RESEARCH ARTICLE

Diachronic data analysis supports and refines conceptual metaphor theory

Marie Teich[1,¤a]*, **Wilmer Leal**[1,2,¤b]*, **Jürgen Jost**[1,3,4]*

1 Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany, **2** Bioinformatics Group, Department of Computer Science, Universität Leipzig, Leipzig, Germany, **3** The Santa Fe Institute, Santa Fe, New Mexico, United States of America, **4** ScaDS.AI, Dresden, Leipzig, Germany

¤a American University Kyiv, Kyiv, Ukraine
¤b Department of Computer Science, University of Florida, Gainesville, Florida, United States of America
* marie.teich@auk.edu.ua (MT); wleal@ufl.edu (WL); jjost@mis.mpg.de (JJ)

Abstract

Our empirical analysis confirms the fundamental assumption in conceptual metaphor theory that metaphors are enduring linguistic and cognitive structures, not merely rhetorical figures. Complex systems tools identified a metaphor network with systematic separation of abstract and concrete categories, and two significant metaphorical processes: mappings from concrete to abstract topics and the emergence of new mappings between concrete domains. Metaphors concentrate on two small sets of everyday topics. One within the concrete group serves as both strong source and target domains, while the other, in the abstract group, primarily acts as targets. These findings can serve as confirmation that metaphor is a creative process emerging primarily from difference and tension between topics which allow (re-)conceptualizations and the display of new similarities.

Author summary

Metaphor theory examines how human thinking, conceptualizations and ideas, arise through new transfers of experiential and earlier acquired structures across thematic domains. These transfers are to some extent reflected in language change over time through metaphor conventionalization. We have carried out a systematic study of such metaphorically generated linguistic changes across the full thematic network of the English language. This analysis revealed several systematic features of metaphorical language change. Two metaphorical processes occur particularly frequently: Metaphors from concrete, often body-related topics, into more abstract topics, which in turn rarely act as metaphor sources, are the most common. The second process are metaphors that transfer and re-conceptualize words within the more concrete topics. Furthermore, we show that metaphorical connections between topics are stable in the long term and investigate the conceptual ordering of topics that results from their metaphorically related neighbors.

Introduction

Metaphorically speaking, there is a wide river between computer-based data analysis methods on one side and Cognitive Linguistics (CL) and especially Conceptual Metaphor Theory (CMT) on the other side. Every now and then, a small boat attempts to cross this river, but our goal is to build a solid and lasting bridge. Less metaphorically, we need and want to address two different research communities simultaneously, each with its own concepts, ways of thinking and arguing, and discursive practices. To metaphor research, we present a statistical, data-based investigation that empirically analyzes long-standing conjectures and provides the first-ever exploration of the systematic structure underlying metaphors. To the Complex System community, we introduce metaphor theory as a basis of meaning emergence that can be quantitatively explored and whose understanding and integration into computational linguistic methodologies hold great potential.

Cognitive linguistics and data analysis

Data driven linguistics is a very active and lively research field, but there exists a blind spot regarding the findings of Cognitive Linguistics [1,2] (CL for short). CL is an actively developing branch of linguistics without an established closed canon. It depends on several widely accepted premises. The most important one is a unity of language and thought. This implies that language is not produced in a separate linguistic module of the brain, but is instead processed and created by general cognitive mechanisms that are responsible for diverse cognitive tasks [3]. This has the important consequence that an understanding of linguistic structure offers a window into general principles of cognition.

Considering language in this way automatically links language research to many other areas of research. For example, with evolutionary anthropology that studies the development and role of human cognition and communication [4]. More importantly for the focus of this work, however, it connects language to human embodiment, since meaning emerges in connection with and from the physical reality of humans, rather than in abstract isolation [5]. Both of these points make computational analysis particularly challenging if its aim is to enable computers to "understand meaning" by analyzing the internal structures of language. The absence of the evolutionary functions of language and the absence of a body as well as of basal human experience mean that, from the perspective of Cognitive Linguistics, a computer can at best inadequately replicate the process of "understanding meaning". This is especially true for the most common computational linguistic method of natural language processing and large language models, where meaning of words is seen in its interrelations with other words according to the distribution hypothesis [6–8]. There, language is studied in isolation as an abstract system of signs, and text corpora are seen as containing meaning entirely within themselves. It is this, why studies of Metaphor in corpus linguistics are non-trivial. They can lead to statistical statements on the occurrence of metaphor [9] but corpus linguistics does not start at the premise of metaphor as a distinguished mechanism. Also studies in linguistic corpora in principle require preliminary delimitation of the analysis a priori, for example, restricting to certain source domains. An investigation across the entire language range, as it is aimed at here, is thus excluded. The approach presented here, on the other side can investigate metaphor across the entire English language and starts at the understanding of metaphor as a mechanism between two domains and from there develops it's further understanding.

Also, we keep our analysis within the framework of Conceptual Metaphor Theory without trying to simulate understanding through formalisation. The mathematical structures

we use do not have the purpose of imitating the structure of meaning, but only of enabling a corpus-based statistical analysis within the framework of an existing model.

Conceptual metaphor theory

A central component of Cognitive Linguistics is Conceptual Metaphor Theory. This theory does not consider metaphor as a rhetorical device in the sense of garnishing meaning packaging. Rather, metaphor is the central cognitive linking mechanism between concrete embodied experiences and abstract linguistic domains. Metaphor is not a simile, but a new transfer—hidden in its Greek root—of experiential structures into abstract topics. Because of the linguistic nature of metaphor, new abstraction can be achieved in linguistic forms, and metaphor, as the instance of transfer, can be seen as the original point of new meaning-creation.

Conceptual Metaphor Theory (CMT) emerged as a distinct field of research in linguistics with the publication of the book "Metaphors We Live By" by Lakoff and Johnson in 1980 [10]. In that book the authors put forward several main theses and developed a terminology. Essentially, the claim is that metaphors are not only a phenomenon of poetic language, but are ubiquitous in everyday language. Metaphoric transfers do not reside in individual rhetorical expressions, but systematically transfer structures between domains. With time, these mappings become part of our cognitive systems of thought and can give rise to a variety of individual metaphoric expressions. These systematic transfers are not comparisons of pre-existing structures, but constitute a model of the abstract domain in which it acquires its meaning. Since its establishment, CMT has undergone a rich development and has found application in numerous fields such as politics, literature, pedagogy, and transcultural studies [11]. However, a constant criticism concerns the lack of systematic research on the theory. The common research method of CMT studies particular language examples, supporting a separate thesis about the nature of metaphoric attributions. The selection of these language examples is based entirely on the researcher's linguistic intuition along with their specific goal, which the example is intended to illustrate. This methodological issue is referred to as the first major problem of CMT [12]. To date, there is no empirical study that reconsiders the theoretical foundations of CMT based on real linguistic data. This is particularly problematic because of the high flexibility of languages. A language forms a complex system with rules and tendencies on one side, but also expectations and competition between rules on the other side. Many phenomena can be discovered in such a system, but the appropriate tool to detect the guiding rules and main tendencies is statistical analysis.

There are attempts to perform corpus-based metaphor analysis with automated metaphor detection. Stefanowitsch and Gries [13] reviewed these attempts, but noted that in principle all attempts at automatic metaphor detection rely heavily on a presupposed metaphor theory. For example, one could search for typical metaphor source or target vocabulary or particular sentence structures, but this requires prior knowledge of the source and target distribution or of triggers indicating metaphors. Consequently, these methods are not immune to yielding highly biased results.

For this reason, this study takes a different approach to analyze a statistically significant set of randomly selected real-world metaphors. We chose to examine metaphors based on diachronic changes over the history of the English language.

Metaphor research through diachronic changes

Some frequently used metaphors experience an increasing conventionalization. They become so-called passive metaphors, where the original metaphoric use of the word is hardly

recognizable as such [14]. Typical examples of this stage are *"Time is running"* or *"Looking up to somebody"*. Continuation of this process can then lead to a new acquisition of word meaning or a transfer of meaning of a word from its source domain to its target domain. This type of exploration of metaphors has been successfully pursued by [15], for example. The examination of metaphors through etymological mappings assumes a homogeneous process of meaning transfer and conventionalization across all relevant source and target domains, contingent solely upon the level of metaphorical language usage. If this is the case, no bias arises from the chosen method, and etymology, like archaeology for a historian, grants access to the study of active language.

Computational methods vs. computational meanings

The metaphor of archaeology further allows us to illustrate an important difference that characterizes the method of this study. The sharp elaboration of this difference is not only central to metaphor research but a generally significant challenge for quantitative methods, especially in the field of digital humanities. It lies in the fact that the research method does not necessarily correspond to the nature of the subject matter and makes no direct statement about it.

It is important to avoid the misconclusion: If semantic novelty can be explored computationally, it is a computational process. If we take the view of metaphor that has been expressed in numerous philosophical contributions [16–19], as well as being influential in the debate on metaphor in cognitive linguistics, then metaphor is a point of semantic innovation. It is, therefore, not a process internal to language but a moment that transcends language. In [18] Paul Ricœur draws a very clear distinction between the living metaphor and the linguistic code. There, the living metaphor is primary and cannot be conceived without the language-transcending function of reference.

The resulting linguistic code consists of the archaeological artifacts of these linguistic events. The investigation of these artifacts, which we can undertake here, approaches the events themselves, i.e., the metaphors, not directly but indirectly. The results of this investigation are thus not an extension of metaphor theory itself, but an analysis of the artifacts that the metaphorical events leave behind in language, and against which the explanatory power of any actual metaphor theory can now be measured.

Similar is the relationship between this study and contributions to the debate on metaphor in cognitive linguistics. Here, the language-transcending moment, i.e., the moment in which something extra-linguistic has a formative influence on language, is often localized in the relationship between metaphor and motor-sensoric experiences. There are different opinions about the exact role of motor-sensoric experiences in the formation of linguistic forms. The debate ranges from perceptuo-motor simulation [20] to a continuous process of multi-modal experiences leading to cognitive representations that can then enrich language [21,22]. In an investigation of language itself, we can only speak of an indirect approach to this process. For this reason, this study does not take its own position in this debate and does not provide a view of what embodied metaphor means. It does, however, analyze the linguistic artifacts against which any such theory can be measured.

Structure of the paper

The paper presents various methodologically distinct analyses of the data from the Mapping Metaphor project [23] exploring different facets of metaphor formation. By adopting the Mapping Metaphor project, we thereby necessarily accept and use the classification of

metaphorical domains that is employed in that data set. Metaphor formation is conceptualized as a network, where topics serve as nodes and edges represent metaphorical connections from sources to targets. First, we investigate the source and target topics of metaphors. Analyzing the distribution of connections among topics provides a global understanding of the dynamics of the metaphor network. A second set of experiments examines the network's division into two anti-communities of topics. This is further refined by a motif and a connection type number analysis. In the third part, we explore a characteristic feature of the network, namely, preferential attachment on edges. The fourth part focuses on the connection between metaphor and semantics by identifying roles within the metaphoric network. A final, fifth part, entails a discrete curvature analysis, which also leads to a critical role analysis of spatial metaphors.

Results

Non-randomness of metaphorical data

We first look at the in- and out-degree distribution of the directed graph representation of the Mapping Metaphor data Fig 1. The plots already include a comparison with the mean of one thousand equally sized Erdős-Rényi graphs (Materials and methods) and one standard deviation up and down. Right away, we find strong deviations of the in-degree and even stronger ones for the out-degree distribution of the metaphoric network from the random case. This already points to the interesting potential of the data for metaphor research, as non-randomness implies structure.

For the Erdős–Rényi model the degree distribution concentrates around the mean connection number of slightly above 30 but for the metaphor data, this scale does not play a particular role. Instead, most categories cluster around low degree numbers and play only a minor role in the metaphor network. In contrast, there are a number of highly active categories, both for in- and out-degree, forming a so called heavy tail of the distribution. They show degree numbers several times higher than the degree maximally found in average random graphs.

Concrete and abstract categories separated in anti-communities

When we compare our metaphor graph with random graphs of the same size and average degree, we see striking differences and can identify important structural properties of the metaphor network. The latter's most marked structural property is that the majority of metaphors can be assigned to one of two classes that can be characterized by connectivity properties. The first class contains mainly spatial, mechanical, concrete and physically perceptible categories while the categories in the second class are mainly temporal, emotional and social. The two classes form anti-communities, meaning that the majority of connections is between instead of inside the classes, although the first class also contains a significant number of internal connections. Fig 2 shows that in comparison to configuration random graphs, there are two systematically dominating metaphor groups, one mapping from the first, more concrete anti-community to the second, more abstract one and another with mappings within the second anti-community, which shows the continuing re-determination and re-linking of concrete themes. The first group is the one most effectively described by CMT, while there is only little research and understand for the second group and it's role in conceptualization. The degree distribution in Fig 1 is heavy tailed. In fact, most categories have low degrees and do not contribute much to the network, while some few highly active categories have in- or out-degrees that are significantly higher than the maximal degrees found in random networks. The most active sources and targets do not match, however, as seen in the plot. The vertices at
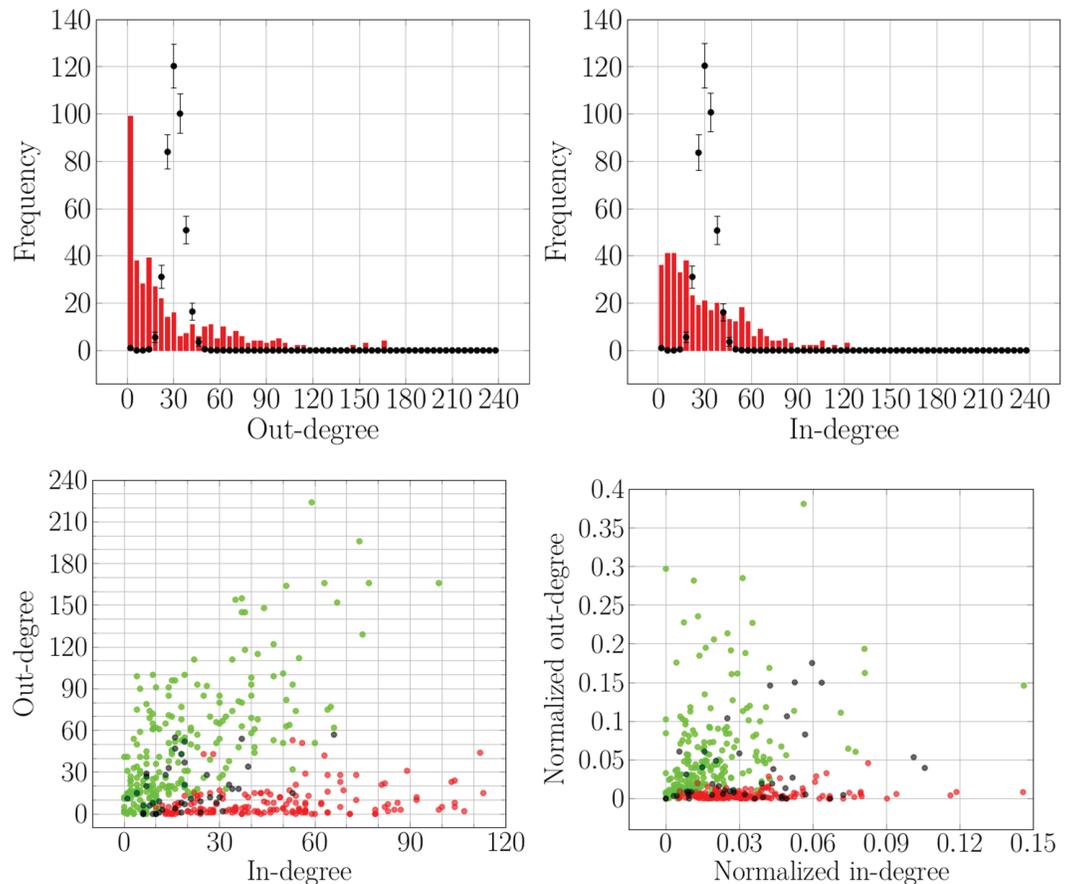
Fig 1. Top: Comparison of out-degree (left) and in-degree (right) distributions of categories in the metaphor network and configuration random model. Bottom: Absolute (left) and normalized (right) in- vs outdegree of categories in the metaphor network. Categories of the two anti-communities are distinguished by colour (categories not assignable to an anti-communities are black).

the bottom right of the graph have a high target probability, but low or at most average source activity. In contrast, the strongest sources also have a high target activity. These two different activity groups – one concentrated at the bottom right, the other on the diagonal – in fact coincide with the two anti-communities identified above. Almost all categories on the diagonal, regardless of their degree, belong to the concrete anti-community. The correlation coefficient for in- and out-degree of this set is 0.709. In contrast, for the abstract anti-community it is only 0.273, evidencing again the difference between the two classes.

The identification of anti-communities and the two particularly common metaphorical processes in relation to them represents an interesting discovery in the context of the embodiment of metaphors. The vast majority of bodily experienceable domains: "tangible," "visible," etc., are found in the first anti-community. One of the statistically relevant processes thus describes metaphors that move from bodily-related domains to less so. The second process, however, repeatedly describes new metaphors between bodily-related domains. In literature, there exists no unified theory of metaphoric embodiment. The concept itself describes the participation of the body and the experience of corporeality in the formation of metaphorical
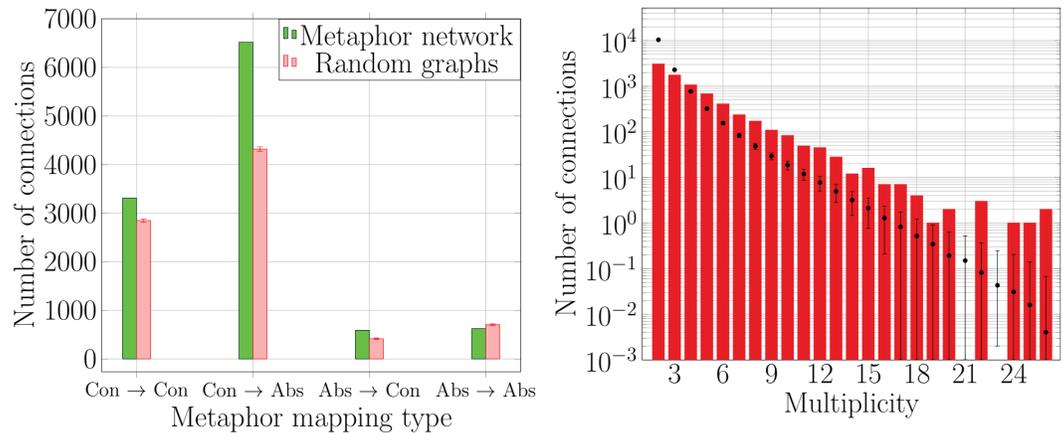
Fig 2. Anticommunities and metaphor persistence. Left: Connectivity patterns in the metaphoric mapping network distinguish two classes of categories that coincide with concrete categories (Con) and abstract categories (Abs). These classes form anti-communities, that is, most of the connections are between the classes and not within them. The bar chart compares the number of connections between classes observed in the metaphorical network (green bars) and the average number of connections observed in an ensemble of 1000 configuration random graphs (red bars with error bars). It shows that there are two systematically dominating metaphor groups, mappings from the concrete to the abstract anti-communities, and another with mappings within the concrete anti-community, which shows the continuing re-determination and re-linking of concrete themes. Right: Comparison of multiplicity values of edges between configuration random model and the metaphor network.

meaning. It roughly contradicts a view where metaphors are pure transfers of abstract cognitive representations [24]. However, there are unanswered questions as to whether metaphor processing involves a one-to-one reactivation of physical sensory perception [20] and, if not, how exactly the continuum from integrated multimodal bodily experience to the imagination and the experience of oneself as mind can be understood [21,22]. The frequency of metaphors within the first anti-community can contribute to further expanding the understanding of embodiment. Here, existing bodily experience is not simply integrated into an abstract domain. Rather, it could point to a multidirectional interweaving of imagination and experience. How bodily perception itself takes place does not remain unchanged but is constantly being linguistically reconnected in relation to other body-related domains. Possibly, the creation of meaning is not so much or not only reached by connecting abstract spaces to experience, but rather a re-experience of the concrete. In this way, metaphor would not only be embodied, but would also be an instrument of integration for our constantly renovating embodiment.

Word transfer density

In order to examine the word transfer density, we now consider the multigraph where each transferred word creates an edge between the corresponding categories. We then normalize the degrees by dividing by the total number of words in each category. Note, however, that this will in general underestimate the word transfer probability, in particular for very active mappings.

We use the same colors for the anti-communities as above in Fig 1. With the expectation of the category "light" all topics with high in-degree density probability belong to the abstract anti-community and accumulate in the lower right corner, as before. The words in the other class, the concrete categories, the predominant sources, however, do not typically

occupy a position on the diagonal, but are just active sources, but not preferred targets. So instead of a diagonal slope, we find that the probability of words in a category being involved in metaphorical mappings may be high either exclusively as a source or as target depending if the category belongs to the concrete or the abstract anti-community.

Motifs and the metaphorical mechanism

The motif count reveals further structure (Table 1). We find five subgraphs of size $n = 3$ with Z-values either above 2 or below –2, and in four cases, the absolute Z-value is even higher than 6. The three motifs of size 3 with one missing edge are clearly more frequent. In contrast, closed triangles are almost absent in the data. Both effects, however, disappear when we restrict the analysis to the internal connections of the anti-communities.

For motifs of size 2, the deviation from randomness is even more pronounced, as symmetric connections dominate. This is true within both anti-communities, however to a much smaller extent within the abstract one. Accordingly, any metaphorical mapping from source to target makes the reverse mapping from target to source significantly more likely.
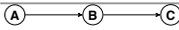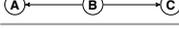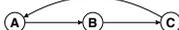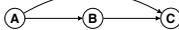
Persistence of metaphorical mappings

When looking at the multiplicity of edges, that is, the number of words transferred between categories, another deviation from a random model emerges (the black lines show the distribution for one thousand random configuration graphs with one standard deviation up and down for each data point.) As can be seen in Fig 2 in the right plot, there is a significantly higher probability of transferring words along already established mappings, so that vertex pairs with a high number of transferred words become significantly more frequent than in random models.

Such a deviation can always be a property of the methodology of the Mapping Metaphor data collection. It may also ultimately be the uneven distribution of a natural connection affinity between pairs of topics. The systematics of the deviation, however, primarily carries the properties of a preferential attachment mechanism on mappings, meaning that an already established metaphorical relation facilitates further word transfer along this relation.

Metaphor and semantics

We now turn our attention to the role that categories play in the metaphor network, which is given by the connections they form. In social network analysis, objects are considered to play

Table 1. Non-symmetric three-Vertices motifs with corresponding z-value. Two-vertex motifs with corresponding z-value. For complete network, only connections within concrete topics block and only connections within abstract topic block.

Metaphor mapping motif	z-value overall	z-value concrete	z-value abstract
(A) → (B) → (C)	6.34	2.88	0.22
(A) → (B) ← (C)	6.43	-0.55	-0.73
(A) ← (B) → (C)	6.44	0.42	-0.81
(A) ⌢ (B) → (C)	-2.46	-2.66	-0.65
(A) ⌢ (B) → (C)	-6.36	-0.52	0.71
(A) → (B)	-9.53	-7.21	-2.92
(A) ← (B)	9.53	7.21	2.92

the same role if they have the same incoming and outgoing neighbors. To examine the roles within the metaphor network, we employed a Hierarchical Cluster Analysis (HCA) algorithm on the collection of categories and their incoming/outgoing neighbors. The distance between categories was measured by the number of non-common connections they possessed. The output of the HCA is a binary tree called dendrogram, whose clusters (branches) represent roles in the metaphor network.

Real-world data sets often feature multiple equidistant objects, leading to numerous clustering solutions and consequently, a multitude of role schemes [25,26]. However, the HCA algorithm yielded only four dendrograms, one of which is illustrated in Fig 3.

These four dendrograms encompass a total of 421 distinct graph clusters, defined as all the different subtrees that can be extracted from the four dendrograms. In S2 Fig, it can be observed that 401 graph clusters, the majority of them (95.3%), consistently appear as subtrees in all four dendrograms. The remaining twenty graph clusters, which are not present as subtrees in all four dendrograms, can be categorized as follows (S2 Fig): the four entire dendrograms (cluster size equal to 415), eight relatively small-sized subtrees (approximately 20 leaves) appearing in two out of the four dendrograms, and twelve subtrees ranging from 8 to 18 leaves, which appear in only one dendrogram.

Additionally, considering clusters as sets of leaves rather than graph structures provides an alternative perspective [25]. For instance, although the binary trees $((a, b),c)$ and $(a,(b, c))$ are distinct graph structures, they share the same set of leaves, namely, a, b, c. Similarly, the four dendrograms obtained share the same set of leaves, making them stable clusters as sets but not as graphs. In this context, we computed the clusters based on their sets of leaves, disregarding the hierarchical graph structure, and determined that there are 410 set clusters, which is 11 fewer than the number of graph clusters. Furthermore, 406 set clusters (99.0%) consistently appear as subclusters in all four dendrograms, while the remaining four set clusters appear in two dendrograms each.

In summary, the results obtained from both graph and set clusters indicate that despite being merged differently during the grouping process, leading to distinct graph structures, the semantic roles exhibit remarkable stability. The similarity between graph and set structure additionally points to the robustness also of the more defined graph structure that was obtained by the HCA.

Through evaluating the metaphor distance as a similarity measure, we aim to assess its ability to provide insights into the semantics of categories. Comparing the metaphor-based similarity distance of all categories with cosine similarity and the euclidean distance derived from fasttext word embeddings reveals a low correlation. The cosine similarity yields a value of –0.24, while the euclidean distance results in 0.23. Similarly, mutual information is present but not dominant, with values of 6.24 for both cosine similarity and euclidean distance. However, the stability of the resulting dendrograms and the consistent ordering of categories vividly demonstrate that the metaphor-based role of a category encodes a semantic position, contributing to the formation of a semantic structure. This structure predominantly goes beyond the semantic information contained in word embeddings, encompassing figurative and conceptual semantic similarities. Consequently, while the acquired semantic structure does not completely diverge from word embeddings, it exhibits semantic aspects that surpass them, underscoring the limitations of word embeddings from a broader semantic perspective.

The role of space and the global network structure

A common and widely accepted assumption in the CMT literature is the specific role of spatial and directional domains for metaphor mappings. In the original set-up of the theory,
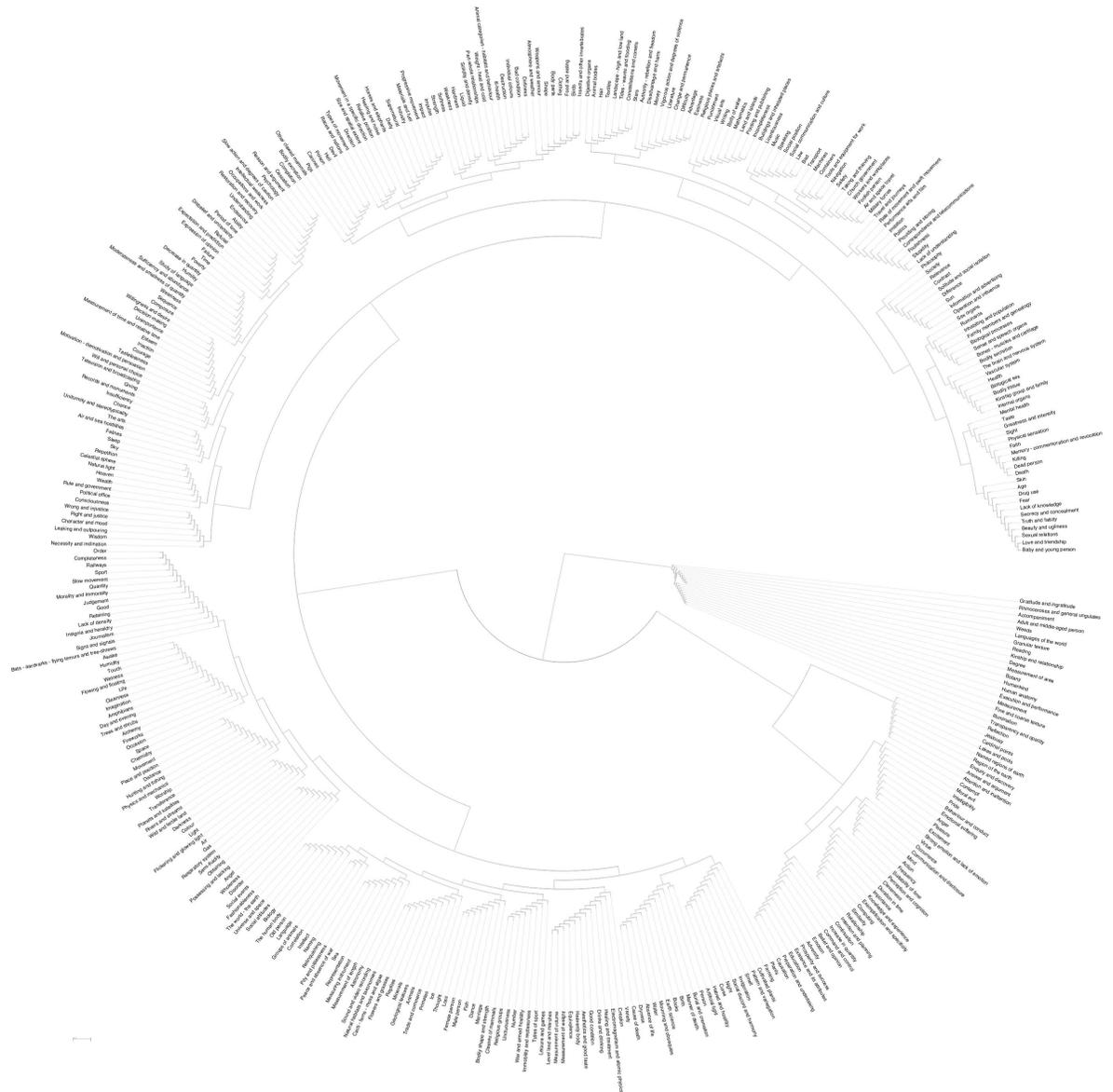
Fig 3. Hierarchical representation of semantic roles in the metaphor network. The figure is one of the four dendrograms obtained from running a Hierarchical Cluster Analysis algorithm on the set of categories from the metaphor network. Categories were represented by their incoming N_{in} and outgoing N_{out} neighborhoods, the distance between categories c and c' was computed by counting the number of non-common connections: $d(c, c') = \frac{\#(N_{in}(c) \ominus N_{in}(c'))}{\#(N_{in}(c) \cup N_{in}(c'))} + \frac{\#(N_{out}(c) \ominus N_{out}(c'))}{\#(N_{out}(c) \cup N_{out}(c'))}$, and Ward's method was used as a grouping methodology.

https://doi.org/10.1371/journal.pcsy.0000058.g003

Lakoff and Johnson elaborated on the so-called orientation metaphors. These do not structure or explain a single abstract concept, but rather perform the higher-level task of organizing the relations between different abstract concepts in terms of the spatial domain. Later, space was increasingly ascribed a special cognitive role in language. Spatial thinking was seen as the platform for schema formation—the universal building blocks of language. Thus, the spatial

domain was considered to be the fundamental source of all concrete structures [27]. Such linguistic considerations went along with neuroscientific research about the role of spatial thinking in the hippocampus for memory and structural thought [28] culminating in an actively pursued research program on cognitive spaces [29,30].

However, none of the analyses conducted on the mapping metaphor data showed a central, highlighted role of the spatial and directional categories in the metaphor network. According to the theory of orientational metaphors, one might expect that mappings starting from spatial categories would influence their neighboring mappings, resulting in higher outward transitivity. In fact, the outwards transitivity of *relative position* and *movement in a specific direction*, is relatively high, but it does not exceed a general tendency found in the graph. This general tendency can be observed in Fig 4, where one sees a clear correlation between the out-degree and the outwards transitivity of the vertices. Apart from being an interesting fact about the metaphorical data in itself, this correlation also shows that spatial regions follow this relation but do not exceed it. They do not play a special role in terms of their transitivity and thus their direct influence on their neighborhood in the network.

In the outlook of further possible special positions of the spatial domains and also with the aim of better understanding the general structure of the network, we consider two types of discrete curvature in the multigraph representation. The statistics of local graph curvatures are known to be good indicators of global network properties. The Forman-Ricci curvature captures the divergence or spreading in a network. Again, our analysis of the metaphor data show systematic differences from random graphs or other real world data sets. In particular, we see a rapid decrease at negative values, meaning that there are only very few edges with strong branching at their starting and/or terminal vertices (Fig 4). Such edges would provide rapid flow access to many different metaphors, but this is definitely not the structure of the metaphor network.

The Ollivier-Ricci curvature, in contrast, captures the presence of triangles and quadrangles in the network and thus the density of local neighborhoods. The existence of several seperated regimes of edges in the distribution indicates density clustering within the network on different scales. In the metaphor data, we do not find such dense local clusters separated from others as seen in (Fig 5).
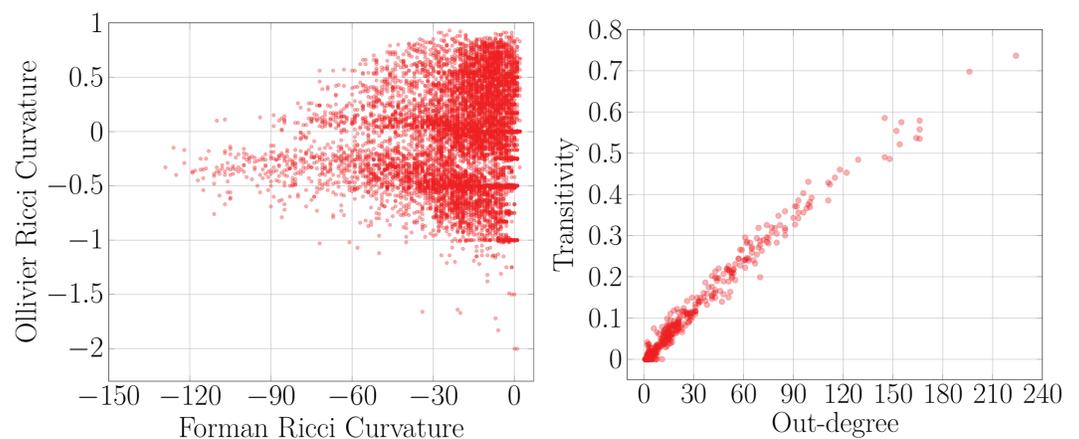

Fig 4. Left: Distribution of Forman Ricci vs Ollivier Ricci curvature of edges in the metaphoric network. Right: Out-degree vs transitivity in the metaphor network.
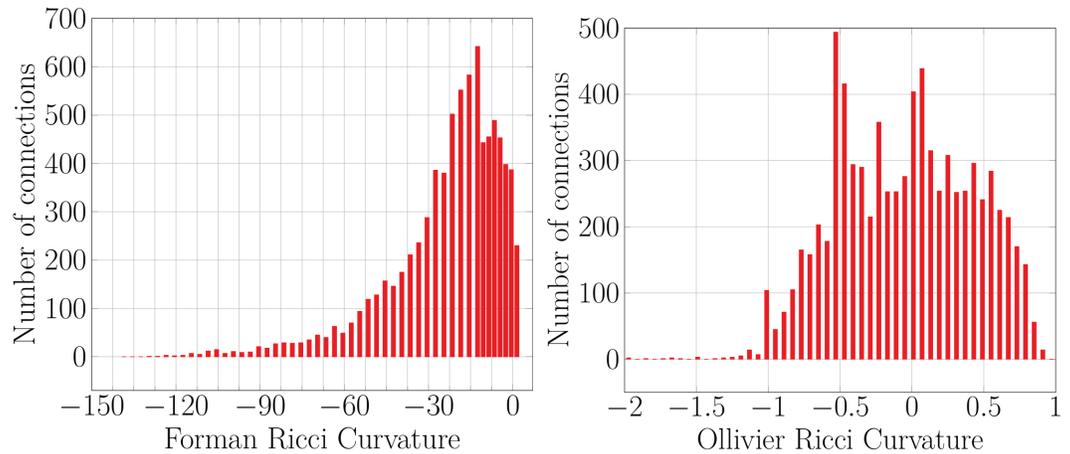
https://doi.org/10.1371/journal.pcsy.0000058.g004

Fig 5. Distribution of Forman Ricci (left) and Ollivier Ricci (right) edge curvature in the metaphoric network.

Discussion

In this study, we have used a wide range of data analysis methods to systematically infer the properties of metaphor. We explore data of diachronic language change that resulted from metaphoric language use to draw back conclusions on the metaphoric language change itself. In doing so, we automatically rely on both the understanding of metaphors, which went into the data selection, and the categorization of the data.

On one hand, the data set that we have employed and the results derived from it support a basic assumption of CMT that metaphor is a phenomenon of everyday language and that the choice of metaphor sources is influenced by experience and human embodiment. On the other hand, we also find global features of the metaphor network that have not been observed before and sometimes even are at variance with some claims of CMT. Summarizing our findings, we find two classes that while forming anti-communities, behave quite differently with respect to their source and target activities. One class consists of the most active metaphor sources: simple and concrete topics such as 'Food and Eating', 'Shape' and 'Relative position'. Their source productivity is simultaneously correlated with a high target activity. The other, structurally very different class contains the strongest metaphor targets. It contains mostly common but abstract topics such as 'Bad', 'Emotional Suffering' and 'Literature', which are little active as metaphors sources. The word density of the sent and received words is anti-correlated. This is very different from random null models. Active topics feature either a high density of sent or of received words. These effects are even stronger for the out-degree distribution than for that of the in-degrees.

Moreover, on the basis of the systematic data of the Metaphor Mapping project, we do not see any special role of spatial and directional domains in the metaphor network, in contrast to a central claim of CMT. In fact, we do not see any category that would occupy a central position in the metaphor network, nor any hierarchies or clustering beyond bilateral edge relationships. Of course, it remains to be seen whether that can be confirmed by other data sets, but we are not aware of any other data set appropriate for a quantitative statistical analysis.

The division into the two anti-communities and the structure of the relationships between them seem to be the only clear global structure of the network. In agreement with an assertion of CMT, mappings from the concrete anti-community to the abstract one play a prominent role. In addition, however, we also see many metaphorical relations between different

concrete domains, indicating a rather fluid structure of that class. Even concrete terms do not make sense in isolation as their meaning conceptualization is embedded and actively renewed in a conceptual network. The consequences of this finding remain to be systematically explored and evaluated.

A review of the two anti-communities shows that the concrete one includes all bodily, spatial and geographic domains. The abstract topics include social constructions and spaces, emotion related domains and, remarkably, temporal domains. The latter is in agreement with the philosophical conception of time as subjective time. Only the category 'period of time' belongs to the concrete anti-community.

And we have found that the metaphor network is globally strongly anti-transitive, as follows from the division into two anti-communities. Within each anti-community, however, such a strong anti-transitivity is not present, nor is the opposite, transitivity. From these findings, we conclude the mapping process is not triggered by a pre-existing structural analogy between two categories [31]. The reason might be an omnipresent potential for analogy in general. But it rather seems in the opposite, that it is a tension between two categories that invites metaphors and thereby reveals structural similarities.

This is further confirmed by the high presence of symmetric metaphorical links between different domains. In more detail, there seem to be at least two possible explanations for the prevalence of symmetric connections. The restructuring of the target domain may make it more likely to become a source of a metaphor in the reverse direction. Or a metaphor may enable the target to re-conceptualize the source domain. This feedback effect on the source may act like a seed for mappings in the opposite direction. That second possibility is supported by the global anti-transitivity that we have observed. The symmetrical component of metaphor surpasses the purely unidirectional understanding of metaphor in CMT.

The experiment in which the number of words per mapping is compared with configuration graph models provides, as far as we know, the first systematic and data-based evidence for persistent metaphoric mapping structures. It thereby confirms CMT's thesis that metaphoric connections are part of the cognitive structure, as opposed to the view that metaphors are individual rhetorical figures. To make this claim, we rely on the basic assumption of Cognitive Linguistics that the mechanisms of language are grounded in cognitive mechanisms. The experiment itself makes a statement on diachronic language evolution and shows that well-established mappings in a speech community also shape future language development. Within the paradigm of Cognitive Linguistics, we can conclude that the language we use, especially the well-established metaphorical mappings of that language, shapes our cognitive system and increases the likelihood of the emergence of new meanings along the typical mappings in our language.

The absence of any flow or clustering or other emergent structure beyond bilateral connections suggests that there is neither a basic source domain from which most metaphors ultimately emerge nor a basic mechanism from the visual or any other domain that is the ultimate basis of conceptual thinking. Rather, there is quite a number of domains, related to different senses, that are productive as metaphor sources. We conclude from this the assumption that new metaphors can nevertheless emerge between every pair of topics with different probabilities following the tendencies of the network.

An important part of our research was devoted to the development of a semantic structure based on the metaphor network. Existing approaches to formalize a semantic structure are mainly word embeddings in the field of computational linguistics and, on the other hand, frame semantics in the framework of the Cognitive Linguistic approach [32]. Our approach is a substantial complement to these two existing approaches. Similar as in word embeddings,

semantic distances are used, but as shown by the comparison experiments, the semantic distance defined by the role in the metaphor network is only weakly correlated and exhibits little mutual information with the semantic distances used in NLP. The crucial difference of our approach to frame semantics is that frame semantics does not provide a method to establish the relationship between topics organized in frames. We can fill precisely this gap by finding that the roles of each category created by all metaphor connections place the categories in a stable semantic net of relations. This net reveals not only the thematic, but above all the conceptual and figurative proximity between the categories.

There are many interesting examples of such figurative and conceptual proximity visible in the dendrograms presented. One of these is the proximity between *dryness* with *absence of life* on the one side and *flowing and floating*, *wetness*, and *humidity* with *life* on the other side. This double-sided proximity has no thematic share, but reproduces the deeply rooted conceptual proximity between fluidity and life, not only through direct metaphorical mappings, but even in an established parallel semantic structure.

Another interesting figurative aspect is revealed in the arrangement of emotion categories. There is a larger collection of strongly moving emotions: *strong emotion and lack of emotion*, *excitement*, *pleasure*, *anger*, *emotional suffering*, *pride*, *contempt*. Additionally the categories *intelligibility*, *behaviour and conduct* and *moral evil* found their way into this group of figuratively similar domains. But other emotions, that are figuratively less active, are not contained in this group, like *fear*, *esteem* and *jealousy*.

Another characteristic observation of the results concerns different categories of movement or motion. These different categories are not grouped together, but on the contrary seem to mark a motif of different related groups. One motion motif is the category *continuation* which finds itself in the group of categories: *intention and planning*, *increase in quantity* and *command and control*, all controllable or constant processes. Similarly the category of *slow action and degrees of caution* finds itself in a set of categories connected to thought processes and continuous change like *completion* and *cessation*. Another motion motif is *repetition*, that is not unexpectedly grouped together with *sky* and *celestial sphere*. *movement in a specific direction* is related to several spatial domains, like *relative position*, *size and spatial extent* and *direction*. Finally *progressive movement* finds itself close to *impulse and impact* and is connected to the topics *industry* but also of *deity*, probably through conceptualizing the scheme of first causation and beginning.

This is only a glimpse of the discoveries that can be made by studying the acquired semantic structure from the perspective of different disciplines, such as psycholinguistics and psychology. Moreover, the metaphor role distance measure can also be used as a tool in discourse analysis. A change in framing can be classified as a more nuanced change if the old and new frames are from nearby categories, but is profound if the categories are not closely related in the semantic structure presented.

This leads to an important task for future research, to combine the information about semantic organization presented by the method of this work with the information provided by word embeddings. This would enrich natural language processing limited through the distributional hypothesis by the figurative and conceptual aspect of semantic distances.

Materials and methods

Mapping metaphor data

We conducted our research based on the data of the Mapping Metaphor project [23,33], a collection of etymological transfers caused by metaphorical mappings. The collection, conducted by the School of Critical Studies of Glasgow University, is based on the Historical

Thesaurus [34], where each entry is assigned to one of the 415 thematic categories (e.g. *movement*, *visual arts*, *weapons and armour*, etc.). The code used to carry out the computational methods is made available at [35].

To detect metaphorical mappings, a computer program searches for words that appear in two or more categories. It is then manually inspected if these lexical overlaps can be seen as metaphorical usage. If this is the case, an entry is created with the transferred word, the source and the target category. The entry also includes the year in which the transferred meaning was detected for the first time as well as the last year of appearance if the usage has not been preserved until today. For our purposes it is important to note that the search for metaphorical mappings was systematically performed, and that not all existing etymological mappings have been included. In particular, since the project's main goal was to detect all possible connections between pairs of categories rather than each possible word transfer, less emphasis was placed on further findings along the same connection if several words had already been registered between a particular pair of categories. Another relevant point is that the categorisation of the lexical items does not coincide with the notion of domain from Conceptual Metaphor Theory. Rather, most categories are broader and can be considered as associations of multiple, thematically close domains. Using the Mapping Metaphor data forces us to adopt its categorisation. Therefore we will mostly use categories in the sense of the Mapping Metaphor project, which allows only for a restricted transfer of the results into the CMT framework.

The etymological circumvention is a methodological choice with several consequences to consider. Since the data, and specifically the categorization, are primarily based on the historical thesaurus and not on a project that aimed to create a metaphor database from the outset, the data is not fundamentally based on a metaphor theory. Only the Mapping Metaphor project requires a metaphor theory to be carried through; however, the influence of this metaphor theory is limited by partial automation in the data annotation process. Still, the criteria used by the MappingMetaphor project to check automatically found words for historical metaphorical use have an uncertain influence on the results of this study. These criteria are not clearly identifiable in the Mapping Metaphor project. Since we question the fundamental assumptions of any metaphor theory and aim to test them empirically, uncertainty regarding the criteria under which words appear as metaphors in the dataset is a significant potential source of error. In this way, the systematic characteristics of metaphorical transfers that are being identified may in part be artifacts of the selection process used in the Mapping Metaphors project.

More generally, the present study is highly dependent on the structure in which the limited Mapping Metaphor data is available. This has the general disadvantage that the results of other studies based on different data cannot be directly compared, tested, and combined with the present results without interpretative steps. A major such limitation is that the data in this structure is only available in English. This means that the large number of important questions that could be asked in a cross-linguistic comparison cannot be investigated.

Ultimately, this data structure only allows a limited view of the metaphorical process. Many aspects which are of interest for metaphor study are not captured in this data set and can thus not directly be inferred about by its analysis: contexts and sentence structures in which the transfer of words occurs; peculiarities of oral metaphors that do not play a role in the dataset; different conceptual mappings between the same domains and the dynamics between these mappings; transfers of partial aspects of domains and combinations of mappings in sentences and speech—to name just a few of the aspects of metaphor that cannot be examined here.

Mapping metaphor network

As the data contains pairwise relations between source and target categories, it very naturally qualifies for a graph representation. A graph in fact represents a set of elements that can be pairwise related to each other. The crucial gain here is that graphs are mathematically very well studied objects for which a richly developed research toolbox already exists. Moreover, this toolbox can be adapted and extended to particular problems in the existing formalism [36].

In the present paper we use two different graph representations of the data, as directed graphs or multigraphs, which will allow us to tackle different research questions.

Directed graph representation. When focusing not on single words, but on the network of mappings between the categories, we can use the following definition.

Definition 1. *A directed graph $G = (V, E)$ consists of a set of vertices V and a set E of ordered pairs (u, v) of vertices called directed edges. We will refer to them simply as edges. Furthermore, u and v are called the source and target of the edge (u, v), respectively.*

In our application, the vertices represent the topic categories from the Mapping Metaphor data. If the data contains at least one word transferred from topic u to topic v, this transfer gets represented by the edge (u, v). This representation ignores the numbers of words transferred. When we also want to include that information, we use the directed multigraph representation.

Directed multigraph representation. Instead of focusing only on the categories as closed units and the connections between these units, we can consider categories as forming sets of words. Each transferred word then induces an individual connection that should be represented by an edge. In this case our representation have to account for multiple edges having the same source and target.

Definition 2. *A directed multigraph $G = (V, E)$ consists of a set of vertices V and a multiset E of ordered vertex pairs.*

Now, different words that are transferred from the same source to the same target are accounted for. Therefore, we do not loose the information about the number of words that were found.

In- and out-degree

The degree is a basic concept of graph and network theory. It is defined for each vertex as the number of edges it participates in. For the directed graph and multigraph we distinguish between in- and out-degree. The out-degree, as the number of outgoing connections, counts all edges in which a vertex appears as a source. Similarly, the number of edges which include a vertex as a target is the number of incoming connections to this vertex.

These simple notions allow us to identify local characteristic properties of the individual categories. On the other hand, the statistics of their distributions also provide global information about the structure of the network, by comparing them with the corresponding distributions of random graph models. That is, we can infer to which extent our networks show, or not, specific structures and features present in random models. This will be a general principle in our graph analysis.

In the multigraph case, where the categories are not viewed as uniform entities, but as word sets of different sizes, it is also interesting to look not only at the number of transferred words, which corresponds to the degree, but also at the transferred word density. This can be achieved by normalizing the in- and out-degree by the category size.

Random graph models

Directed Erdős–Rényi model.   To get a first idea about the metaphoric network, that is the graph from the Mapping Metaphor data, we compare it with the most simple random directed graph model, the Erdős–Rényi model [37].

Definition 3. *In the directed Erdős–Rényi model, a graph is chosen uniformly from the collection of all directed graphs with a given number of vertices n and edges N. This means that all edges can appear with the same probability.*

Directed configuration model.   The configuration model describes random graph models with a prescribed degree distribution. In the directed case, not only the in- and out-degree distribution are fixed but also the degree sequences and by this way, the in- and out-degree combination for each vertex. The outgoing and incoming edge stubs then are randomly fused [36].

Definition 4. (Directed configuration model). *Let $n \geq 1$ and $N \geq 0$ be a fixed number of vertices and edges. We denote by $i = (i_1, i_2, \ldots, i_n)$ and $o = (o_1, o_2, \ldots, o_n)$ the in- and out-degree sequences of the n nodes. It holds that $\sum_{v=1}^{n} i_v = \sum_{v=1}^{n} o_v = N$. To each vertex v, we attach i_v ingoing and o_v outgoing edge connections such that each incoming stub is matched with equal probability to each outgoing stub. This way the probability for an edge to exist going from some vertex v_1 to v_2 is proportional to*

$$p_{v_1,v_2} \sim i_{v_1} \cdot o_{v_2} \tag{1}$$

The following algorithm implements a configuration model starting from a real world data model and keeping its degree sequences.

Two random connected pairs $(v_1, v_2), (v_3, v_4)$ are cut and the connections are swapped, becoming $(v_1, v_4), (v_3, v_2)$. This step is repeated a given number of times k. The output network approaches the directed configuration random multigraph. One can also prevent the swapping in case one of the edges $(v_1, v_4), (v_3, v_2)$ already exists. This modification of the algorithm makes the resulting network a **directed configuration random graph** instead of a multigraph. The drawback is that one is no longer able to analytically compute the edge probability, but this modified algorithm will be important for some of our experiments presented to create comparable random graphs [41].

Anti-community detection

The principle of anti-communities is exactly the opposite of clusters, to minimize the internal connectivity and maximize the connectivity across in- and outside of the group. Accordingly, the search can be performed using the negative graph in which disconnected nodes are connected to each other and vice versa. On this one of the well developed cluster detection algorithms can be applied. We use the Louvain cluster detection method based on greedy modularity maximization:

$$Q = \frac{1}{2m} \sum_{i,j} \left[A_{ij} - \frac{k_i k_j}{2m} \right] \delta \left(c_i, c_j \right), \tag{2}$$

where A_{ij} is the Adjacency matrix entry of the graph, k_i, the degree of node i, m the total number of connections and c_i the cluster of node i while the Kronecker-delta function equals one

only if $c_i = c_j$ and 0 otherwise [38,39]. For each iteration, the algorithm produces slightly different results. In our case, the vast majority of categories are stably assigned to the same cluster. We run the assignment 10 times and keep only the overlapping sets as anti-communities while categories that change their assignment at least once are placed in the remainder set.

Edge preferential attachment

In the directed multigraph case, we can not only count the in- and out-degree for each vertex, but we can also count the edges between each two connected categories in each direction. For our data, this represents the number of words that were found to be mapped from a specific source category to some specific target.

The purpose of counting words along connections is to test whether metaphorical mappings do form persistent structures in language and mind, which is one of the main claims of Conceptual Metaphor Theory. In the Aristotelian view of metaphor as an individual rhetoric figure, individually mapped words should not influence the probability of other words being mapped along the same connection. In CMT, on the other hand, the phenomenon of metaphor is rooted not in the individual words and expressions, but in the conceptual transfer between two domains. Further, this transfer structure does not arise in the short term and temporally for production or processing of metaphoric expressions. Instead, it forms a persistent cognitive structure, which makes the transfer of many different partial aspects possible and becomes activated again and again. Once such a transfer structure gets established, the probability of further metaphoric expressions making use of this connection should increase.

The case of the random configuration model thus corresponds to the Aristotelian case. Each new formed edge of the multigraph is equally likely to form, given the degree sequences, and does not depend on previously existing identical edges. On the other hand, in the case of persistent mappings, the probability to form equal-directional edges between already connected vertices should increase.

It is possible to formulate exact models for which the probability of a new edge from u to v has a term proportional to the already existing number of edges along this possible connection:

$$p_{u,v} \sim w_{u,v} \tag{3}$$

and by comparison with the data, measure the influence of this term. However the collection method of the Mapping Metaphor data does not allow for this, as exploring further words along already detected connections has been systematically reduced. The aim of the project has not been to find all possible metaphorically transferred words but to focus on finding the mappings themselves. Thus, a quantitative examination of persistence is not possible. But we can test whether, despite the systematically lower word count, a confirmation of persistent metaphorical mappings can be found. This is done by comparing the distribution number of edges along a potential connection of the data against configuration random graph models (see).

Motif analysis

The analysis of graph motifs is a method for investigating the nature of processes behind the connections of a graph [40]. Graph motifs are defined in terms of small subgraphs that appear much more or much less often than expected from comparable configuration graphs. They reveal the logical properties according to which connections in a network emerge, and also those that do not play a role. We can then identify those different processes or mechanisms behind connection formation that possess or lack these logical properties.

We will conduct a motif analysis on the network to test if the two such logical properties drive the formation of the metaphoric network, namely, transitivity and symmetry.

Transitivity. Transitivity can be tested by looking at motifs of size $n = 3$ that hold the following property:

$$u \to v, \quad v \to w \quad \Rightarrow \quad u \to w, \tag{4}$$

where u, v and w are categories of the metaphoric network and $u \to v$ represents an edge between two categories u and v. Transitivity is a logical property of every equivalence relation, for instance of analogy.

We computed the *outward transitivity of a vertex u* as the fraction of transitively closed directed triangles starting and ending at u, e.g., $(u \to v, v \to w, w \to u)$. In other words, starting at one vertex, we consider the fraction of triangles that are transitively closed.

Definition 5. (Outward transitivity). *Let $G = (V, E)$ be a graph, H_v the set of all vertices with an outgoing edge from v, and let K be the set of paths $(v \to h, h \to w)$ of length two, with $w \neq v$. The outward transitivity T(v) of v in G is given by:*

$$T(v) = \frac{\#\{w \to v \in E : (v \to h, h \to w) \in K\}}{\#K} \tag{5}$$

Symmetry. In a directed graph $G = (V, E)$ we can distinguish between symmetric and non-symmetric relations. A relation is non-symmetric if $u \to v$ is an edge of the graph, but $v \to u$ is not, while a relation is symmetric if both $u \to v$ and $v \to u$ are in E. This property can be expressed as:

$$u \to v \quad \Rightarrow \quad v \to u. \tag{6}$$

Consequently, symmetry can be quantified by computing the fraction of motifs of size $n = 2$ that hold 6. To find all motifs of size $n = 2$ and $n = 3$ in order to check the metaphorical mechanism for symmetry and transitivity, we make use of the motif detection software by [41].

Counting motif appearance. Of course, when looking at our metaphorical mappings data, we cannot expect that transitivity and symmetry always hold, in particular since the topic categories are rather broad and most probably combine several semantic domains so that different mappings can involve also quite distant aspects of one and the same category. Thus, we rather study tendencies, that is, a significant over- or under-representation of particular subgraph motifs when compared to random models.

To find out which subgraphs are frequent or rare motifs of a network, we need the following definitions [42]:

Definition 6. *A **subgraph** of $G = (V, E)$ is a graph $G' = (V', E')$ that consist of a subset V' of V and a connected subset E' of edges whose nodes belong to V'. Furthermore, G' is called an **induced subgraph** of G if all edges $(u, v) \in E$ such that $u, v \in V'$, are also in E'.*

Definition 7. *Two graphs $G = (V, E)$ and $G' = (V', E')$ are called **isomorphic** if there exists a bijection f between V and V' that preserves adjacency, that is, the bijection f is an isomorphism of G and G' if $(u, v) \in E$ implies that $(f(u), f(v)) \in E'$.*

Isomorphic subgraphs represent of course the same motif. To count how many times a motif of n vertices appear, one checks all non-identical connected combinations of n vertices, including the overlapped ones. The exact counting algorithm can be viewed in [41]. To assess the count $N_{\text{real}}(\Omega)$ and decide if a subgraph Ω is significantly over- or underrepresented in

the metaphor network, it is compared with the count of the same subgraph in corresponding configuration graph models $N_{\mathrm{rand}}(\Omega)$. The difference between the empirical network count and the mean random model count, $\langle N_{\mathrm{rand}}(\Omega)\rangle$, is measured in units of standard deviation $\sigma(N_{\mathrm{rand}}(\Omega))$:

$$Z(\Omega) = \frac{N_{\mathrm{real}}(\Omega) - \langle N_{\mathrm{rand}}(\Omega)\rangle}{\sigma(N_{\mathrm{rand}}(\Omega))}. \tag{7}$$

A subgraph is considered a significantly over- or underrepresented motif if $|Z(\Omega)| > 2$.

Similarity structure underlying the metaphoric network

Two categories are structurally equivalent if they have the same incoming and outgoing neighbors. Sets of structural equivalent categories are said to form a position in the network, since they play the same role. In practice, though, sets of structurally equivalent vertices are rare and it is therefore necessary to relax this notion and to quantify how similar vertices are regarding their connections [43]. In order to identify roles within the metaphor network, we utilized Hierarchical Cluster Analysis (HCA), which is an unsupervised method. HCA takes a set of categories, along with their incoming and outgoing neighbors, as input, and generates a classification scheme. This scheme, called dendrogram, is a binary tree that visually represents the underlying role structure of the categories in the metaphor network. HCA leverages the given metric structure of a dataset to construct the cluster hierarchy represented by the dendrogram. Agglomerative algorithms, in particular, follow an ascending grouping methodology. This process initiates with clusters consisting of single elements, referred to as leaves. At each subsequent step, the algorithm identifies the pair of closest clusters and merges them into a new cluster. The distances between this merged cluster and other clusters are then recalculated based on pairwise distances between the leaves, thereby progressively building the cluster hierarchy depicted in the dendrogram.

Clustering large data sets with, for example, distances limited to narrow ranges, may produce several different dendrograms if two or more pairs of clusters are equidistant [26]. Therefore, it is necessary to consider these *ties in proximity* and look for clusters, if any, that appear in a large fraction of the resulting dendrograms [25].

We run an HCA algorithm to classify categories by structural similarity in the network based on the method introduced in [43]. The set of categories is endowed with a metric structure by considering the following distance:

$$d(c, c') = \frac{\#(N_{\mathrm{in}}(c) \ominus N_{\mathrm{in}}(c'))}{\#(N_{\mathrm{in}}(c) \cup N_{\mathrm{in}}(c'))} + \frac{\#(N_{\mathrm{out}}(c) \ominus N_{\mathrm{out}}(c'))}{\#(N_{\mathrm{out}}(c) \cup N_{\mathrm{out}}(c'))}$$

where c, c' are categories, $N_{\mathrm{in}}(c)$ and $N_{\mathrm{out}}(c)$ are the incoming and outgoing neighbors of category c, and $N_{\mathrm{in}}(c) \ominus N_{\mathrm{in}}(c')$ is their symmetric difference.

Then, the Ward grouping methodology is used to produce all dendrograms resulting from ties in proximity. Cluster distances are updated at each step according to the recursive Lance–Williams formula:

$$d(C_i \cup C_j, C_k) = \alpha_i d(C_i, C_k) + \alpha_j d(C_j, C_k) + \beta d(C_i, C_j)$$

where C_i, C_j, and C_k are disjoint clusters with size n_i, n_j, and n_k, respectively, and the coefficients are given by

$$\alpha_i = \frac{n_i + n_k}{n_i + n_j + n_k}, \qquad \alpha_j = \frac{n_j + n_k}{n_i + n_j + n_k}, \qquad \beta = \frac{-n_k}{n_i + n_j + n_k}$$

.

Then, we compute the fraction of dendrograms in which each of the resulting clusters appears, using the four contrast functions introduced in [25].

Fasttext word embedding

We want to investigate if the distances defined through the network neighborhood contain information similar or different to the semantic distance of a typical word embedding. For this we look for correlations and mutual information between this network role distance and the embedding distances of words, using fastText word embedding in 300 dimensions. The word embedding is trained on Wikipedia and based on the skip-gram model described in [44].

Typically, two different measures of semantic similarity and distance are used in word embeddings: the Euclidean distance between two words in space, or the cosine similarity between the two vectors from the origin to the words of interest. The Euclidean distance between points $x = (x_1, x_2, \dots, x_{300})$ and $y = (y_1, y_2, \dots, y_{300})$ in 300 dimensions is given by

$$|x - y| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_{300} - y_{300})^2} \tag{8}$$

and the semantic dissimilarity according to the cosine similarity is computed as

$$d_c = 1 - \cos(\vec{x} - \vec{y}) = 1 - \frac{\vec{x} \cdot \vec{y}}{|\vec{x}||\vec{y}|}. \tag{9}$$

To test whether these two semantic measures are related to the metaphor network role similarity we not only look for possible correlations but also for possible mutual information [45]. The mutual information between two different measurements U and V is given by

$$MI(U, V) = \sum_{i=1}^{\#U} \sum_{j=1}^{\#V} \frac{\#(U_i \cap V_j)}{N} \log \frac{\#(U_i \cap V_j) \cdot N}{\#U_i \cdot \#V_j} \tag{10}$$

where $\#U$ and $\#V$ are the number of samples. Compared to the linear correlation coefficient, the mutual information coefficient reveals in a much more general way the extent to which knowing the result of a measure on a data point informs the value of the second measure.

Graph curvature

To understand the structure of the metaphor network in a more complete and general way, it is useful to learn about its "shape" just as we do to investigate geometric objects. In this section we design experiments to reveal the overall "shape" of the metaphor network. The local geometry of an object is primarily described by its curvature and the the distribution of those local observations informs about the global shape as a whole. This principle of local attributes that together allow inference on global properties was already applied when we looked at the degree distribution. But unlike the degree, discrete curvature is defined not for a vertex but for an edge [46], and therefore it directly informs about the structure of relations.

There exist several different translations of the continuous geometrical notion of curvature into discrete graph spaces. We shall look here into those two that have so far proved most useful in empirical data analysis, the Forman Ricci curvature and the Ollivier Ricci curvature [47, 48]. Both concepts have their origin in Riemannian geometry, but allow us to study also structural properties in graphs that have no direct analogue in Riemannian geometry [47].

Forman Ricci curvature. The Forman Ricci curvature of a directed multigraph measures the prominence of flow structures in a network. For each connected pair of vertices (v_i, v_j) it takes all uni-directed edges e.g. $e = (v_i, v_j)$ and counts the number of incoming edges to its source I_{v_i} and outgoing edges of its target O_{v_j} thus quantifying to which degree this edges contributes to a flow structure. Following [49], we define:

Definition 8. (Forman Ricci curvature). *Let e be the non-empty set of edges from vertex v_i to vertex v_j. Let further $in(v_i)$ be the set of incoming edges to v_i and similarly $out(v_j)$ the set of outgoing edges from v_j. Then the Forman Ricci Curvature of e is defined as*

$$\mathbf{F}(e) = 2 - \#in(v_i) - \#out(v_j) \tag{11}$$

A more negative Forman Curvature indicates strong flow structure, while values around zero show the absence of flow.

Ollivier Ricci curvature. The Ollivier Ricci Curvature of an arc (u, v) quantifies how close the incoming neighbors of u are from the outgoing neighbors of v [47,48,50].

To define the Ollivier Ricci Curvature of an edge e from u to v, we compute an optimal transport plan between a probability measure μ_{in}, defined on a neighborhood of u called masses, and a probability measure μ_{out}, defined on a neighborhood of v called holes. The sets of masses and holes are defined as follows [46,49,51]:

Definition 9. (Masses and Holes of an Arc). *Given an arc $e = (u, v)$ of a directed graph $G = (V, E)$, we define the set of masses $\mathcal{M}$ of e to be the singleton $\{u\}$ if u does not have incoming neighbors. Otherwise, $\mathcal{M}$ is the set of incoming neighbors of u. Similarly, the set of holes $\mathcal{H}$ of e is the singleton $\{v\}$ if v does not have outgoing neighbors. Otherwise, $\mathcal{H}$ is the set of outgoing neighbors of v.*

The probability measures are the following: If the tail u of e has no incoming neighbors, we set $\mu_{\text{in}}(u) = 1$. Otherwise, for each mass $m \in \mathcal{M}$, we define $\mu_{\text{in}}(m) = \frac{1}{\#\text{in}(u)}$, where $\#\text{in}(u)$ is the number of incoming neighbors of u. Similarly, if the head v of e has no outgoing neighbors, we set $\mu_{\text{out}}(v) = 1$. Otherwise, for each hole $h \in \mathcal{H}$, we define $\mu_{\text{out}}(h) = \frac{1}{\#\text{out}(v)}$, where $\#\text{out}(v)$ is the number of outgoing neighbors from v.

We can now define the Ollivier Ricci curvature of an arc $e = (u, v)$.

Definition 10. (Ollivier Ricci curvature). *Given an arc $e = (u, v)$ of a directed graph $G = (V, E)$, we define the Ollivier Ricci curvature $\mathcal{O}(u, v)$ as*

$$\mathcal{O}(e) = 1 - W_1(\mu_{in}, \mu_{out})$$

where $W_1(\mu_{in}, \mu_{out})$ is the transportation distance defined as

$$W_1(\mu_{in}, \mu_{out}) := \inf_{p \in \Pi(\mu_{in}, \mu_{out})} \sum_{(e_1, e_2) \in E \times E} d_e(e_1, e_2) p(e_1, e_2)$$

Here, $d_e(e_1, e_2)$ represents the directed distance between e_1, an incoming edge to u, and e_2, an outgoing edge from v. It is defined as the minimal number of edges needed to travel from the tail

of e_1 to the head of e_2. $\Pi(\mu_{in}, \mu_{out})$ represents the set of measures on $E \times E$ that project to μ_{in} and μ_{out}.

In [51], it is proven that the Ollivier Ricci curvature of (u, v) can be expressed as

$$\mathcal{O}(u, v) = \mu_0 - \mu_2 - 2\mu_3$$

where μ_i quantifies the fraction of incoming neighbors of u that are at distance $0 \leq i \leq 3$ from v in an optimal transport plan. Positively curved edges imply that a large fraction of incoming neighbors of u are also outgoing neighbors of v, therefore they are involved in directed triangles $(x, u)(u, v)(v, x)$. Flat edges (u, v) are, for instance, those whose incoming neighbors of u are at distance one from the outgoing neighbors of v, forming directed quadrangles $(x, u)(u, v)(v, y)(y, x)$. Negatively curved arcs are those for which the incoming neighbors of u are at distance larger than one from the outgoing neighbors of v.

Conclusion

We have provided a systematic and multifaceted corpus-based metaphor analysis. Metaphors are novel acts of language usage that arise at a particular time, cause changes in conventional meaning and then can become conventionalized themselves. Therefore, data on diachronic language change can provide insight into this fundamental semantic mechanism. The analysis empirically confirmed a basic assumption of conceptual metaphor theory, namely, that metaphors are persistent linguistic and cognitive transfer structures rather than discrete rhetorical figures. Our results show that the metaphorical network lacks a rigid global structure and that metaphors emerge from dynamic tensions between domains. Although we cannot directly access the causes of the emergence of individual metaphors, the complex patterns of connectivity present metaphor as an interplay of sensorimotor, phenomenological, and socio-cultural dimensions, none of whose influences can be isolated.

The metaphor network encompasses a wide range of language topics, each exhibiting varying levels of activity. While the majority of topics demonstrate low activity, the activity is strongly concentrated on two groups of topics: on the one hand, concrete, sense-related domains, which are by far the richest sources but also serve as target domains for metaphors, and on the other abstract but nevertheless everyday topics, which are almost exclusively targets of metaphorical transmissions. This division is reflected in the global structure of the network. Most categories can be grouped into two anti-communities, a concrete and an abstract one. Along these anti-communities there are two systematically important metaphor processes. Mappings from the concrete to the abstract categories and mappings between concrete categories. Beyond that, no global structure is found, no clusters of thematic groups emerge that could be separated from each other. Also there is no flow of structure but only bilateral relations between two topics. Structurally, spatial metaphors do not occupy any special rank, they show a similar behavior as other concrete domains, which are not connected to the sense of orientation and vision. The absence of any more particular structure in the network of metaphorical relations may suggest that this reflects a rather universal cognitive mechanism that can flexibly connect different domains and capture abstract structures through more concrete images, but at the same time also reorganizing the relations between concrete domains.

Our findings suggest that analogy or similarity is unlikely to be the primary selection criterion for new metaphor formation. Instead, they support the theory that metaphor is a creative process where the target domain is conceptualized through emerging similarity, exploiting the tension between the source and target domains. Additionally, there is evidence of

feedback re-conceptualization of the source domain, indicating that metaphor is not a purely one-directional process.

Finally, it was found that the role of a topic in the metaphor network reveals in a stable way the semantic position of this topic relative to other topics. There is thus an explicit relation of the metaphor to semantic structure. We extracted the information on the semantic role from the metaphor network and constructed the resulting structure in form of a dendrogram. This structure does hold some mutual information as word embeddings based on the distributional hypothesis. At the same time, it contains information about conceptual and figurative semantic relations. This is a valuable and new addition to previously computed semantic structures, which allows for extensive further research.

Supporting information

S1 Fig. Cluster frequencies. Distribution of cluster contrast values for clusters within the four dendrograms obtained from hierarchical cluster analysis of categories in the metaphor network, using $d(c, c') = \frac{\#(N_{\text{in}}(c) \ominus N_{\text{in}}(c'))}{\#(N_{\text{in}}(c) \cup N_{\text{in}}(c'))} + \frac{\#(N_{\text{out}}(c) \ominus N_{\text{out}}(c'))}{\#(N_{\text{out}}(c) \cup N_{\text{out}}(c'))}$ as a distance between categories c and c', and Ward's grouping methodology.
(TIF)

S2 Fig. Size of persistent clusters. The Hierarchical Cluster Analysis (HCA) algorithm ran on the collection of categories in the metaphor network generated four different dendrograms. Most cluster consistently appeared in all four dendrograms. These scatter plots show the distribution of pairs (cluster size, fraction of dendrograms) corresponding to the clusters found within the four dendrograms. The size of a cluster is defined as its number of leaves, that is, the number of categories within the cluster. The ball size in the plot represents the frequency of occurrence for a specific pair (cluster size, fraction of dendrograms). The scatter plot on the left defines clusters as subtrees of the dendrograms, while the scatter plot on the right defines clusters as the collection of leaves of subtrees of the dendrograms. The HCA algorithm was ran using: $d(c, c') = \frac{\#(N_{\text{in}}(c) \ominus N_{\text{in}}(c'))}{\#(N_{\text{in}}(c) \cup N_{\text{in}}(c'))} + \frac{\#(N_{\text{out}}(c) \ominus N_{\text{out}}(c'))}{\#(N_{\text{out}}(c) \cup N_{\text{out}}(c'))}$ as a distance between categories c and c', and Ward's grouping methodology.
(TIF)

S1 Table. Two Anti-community category blocks and remaining categories. The metaphor network globally separates into two anti-communities which together encompass 370 categories, 217 in the first and 153 in the second. Thematically, there is a clear tendency in the split: more experience and body-related categories are likely to be found in the first anti-community and more abstract categories in the second. The splitting was done with the help of the Louvain cluster detection method algorithm on the negative graph.
(TIF)

Acknowledgments

Author contributions

Conceptualization: Marie Teich.

Formal analysis: Marie Teich, Wilmer Leal.

Investigation: Marie Teich.

Methodology: Marie Teich, Wilmer Leal.

Project administration: Marie Teich.

Software: Marie Teich, Wilmer Leal.

Supervision: Jürgen Jost.

Visualization: Marie Teich, Wilmer Leal.

Writing – original draft: Marie Teich, Wilmer Leal.

Writing – review & editing: Wilmer Leal, Jürgen Jost.

References

1. Bisk Y, Holtzman A, Thomason J, Andreas J, Bengio Y, Chai J. Experience grounds language. arXiv preprint 2020. https://doi.org/10.48550/ARXIV.2004.10151

2. Tamari R, Shani C, Hope T, Petruck MRL, Abend O, Shahaf D. Language (Re)modelling: towards embodied language understanding. arXiv preprint 2020. https://doi.org/10.48550/ARXIV.2005.00311

3. Croft W, Cruse DA. Cognitive linguistics. Cambridge University Press; 2004.

4. Tomasello M. Constructing a language: a usage-based theory of language acquisition. Harvard University Press; 2005.

5. Rohrer T. Image schemata in the brain. From perception to meaning: Image schemas in cognitive linguistics. 2005;29:165–96.

6. Firth J. Modes of meaning, papers in linguistics; 1957.

7. Harris ZS. Methods in structural linguistics; 1951.

8. Le Q, Mikolov T. Distributed representations of sentences and documents. In: Xing EP, Jebara T, editors. Proceedings of the 31st International Conference on Machine Learning. vol. 32 of Proceedings of Machine Learning Research. Bejing, China: PMLR; 2014. p. 1188–96.

9. Deignan A. Metaphor and corpus linguistics. J. Benjamins Pub.; 2005.

10. Lakoff G, Johnson M. Metaphors we live by. University of Chicago Press; 2008.

11. Low G, Cameron L, Deignan A, Todd Z. Researching and applying metaphor in the real world. 2010; p. 1–400.

12. Kövecses Z. Conceptual metaphor theory: some criticisms and alternative proposals. Annu Rev Cogn Linguist. 2008;6(1):168–84.

13. Stefanowitsch A, Gries ST. Corpus-based approaches to metaphor and metonymy. Mouton Select: Mouton de Gruyter. 2007.

14. Bowdle BF, Gentner D. The career of metaphor. Psychol Rev. 2005;112(1):193–216. https://doi.org/10.1037/0033-295X.112.1.193 PMID: 15631593

15. Sweetser E. From etymology to pragmatics: metaphorical and cultural aspects of semantic structure. Cambridge: Cambridge University Press; 1990.

16. Black, Black, M. Models and metaphors: Studies in language and philosophy. Cornell University Press; 2019.

17. Blumenberg H. Paradigmen zu einer Metaphorologie. Archiv für Begriffsgeschichte. 1960:7–142.

18. Ricoeur P. The rule of metaphor: the creation of meaning in language. Routledge; 2004.

19. Soskice J. Metaphor and religious language. Int J Philos Religion. 1985;23.

20. Barsalou LW. Perceptual symbol systems. Behav Brain Sci. 1999;22(4):577–609; discussion 610-60. https://doi.org/10.1017/s0140525x99002149 PMID: 11301525

21. Casasanto D, Gijssels T. What makes a metaphor an embodied metaphor?. Linguist Vanguard. 2015;1(1):327–37. https://doi.org/10.1515/lingvan-2014-1015

22. Torres-Martínez S. Metaphors are embodied otherwise they would not be metaphors. Linguist Vanguard. 2022;8(1):185–96. https://doi.org/10.1515/lingvan-2019-0083

23. Mapping Metaphor. http://mappingmetaphor.arts.gla.ac.uk

24. Gibbs RW Jr, Costa Lima PL, Francozo E. Metaphor is grounded in embodied experience. J Pragmatics. 2004;36(7):1189–210. https://doi.org/10.1016/j.pragma.2003.10.009

25. Leal W, Llanos EJ, Restrepo G, Suárez CF, Patarroyo ME. How frequently do clusters occur in hierarchical clustering analysis? A graph theoretical approach to studying ties in proximity. J Cheminform. 2016;8:4. https://doi.org/10.1186/s13321-016-0114-x PMID: 26816532

26. MacCuish J, Nicolaou C, MacCuish NE. Ties in proximity and clustering compounds. J Chem Inf Comput Sci. 2001;41(1):134–46. https://doi.org/10.1021/ci000069q PMID: 11206366

27. Jackendoff R. The architecture of the linguistic-spatial interface. 1996.

28. Deuker L, Bellmund JL, Navarro Schröder T, Doeller CF. An event map of memory space in the hippocampus. Elife. 2016;5:e16534. https://doi.org/10.7554/eLife.16534 PMID: 27710766

29. Bellmund JLS, Gärdenfors P, Moser EI, Doeller CF. Navigating cognition: spatial codes for human thinking. Science. 2018;362(6415):eaat6766. https://doi.org/10.1126/science.aat6766 PMID: 30409861

30. Gardenfors P. Conceptual spaces: the geometry of thought. MIT Press; 2004.

31. Gentner D. Metaphor as structure mapping: the relational shift. Child Developm. 1988;59(1):47. https://doi.org/10.2307/1130388

32. Fillmore CJ, et al. Frame semantics. Cognitive linguistics: basic readings. 2006. p. 373–400.

33. Anderson W, Bramwell E, Hamilton R. Mapping metaphor with the historical thesaurus. 2015.

34. Kay C, Roberts J, Samuels M, Wotherspoon I. Historical thesaurus of the Oxford English dictionary. Oxford University Press; 2009.

35. Teich M, Leal W, Jost J. Diachronic data analysis supports and refines conceptual metaphor theory. 2023. https://doi.org/10.5281/zenodo.8383019

36. Newman M. Networks. OUP Oxford; 2018.

37. Barabási A-L. Network science. Philos Trans A Math Phys Eng Sci. 2013;371(1987):20120375. https://doi.org/10.1098/rsta.2012.0375 PMID: 23419844

38. Aynaud T. Python-louvain x.y: Louvain algorithm for community detection. 2020.

39. Blondel VD, Guillaume J-L, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. J Stat Mech. 2008;2008(10):P10008. https://doi.org/10.1088/1742-5468/2008/10/p10008

40. Alon U. Network motifs: theory and experimental approaches. Nat Rev Genet. 2007;8(6):450–61. https://doi.org/10.1038/nrg2102 PMID: 17510665

41. Kashtan N, Itzkovitz S, Milo R, Alon U. Network motif detection tool Mfinder tool guide. Rehovot, Israel: Weizmann Institute of Science; 2005.

42. Patra S, Mohapatra A. Review of tools and algorithms for network motif discovery in biological networks. IET Syst Biol. 2020;14(4):171–89. https://doi.org/10.1049/iet-syb.2020.0004 PMID: 32737276

43. Leal W, Restrepo G, Bernal A. A network study of chemical elements: from binary compounds to chemical trends. Match (Mulh). 2012;68(2):417–42.

44. Bojanowski P, Grave E, Joulin A, Mikolov T. Enriching word vectors with subword information. Trans Assoc Comput Linguist. 2017;5:135–46. https://doi.org/10.48550/arXiv.1607.04606

45. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O. Scikit-learn: machine learning in python. J Mach Learn Res. 2011;12:2825–30.

46. Eidi M, Farzam A, Leal W, Samal A, Jost J. Edge-based analysis of networks: curvatures of graphs and hypergraphs. Theory Biosci. 2020;139(4):337–48. https://doi.org/10.1007/s12064-020-00328-0 PMID: 33216293

47. Leal W, Eidi M, Jost J. Ricci curvature of random and empirical directed hypernetworks. Appl Netw Sci. 2020;5(1):65. https://doi.org/10.1007/s41109-020-00309-8

48. Samal A, Sreejith RP, Gu J, Liu S, Saucan E, Jost J. Comparative analysis of two discretizations of Ricci curvature for complex networks. Sci Rep. 2018;8(1):8650. https://doi.org/10.1038/s41598-018-27001-3 PMID: 29872167

49. Leal W, Restrepo G, Stadler PF, Jost J. Forman–Ricci curvature for hypergraphs. Advs Complex Syst. 2021;24(01):2150003. https://doi.org/10.1142/s021952592150003x

50. Ollivier Y. Ricci curvature of Markov chains on metric spaces. J Funct Anal. 2009;256(3):810–64.

51. Eidi M, Jost J. Ollivier Ricci curvature of directed hypergraphs. Sci Rep. 2020;10(1):12466. https://doi.org/10.1038/s41598-020-68619-6 PMID: 32719341