Metaphor is a Key to Abstraction: Modelling Figurative Abstraction with Distributional Semantics

Hamad Al-Azary (halazary@ltu.edu)[1] & J. Nick Reid[2]
[1]Humanities, Social Sciences & Communication, Lawrence Technological University
[2]Department of Psychology, University of Northern British Columbia

Abstract

Metaphorical abstraction is the process wherein concrete concepts (e.g., *key*) develop lexicalized abstract meanings (e.g., *important*) as a result of frequent figurative contexts (e.g., *communication is a key*). In this study, we examine the semantic properties associated with metaphorical abstraction. Participants interpreted a series of low-familiar metaphors (e.g., *religion is a raft*), and generated their own related metaphors (e.g., *spirituality is a raft*). After this familiarization procedure, participants were presented with familiarized metaphors (e.g., *faith is a raft*) and non-familiarized metaphors (e.g., *highways are snakes*) and listed semantic properties associated with both the familiarized and non-familiarized metaphorical concepts (e.g., *raft, snakes*). The results showed that familiarized metaphorical concepts evoked semantic properties that were more semantically distant and more abstract than those evoked by non-familiarized metaphors or literal sentences. These results demonstrate how metaphorical abstraction results in the growth of novel and abstract semantic connections.

Keywords: metaphor; abstraction; psycholinguistics; distributional semantic models

Introduction

Some words like *key*, although highly concrete, nonetheless have additional lexicalized abstract meanings (i.e., *something of critical importance*). According to the *career hypothesis*, concrete concepts undergo a career wherein they initially refer to a basic-level concept and may later develop abstract senses as a result of frequent figurative use, wherein the abstract meaning is initially computed on the fly but eventually becomes stored in the semantic lexicon (Bowdle & Gentner, 2005; Jamrozik et al., 2016). Indeed, *key* is an example of a concrete concept that has undergone such a career; its concrete meaning (i.e., *a metal tool to open locks*) dates to Middle English, whereas the abstract sense of *key* meaning *critical importance* dates to 1913 (Online Etymology Dictionary, n.d.). Moreover, many polysemous concrete words have historically older concrete meanings established prior to abstract senses. Although there is diachronic evidence for the career hypothesis, no study to our knowledge has yet to experimentally test whether metaphorical abstraction results in the generation of new abstract associations. To that end, in this study, we experimentally test whether repeated metaphorical use of concrete concepts results in the growth of semantically distant and abstract associations, as is observed with concepts such as *key*.

By its nature, metaphor processing involves abstraction. For example, understanding *language is a bridge*, involves the juxtaposition of two semantically distinct concepts to produce a figurative meaning, such that language *connects* people socially and mentally. The meaning is fundamentally abstract, because language connects people in a different way than how literal bridges physically connect land-masses. Moreover, the concrete features of bridge are not primarily involved in the meaning of the metaphor. In the example of *language is a bridge*, the topic of the metaphor, which is often the first concept and the literal concept that the metaphor is about, is *language*, whereas the vehicle of the metaphor, which is the concept that is being used metaphorically, is *bridge*. Often, metaphor vehicles are concrete concepts, whereas metaphor topics may be either abstract (e.g., *time is money*) or concrete (e.g., *that lawyer is a shark*). Numerous psycholinguistic models of metaphor propose that processing involves semantic properties of the vehicle being conveyed to the topic, but differ regarding whether semantic properties are derived from comparison (Gentner & Bowdle, 2008), categorization (Glucksberg, 2008), lexical spreading activation (Kintsch, 2000), or embodied simulation processes (Gibbs, & Matlock, 2008).

The career-of-metaphor framework proposes that the nature of metaphor processing depends on the familiarity[1] of the figurative meaning (Bowdle & Gentner, 2005). When initially encountered, a concept's figurative meaning must be computed on the fly. Accordingly, similes (e.g., *A.I. is like an assistant*) are particularly effective for conveying novel figurative relationships, due to their explicit comparison (i.e., the word *like* overtly invites comparison processes). When a term becomes frequently used in figurative contexts, its abstract meaning becomes lexicalized, and hence stored in long-term memory. At this point, it can be directly retrieved and meaningfully expressed as a metaphor where the

[1] The term *familiarity* is often used interchangeably with *conventionality*, which is the degree to which a word has a conventional figurative meaning.

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

figurative relationship is directly asserted (e.g., *A.I is an assistant*). A related view, the *sensorimotor shedding hypothesis,* further argues that familiarization affects the nature of the semantic representation of the figurative concept (Jamrozik et al., 2016). When a concrete concept is initially used figuratively, its sensorimotor properties are accessed to compute a figurative meaning. For example, in *A.I. is an assistant,* sensorimotor and experiential information associated with *assistant* is necessary to compute the figurative meaning (such as speaking and interacting with someone else). Over time, sensorimotor information is "shed" such that an abstract meaning is the predominant representation of the expression (i.e., *helps make life easier*). Thus, both views maintain that familiarity results in stored abstract representations and are considered part of the general career camp.

Experimental support for the career-of-metaphor hypothesis comes from *in-vitro conventionalization*, wherein a novel expression is repeatedly encountered in a lab setting, resulting in a rapidly formed abstraction. In one variation of this technique, Bowdle and Gentner (2005) presented participants with sentences such as *An acrobat is like a butterfly* and *A figure skater is like a butterfly*, along with an incomplete sentence frame such as ____ *is like a butterfly.* Participants produced a concept to make a semantically similar sentence (e.g., *A dancer is like a butterfly*). Afterwards, participants were presented with sentences containing previously encountered metaphorical words (*A ballerina is like a butterfly*) along with sentences containing unpresented concepts, such as *A job is like a jail*, and judged whether such sentences are best expressed as similes or metaphors. The results showed that participants judged familiarized sentences to be better as metaphors rather than similes, and non-familiarized sentences to be better as similes rather than metaphors. Bowdle and Gentner (2005) argued that this result is due to an abstraction that formed – when concepts such as *butterfly* are frequently encountered in figurative contexts, their abstract sense accrues, and as a result, serves as lexicalized abstraction that is directly understood (as a metaphor), rather than as a simile wherein it needs to be computed from scratch.

Thibodeau and Durgin (2011) also used an in-vitro conventionalization technique and demonstrated priming effects consistent with the career hypothesis, such that, as a result of familiarization, a figurative sentence like *Education is a lantern* primes *A mentor is a lantern* rather than an alternative figurative sense, such as *A uniform is a lantern*. Thus, metaphorical abstraction appears to demonstrate a change in semantic representation of the vehicle affecting comprehensibility. Cardillo et al. (2012) found that in-vitro familiarized metaphors resulted in less activity in brain regions associated with semantic (i.e., left and right inferior frontal gyri; left middle temporal gyrus) and visual processing (i.e., right postero-lateral occipital cortex), demonstrating neural correlates of metaphorical abstraction.

Although the aforementioned experiments demonstrated that metaphorical familiarization resulted in semantic priming effects, grammatical shifts, and changes in brain areas involved in semantic and visual processing, it remains unclear what the semantic effects of figurative abstraction are. In other words, how does the semantic content of concepts, such as *key*, change as a result of metaphorical familiarization? In this paper, we examine directly whether metaphorical abstraction results in new, semantically distant associations, which is how we operationally define metaphorical abstraction. Thus, if metaphorical abstraction depends on familiarization, we expect that participants will list more semantically distant properties for familiarized metaphors than for non-familiarized metaphors. In addition to being more semantically distant, we also expect properties of familiarized metaphors to be more abstract than properties of non-familiarized metaphors. Characterizing the semantic representations of familiarized metaphor vehicles would therefore confirm whether such concepts do in fact develop abstract senses, as predicted by the career account.

In addition to metaphorical familiarity, we also explore the role played by word-level semantic variables on abstraction. In particular, experimental work has demonstrated words with many nearby semantic associations, and hence residing in dense semantic spaces, are associated with less comprehensible metaphors (Al-Azary et al., 2019; Al-Azary & Buchanan, 2017; Al-Azary & Katz, 2023). That is, when words have many near neighbors, they have less means to form new figurative connections, resulting in processing difficulties and less overall comprehensibility. Thus, semantic neighborhood density is a relevant variable for abstraction, leading to the prediction that concepts from sparse semantic spaces will be more amenable for the growth of new associations, and hence, abstraction, than concepts from dense semantic spaces.

The Current Study

Here, we use a variation of the in-vitro conventionalization method to examine the semantic representations of concrete concepts used metaphorically. To that end, we used a property listing technique (McRae et al., 2005; Roncero et al., 2015; Utsumi, 2005) to characterize the content of familiarized metaphorical vehicles relative to non-familiarized metaphorical vehicles. In particular, we operationally define abstraction in two complementary ways; (1) as the semantic distance (measured by vector cosines derived from distributional semantic models) between the vehicle word and its semantic properties and (2) the concreteness ratings of the semantic properties.

We used a two-part familiarization phase wherein we presented participants with low-familiar metaphors (e.g., *streets are snakes; roads are snakes*) and literal sentences (e.g., *tea is a drink; juice is a drink*), and asked them to list three words that the vehicle concept expresses about the topic concept. Subsequently, participants were presented with the previously encountered sentences (e.g., *streets are snakes*; *roads are snakes*) in addition to an incomplete sentence frame (e.g., ____ *are snakes*) they were asked to complete to be semantically similar (e.g., *paths are snakes*). Finally, in the test-phase, participants were presented with sentences that contained familiarized vehicles (e.g., *highways are snakes*) and non-familiarized vehicles (e.g., *faith is a raft*), and once again, listed properties that the vehicle was expressing about the topic.

To foreground our results, we found that familiarized metaphors resulted in (1) more semantically distant properties; (2) more abstract properties, and (3) that vehicles from sparse semantic spaces resulted in more semantically distant properties. For literal sentences, familiarization resulted in more concrete properties; moreover, familiarized literal sentences with semantically dense vehicles evoked more proximal semantic properties.

Method

Participants

Eighty-eight participants (48 females, 40 males; *M* age = 22.11 years old; range 18-55 years old) from the Lawrence Technological University community who identified as native speakers of English participated (in-person) for course credit or monetary compensation (i.e., $10 Amazon gift card).

Materials

Metaphorical and literal sentences in the form of *X(s) is/are Y(s)* were the primary items in this experiment. The items are broken down by *test items*, which are the main stimuli that our analyses are based on, and *study items*, which are presented to participants so that they can become familiarized. See Table 1 for a sample of the items.

Table 1. Example items

Sentence Type	Example
Metaphorical	Highways are snakes
Metaphorical	Jobs are jails
Metaphorical	Faith is a raft
Literal	A car is a vehicle
Literal	Basketball is a game
Literal	A painter is an artist

Test Items Twenty-four metaphorical (e.g., *highways are snakes*) and twenty-four literal (e.g., *a canary is a bird*) sentences were used as the primary stimuli serving as test items. The metaphors were sourced from Roncero and de Almeida's (2015) norms. As a starting point in stimulus selection, the 12 low-familiar metaphors used in Al-Azary and Katz (2021) were used, as those items primed literal-level bodily actions during processing and were therefore appropriate items for potentially demonstrating a shift to abstraction. Twelve additional items were selected from the norms that were similar in terms of aptness and familiarity. The mean familiarity rating of the metaphors is 2.48 whereas the mean aptness rating is 5.32, both on ten-point scales; therefore, the metaphors are low on familiar yet moderately apt. Literal sentences in the same format as the metaphors were constructed to serve as experimental controls; however, literal sentences were only in singular format. Two lists were created, such that half the items (12 metaphorical and 12 literal sentences) were familiarized in one version, whereas they were non-familiarized in another version.

Study Items For each test item (e.g., *highways are snakes*), two semantically similar sentences (i.e., *roads are snakes; streets are snakes*), were created to serve as study items.

Procedure

This experiment involved three phases. In the pre-familiarization phase, participants interpreted the study items (24 metaphorical and 24 literal sentences that were presented in a random order (e.g., *streets are snakes*; *a robin is a bird*) by listing three semantic properties that the vehicle concept (i.e., *snakes*; *bird*) was expressing about the topic concept (i.e., *highways*; *robin*). In the familiarization phase, participants were presented with the previously encountered study items (e.g., *streets are snakes; roads are snakes*), and were asked to complete a sentence frame (i.e., ______ *are snakes*) to produce a semantically similar sentence. Finally, in the test phase, participants were presented with the 24 metaphorical test items, 12 being based on familiarized vehicles, such as *highways are snakes*, and 12 being based on non-familiarized vehicles, such as *faith is a raft*. In addition, participants were presented with 24 literal sentences, with 12 based on familiarized vehicles and 12 based on non-familiarized vehicles. All items in the study were presented in random order for both the familiarization and test phases. Two lists were created such that familiarized items in one list were the non-familiarized items in another list. Participants received only one of the two lists.

Results

Pre-processing

Across the 88 participants, there was a maximum potential number of 12,672 (3 properties for 48 sentences across 88 participants) properties to be obtained. However, 125 of the possible properties were not generated (either intentionally ignored, or accidentally skipped by the participants, or were not semantic properties, such as 'what', 'is', 'this'), leaving a total of 12,547 items. Of these 12,547 items, Excel spellchecker was used to flag potential typos, which we corrected with consideration to the context of the feature (the stimuli and other features the participant generated; for example, for *A canary is a bird,* one property listed was *nimple*, which was corrected to *nimble*.). We corrected 455 typos, amounting to 3.6% of the data. See Table 2 for a set of properties listed for a metaphorical and literal sentence across levels of familiarization.

The vector-space model, FastText (Bojanowski et al., 2017), was used to compute semantic distances between the vehicle words and the semantic properties participants produced. FastText is advantageous for this dataset because it can handle out-of-vocabulary words due to its focus on sub-lexical information. The semantic distance is inferred from the cosine-similarity between the vectors of the words, with higher values representing more semantic similarity, or proximity. Concreteness was determined from (Martínez et al., 2025) ratings, which are an extensive dataset generated by AI and based on earlier concreteness ratings by human participants (Brysbaert et al., 2014), with higher values representing more concrete words. To analyze the data, the *lme4* package in R was used to fit mixed-effects models using the *mixed* function whereas the *emmeans* package was used to derive mean estimates and the *pairs* function to contrast condition means. Generally, we fit linear models with Familiarity (Familiarized vs Non-Familiarized) and Sentence Type (Metaphor vs Literal) as fixed-effects, along with the Familiarity by Sentence Type Interaction. We also included random intercepts for participants and items (i.e., the vehicles of the sentences). Model convergence permitting, we included random slopes for Sentence Type by participant and for Familiarity by item (i.e., vehicle).

Semantic Distance of Properties

With regards to the average semantic distance between the vehicles and the generated properties, we employed a mixed model with Familiarity (Familiarized vs Non-Familiarized) X Type (Metaphor vs Literal) fixed factors and participants as random intercepts with sentence type as random slopes and vehicles as random intercepts with familiarity as random slopes. A Familiarity by Type interaction was observed, $\chi^2(1) = 8.37$, $p = .004$, such that familiarized metaphors resulted in more semantically distant properties ($M = .188$, $SE = .009$) than Non-familiarized metaphors ($M = .20$, $SE = .009$; $z = 3.06$, $p = 0.002$). However, there was not a significant difference in the semantic distances of properties for familiarized literal sentences ($M = .286$, $SE = .009$) compared to non-familiarized literal sentences ($M = .281$; $SE = .009$; $z = 1.22$, $p = 0.22$).

Table 2. Properties for a metaphorical and literal sentence.

Sentence	Non-Familiarized Properties	Familiarized Properties
Faith is a raft	*drifting* *bobbling* *still*	*comforting* *supportive* *guidance*
A painter is an artist	*color* *vision* *imagination*	*creative* *innovative* *colorful*

To explore the nature of the observed familiarization effect, we repeated the analysis on the first, second and third properties individually, rather than the pooled properties. With regards to the first property listed, the familiarity by type interaction was significant, $\chi^2(1) = 4.16$, $p = 0.04$; the semantic distance of properties produced for familiarized metaphors ($M = .207$, $SE = .01$) were more distant than the properties for non-familiarized metaphors ($M = .234$, $SE = .01$), $z = 3.89$, $p < .001$. No effect was observed for the difference between familiarized literal ($M = .30$; $SE = .012$) and non-familiarized literal sentences ($M = .307$; $SE = .012$; $z = 1.09$, $p = 0.27$). With regards to the second property, A familiarity by type interaction was observed, $\chi^2(1) = 5.98$, $p = 0.014$; however, follow-up pairwise tests were non-significant (ps $< .10$). Finally, for the third feature, the familiarity by sentence type interaction was non-significant ($p = .10$). Taken together, the results demonstrate that familiarization of metaphor vehicles results in more semantically distant properties; moreover, this effect was most pronounced in the first properties that were produced.

Finally, we examined the role of the vehicle's semantic neighborhood density (SND) in the semantic properties by adding a fixed-effect for the vehicle's semantic density (semantic distance of 10 nearest neighbors). Semantic neighborhood density is a continuous variable reflecting the average semantic distance between a vehicle and its ten nearest neighbors, with less distance reflecting greater neighborhood density.

A three-way interaction between Familiarity, Sentence Type, and Vehicle SND was observed, $\chi^2(1) = 4.04$, $p = 0.044$. As Figure 1 shows, the effect of vehicle SND is visibly positive for each sentence type. However, the effect of vehicle SND differed across levels of familiarity and sentence type. In particular, the contrast between the slopes (representing the effect of vehicle SND) for familiarized literal and non-familiarized literal sentences was significant ($z = 1.97$, $p = 0.049$), such that familiarized sentences with denser vehicles elicited semantically closer (i.e., greater cosines) properties than non-familiarized literal sentences, whereas for metaphors, the slopes did not differ significantly ($z = 0.735$, $p = 0.46$). Greater vehicle SND elicited semantically closer properties for both familiarized and non-familiarized metaphors. In terms of individual slopes, only the non-familiarized metaphor condition was significant ($M = .304$; $SE = .15$, $z = 2.058$, $p = 0.04$.

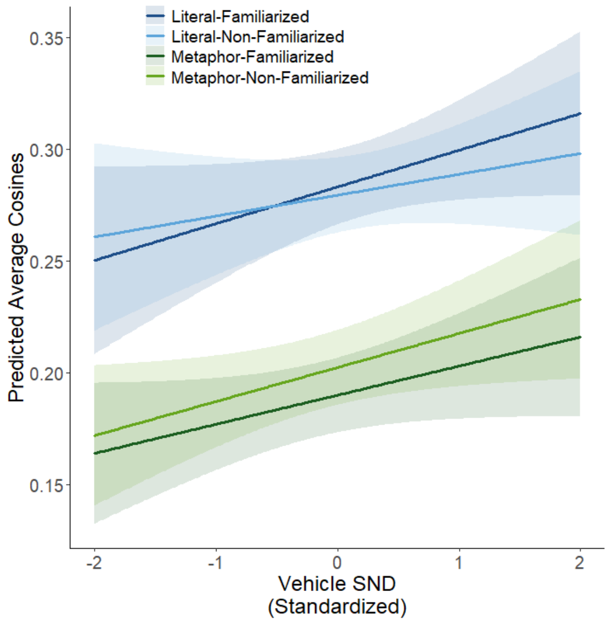

Figure 1: Familiarity by Sentence Type by Vehicle SND Interaction. The y-axis represents the average cosines between the vehicle words and the listed three semantic properties. Higher cosines indicate greater semantic similarity. Bands reflect 95% Confidence Intervals.

Concreteness of Properties

A similar linear-mixed effects model was run on the concreteness ratings for the semantic properties. With regards to the average concreteness score across all properties (regardless of the order they were listed), a significant Familiarity by Type interaction was observed, $\chi^2(1) = 18.32$, $p < .001$. Familiarized literal sentences resulted in more concrete ($M = 2.68$, $SE = .08$) properties than non-familiarized literal sentences ($M = 2.56$, $SE = .08$), $z = 4.35$, $p < .0001$. However, the concreteness of properties evoked by familiarized metaphors ($M = 2.02$, $SE = .08$) did not significantly differ from those evoked by non-familiarized metaphors ($M = 2.07$, $SE = .08$), $z = 1.724$, $p = .08$.

With regards to the first property listed, a familiarity by type interaction was again observed, $\chi^2(1) = 7.09$, $p = 0.008$, with the effect driven by metaphors, such that the properties evoked by familiarized metaphors ($M = 2.01$, $SE = .086$) were less concrete (i.e., more abstract) than those evoked by non-familiarized metaphors ($M = 2.11$, $SE = .086$), $z = 2.64$, $p = .007$. However, for literal sentences, there was no significant difference in concreteness between the properties from familiarized ($M = 2.66$, $SE = .086$) and non-familiarized ($M = 2.62$, $SE = .086$) sentences, $z = 1.08$, $p = .28$. For the second property, the familiarity by type interaction, $\chi^2(1) = 5.54$, $p = 0.02$, revealed that familiarized and non-familiarized metaphors were equal (both Ms = 2.04, SEs = .084); $z = .095$, $p = 0.92$, but familiarized literal sentences ($M = 2.70$, $SE = .08$) evoked more concrete properties than non-familiarized sentences ($M = 2.57$, $SE = .08$), $z = 3.25$, $p = .0012$). An interaction was also observed for the third property, $\chi^2(1) = 11.13$, $p < .001$, again with the effect driven by greater concreteness for familiarized literal sentences ($M = 2.67$, $SE = 0.08$) relative to non-familiarized literal sentences ($M = 2.50$, $SE = .08$), $z = 4.074$, $p < .0001$, with no significant difference between the familiarized metaphors ($M = 2.01$, $SE = .0845$) and non-familiarized metaphor ($M = 2.04$, $SE = .084$), $z = 0.66$, $p = 0.51$.

Discussion

Three main findings emerged from the results of our study indicating that metaphor familiarization results in semantic abstraction processes. First, participants produced more semantically distant properties to familiarized metaphor vehicles than non-familiarized metaphor vehicles. This effect was observed across the average set of properties (total three) but was driven by the first property participants produced. We interpret greater semantic distances between the vehicle words and their evoked properties to reflect abstraction processes, wherein distant connections are formed in semantic space. Importantly, this effect was not observed for literal sentences, which, by their nature, do not evoke semantically distant associations as a result of familiarization.

Second, we observed that the semantic neighborhood density (SND) of the vehicle affected the ease with which metaphorical abstractions were formed. In particular, vehicles from sparser semantic spaces resulted in more distant semantic properties than metaphor vehicles from denser semantic spaces, a finding consistent with prior work demonstrating the detrimental effects of semantic neighborhood density on computing metaphorical meaning. Accordingly, it has been argued that semantically dense words have too many nearby associations, and therefore less "room" in semantic space to develop new associations (Al-Azary & Buchanan, 2017; Al-Azary & Katz, 2023). Thus, vehicle density restricts the generation of distant semantic properties in metaphors. The observed results here converge with previous work, illustrating that one reason higher SND values of a vehicle are associated with less comprehensibility is due to the difficulties with drawing connections to semantically distant associations. With regards to literal sentences, we also observed that higher vehicle SND was associated with more proximal semantic properties; however, this effect was more pronounced with familiarized literal sentences. With literal meaning, high vehicle SND can facilitate processing more proximal associations as a concept is familiarized.

Third, with regards to the first properties participants produced, familiarized metaphors resulted in more abstract words than non-familiarized metaphors, which we interpret to be converging evidence in support of metaphors producing abstractions in memory. Moreover, for literal sentences, familiarization led to more concrete semantic properties being generated, which, similar to the SND effect for familiarized literal sentences, further illustrates different processes underlying literal and metaphorical familiarization.

Metaphorical familiarization results in more distant and abstract semantic associations whereas literal familiarization results in more concrete associations and more proximal semantic associations in the context of vehicle SND.

Taken together, our results flesh out the semantic representations that result from metaphorical abstraction. Prior work demonstrated metaphorical abstraction in grammatical form (Bowdle & Gentner, 2005), semantic priming (Thibodeau & Durgin, 2011) and neural responses (Cardillo et al., 2012), whereas our study characterized the content of metaphorical abstraction; namely that metaphorical abstraction resulted in semantic properties that were both more distant and more abstract. In addition, concrete concepts differ in their potential for abstraction, such that concepts with denser semantic neighbors were less amenable to generating novel, distant associations.

While we interpret the observed familiarization effects within the career framework (Bowdle & Gentner, 2005; Jamrozik et al., 2016), we acknowledge that the data may be interpreted from other perspectives, such as the *quality of metaphor hypothesis* (Glucksberg & Haught, 2006). According to this framework, metaphorical abstraction is derived from *aptness*, or the degree to which the vehicle captures properties of the topic, rather than familiarity or conventionality. Apt metaphors are processed abstractly whereas inapt metaphors are processed primarily through comparison, which results in superficial properties. Thus, from this perspective, participants may have produced more semantically distant and abstract properties for familiarized metaphors due to increased perception of aptness, or ways in which the vehicle captured properties of the topic, rather than familiarization *per-se*. Consistent with this, Thibodeau and Durgin (2011) reported that aptness ratings increased as a result of in-vitro familiarization. Although this account does not lead to *a priori* predictions regarding familiarity in the same way as the career hypothesis, future experiments should nonetheless tease apart familiarity and aptness with regards to the abstraction effects observed here.

Conclusion

In this study, we sought to characterize the semantic representations underlying metaphorical abstraction. We used an in-vitro familiarization technique to demonstrate that figurative abstraction processes result in concepts forming new, more distant, and abstract semantic connections. Importantly, literal familiarization did not result in such changes. These results have implications for metaphor processing, abstraction, and long-term, polysemous connections in the semantic lexicon.

Acknowledgements

Generative artificial intelligence (Open AI's ChatGPT) was used to assist with Python coding for data pre-processing and R coding for statistical analysis and figures.

References

Al-Azary, H., & Buchanan, L. (2017). Novel metaphor comprehension: Semantic neighbourhood density interacts with concreteness. *Memory & Cognition 2016 45:2*, *45*(2), 296–307. https://doi.org/10.3758/S13421-016-0650-7

Al-Azary, H., & Katz, A. N. (2021). Do metaphorical sharks bite? Simulation and abstraction in metaphor processing. *Memory and Cognition*, *49*(3), 557–570. https://doi.org/10.3758/S13421-020-01109-2/TABLES/5

Al-Azary, H., & Katz, A. N. (2023). On choosing the vehicles of metaphors 2.0: the interactive effects of semantic neighborhood density and body-object interaction on metaphor production. *Frontiers in Psychology*, *14*, 1216561. https://doi.org/10.3389/fpsyg.2023.1216561

Al-Azary, H., McAuley, T., Buchanan, L., & Katz, A. N. (2019). Semantic processing of metaphor: A case-study of deep dyslexia. *Journal of Neurolinguistics*, *51*, 297–308. https://doi.org/10.1016/j.jneuroling.2019.04.003

Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T. (2017). Enriching Word Vectors with Subword Information. *Transactions of the Association for Computational Linguistics*, *5*, 135–146. http://arxiv.org/abs/1607.04606

Bowdle, B. F., & Gentner, D. (2005). The career of metaphor. *Psychological Review*, *112*(1), 193–216. https://doi.org/10.1037/0033-295x.112.1.193

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods 2013 46:3*, *46*(3), 904–911. https://doi.org/10.3758/s13428-013-0403-5

Cardillo, E. R., Watson, C. E., Schmidt, G. L., Kranjec, A., & Chatterjee, A. (2012). From novel to familiar: Tuning the brain for metaphors. *NeuroImage*, *59*(4), 3212–3221. https://doi.org/10.1016/j.neuroimage.2011.11.079

Gentner, D., & Bowdle, B. (2008). Metaphor as structure-mapping. In R.W. Gibbs, Jr. (Ed.), *The Cambridge handbook of metaphor and thought* (pp. 109–128). Cambridge University Press.

Gibbs, R. W., & Matlock, T. (2012). Metaphor, imagination, and simulation. *The Cambridge handbook of metaphor and thought*, 161–176. https://doi.org/10.1017/cbo9780511816802.011

Glucksberg, S. (2008). How metaphors create categories—quickly. In R.W. Gibbs, Jr. (Ed.), *The Cambridge handbook of metaphor and thought* (pp. 67–83). Cambridge University Press.

Glucksberg, S., & Haught, C. (2006). Can Florida become like the next Florida? When metaphoric comparisons fail. *Psychological Science*, 17(11), 935-938.

Jamrozik, A., McQuire, M., Cardillo, E. R., & Chatterjee, A. (2016). Metaphor: Bridging embodiment to abstraction. *Psychonomic Bulletin & Review 2015 23:4*, *23*(4), 1080–1089. https://doi.org/10.3758/s13423-015-0861-0

Kintsch, W. (2000). Metaphor comprehension: A computational theory. *Psychonomic Bulletin & Review*, *7*(2), 257–266. https://doi.org/10.3758/bf03212981

Martínez, G., Molero, J. D., González, S., Conde, J., Brysbaert, M., & Reviriego, P. (2024). Using large language models to estimate features of multi-word expressions: Concreteness, valence, arousal. *Behavior Research Methods*, *57*(1), 5. https://doi.org/10.3758/s13428-024-02515-z

McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods 2005 37:4*, *37*(4), 547–559. https://doi.org/10.3758/BF03192726

Online Etymology Dictionary, (n.d.) Key. In *Online Etymology Dictionary*. Retrieved February 2, 2026, from https://www.etymonline.com/search?q=key

Roncero, C, & de Almeida, R. (2015). Semantic properties, aptness, familiarity, conventionality, and interpretive diversity scores for 84 metaphors and similes. *Behavior Research Methods 2014 47:3*, *47*(3), 800–812. https://doi.org/10.3758/s13428-014-0502-y

Thibodeau, P. H., & Durgin, F. H. (2011). Metaphor aptness and conventionality: A processing fluency account. *Metaphor and Symbol*, *26*(3), 206–226. https://doi.org/10.1080/10926488.2011.583196

Utsumi, A. (2005). The role of feature emergence in metaphor appreciation. *Metaphor and Symbol*, *20*(3), 151-172.