factors that may impact the extent to which language users are able to draw on primary experience would modulate novel metaphor use in recalled and imagined stories. We analyzed three factors that were collected in the Hippocorpus (1) openness to experience—one of the Big Five personality traits; (2) emotionality or how memorable/important the event described is perceived to be; and (3) similarity or how similar the event

described is to one's own experience (only relevant to imagined events). Openness to experience is one of the Big Five personality traits (Donnellan, Oswald, & Lucas, 2006) and is associated with individuals typically described as creative, cognitively open, introspective, and imaginative. To this extent, openness may closely reflect a language user's ability to creatively use language to describe events drawing on both imagination and episodic memory. On the other hand, how important an event is perceived to be may impact recall of relevant primary experience. Critically, however, we expect that overall novel metaphor use should depend on the ability to more directly access relevant primary experience which should be more closely associated with recalled compared to imagined stories and associated to a greater extent with imagined stories that describe an event highly similar to one's own experience.

Hippocorpus

The Hippocorpus is a dataset of 6,854 English diary-like short stories. This involves pairs of recalled and imagine stories in the first person perspective that are matched for topic (Sap et al., 2020). Specifically, the data was collected in two main stages using MTurk. First, workers were asked to recall memorable/salient events that happened to them. Secondly, a separate set of workers were asked to imagine events using a summary sentence. Importantly, they were instructed to describe the event as if it had happened to them but not describe an experience that actually happened to them. An example of a recalled event containing metaphor is shown in (1). An example of a summary of the above recalled story is shown in (2). Lastly, an example of an imagined story containing metaphor and using the above summary as a topic is shown in (3).

1. **Recalled:** "Five months ago, my niece and nephew were born. They are my sister's first children, and I was so excited when she announced she was pregnant and I would be an aunt... They were both so perfect an tiny. *I was like holding tiny dolls...*"

2. **Summary:** My sister gave birth to my twin niece and nephew. I visited them in the hospital and held them just after they were born. They were premature and tiny, but strong and healthy.

3. **Imagined:** "...When I arrived my sister had just delivered and I just was in awe. Even though they were a bit small they were mighty. *Seeing the tears of joy my sister had was the most warming thing...*"

The Hippocorpus is accompanied by data on variables such as openness to experience, emotionality or how memorable or important the event was, and similarity to one's own experience (relevant mainly to imagined events), among other demographic information (Sap et al., 2020). Openness to experience was assessed using Mini-IPIP personality questionnaire (Donnellan et al., 2006). Emotionality and similarity were assessed using five-point Likert scales.

Automatic Metaphor Identification

To automatically identify metaphors in the Hippocorpus, we extract predictions from a neural network pre-trained on a corpus annotated for metaphor use at word level. State-of-the-art metaphor identification methods cast the task in a sequence labelling paradigm and fine-tune a pre-trained language model on sentences from a metaphor corpus, predicting a binary metaphoricity label for each word (Dankers et al., 2019; Chen et al., 2020). We use the model of Dankers et al. (2020), who extend that setup by including discourse information in the prediction. The Hippocorpus contains sentences in the context of a document, which makes the discourse-aware neural model particularly suitable. The metaphor identification model is initialized with the BERT model (Devlin et al., 2019), that contains 12 self-attention layers, pre-trained on large corpora with a language modeling objective. BERT is augmented with a metaphor classification layer that takes two components as input at every step: BERT's representation of the word to label, and a representation that is a weighted average of the words contained in the wider discourse, consisting of the sentence itself, and one preceding and one succeeding sentence.

The model is trained using data from the FigLang 2020 metaphor shared task (Leong et al., 2020), that provides binary labels for verbs, adjectives, nouns and adverbs from a subset of the VU Amsterdam metaphor corpus (Steen et al., 2010). The VU Amsterdam metaphor corpus contains documents from the British National Corpus (Clear, 1993) that fall into four genres: academic, news, conversation, and fiction. The shared task data contains a training dataset with 73 thousand labelled words, of which 15% are metaphorical. We refer the reader to Dankers et al. (2020) for further details on the training procedure of the model. We apply a model pre-trained by Dankers et al. (2020)—specifically, BERT with *general attention* and ($k = 1$)—to the Hippocorpus, that obtained an F1-score (i.e. the harmonic mean of the precision and recall) of 71.1 on the shared task test set. A total of 53,375 words in the Hippocorpus were tagged as being potentially metaphorical.

Novel Metaphor Identification

The VU Amsterdam corpus provides binary labels that indicate whether a word is used metaphorically. Do Dinh, Wieland, and Gurevych (2018) extended the corpus with novelty scores for 15,180 metaphors. The scores range from 1 (novel) to -1 (conventionalized), and were collected through best-worst scaling of four items at a time by crowd-workers from Amazon Mechanical Turk, resulting in approximately normally distributed scores.

We use these novelty annotations to train a neural network that jointly predicts metaphor and novelty, and apply it to the Hippocorpus after training. The network uses BERT as neural encoding module and has both a linear classifier with a sigmoid function to predict metaphoricity, and a linear classifier with a tanh (hyperbolic tangent) function to predict

metaphor novelty. To learn the model parameters, the training data is presented to the network in batches of size 32 over the course of 2000 training steps. We use the AdamW optimizer (Loshchilov & Hutter, 2018) with an initial learning rate of $5e-5$ and a cosine-based learning rate schedule with 10% warmup. For the tasks of metaphor identification and metaphor novelty prediction, the binary cross-entropy and L1 loss functions are used, respectively. The multi-task loss is a weighted sum of the two components, with weights .05 and .95, respectively.

The training procedure is repeated ten times with different random seeds, that determine the random initialization of the classifiers and the order in which the VU Amsterdam data samples are presented to the model. Do Dinh et al. (2018) report a mean absolute error (MAE) of .163 for a baseline Bi-LSTM model. For the metaphors contained in the test set of the FigLang 2020 shared task, we obtain a MAE of .142. The correlation between the predictions and the target is .705 (Pearson's r). We combine the predictions on the Hippocorpus of the ten models by using the maximum score assigned to improve the recall of potentially novel metaphors in the Hippocorpus.

Do Dinh et al. (2018) in their analysis and reporting of novel metaphor chose an arbitrary threshold of T = 0.5 (with -1 being very conventional and 1 being very novel) and labeled everything above this threshold as novel. In the Hippocorpus we used a threshold of T = 0.45 and labeled everything above this threshold as novel. We used a slightly less conservative threshold in order to have a minimum of 1% possible novel metaphor out of all metaphor detected. This resulted in $n = 468$ novel metaphor in recalled stories and $n = 354$ novel metaphor in imagined stories in the Hippocorpus.

Results

A Wilcoxon rank-sum test showed significantly higher normalized frequency counts for novel metaphoric words in recalled stories compared to imagined stories $z(2611) = 2.18$, $p < .05$. Relatedly, we found that a Wilcoxon rank-sum test showed higher values for the max novelty score of metaphoric words in each story for recalled compared to imagined stories $z(2611) = 2.65$, $p < .01$. We also showed significantly higher average novelty scores for metaphoric words in recalled stories ($m = 0.07$, $std = 0.17$) compared to imagined stories ($m = 0.05$, $std = 0.15$), $t(2611) = 3.10$, $p < .01$. In the latter case, we set all stories with no novel metaphors to have a novelty score of zero. The results suggest that language users are more likely to use novel metaphor when they are able to directly draw on primary experience during event narration.

A regression analysis showed that factors that impact the degree to which language users are able to draw on primary experience during event narration (i.e., openness to experience—one of the Big Five personality traits, similarity or how similar the event described is to one's own experience [only relevant to imagined stories], and emotionality or how memorable/important the event described is perceived to be)

may be subtly associated with novel metaphoric word counts (and novelty scores for metaphoric words) in different ways in recalled versus imagined stories. A multiple regression analysis was run to predict the normalized frequency counts for novel metaphoric words in each story from openness and emotionality with type of story (recalled = 1 versus imagined = 0) coded as a dummy variable. We also included all possible interactions with the type of story as predictor variables in the regression. We found that openness $\beta = 0.0003$, $t(5219) = 3.96$, $p < .001$ and the interaction between openness and story type $\beta = 0.0003$, $t(5219) = 3.39$, $p < .01$ added statistically significantly to the prediction explaining a small proportion of variance in normalized frequency counts for novel metaphoric words with $R^2 = 0.016$, $F(5,5219) = 17.44$, $p < .001$. We also ran a multiple regression analysis to predict the average novelty scores for metaphoric words, which also showed a significant effect for openness $\beta = 0.03$, $t(5219) = 4.047$, $p < .001$ and the interaction between openness and story type, $\beta = 0.03$, $t(5219) = 3.00$, $p < .01$ explaining a small proportion of variance in novel metaphor counts with $R^2 = 0.018$, $F(5,5219) = 19.23$, $p < .001$. Taken together these result suggest that the personality trait of openness is more closely associated with novel metaphor use in recalled compared to imagined stories.

In order to test the extent to which similarity (only relevant for imagined stories) impacted novel metaphor use in each story we conducted an additional regression analysis. A multiple regression analysis was run to predict the average novelty scores for metaphoric words in each imagined story from openness, emotionality, and similarity. We also included all possible interactions with similarity as predictor variables in the regression. We found that openness $\beta = 0.03$, $t(2606) = 1.97$, $p < .05$ and the interaction between openness and similarity $\beta = 0.01$, $t(2606) = 2.10$, $p < .05$ added statistically significantly to the prediction, however, only explaining a very small proportion of variance in average novelty scores for metaphoric words with $R^2 = 0.005$, $F(5,2606) = 2.86$, $p < .05$. Nevertheless, this result is in line with our intuition that when language users are able to draw on primary experience (e.g., if they have had a similar experience to the event described) they may be more likely to use novel metaphor during event narration.

Lastly, a multiple regression analysis was also run to predict the normalized frequency counts for all metaphoric words (irrespective of novelty) in each story from openness and emotionality with type of story (recalled = 1 versus imagined = 0) coded as a dummy variable. We also included all possible interactions with type of story as predictor variables in the regression. We found that emotionality added statistically significantly to the prediction $\beta = 0.001$, $t(5219) = 2.355$, $p < .05$ but only explaining a very small proportion of variance in normalized frequency counts for metaphoric words with $R^2 = 0.003$, $F(5,5219) = 2.65$, $p < .05$. A similar but slightly stronger result was found when looking at predicting the normalized frequency counts for conventional

metaphoric words. Taken together, these results may suggest that emotionality (or how important the event was perceived to be) may be subtly associated with metaphor use overall. This is in line with prior research showing a link between metaphor and emotion but this will need to be further investigated in a larger dataset that more directly assesses the role of emotion.

Discussion

In this study we used NLP techniques to analyze metaphor in a corpus of English diary-like stories of recalled and imagined events. Specifically, we sought to investigate whether episodic memory demands (or cognitive factors that impact a language users ability to draw on primary experience) during event narration modulate novel (conventional) metaphor use. Importantly, imagined events were matched in topic with recalled events in the corpus but included descriptions of events as if they had happened to the authors of the stories but not describing an event that actually happened to them. For the stories in the corpus, Sap et al. (2020) showed that describing recalled events places larger demands on episodic memory compared to imagined events, which largely involve semantic memory. Furthermore, stories were also accompanied by variables or factors that may impact a language user's ability to draw on primary experience to describe events including (1) openness to experience—one of the Big Five personality traits; (2) emotionality or how memorable/important the event described was perceived to be; and (3) similarity or how similar the event described was to one's own experience (only relevant to imagined events). We hypothesized that greater episodic memory demands in recalled compared to imagined stories (or cognitive factors that may modulate a language user's ability to draw on primary experience) should positively modulate novel metaphor use during event narration.

The results showed significantly greater novel metaphor count (and overall higher novelty scores for metaphoric words) in recalled compared to imagined stories supporting the idea that access to primary experience may be one important cognitive factor driving novel metaphor use at least in event narration. Put another way, novel metaphor use is associated with recalled events to a greater extent than imagined events as they involve describing or reflecting on a private experience that may be more difficult to fully convey using only conventional metaphor. Filling these 'lexical gaps' involves the creation of novel semantic representations through the association of novel ideas (creative use of language). It is believed that this process is mediated in part through relational processing in the hippocampus that allows for the interaction of information associated with episodic and semantic memory (Duff, Covington, Hilverman, & Cohen, 2020). In contrast, while imagined events may also draw on these processes—including to flexibly organize semantic representations in terms of temporal-spatial events—it is largely dependent on semantic memory. This is in line with the find-

ing that imagined compared to recalled events were found to rely more on commonsense knowledge and overall had more linear narrative flow (i.e., likelihood of sentences under generative language models conditioned on varying amounts of history) (Sap et al., 2020).

Importantly, we found that factors that impact the extent to which language user's are able to draw on primary experience to describe events were associated with novel metaphor use in distinct ways in recalled compared to imagined events. On the whole, we found evidence to suggest that the Big-Five personality trait of openness to experience was significantly associated with novel metaphor use (i.e., novel metaphor counts and novelty scores for metaphoric words) irrespective of story type. Notably, however, we found that openness was significantly more associated with novel metaphor use in recalled compared to imagined stories. Moreover, when only looking at imagined stories we found a small but significant interaction effect between openness and similarity suggesting that openness may be more associated with novel metaphor use for imagined events that describe an experience very similar to the author's own personal experience. Openness to experience in the Big Five personality traits has been associated with being creative, cognitively open, imaginative, and introspective (Donnellan et al., 2006). These qualities would certainly facilitate the ability to describe a private experience that may be more difficult to convey using only conventional metaphor (an important component of linguistic introspection). The results add further support to the idea that novel metaphor use may be modulated by episodic memory demands (or factors that directly impact a language user's ability to draw on primary experience) during event narration.

Interestingly, when looking at metaphor counts irrespective of novelty we found a very small positive association with emotionality (or the perceived importance of the event described) and not openness. This may suggest that emotionality is to some extent associated with metaphor use, more generally. This would be in line with prior research suggesting that language users draw on emotion metaphors to talk about emotion (Kovecses, 2000), as well as metaphor to conceptualize their emotional experiences (Lakoff & Johnson, 1980). This is further supported by work showing that metaphor may be more emotionally evocative compared to equivalent literal expressions (Citron & Goldberg, 2014; Mohammad, Shutova, & Turney, 2016), which supports the idea that metaphor may be used for added rhetorical effect. Critically, the task involved depictions of memorable life events (e.g., the birth of a child or the death of a loved one) and, therefore, a good number of these events were rated as having high perceived importance overall, which may have prevented finding a larger effect of emotion on metaphor during event narration. Future work will need to more directly assess the impact of emotion on metaphor (also novel versus conventional metaphor use) during event narration.

There are important limitations to the current work. The results presented depend on the accurate identification of

metaphor and, additionally, detection of novelty—a graded concept difficult to pinpoint (Do Dinh et al., 2018). The model we used for automatic metaphor identification suffers from both false positives and negatives but at a low rate with an F1-score of 71.1 (i.e. the harmonic mean of precision and recall) on the FigLang 2020 metaphor shared task set (Leong et al., 2020). We also obtained a low mean absolute error performance (MAE = .142) for novelty score prediction in the test set of the FigLang 2020 shared task when compared to Do Dinh et al. (2018) human annotation best-worst scaling of four metaphors at a time for novelty. Nevertheless, these measures clearly indicate that a certain amount of noise (e.g., false positives and negatives) may have impacted our results possibly resulting in small effect sizes. On the other hand, it is possible that systematic biases in the datasets and the task of learning binary classification of metaphoric (non-metaphoric) words and novelty scores may have impacted our results. Human annotation of the Hippocorpus for metaphor and novelty scores are ultimately needed to corroborate our findings. This work will also help to evaluate and improve models for metaphor (and novel metaphor) identification.

Finally, the work depends on the assumption that the task in the Hippocorpus effectively manipulated episodic memory demands which was corroborated through various linguistic markers that were found (Sap et al., 2020). However, it is entirely possible that the linguistic differences observed by Sap et al. (2020) between recalled and imagined stories are due to more general differences in the cognitive demands between imagined versus recalled narration due to the nature of the task itself. In particular, in the imagined cases the authors were asked to describe events given a topic which may be a more difficult task compared to the task of merely retelling a story. Moreover, participants were additionally asked to describe the event as if it happened to them (i.e., in the first-person) but *not* describe an event that actually happened to them. This may be somewhat of an unnatural task and, therefore, an overall more difficult task cognitively. Not to mention it is unclear what motivation if any the participants might have to do a good job at imagined storytelling. There might be more communicative goals to consider such as imagined storytelling for the purpose of deception or entertainment. Thus, differences in the use of novel metaphor may merely reflect the fact imagined storytelling placed greater overall cognitive demands compared to recalled storytelling. Still, it is not entirely evident how this could explain our findings showing a small but significant interaction between the personality trait of openness and type of story (recalled versus imagined), as well as an interaction between openness and similarity in imagined stories. These findings rather suggest an association between the ability for creative language introspection and access to relevant primary experience irrespective of cognitive demands. Nevertheless, future work may consider giving all participants a topic to narrate with more explicit goals (communicative goals or otherwise) and later assessing the extent to which the story closely reflects a spe-

cific event in memory or is more fictitious in nature. This might better control for underlying differences in the difficulty of narrating imagined versus recalled events.

In conclusion, we found evidence to suggest that episodic memory demands (or cognitive factors that influence the extent to which language users are able to draw on primary experience) modulate novel metaphor use in event narration in subtle ways. Future studies will need to more carefully tease apart the exact cognitive factors associated with novel (conventional) metaphor use and the mechanisms by which they may be driving metaphor use in event narration. Attention should also be given to the underlying communicative (interpersonal) or other goals that may involved and its potential impact on more novel metaphor use (Gerrig & Gibbs, 1988; Horton, 2007). This analysis paves the way for using large-scale corpora to analyze underlying cognitive processes that modulate metaphorical language use.

References

Chen, X., Leong, C. W. B., Flor, M., & Beigman Klebanov, B. (2020, July). Go figure! multi-task transformer-based architecture for metaphor detection using idioms: ETS team in 2020 metaphor shared task. In *Proceedings of the second workshop on figurative language processing* (pp. 235–243). Online: Association for Computational Linguistics. doi: 10.18653/v1/2020.figlang-1.32

Citron, F. M., & Goldberg, A. E. (2014). Metaphorical sentences are more emotionally engaging than their literal counterparts. *Journal of Cognitive Neuroscience*, *26*(11), 2585-2595.

Clear, J. H. (1993). The british national corpus. In *The digital word: Text-based computing in the humanities* (p. 163–187). Cambridge, MA, USA: MIT Press.

Dankers, V., Malhotra, K., Kudva, G., Medentsiy, V., & Shutova, E. (2020, July). Being neighbourly: Neural metaphor identification in discourse. In *Proceedings of the second workshop on figurative language processing* (pp. 227–234). Online: Association for Computational Linguistics. doi: 10.18653/v1/2020.figlang-1.31

Dankers, V., Rei, M., Lewis, M., & Shutova, E. (2019, November). Modelling the interplay of metaphor and emotion through multitask learning. In *Proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (emnlp-ijcnlp)* (pp. 2218–2229). Hong Kong, China: Association for Computational Linguistics. doi: 10.18653/v1/D19-1227

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019, June). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the north American chapter of the associa-*

tion for computational linguistics: Human language technologies, volume 1 (long and short papers) (pp. 4171–4186). Minneapolis, Minnesota: Association for Computational Linguistics. doi: 10.18653/v1/N19-1423

Do Dinh, E.-L., Wieland, H., & Gurevych, I. (2018, October-November). Weeding out conventionalized metaphors: A corpus of novel metaphor annotations. In *Proceedings of the 2018 conference on empirical methods in natural language processing* (pp. 1412–1424). Brussels, Belgium: Association for Computational Linguistics. doi: 10.18653/v1/D18-1171

Donnellan, M., Oswald, F. L., & Lucas, R. (2006). The mini-ipip scales: tiny-yet-effective measures of the big five factors of personality. psychological assessment. *Psychological assessment*, *18*(2), 192–203.

Duff, M. C., Covington, N. V., Hilverman, C., & Cohen, N. J. (2020). Semantic memory and the hippocampus: Revisiting, reaffirming, and extending the reach of their critical relationship. *Frontiers in Human Neuroscience*, *13*, 471. doi: 10.3389/fnhum.2019.00471

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

Gerrig, R. J., & Gibbs, R. W. (1988). Beyond the lexicon: Creativity in language production. *Metaphor and Symbolic Activity*, *3*, 1-19.

Horton, W. S. (2007). Metaphor and reader's attributions of intimacy. *Memory and Cognition*, *35*(1), 87-94.

Kovecses, Z. (2000). *Metaphor and emotion: Language, culture, and body in human feeling*. Cambridge University Press.

Lakoff, G., & Johnson, M. (1980). *Metaphors We Live By*. Chicago: University of Chicago Press.

Leong, C. W. B., Beigman Klebanov, B., Hamill, C., Stemle, E., Ubale, R., & Chen, X. (2020, July). A report on the 2020 VUA and TOEFL metaphor detection shared task. In *Proceedings of the second workshop on figurative language processing* (pp. 18–29). Online: Association for Computational Linguistics. doi: 10.18653/v1/2020.figlang-1.3

Loshchilov, I., & Hutter, F. (2018). Decoupled weight decay regularization. In *International conference on learning representations*.

Mohammad, S., Shutova, E., & Turney, P. (2016, August). Metaphor as a medium for emotion: An empirical study. In *Proceedings of the fifth joint conference on lexical and computational semantics* (pp. 23–33). Berlin, Germany: Association for Computational Linguistics. Retrieved from https://www.aclweb.org/anthology/S16-2003 doi: 10.18653/v1/S16-2003

Sap, M., Horvitz, E., Choi, Y., Smith, N. A., & Pennebaker, J. (2020, July). Recollection versus imagination: Exploring human memory and cognition via neural language models. In *Proceedings of the 58th annual meeting of the association for computational linguistics* (pp. 1970–1978). Online: Association for Computational Linguistics. Retrieved from https://www.aclweb.org/anthology/2020.acl-main.178

doi: 10.18653/v1/2020.acl-main.178

Steen, G. J., Dorst, A. G., Herrmann, J. B., Kaal, A. A., Krennmayr, T., et al. (2010). Vu amsterdam metaphor corpus. *Oxford Text Archive Core Collection*.

Tulving, E. (1972). *Episodic and semantic memory, in organization of memory eds e. tulving and w. donaldson*. New York, NY: Academic Press.