Maria Berger and Tatjana Scheffler

Can we Operationalize Conceptual Metaphor Cross-Lingually?

Abstract

The conceptual nature of metaphorical expression is a long-discussed phenomenon, highly investigated by linguists, psychologists, translators, and philosophers, amongst others. In theoretical work, distinctions are made between conceptual metaphors (a phenomenon of human cognition) and linguistic metaphors (their concrete realizations in language), while most computational approaches have only addressed the latter. In the age of massive language models, metaphor and other phenomena of figurative speech are earning new attention as more and more textual analyses are built on top of neural-networking tools that do not necessarily make a distinction between the lexicalization of a concept and the concept itself. Hence, an investigation of conceptual metaphor using a more linguistics-driven perspective is of much importance. In this work, we investigate the conceptuality of metaphoric expressions across two languages utilizing a parallel corpus of news commentaries from the web. We assume that a conceptual metaphor is represented by many instances of linguistic metaphors. This idea presupposes linguistic metaphor as an operationalization of conceptual metaphor. We perform several tests on how metaphors are translated between the languages, to assess whether distinct lexicalizations of a metaphor form conceptual clusters, and whether the usage of words in a metaphorical context is distinguishable from their usage in literal contexts.

We find that we are able to group linguistic metaphors in one language into semantically related sets by clustering their translations in another language. We argue that these semantically related sets constitute an operationalization of conceptual metaphors. In English, the clusters are formed by fewer, but more diverse lexelts (linguistic types), while in German we find more and bigger clusters composed primarily of derivatives and compounds. We also find that when a lexelt is translated similarly in unannotated instances to known metaphoric usages, then its contextual sense tends to be figurative as well.

1 Introduction

Among experts, there is consensus that the phenomenon of metaphor requires that the interpreter of the figure creates a link to the metaphoric concept (overtly denoted by the expression of the figure). This phenomenon is not only considered a linguistic one (Lakoff & Johnson, 1980), it also is well-approved as a mental one (Borbely, 1998). The connection between metaphors and concepts actually goes back to (Ricoeur, 1978), who sees language itself as a resource to regenerate ideas in new ways, and argues that metaphor is a driver of that process.

1.1 Conceptual Metaphor Theory

Following Lakoff and Johnson (1980), conceptual metaphor theory, at its core, understands metaphor as a matter of cognition. Metaphors (e.g., “LIFE IS A JOURNEY”) are processed by mapping “concrete” source domain concepts (e.g., journey) onto “abstract” target domain concepts (e.g., life). Zoltán Kövecses writes in Semino and Demjén (2016): “A conceptual metaphor is understanding one domain of experience (typically abstract) in terms of another (typically concrete).” In this definition, conceptual metaphor is, on the one hand, a cognitive process of understanding a domain, on the other hand, the concept is a product. In the linguistics literature, these two are not necessarily differentiated. In this article, we won’t distinguish them either as we work based on textual lexicalization of metaphors “in the wild”.

Experts in general agree on the opinion that metaphors play a role in the process of verbal conceptualization. In the field of cognitively influenced research on conceptual metaphor, many researchers also argue that conceptual metaphor resides in thought (Lakoff, 2009) and that metaphorical concepts can be built and enriched by linguistic instances (Semino & Demjén, 2016). Lakoff even claims that the human brain works effortlessly, precisely because it is processing language in a metaphorical way, hence, metaphor must be established at a cognitive level (cf. Holyoak & Stamenković, 2018). Lakoff and Johnson (1980) see metaphors not as a pure linguistic phenomenon, but instead argue that language is the means that exhibits metaphor’s work. In their view, metaphor is an embodied understanding of the world, relating the mind to the body, and the body to the world. The relationships in this scenario (e.g., space, orientation, movement) are driven by metaphors while the embodied concepts (e.g., container, path) are language independent (cf. Veale, Shutova, & Klebanov, 2016).

Borbely (1998) moves from a linguistic metaphor concept to a supra-linguistic one, a psychoanalytic concept of metaphor. Starting from Ricoeur (1978), who defines metaphor as “seeing something in terms of something else”, a psychoanalytic approach sees the present in terms of the past and the past in terms of the present. That understanding broadens the linguistic metaphor concept. Metaphor is not only a phenomenon of language, but of mentation. In fact, Borbely (1998) highlights a specific psychoanalytic concept of metaphorical process between primary (unconscious) and secondary (rational) thinking. Experiments showed that the mind functions metaphorically when faced with challenges requiring creativity.

Finally, in a comprehensive, empirical study, Holyoak and Stamenković (2018) elaborate three main theories in the research field of metaphor comprehension: analogy, categorization, and conceptual mapping. While analogy and categorization are simply represented by a model of information processing; conceptual mapping calls for an embodied understanding of metaphor comprehension. Holyoak and Stamenković wish for a greater consideration of a metaphors’ contextual function and its emotional impact. They propose ways to combine analogy and conceptual mapping to broaden metaphor understanding. One way, they suggest, is considering conceptual metaphor based on individual representations of word pairs (source domain and target domain of a metaphor). Following Kintsch (2000), who says that metaphor comprehension is a real-time generation of a specific meaning adapted to the specific context in which a metaphor is used, a metaphor instance might also be understood as an analogical alignment process.

1.2 Computational Approaches to Conceptual Metaphor Theory

Work by Tsvetkov, Boytsov, Gershman, Nyberg, and Dyer (2014); Tsvetkov, Mukomel, and Gershman (2013) and Bizzoni and Ghanimifard (2018) shows that the conceptual metaphor idea is made use of in computationally-driven cross-lingual metaphor detection setups. Especially, Schuster and Markert (2023) build their work in cross-lingual metaphor detection on the idea that a concept remains stable across languages while its lexicalization does not necessarily do so.

Kintsch (2000) develops an approach to determine the semantic representation of a metaphor computationally. Kintsch argues that when individual instances of the source and target domains of a metaphor[1] are "merged", we receive a set of paired "concepts" (corresponding to lexicalized words). These pairs could then be "conceptually combined" via a so-called metaphor vector (such as the procedure described in Kintsch, 2000). This again, creates a context-specific semantic representation of a metaphor. Kintsch (2000) further makes an attempt to distinguish target domains used in a figurative meaning from a literal one: A comparison between the metaphor vector "my lawyer is a shark" towards typical attributes of the source domain "shark" (e.g., "fish") shows a strong distance between the metaphor vector and the "fish"-attribute. In contrast, the non-metaphor "my lawyer is young" shows a strong similarity between the phrase vector and the source-domain attribute ("young"). This shows that the "source domain" can have an impact to the contextual meaning of a (metaphoric) phrase.

Based on the understanding that the individual instances of conceptual metaphors are used in similar situations, they build similar contexts across languages. We can represent these contexts by using distributional semantics, a theory that has been pivotal in computational processing of language (Mikolov, Sutskever, Chen, Corrado, & Dean, 2013). We can adapt distributional embedding representations for multiple languages for example by using multi-lingual embedding models. Bilingual vector models might enable us to measure/compute these metaphoric contexts. They were successfully applied in NLP before (Luong, Pham, & Manning, 2015; Upadhyay, Faruqui, Dyer, & Roth, 2016; Vulić & Moens, 2015) and can help to investigate the phenomenon more deeply.

Related work often only considers the extracted linguistic metaphors in fixed POS patterns such as NV, VN, and AN (Köper & Schulte im Walde, 2017; Shutova, 2010;

[1] An instance of a metaphor for Kintsch always contains a metaphor target and a metaphor source, e.g., "my lawyer (target) is a shark (source)".

Tsvetkov et al., 2014, 2013) and translated independently of their context. Alternatively, metaphors are represented based on external knowledge (Bizzoni & Ghanimifard, 2018; Tsvetkov et al., 2014) while language-specific lexicalization of the metaphor is not accounted for on a precise level. The applied techniques also do not necessarily consider strongly divergent cultures between source and target languages.

In this research, we do not focus on automated metaphor detection in general, instead, we make use of the comparable lexicalization of conceptual metaphor across languages. We therefore do not review and present relating literature of automated metaphor recognition. The interested reader, however, may be referred to an extensive study on neural automated metaphor recognition (Choi et al., 2021), a procedure that also considers multi-lingual metaphor recognition (Sanchez-Bayona & Agerri, 2022), a contrastive learning approach (Lin, Ma, Yan, & Chen, 2021) that is based on the literal and figurative context of a given lexelt considering pre-trained neural models and a silver data augmentation procedure, and a verb metaphor detection procedure treated as a relation extraction problem (Song, Zhou, Fu, Liu, & Liu, 2021) in which the authors investigate a metaphoric verb in relation to its context.

1.3 Non-universal Conceptual Metaphors

Conceptual metaphor, not only its lexicalization, does not necessarily need to be universal. Concepts may be modified or lost, in particular across different cultures. As such, Kövecses (2018) writes that the universality of conceptual metaphors can vary from culture to culture, since they experience different contextual influences and are elaborated in different ways, which also can be influenced by the given context. Also, Yan, Noël, and Wolf (2010) investigate cross-lingual metaphor transfer asking how a "target language culture" affects the lexicalization of a metaphor and how that lexicalization, again, affects the concept of the metaphor to diverge from the source language. Yan et al. (2010) find that the way a metaphor concept is modified in a target language, relies strongly on the entrenchment rank. This means that a metaphor either remains preserved, or, when not shared in the target language's system, it might be replaced by another metaphor of a higher entrenchment rank.

1.4 Our Contribution - Investigate Conceptual Metaphor via Translation

We approach the problem differently, from an empirical perspective, to find evidence for conceptual metaphor across languages. We are interested in how a metaphor's lexicalization is translated. We build on the hypothesis that similarly translated linguistic metaphors share or at least somehow delimit a concept. We use the translations' contextual representations and attempt to form clusters of words that share similar meanings, and therefore represent separate instances of the same conceptual metaphor. We present several experiments to evaluate translation embeddings and study whether translation representation can help to distinguish metaphorical and literal context for metaphor candidates.

We also build our work on the understanding that conceptual metaphor tends to remain rather preserved across languages, even though this is not always the case, while this does not necessarily apply to linguistic metaphor. Our overarching research question reads:

> **RQ:** "How is conceptual metaphor represented by means of linguistic metaphor?"

As we understand conceptual metaphors as essentially universal (Kövecses, 2018; Lakoff & Johnson, 1980) and thus language independent, we can use data from different languages to investigate to what extent a conceptual metaphor is expressed using the same or different linguistic metaphors. In doing so, we gain new insights into how well contextualized representations support capturing conceptual metaphor.

In the following, we refer to a source language (SL) and a target language (TL) as is common in machine translation studies, and we refer to the metaphoric source domain (SD) and the metaphoric target domain (TD).

2 Background

2.1 Conceptual Metaphor in Translation Studies

The conceptuality of metaphor, in fact, is critical to analyzing metaphor across languages. Finding and automatizing ways to cross-language metaphor detection, hence, represents an important complement to studies in conceptual metaphor theory. As such, one of the major problems of current approaches is that cross-lingual metaphor detection, in languages culturally divergent, does not necessarily work based on simple dictionary translation. Investigating this phenomenon more deeply, Yan et al. (2010) investigate the linguistic perspective of metaphor transfer, i.e., the effect of the target language's culture on the translated expression, and the conceptual perspective, i.e., the effect of the translation on the source language's metaphoric concept. In their empirical categorization of FEAR metaphors, the authors make use of Vukovic (2007)'s assumption that a metaphor forms a general concept, which appears as specific lexical instances, e.g., FEAR IS AN OPPONENT vs. "conquered by fear", "besieged by fear". To retrieve metaphors in native/translated English and translated/native Chinese, the authors apply metaphorical patterns assuming that a metaphor forms a multi-word expression from a given SD towards more specific lexical items of a TD (Bisang, Hock, Winter, Stefanowitsch, & Gries, 2006). The authors find that whether a metaphor remains lexically stable or is replaced by a conceptually similar one, highly relies on its familiarity (entrenchment, Langacker, 1987) in the target language. Piccioni (2013) also finds some oddities in lexicalization during the translation of metaphors. The author compares the collocation variety of metaphors in Spanish original texts and English-to-Spanish translations and finds several unconventional metaphors in the translations, e.g., "see black" vs "see gray". The author refers to that effect as "shining through", cf. (Teich, 2003).

Piccioni (2013) performs a corpus study to derive a methodology for the analysis of translations via a comparison of metaphor use in original and translated texts. It uses a monolingual corpus of corporate sustainability reports consisting of (i) a sub-corpus of Spanish originals and (ii) a sub-corpus of translations from English to Spanish. The author analyses verb-noun metaphors to compare collocation variety, collocation regularity and metaphorical conventionality (depending on the appearance of a verb and a noun in its basic meaning or in its metaphoric usage) of the verb-noun pairs in original and translated texts. Results show that metaphors in translated texts show a tendency towards normalization (frequent metaphors make up a very high portion of all metaphors) as well as a preference for unconventional uses, which means that metaphors that do not necessarily appear in an original language text, do appear in a translation because words might be translated too literally. Very recently, Y. Liu et al. (2023) used conceptual features to help predict which language family a metaphor belongs to. As the conceptual features used are culture-independent, this approach successfully predicted the correct TL.

2.2 Representing Conceptual Metaphor by Embedding Vectors

Approaches to automated metaphor detection often utilize a form of vector representation to encode the meaning of words, individually or in context. Today, semantic information mainly is encoded by word embeddings (Mikolov et al., 2013). For example, Bizzoni and Ghanimifard (2018) and Maudslay, Pimentel, Cotterell, and Teufel (2020) present architectures of artificial neural networks that process the semantic representation of texts to recognize metaphors. Input representations might also be enriched with concreteness scores from relevant dictionaries and together passed into the network as features. However, these embeddings often come with a certain degree of bias (Papakyriakopoulos, Hegelich, Serrano, & Marco, 2020) and cannot necessarily distinguish metaphoric from literal use of lexical items. In earlier work, Kintsch (2000) says that when individual instances of metaphor are “merged”, we receive a set of paired “concepts” (corresponding to lexicalized words). These pairs could then be “conceptually combined” via a so-called metaphor vector to create a context-specific semantic metaphor representation.

3 A Data-driven Perspective on Conceptual Metaphor

It is tricky to study a cognitive notion such as conceptual metaphor directly on the basis of textual evidence. Indeed, a linguistic metaphorical expression in one text in one language does not give good evidence for an analysis of conceptual metaphor. However, we propose that parallel corpora may enable an operationalization for identifying conceptual metaphors by using cross-lingual translations of individual metaphoric expressions. Precisely, we imagine the process of metaphor translation as one of transferring the meaning of a metaphor represented by many events of lexicalization.

Our overarching research questing is: "How can conceptual metaphor be operationalized by means of linguistic metaphor?"

As shown in the literature, there is a tendency to view metaphor as something language-independent. However, from a lexicalized perspective, metaphors are not always translated literally, nor is the same linguistic metaphor necessarily used in the same context in all languages. This contributes to a range of challenges (variability, uncertainty, conceptual preservation) when metaphors are translated across languages. Descriptionist Van den Broeck (1981), for example, addresses this by introducing substitution and paraphrasing next to literal translation of a metaphor.

In the following, we list a more granular operationalization based on the different data types that helps us to address our main research question. We address metaphor variation across languages in a parallel corpus, consider their figurative and literal contexts and look at metaphorical compound nouns, which are very frequent in German.

3.1 Research Questions

When translating metaphor, we encounter two problems: (i) Translating the linguistic metaphor, and (ii) translating the conceptual metaphor (i.e., that also carries the metaphoric act; cf. Yan et al. (2010)). Following Schäffner (2004), translation studies experienced a strong shift from pure linguistics towards focusing on more contextual factors that can affect a translation.

For example, equivalence-based approaches say that a metaphor should be transferred intact from a SL to a TL, but different cultures can have a strong impact on such a transfer. For this purpose, Schäffner (2004) establishes conceptualization as relevant to translating an individual metaphor, hence, raises the problem of translatability from the lexical/linguistic level to the level of the cultural system of SL and TL.

Van den Broeck (1981) reflects on the boundaries of translatability exemplified by the translation of metaphor. Borrowing from (Even-Zohar, 1971), Van Den Broeck lists some rules that need to be considered when translating metaphor. One such rule says that "translatability is high when the general cultural evolution of SL and TL" is similar. Another rule correspondingly says that translatability is higher the fewer kinds of information need to be delivered. Concerning the first point, we shall be fine concerning the translatability of metaphor from English to German in the news commentary domain. The second point can be applied to the diversity of possible translations of a word in another language. We can infer that depth-2 entropy of a metaphor makes translatability harder.

In our dataset, it further often happens that a metaphor's SD and TD (Lakoff & Johnson, 1980) from the English language text are translated together as the same unique compound word in the German language part of our corpus. Examples are:

1. gun(TD) control(TD) battle(SD) : Waffenkontrollkampf
2. gun(TD) war(SD) : Waffenkrieg
3. poverty(TD) card(SD) : Armutskarte

Van den Broeck (1981) also points to the challenges that accompany the translation of compounds. As metaphor as such typically violates contextual preference in the SL, such a violation in the TL's system might also be expected. Van den Broeck (1981) argues this is not always the case as, for example, the condensed nature of the German compound metaphor "Tränenbäume" (tear trees) is translated in French as "arbres de larmes" (trees of tears), however, larmes has other translations (droplet, drop, water), hence it is "rendered more analytically" (Van den Broeck, 1981, p. 80). Even though, in our data, poetic jewellery is not predominant, we still are interested in how compounds affect translation across two languages and whether the distribution of the translations of compounds is similar to that of single words.

We refine our research to the following **Questions:**

RQ1: To what extent are linguistic metaphors preserved during translations (cf. Experiment 1)?

RQ2: How does linguistic metaphor volatility contribute to the same conceptual metaphor, i.e., can conceptual metaphor be encoded by contextualized representations of linguistic metaphor (cf. Experiment 2)?

RQ3: Can contextualized representations of conceptual metaphor help to distinguish metaphorical and literal contexts of linguistic metaphor candidates (cf. Exp. 3)?

RQ4: Does lexicalization of linguistic metaphor differ between German compounds and simplicia (cf. Experiment 4)?

Experiments and Data Overview: We measure how a specific metaphoric word is translated, how diverse the German translation set, cf., the different possible translations of the metaphor in the TL is, and measure how dominant one specific translation is in our corpus compared to others.

The data used in the study are metaphoric words from our metaphor corpus together with their German translations. To control the experiment, we also look into the translation distributions of the metaphoric word when it is not used in a metaphoric context, i.e., not annotated for metaphor using alignment data. The alignment data reveal information on how words are associated with each other across two languages. When a lexelt is a compound, the associated translation is a multi-word-unit (cf. Fig. 1).

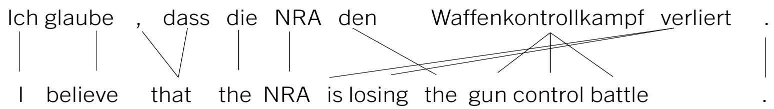


Figure 1: Example of bilingual alignment containing compound words

3.2 Metaphor Data

Our dataset of parallel metaphors initially originates from the English language corpus by Gordon et al. (2015)[2]. Gordon et al. (2015) semi-automatically gathered metaphors pertaining to target concepts in the domains of governance, economic inequality, and the gun control-debate. The sentences collected for that purpose are from different genres of writing including press releases, news articles, weblog posts, online forum discussions, and social media. The authors justify their choice of sentences in that they "exhibit the diversity of ways that people can conceptualize the same set of target concepts".[3]

Most metaphors in our data are of the form where "source + target" are directly syntactically related (e.g., "fighting poverty", or "kills democracy"). They very frequently are formed by words that belong to the "fighting" domain, such as "combat hunger" and "assault of poverty", but also the "securing" domain is represented often: "poverty alleviation" and "protecting democracy"

We use a machine-translated and manually evaluated subset of that corpus (Berger, 2022) for our study. It consists of 1000 English sentences. In every sentence, one metaphoric SD (one lexelt) and one TD (one lexelt) was annotated.[4] For every lexelt that was used in a metaphoric context, we also retrieved contexts in which it has been used literally from the same corpus. This explains why the (literally used) types are a subset of the metaphorically used types. Tab. 1 gives an overview of the corpus statistics.[5]

Table 1: Corpus statistics

(a) Words used metaphorical in the corpus

contexts	**EN**	**DE**
types metaphorical	276	384
out of which literal	101	102
tokens metaphorical	961	961
tokens literal	209	396

(b) Metaphor tokens distributed by part-of-speech using spacy's POS-tagger (`https://spacy.io/`)

	EN	**DE**
VERB	583	525
NOUN	318	376
ADJ	34	24
ADV	0	21
PROPN	16	9
X	0	1
tokens (total)	951	956

[2]Accessed: July 2025 from: `https://figshare.com/articles/dataset/A_Corpus_of_Rich_Metaphor_Annotation/6179210?file=11188490` This effort makes use of the IARPA-funded Metaphor Program USC/ISI annotated metaphorical language collection, release iarpa_metaphor_isi.edu_metaphor_corpus_2015-04-03. `http://purl.org/net/metaphor-corpus`

[3]The exact references are not provided. The annotation procedure the authors are following originates from Mohler, Tomlinson, and Rink (2015).

[4]The entire parallel corpus is published in Berger, Kiwitt, and Reimann (2024).

[5]Note that token frequency differs between the label distribution and the corpus representation, because of translation quality, labeling procedure and multi-word expressions.

3.3 Contextualized Representations of the Data Used

We use contextual embedding representations from a pre-trained multilingual BERT model (Vaswani, 2017). This way, we retrieve a vector for each word that is used metaphorically in our corpus, and also for each of its contextual translations. Contextual embeddings can represent all words of the vocabulary of a language within one high-dimensional numerical space (Mikolov et al., 2013). In our work, these embeddings allocate the vocabulary of English and German words into the same vector space. We obtained the representation by aggregating the last four layers of a pre-trained multilingual BERT model hosted on hugginface.[6]

4 Metaphor Variation Across Languages

4.1 Entropy & Stability of Linguistic Metaphor (Exp. 1)

Y. Liu et al. (2023) propose a method to define the degree of conceptual similarity across languages. The authors create a bipartite alignment graph between the concepts of a SL and a set of TL tokens. Based on that graph, the authors align cross-lingual conceptuality in over 1,000 languages. Our study focuses on two languages only, and presumes the language agnostic conceptuality of a metaphor. Working based on a rather small dataset, we can also consider the translation procedure as a graph (bipartite or tripartite depending on the data format), and then retrieve translations in individual events. Inspired by Y. Liu et al. (2023), we define *stability* $S(M)$ of a linguistic metaphor. We consider translating a metaphor into one language and back as a tripartite graph $\mathcal{G}$ starting from one given lexelt l. Then, the number of paths of depth two—represented as a triple of nodes (y,x,z)—starting from that lexelt and ending at the same lexelt is divided by all other paths starting from that lexelt, and ending up in any lexelt of the SL.

$$S(M) = \frac{|\{(l, x, l) \in \mathcal{G}\}|}{|\{(l, x, y) \in \mathcal{G}\}|} \tag{1}$$

We measure depth-2 *entropy* $E_{depth-2}(M_l)$ of a linguistic metaphor by retrieving metaphoric words that still are contained in the translation set after one translation forward, and one backward pass (again, path depth two, cf. Y. Liu et al. (2023)). We calculate $E_{depth-2}(M_l)$ in the following way: We first calculate the entropy $E(l)$ of a given lexelt l in a SL by considering all ratios of all its translations (TL) in the total of possible translation as they appear in our data. We build the total $\sum$ over all these ratios (cf. Formula 3). Then, we calculate the entropy of all the translations of a lexelt l, namely $E(T(l))$. Here, we follow exactly the same procedure as above, simply starting with words in the TL. We build the average of these values $E_{av}(T(l))$. We balance $E(l)$ with .5 and the averaged entropy of all its translations $E_{av}(T(l))$ also with .5. This way, we obtain $E_{depth-2}(M_l)$. This measure especially helps us to gain insights into

[6] `https://huggingface.co/google-bert/bert-base-multilingual-cased`, acc. Nov. 2024

how diversely a lexelt and its translations are translated, hence how volatile metaphoric words are in our dataset.

$$E_{depth-2}(M_l) = \frac{E(l) + E_{av}(T(l))}{2}, l \in SL, T \in TL \tag{2}$$

While the entropy E(x) of many events for x is defined as (Gray, 2011):

$$E(x) = -\sum_{x \in X} p(x) * \log p(x) \tag{3}$$

We present the stability-depth-2 *entropy* matrix in Figure 2. At first glance, the inverted correlation between depth-2 *entropy* and stability is especially visible for the German data rather than for the English data. Further, especially the (y=1) and (y=.5) lines show that there are many cases in which only one or two single translations exist for a given metaphor in both languages. We can also see that English and German metaphors are located a bit apart from each other in the matrix. German words are located more towards lower stability, while English words have higher stability. German intuitively has higher depth-2 *entropy* values due to the higher number of derivatives and compounds (cf. Sec. 4.2).

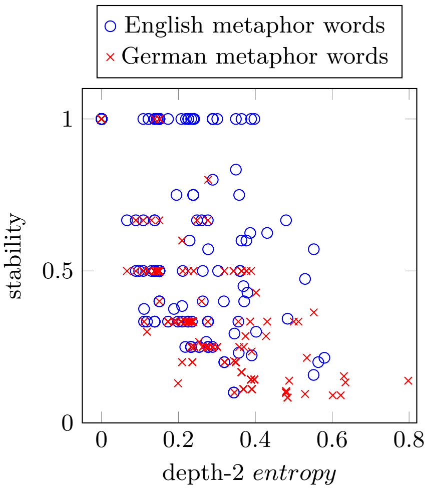


Figure 2: depth-2 *entropy* versus stability in English and German metaphoric words

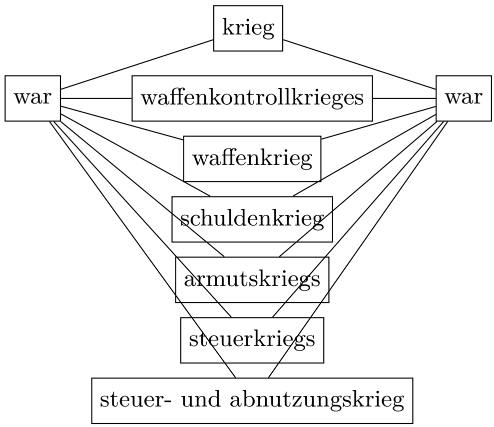


Figure 3: Translations graph of "war"

Figure 3 (and Fig. 8 to 10 in Appendix A) show samples of translation graphs of some of our English language linguistic metaphors. Often (about 25% of cases) an

English linguistic metaphor is alternatively translated by two different German verbs. This mainly happens because the distribution of words used in a text follows a power distribution (cf. Sec. 6.2).

Further, the English metaphoric words are often translated with noun compounds in German. Because we only consider metaphoric source words, German translations form a more diverse translation set, because the source and the target of an English language metaphor are combined into the same word in the German translation. This influences the stability and depth-2 *entropy* figures of English linguistic metaphors, because it leads to a high depth-2 *entropy* of words such as “war”, with a high stability at the same time (cf. Fig. 2).

4.2 Metaphor Distribution across Languages (Exp. 2)

Using the annotated English language metaphors and their contextual German translations, we investigate the translation distributions of lexelts in a language. The translation distribution of a lexelt is represented by a set of words in the TL. For example, the translation distribution of the English lexelt “oppose” is represented by the German lexelts “lehne (ab)” and “widersetzen”. For this set of words, we retrieve contextual embedding representations (CER) from a pre-trained multilingual BERT model[7]. When we average the vectors built by the contextual embedding representations V_{CER} of each word in the translation set into one single vector, we retrieve a translation vector TV for each word that is used metaphorically in our corpus:

$$TV_{oppose} = 0.5 * (V_{CER_{lehne}} + V_{CER_{widersetzten}}) \quad (4)$$

A clustering based on the 768 parameters of each vector of each lexelt can now help us to find lexelts that belong together in a—as we assume—conceptual sense. Throughout our experiments, when we talk about metaphor or metaphoric words, we only consider lexelts that were originally annotated as metaphor in the English language corpus, and their direct translations.

We use k-means clustering with a cluster size of about 10% of the total number of data-points resulting in 33 clusters for English and 38 clusters for German.[8] We use initial centroid sampling based on the data points’ weights to the overall distribution, which is run one hundred times ($n\_init = 100$). We cluster according to the 768 dimensions of the embedding representations. That is, because the 768 dimensions enable us to make use of a much more detailed parameter space, distinguishable and unifying characteristics are better usable. Tables 2 and 3 show all the clusters of metaphoric words. Clusters are often formed by inflections and compounds of the same word (for German). We also encounter a lot of synonymous and conceptual relations within the same cluster. We cannot determine the metaphor TD from the clusters, but, we can delimit the conceptual SD.

[7] `https://huggingface.co/google-bert/bert-base-multilingual-cased`, acc. Nov. 2024

[8] For English we can allow more clusters (.12) as vocabulary diversity is lower than in German.

Manual analysis shows that about 75% of the English and about 50% of the German language clusters represent moderately stable concepts in the sense that the vast majority of terms contained belong to an overarching topic. We "allow" this fuzziness, because we consider high contextual flexibility in today's language usage. The other half however, contains terms that appear too random considering semantic and domain relatedness. Be aware that the purity of clusters strongly depends on their granularity. Hence, there always will be words that need to be assigned to clusters they "rather" belong to even if they do not match the concept very well. This trend especially becomes clear as smaller, more distinguishable conceptual clusters are much cleaner. For example, "Protect" is wrongly assigned to the "attack" cluster (cf. Tab. 2, cluster no. 11). An explanation can be that "Protect" is upper-cased, so that a very specific context, such as the beginning of a sentence or a heading, probably misguided the model (cf. Tables 2 and 3).

For visualizing purposes, we transform the 768 dimension down to a 2D plot, as shown in Figure 4a and 4b. Discarding the long tail of clusters, to have a clearer view at the data[9], we encounter a representation that gives us some understanding of the arrangement of clusters in the space of contextual representations. The main purpose of visualizing the results of a clustering technique is to demonstrate that the properties (translations) used to describe an item (lexelt) are suitable to locate/arrange the semantics of that lexelt in relation to others. In our data, the plotted clusters show well-defined boundaries and the translations are thus suitable to describe the meaning of a lexelt.

We use principal component analysis (PCA Jolliffe, 2005) for dimensionality reduction, which aims to reduce dimensionality while retaining as much data variance as possible making it best suitable for our problem. In principle, PCA identifies the principal components in which the data varies the most.

Another common dimensionality reduction algorithm is multidimensional scaling. However, it is less suitable for our purpose. In contrast to PCA, MDS mainly focuses on preserving initial distances between each two data points. It also commonly is applied to problems situated in a lower-dimensional space. We also applied MDS to our clusters to visualize them in a 2D plot. Figure 5a and 5b show that while comparable clusterings are created, their visualization is less structured, hence less practical in our use-case.

4.3 Metaphoric Words in Literal Contexts (Exp. 3)

To operationalize an evaluation technique that helps us to judge whether and how clusters formed by linguistic metaphor can represent metaphorical concepts, we investigate linguistic metaphors (word types) used in both meanings, metaphorical and literal. We measure how both meanings' translation distributions differ from each other.

To measure the conceptual divergence between words that are used in both contexts, metaphorical and literal, we consider all translations from each token, used in both contexts separately, and sum the embeddings representations of the translations for

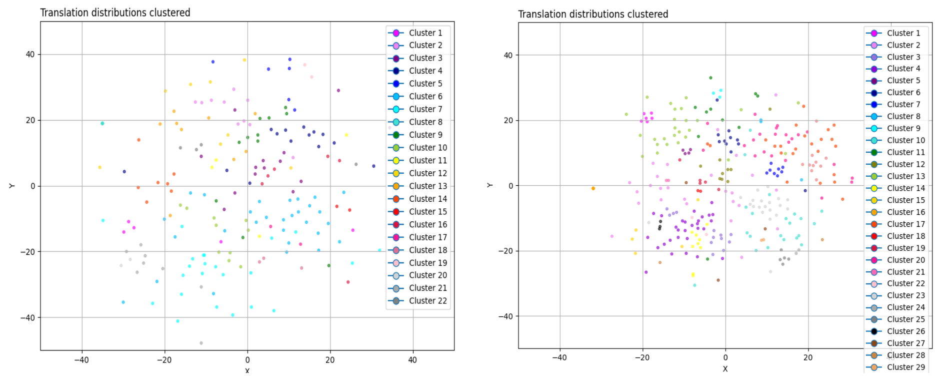
[9] 1-word clusters and "garbage" clusters; ca. 25% for German and ca. 33% for English

Table 2: Clusters of EN lexelts; high saturation indicates clean clusters; gray indicates noisy clusters

no.	lexelts
1	'circulation'
2	'flow', 'appetites', 'ceiling', 'competition', 'Race'
3	'appetite', 'parasite', 'cure', 'cured', 'remedy', 'decreased', 'bulwark', 'guardian', 'shield', 'addiction', 'inhibiting', 'corrupt', 'infection', 'evil', 'destructive', 'casualties', 'collateral damage', 'damage', 'killed', 'feeding', 'loss', 'disaster', 'disintegration', 'abyss', 'advanced', 'hole', 'hurdles', 'bondage', 'vehicle', 'burden', 'trajectory', 'sails', 'bludgeon', 'sword', 'challenged', 'competitive', 'conflicts', 'struggle', 'resist', 'defeated', 'tragedy', 'won', 'played', 'wrestle'
4	'feed', 'tastes', 'alleviate', 'relieve', 'addict', 'addicted', 'overdoses', 'escape', 'corrupts', 'cripples', 'damages', 'harms', 'hurt', 'hurts', 'weakens', 'eroding', 'harm', 'Hurt', 'impacts', 'fly out', 'fallen', 'encumbers', 'overcame', 'decreasing', 'trump', 'competes', 'fail', 'failed', 'gives up', 'Lose', 'give up', 'oppose', 'wins', 'occupy'
5	'feeds', 'decrease', 'inhibit', 'reduce', 'Reduce', 'impairs', 'shrink', 'compromises', 'impact', 'threaten'
6	'infests', 'protect', 'preserves', 'protected', 'Protecting', 'secure', 'Destroy', 'destroy', 'defeat', 'guard', 'protecting', 'protects', 'tackle', 'opposing'
7	'alleviates', 'relieves', 'restrains', 'handicaps', 'greedy', 'injures', 'threatens', 'attacked', 'attacks', 'prevents', 'advances', 'hinders', 'burdened', 'burdens', 'trumps'
8	'cures', 'heals', 'harbors', 'retreated', 'travels', 'retreats'
9	'relief', 'Guardian', 'protection', 'alleviation', 'Alleviation', 'devastation', 'destruction', 'collapse', 'Poisoning', 'diagnosis', 'annihilation', 'degradation', 'defender', 'safeguarded', 'Surrender', 'surrender', 'advancement', 'progression', 'quest', 'chasm', 'retreat', 'liberation', 'tackling', 'revolution', 'combat', 'downfall', 'overthrow', 'overcome', 'overcoming', 'Tackling'
10	'pill', 'haven', 'addicts', 'fever', 'Kills', 'cliff', 'sojourn', 'Trends', 'block', 'obstructive', 'walls', 'door', 'route', 'baseball', 'card', 'conflict', 'crisis', 'debate', 'issues', 'wars', 'team', 'community', 'blitzkrieg'
11	'Protect', 'kill', 'attack', 'fight', 'Attacking', 'fighting', 'war'
12	'overdose'
13	'kills', 'reduced', 'reduces', 'destroyed', 'destroying', 'destroys', 'recovery', 'lost', 'win', 'loses', 'winning'
14	'assault', 'invasion', 'death', 'attacking', 'evolution', 'mountain', 'engine', 'ball', 'opponent', 'foes'
15	'dangerous', 'endanger', 'endangers', 'jeopardizes'
16	'erode', 'challenge', 'safeguard', 'enforce', 'advance', 'restrain', 'eradicate', 'play'
17	'killing', 'Besieges', 'Killing', 'fought', 'Kill', 'Fighting', 'Fights', 'combating', 'vie'
18	'threatened', 'overdosed', 'barrier', 'wall', 'compete', 'resists', 'battles'
19	'preserving', 'preserve'
20	'defend', 'Defend', 'defended', 'defending', 'opposes', 'defends'
21	'growth', 'progress'
22	'journey', 'path', 'pathway'
23	'pit'
24	'barriers'
25	'Chasm'
26	'recedes', 'decreases'
27	'trail', 'season'
28	'football', 'contest', 'race', 'races', 'victory', 'duel', 'battle', 'War', 'triumph'
29	'resisted', 'opposed', 'Opposing', 'faced'
30	'resisting'
31	'failure'
32	'scored'
33	'defeating'

Table 3: Clusters of DE lexelts; high saturation indicates clean clusters; gray indicates noisy clusters

no.	lexelts
1	'Kreislauf'
2	'Fluss', 'Strom', 'Waffenfluss', 'Steuerfluss', 'Geldfluss', 'Einkommensflusses'
3	'Appetit', 'schmeckt', 'geheilt', 'Geldkuren', 'Wahlschutz', 'Überdosis', 'Infektion', 'gierig', 'böse', 'böser', 'gefährlich', 'Kills', 'ernährt', 'verletzt', 'angreift', 'anzugreifen', 'angreifen', 'Barrieren', 'Steuerlast', 'übertrumpfen', 'konkurrieren', 'greife', 'scheitern', 'scheiterte', 'versagt', 'Versagen', 'stellen', 'anzupacken'
4	'ernähren', 'verringert', 'verringern', 'verringerten', 'entlasten', 'reduzieren', 'reduziert', 'reduzierte', 'geschadet', 'zurückgegangen', 'nimmt', 'abnehmen', 'sinken'
5	'einfließt', 'befällt', 'lindert', 'heilt', 'mildert', 'entlastet', 'bewahrt', 'überdosiert', 'hindern', 'hemmend', 'zurückhält', 'korrumpiert', 'lähmt', 'schadet', 'gefährden', 'gefährdet', 'beeinträchtigt', 'trägt', 'untergräbt', 'Erhaltung', 'reist', 'belasteten', 'belastet', 'lastet', 'überwunden', 'geht', 'übertrumpft', 'konkurriert', 'nehmen'
6	'Steuerhungers', 'Umsatzsteuersucht', 'Lohnsteuerobergrenze', 'Steuerstreit'
7	'Parasit', 'Heilmittel', 'Demokratieheilmittel', 'Mittel', 'Hort', 'Bollwerk', 'Hüter', 'Übel', 'Frontalangriff', 'Hindernis', 'Last'
8	'lindern', 'heilen', 'Waffenkontrollpille', 'Schutz', 'schützen', 'geschützt', 'sichern', 'Waffensucht', 'korrumpieren', 'vernichtet', 'zerstört', 'zerstören', 'schaden', 'gefährliche', 'auszuhöhlen', 'töten', 'getötet', 'Vergiftung', 'zunichte', 'besiegen', 'Zerstöre', 'schützt', 'verletzen', 'auswirkt', 'bedrohten', 'Hüterin', 'Wächter', 'Protect', 'beschütze', 'durchzusetzen', 'belasten', 'einzuschränken', 'voranzutreiben', 'Töten', 'Widerstands', 'beseitigen', 'Überwindung', 'überwinden'
9	'Doppelbesteuerungserleichterungen', 'Erleichterung', 'Steuerentlastung', 'Grundsteuererleichterungen', 'Linderung', 'Vermögenssteuererleichterungen', 'Steuererleichterungen', 'Doppelbesteuerungsentlastung'
10	'Waffenpille', 'Armutspille', 'Armutsaufenthalt'
11	'Steuerparadies', 'Hafen', 'Waffenkontrollhafen'
12	'Entlastung', 'Grundsteuersucht', 'Armutsbekämpfung', 'bösen', 'Opfer', 'Tod', 'gesicherter', 'Abgrund', 'Waffenrechtsförderung', 'vorankommen', 'Waffenbesitztrends', 'Hürden', 'Zollmauer', 'Knechtschaft', 'Berg', 'vorangekommen', 'verloren', 'Win-Win-Antwort', 'Knüppel', 'wettbewerbsintensiver', 'Rennen', 'Wahlverlust', 'Konflikte', 'Kampf', 'Waffenproblemen', 'Krieg', 'Schuldenkrieg', 'Armutskriegs', 'Steuer- und Abnutzungskrieg', 'Zollkriegen', 'verliert', 'erzielte', 'Tragödie', 'gewinnen', 'gewinnt', 'gewonnen', 'hineinspielen', 'Pro-Gun-Teams', 'Schusswaffen-Community', 'verteidigt', 'Waffenkontrolle', 'ringen', 'Waffenkontroll-Blitzkriegs', 'konfrontiert', 'kämpften'
13	'Guardian'
14	'Waffenschützer', 'Angriff', 'herauszufordern', 'verteidigen', 'Verteidigerin', 'verteidigst', 'Verteidigung', 'greifen', 'greift', 'fördert', 'Gegner', 'widersetzen', 'Ablehnung', 'widersetzten', 'widersetzt', 'lehne', 'Gegners', 'ablehnt', 'lehnen', 'widersetzte', 'stärken'
15	'Schild', 'Opfern', 'Billionen-Dollar-Regierungsberg', 'Vehikel', 'Motor', 'Segeln', 'Vermögensträger', 'Schutzschild', 'Regierungsschild', 'Schwertes'
16	'Waffenkontrollsüchtigen', 'süchtig', 'Sucht', 'Wahlsucht', 'Ausgabensucht', 'Steuersucht', 'Waffenkontrollsüchtige', 'Waffensüchtige', 'vorgeschossen'
17	'entgangen', 'Ansturm', 'Zerstörung', 'Vermögensvernichtung', 'Verwüstung', 'Schaden', 'Invasion', 'Zusammenbruch', 'Verlust', 'Diagnose', 'Ende', 'Verschlechterung', 'Bürokratiezerfall', 'Entwicklung', 'Wachstum', 'Weiterentwicklung', 'Fortschritt', 'Rückzug', 'Befreiung', 'Förderung', 'Armutskonflikt', 'Revolution', 'Untergang', 'Sturz', 'Kapitulation', 'Übergabe'
18	'hindert', 'behindert', 'bedroht', 'droht', 'schwächt', 'schädigt', 'bedrohen', 'treibt', 'bringt'
19	'schränkt', 'belagert', 'Schützen', 'aufzugeben', 'verlieren'
20	'tötet', 'Kollateralschaden', 'Money$kills', 'aufgibt', 'fliegen', 'aufgeben', 'Niederlage'
21	'Waffenkontrollfiebers', 'Grube', 'Wiederaufbau', 'Waffenkontrolltür', 'Ball', 'Waffenkontroll-Baseball', 'Waffenkontrollkarte', 'Fußball', 'Waffenkontrollrennen', 'Waffenkrise', 'Waffendebatte', 'Waffenkontrollkampf', 'Waffenrechtskampfes', 'Wahlrevolution', 'Waffenkontrollkrieges', 'Waffenkrieg', 'Anti-Waffen-Sieg', 'Waffenteam', 'Waffenrechten', 'Waffenkontrollteam', 'Waffenrechtsgemeinschaft', 'Waffenkontrollgemeinschaft', 'Waffenkontrollangriffen'
22	'Verringerung'
23	'zerstörerischsten', 'zerstörerische'
24	'gefährlichen', 'korruptes', 'korrupten', 'korrupte', 'Steuerkonflikt'
25	'drohten', 'Vermögenskatastrophe', 'Waffenkontrollklippe', 'Geldwachstum', 'Weg', 'Suche', 'Armutssuche', 'Vermögenskluft', 'Loch', 'Armutsloch', 'Geldloch', 'Straßensperre', 'Hochzollmauern', 'Belastung', 'Waffentür', 'Waffenroute', 'Mindeststeuerpfad', 'Wahlspur', 'Waffenhasser-Karte', 'Armutskarte', 'Steuerwettbewerb', 'Steuerkriegs'
26	'Vernichtung', 'abgenommen'
27	'beeinträchtigen', 'auswirken'
28	'bekämpfen', 'kämpfen', 'Bekämpfung', 'kämpft', 'bekämpft', 'Bekämpfen', 'griff', 'bestreiten', 'antreten', 'bestritten', 'bekämpfte', 'Kampfes', 'Ungleichheiten', 'Armutsbekämpfungsindex', 'besiegte', 'besiegt', 'gehalten', 'gekämpft', 'Schwierigkeiten', 'vorgeht'
29	'verteidigte', 'verteidigten'
30	'angriffen', 'angefochten', 'forderten', 'stellten', 'herausgefordert'
31	'zurückging', 'zog', 'zieht'
32	'voranschreiten'
33	'hinausgekommen', 'weiterentwickelt', 'verfallen'
34	'Fortschritte', 'voranschreitet'
35	'obstruktive'
36	'mithalten', 'bestehen'
37	'Wahlkampfs', 'Steuerrennens', 'Wahlrennen', 'Hauptrennen', 'Wahlsieg', 'Sieg', 'Präsidentschaftswahlduell', 'Wahlsaison', 'Regierungskampf', 'Wahlkampf'
38	'US-Waffenkontrollkampfes'



(a) English **(b)** German

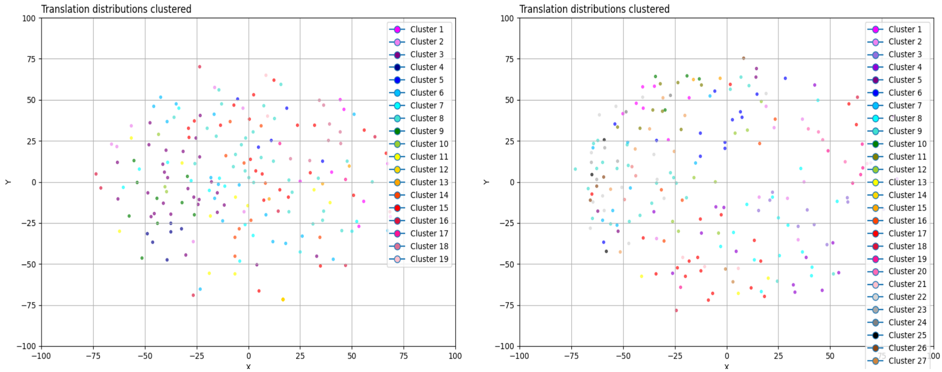
Figure 4: Translation distributions of metaphoric words using k-means after pca dimension reduction

both contexts. In this experiment, we only consider those words that have at least one literal as well as at least one metaphorical translation in our dataset. We apply several similarity statistics based on these two vectors. Appendix B Table 6 and 7 show the cosine similarity, the Z-test value and the significance level (α) based on the Kolmogorov–Smirnov test (KS-test) (Kolmogorov, 1933) under a given p-value.

The cosine value determines how similar both vectors are. That means a cosine value of 1.0 indicates that both contexts, literal and metaphorical, use the same translations. For matters of visualization, we color code higher cosine similarity values using higher color saturation. The Z-test value is a measure that is based on the mean error and the dispersion between two distributions, and, hence, measures whether two distribution are the same or different. A high absolute Z-value means that the distributions are different, a low absolute Z-value that they are similar. For matters of visualization, we color code higher Z-values using higher color saturation. Finally, the KS-test[10] shows the statistical evidence (α) to which degree the null hypothesis can be rejected under the displayed p-value, which means that a certain variable, or distribution, does not follow a given distribution. If the p-value is smaller than the significance level, both distributions cannot be considered the same or similar, hence the null hypothesis can be rejected. In other words, the higher α the more likely it is that both contexts are different from each other. While a very high α correlates with very low p-values. We can see from the table that many metaphor distributions show a clear difference between literal and metaphorical translations at least in some of the similarity metrics.[11] In the German data, this evidence appears even more frequently.

[10] We use the Python implementation as delivered with the Scipy package.

[11] The number of dimensions we work with in this experiment could be considered somewhat high (300). Hence, the significance of the KS-test might be impacted over-proportionally.



(a) English **(b)** German

Figure 5: Translation distributions of metaphoric words using k-means after mds dimension reduction

Looking at sample evaluations, we find that, when a strong difference in the distribution of literal and figurative meaning is observed, often this is caused by the very rare use of the word in the "literal" context or meaning. We look at samples from English and German words according to Tab. 6 and 7. We distinguish their usage in a literal or metaphorical context and evaluate them according to very low and very high statistically estimated difference (visualized by a low versus high saturation of blue respectively). We count how many lexelts and instances thereof from the, as we presume, literal usages are actually used in a literal sense. When the majority of all instances of a lexelt is used metaphorically and the p-value indicates strong correlation, cf. similar contextual usage, then the manual evaluation is scored +, otherwise −. In the case, the p-value indicated the opposite, we score the exact negation.

Strong statistical correlation indicates similar contexts:

"play", −:

"play" was once annotated as metaphor and translated to the German "hineinspielen". The unannotated appearances of "play" are aligned with "beitragen" (*contribute*) in the context "[...] dass höhere Strafen [...] dazu beitragen könnten [...] den Waffenfluss von Händlern zu Kriminellen einzudämmen" *([...] that increased criminal penalties [...] could help play a role in stanching the flow of guns from traffickers to criminals)*. However, in this context "beitragen" is used literally. Twice "play" translates to "spielen": Once, in the context "[...] dass er einen großen Jackpot im Bingo gewinnen würde, da er es liebt, es zu spielen!" *([...] that he would win big jackpot in bingo as he love to play it!)*, also used in the literal context. Once in: "[...] eine führende Rolle [....] spielen" *([...] could also play a leading role [...])*. In this context, "spielen" is again used metaphorically. For "play", the statistics indicate comparable contexts between metaphorical and literal cases, but the evaluation shows different usage.

"revolution", −:
"revolution" was used three times figuratively ('revolution' (1), 'wahlrevolution' (2)), and 4 times literally. Twice in "revolution by the working class", and twice in the context of the colonists' resistance against Britain, hence, all four times literally. Again, while statistics indicates similar contexts, manual evaluation show a different tendency in used contexts.

Weak statistical correlation - indicates different contexts:
"hole", −:
"hole" is used four times metaphorically (translations: 'loch' (1), 'armutsloch' (2), 'geldloch' (1)), and three times "literally" ("hole in the social safety net"), which actually, also represents a metaphorical usage even if it forms a literal relationship to "net".[12] Again, the distinguished contexts still turn out to belong to the same, metaphorical, sense.

"progress", +:
"progress" twice translated to "Fortschritt" in the German version. Both appearances are used in a personification context *("progress of their guns")*, hence it was annotated as metaphor. In the "literal" contexts, it was twice translated to "Fortschritte" out of which, once its meaning was literal: "[...] hat das FBI erhebliche Fortschritte [...] *(the FBI has made substantial progress)*". It was further once translated to "in Bearbeitung" ("[...]die Anzahl der verabschiedeten Landesgesetze und der noch in Bearbeitung befindlichen Gesetzentwürfe [...]"), again used literally. For "progress" our statistical estimate predicts correctly that it is used in a different context, when it was not annotated for metaphor in the first place.

Strong statistical correlation indicates similar contexts:
"böse", +:
"böse" appeared 11 times in the metaphorical context translated from "evil". It was also used five times in the "literal" context, again translated from "evil", always used in the metaphorical meaning ("evil gun pill", "evil wealth", "evil NRA" (2), "evil tyranny of concentrated wealth"). Statistic correlation here indicated similar context and manual evaluation proves it to be true.

"Weg", −:
"Weg" translates in the metaphoric context twice to "journey", three times to "path", once to "pathway", and once to "trajectory". In the "literal" context it translates 8 times to "way". In six of these contexts it is used literally within a lexical phrase ("to find a way to", "the [best] way to ..." (5)); twice metaphorically ("have lost the way", "along their way"). Once to "journey" (also literally, "[...] Jesus on his journey to the cross"), and once to "path" (used metaphorically, "what 's down the path"). Again statics indicate similarity, but actual usage shows distinguishing contexts.

[12] According to `collinsdictionary.com`"safety net" describes (1) "a net used in a circus to catch high-wire and trapeze artistes if they fall", or (2) "any means of protection from hardship or loss" like a social insurance can provide. Because there is a more basic meaning to "safety net", "net" and "hole" are considered metaphorical in this sense according to Steen, Dorst, Herrmann, Kaal, and Krennmayr (2010).

Weak statistical correlation - indicates different contexts:
"Ende", +:
"Ende" was once metaphorically annotated in the English version "death". In the un-annotated (tendentiously literal) appearances "Ende" was five times translated from "end", out of which four times it was used in a literal sense ("at the end of President Obama 's first term", "election [...] to an end", "defend my gun rights [...] to the end" (2)), and once metaphorically in a personification ("the end of [...] gun control efforts"). Here, again statistics indicate different usage context, and also the evaluation of the appearances in the corpus prove a tendency to different usage between literal and metaphorical senses.

"gewinnen", −:
"gewinnen" was used three times figuratively and six times not annotated, hence assumed to be used in a literal context (translated: 'win' (3), 'winning' (3)). Evaluation showed all six appearances are also in a figurative context.

We can cautiously claim that positive correlation between the contextual embedding vectors of a lexelt's translation and other metaphorical instances of a given lexelt can be a way to detect more metaphorical uses especially in German, even when they are not annotated. Positive correlation (low saturation of blue in Table 6) means that the word is translated with the same words in both contexts, which indicates that indeed the word is used in only one meaning.

Negative correlation can help to detect literal contexts of "metaphorical" words (i.e., contexts, in which a lexelt from the words annotated as metaphor is used literally) when translation variation is far off. However, translation variation does not ensure different contextual usages. This applies for example for compounds. We find that metaphorically used words translated with compounds in German, most probably, are metaphors. We present a comprehensive evaluation of computationally clearer cases in Tab. 4.

4.4 Translations of Compounds (Exp. 4)

Figure 6 shows the distribution in terms of frequency of translations of a word that is a simplicium or a compound in our dataset. Our dataset contains 41 compound-simplicium couples. Whereas 21 translations (tokens) are found for the simplicia, 37 translation tokens are found in total for the compounds. We can see that compounds and simplicia are translated slightly differently.

Table 5 shows that the cosine measures are almost always near 1.0. The German compound's determinant usually represents the metaphor's TD while the base word is the metaphor's SD. Because we work with alignment data that associate a metaphor's target, there is not necessarily a difference between the translation from German compounds/simple words to English pendants.

Figure 7 confirms that some of the compounds' translations distributions differ from those of their simplex pendants, even though the vast majority shares the same translations between compound and simplicia (cf. Figure 7).

Table 4: Sample evaluation of English lexelts (top) and German lexelts (bottom); samples showing either significant difference or no difference in metaphorical vs. literal contexts according to p-value; manu(al) evaluation summarizes whether the actually usage of lexelts matches this prediction (Exp. 3)

lexelt	p-val	figurative use	literal use (actually literally)	manu.
p-value indicates strong correlation - similar contexts:				
burden	.73	5x“last”, 1x“steuerlast” 1x“belasten”, 1x“belastung”	1x“steuerlast” (0), 1x“entlastung” (0), 1x“belastung” (0)	+
feeding	1.0	1x“ernährt”	1x“stillen” (0)	+
hurdles	.34	1x“Hürden”	1x“Hürden”(0)	+
killed	.85	1x“getötet”, 1x“töten” 1x“zunichte gemacht”	3x“getötet” (3), 1x “tötete” (1)	-
oppose	.97	4x“lehne(n) ab”, 2x“widersetzen” 1x“gegen stellen”	3x“widersetzen” (2) 2x“entgegenzustellen” (0)	-
played	.96	2x“zum Narren gehalten” (.. rubes for fools)	3x“gespielt” (2), 1x“spielten” (0)	o
play	.88	1x“hineinspielen”	1x“beitragen” (1), 2x“spielen (1)	-
protected	.48	1x“geschützte”, 1x“schützt”, 1x“schützen”	1x“geschütztes” (0) , 2x“geschützt” (0)	+
reduce	.73	4x“reduzieren”, 4x“verringern”	1x“reduzieren” (0), 1x“verringern” (0)	+
revolution	.40	1x“revolution”, 2x“wahlrevolution”	4x“revolution” (4)	-
p-value indicates no correlation - different contexts:				
block	.052	2x“Straßensperre”	1x“Barrieren” (0)	-
bulwark	.025	3x“Bollwerk”	5x“Bollwerk” (0)	-
defended	.001	3x“verteidigte”, 3x“verteidigten”	1x“verteidigen” (0)	-
feeds	.001	1x“einfließt”	1x“ernährt” (0), 1x“nährt” (0)	-
hole	.013	1x“loch”, 2x“armutsloch”, 1x“geldloch”	3x“loch” (0)	-
protects	.000	6x“schützt	2x“schützt’ (0)	-
progress	.005	2x“Fortschritt”	2x“Fortschritte” (1) 1x“in Bearbeitung” (1)	+
race	.011	1x“waffenkontrollrennen”, 1x“wahlkampfs” 1x“steuerrennens”, 1x“hauptrennen” 1x“rennen”	1x“rasse” (1) 3x“rennen” (0)	-
win	.000	1x“Win-Win-Antwort” 2x“gewinnen”, 1x“gewonnen”	3x“gewinnen” (0), Heller-Sieg (1)	+
winning	.113	1x“gewinnen”	3x“gewinnen” (0)	-
p-value indicates strong correlation - similar contexts:				
ablehnt	.811	1x“opposes”	1x“rejecting” (1)	-
böse	.730	11x“evil”	5x“evil” (0)	+
Belastung	.811	1x“burden”	1x“burden” (0)	+
erzielte	.848	1x“scored”	1x“earned” (1)	-
Hürden	.984	1x“hurdles”	1x“hurdles” (0)	+
konfrontiert	.771	1x“faced”	1x“faced” (0)	+
konkurrieren	.882	2x“compete”	3x“compete” (0)	+
Opfern	.848	1x“casualties”	1x“victims” (0)	+
Revolution	.730	1x“revolution”	4x“revolution” (4)	-
Weg	.811	2x“journey”, 3x“path”, 1x“pathway” 1x“trajectory”	8x“way” (6), 1x“journey” (1) 1x“path”(0)	-
p-value indicates no correlation - different contexts:				
bringt	.000	1x“advances” (.. voran)	1x“expresses” (.. zum Ausdruck) (0) 2x“puts further” (.. voran) (0) 2x“(doesn’t) do (much)” (2) 1x“was (a quiet one)” (0) 1x“attract” (0)	-
Ende	.000	1x“death”	5x“end” (4)	+
Entwicklung	.000	1x“evolution”	2x“Development” (2), 4x“development” (4) 3x“developing” (3)	+
Gegner	.001	6x“opponent”, 1x“foes”	1x“opponent” (1)	+
gewinnen	.001	1x“winning”, 2x“win”	3x“winning” (0), 3x“win” (0)	-
gewonnen	.001	1x“won”, 1x“(did) win”	2x“won” (0)	-
kämpfen	.000	6x“fight”, 4x“fighting”, 1x“vie”	2x“fight” (0), 1x“campaigning” (0) 1x“struggled” (0)	-
Sucht	.000	4x“addiction”	1x“addiction” (0)	-
töten	.000	9x“kill”, 1x“killed”, 1x“kills”	1x“kill” (1)	+
Tod	.000	4x“death”	1x“deaths” (1), 1x“death” (1)	+

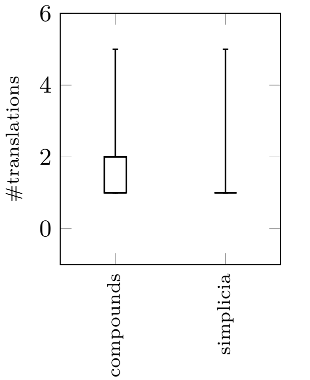


Figure 6: DE compounds' and simplex' translation distr. (En) accord. to vocab-diversity (Exp. 4)

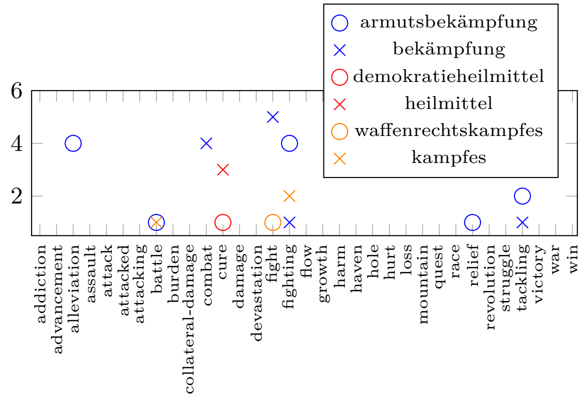


Figure 7: Distribution of samples of German compounds/simplicia with respect to their English translations and number of occurrence (Exp. 4)

5 Discussion

In Semino and Demjén (2016)'s overview of features of conceptual metaphor theory, Kövecses (2016) lists the following characteristics of conceptual metaphor that we discuss here with respect to how our experiments utilize them.

1. Concrete conceptual source domain & abstract conceptual target domain: A conceptual metaphor is a systematic set of correspondences between two domains of experience; the mapping helps to resolve a metaphor, in other words, the source domain conceptualizes the target domain, as in "fire vs anger". Our metaphor data stem from contemporary ordinary news texts and interviews that contain terms, for example, from the fighting domain, such as "assault", "attack", "battle", "fight", "hurt". We do not analyze concreteness here, but (J. Liu, O'Hara, Rubin, Draelos, & Rudin, 2020) show a clear correlation between stability and abstractness across many languages. Our experiments on **stability and depth-2** *entropy* give us a numerical understanding of how stable translations of metaphor are, and which words rather tend to be translated back to the same source word, and whether stable or volatile words are more frequent in our corpus. We find evidence that depth-2 *entropy* plays a role in distinguishing literal and figurative context of a metaphor. Further, metaphors are very stable when they are translated with German compounds. However, our metaphors often use political vocabulary, which is not especially concrete or abstract, such as "protect", "war", "decrease", "secure", "struggle".

2. Metaphor resides in thought: Kövecses (2016) writes, "When we conceptualize an intangible domain as a more tangible one, we can view a domain from different

Table 5: Corpus-specific cosine similarity between German compounds' and simplicia' translations in English according to vocab-diversity (Exp. 4)

compound-simplicia	cosine similartiy
waffenfluss-fluss	1.0
steuerfluss-fluss	1.0
geldfluss-fluss	1.0
demokratieheilmittel-heilmittel	1.0
steuerentlastung-entlastung	1.0
wahlsucht-sucht	1.0
ausgabensucht-sucht	1.0
waffensucht-sucht	1.0
grundsteuersucht-sucht	1.0
umsatzsteuersucht-sucht	1.0
steuersucht-sucht	1.0
armutsbekämpfung-bekämpfung	0.148
doppelbesteuerungsentlastung-entlastung	1.0
kollateralschaden-schaden	0.0
frontalangriff-angriff	0.120
waffenkontrollhafen-hafen	1.0
geldwachstum-wachstum	1.0
waffenrechtsförderung-förderung	1.0
armutssuche-suche	1.0
armutsloch-loch	1.0
geldloch-loch	1.0
billionen-dollar-regierungsberg-berg	1.0
steuerlast-last	1.0
waffenkontrollrennen-rennen	1.0
wahlverlust-verlust	1.0
hauptrennen-rennen	1.0
wahlsieg-sieg	0.914
waffenkontrollkampf-kampf	0.729
waffenrechtskampfes-kampfes	0.0
regierungskampf-kampf	0.724
us-waffenkontrollkampfes-kampf	0.724
wahlrevolution-revolution	1.0
waffenkrieg-krieg	1.0
schuldenkrieg-krieg	1.0
steuerkriegs-krieg	1.0
steuer-und-abnutzungskrieg-krieg	1.0
waffenkontrollangriffen-angriffen	0.0
steuersieg-sieg	0.914

perspectives", in metaphorical words, different sources. It is even possible to create a non existing TD by adding more sources. **Bilingual embeddings** enable us to form a cognitive space that is supposed to map our conceptual perception represented by massive lexicalization. In our experiments we make use of these representations.

3. Conceptual metaphors are grounded: "Why are certain sources paired with certain targets?" Kövecses (2016) asks. Following the author, often, their objective/subjective similarity, such as "life" = "gambling game", "human life cycle" = "plant life cycle", lies in certain peculiarities the TD shares with the source. In static word embedding models, the representation of a word in the embedding space does not differentiate figurative and literal meaning nor do embeddings encode whether a word is supposed to be a metaphor's SD or TD. The model represents those contexts of a word that are typical for a given domain. Say, the context of "fire" from a political news corpus would find it most often to be used metaphorically only, hence the representation reflects that. Our experiments on **literal & figurative contexts of metaphors** make use of contextual embeddings representing a word's translation set in both meanings across two languages to address ambiguity. Distribution significance measures further can help to differentiate contexts.

4. Provenance of sources: Provenance of sources means that metaphor sources come from comparable areas and physical domains such as body parts, or other concrete objects. They are often supposed to be culturally and language independent, and instead are driven by the context (Semino & Demjén, 2016). We encode our linguistic metaphor sources by word vectors from contextualized representations. Hence, we derive a language-agnostic representation of them that can be used in several experimental setups, such as **concept clustering** (Exp. 1) and **literal versus figurative context deviation** (Exp. 2).

6 Discussion of Limitations

While our study presents an empirical approach towards linking linguistic and conceptual metaphor, we acknowledge the following potential limitations of our methodology.

6.1 Human versus Automated Translation

Some may raise concerns that our machine-translated corpus is not well suited to investigate behavior in metaphor translation. We do not think this is critical. First of all, automated approaches are frequently applied in retrieving and analyzing metaphor cross-lingually. The reader may be referred to related work, such as Maudslay et al. (2020); Shutova (2010); Tsvetkov et al. (2014), see also literature in Sec.1.2. Further, neural machine translation in many domains approaches human translation quality (Wu et al., 2016), and translation quality in particular does not necessarily affect the metaphoric expressions. In addition, we ensured the quality of the translation of our parallel corpus by manual evaluation (Berger, 2022).

6.2 Power-law and Asymmetry in the Corpus

Natural language text follows a power law distribution (Zipf, 2016). Our experiments on stability, depth-2 *entropy*, as well as the translation distributions in the TL that our methodology builds upon, are affected by such distributions. The natural power-law distribution of words in a textual context in particular caused that we often see many data points near the zero, one or .5 y-axis (Exp. 1) or that we find a few bigger, a moderate number of medium-sized, and many quite small clusters (Exp. 2). However, we do not want to remove these many data points as they show empirical facts specific to the data sample we use.

On the other hand, as we do work with a one-to-one translation of one single corpus, we encounter a form of asymmetry in the metaphors detected, which might has an effect on the translation distribution of a given lexelt in our dataset.

6.3 Non-universal Conceptual Metaphors

Conceptual metaphor does not necessarily need to be universal. Concepts may be modified or lost, in particular across cultural differences. As such, Kövecses (2018) writes that the universality of conceptual metaphors can vary from culture to culture, since they experience different contextual influences and are elaborated in different ways, which also can be influenced by the given context. Also, Yan et al. (2010) investigate cross-lingual metaphor transfer asking how a "TL culture" affects the lexicalization of a metaphor and how that lexicalization, again, affects the concept of the metaphor to diverge from the SL. Yan et al. (2010) find that the way a metaphor concept is modified in a TL relies strongly on the entrenchment rank. This means that a metaphor either remains preserved, or, when not shared in the TL's system, it might be replaced by another metaphor of a higher entrenchment rank.

7 Conclusion

Our paper addresses conceptual metaphor via its expression in parallel linguistic data. We assume, following the literature, that a conceptual metaphor can be observed via its different linguistic instances. Conversely, we hypothesize that we can gain insight into conceptual metaphor by observing several linguistic metaphors in a language that belong together, as operationalized by their common translations in a parallel corpus. We test this hypothesis in four experiments carried out in a parallel news corpus annotated for linguistic metaphors, using multilingual BERT-based embeddings.

We find that a shared vocabulary space offers an opportunity to investigate how a cross-lingual phenomenon can be measured. Utilizing word distributions, we find that, first, many metaphors remain stable when they are translated, which is confirmed by measures of bilingual translation stability as well as by numeric clustering. Second, we found that words used metaphorically and literally differ in terms of their translation distributions, and that metaphor sources with the same translation set tend to be used only in one sense, the figurative one. Third, we observe that the translations of metaphors expressed in German compound nouns behave very similarly to the translations of simplicia, but our dataset containing a meaningful number of German compounds has an impact on the stability and depth-2 *entropy* of English linguistic metaphors. In the future, we plan to build on these empirical findings by designing a discrimination technique for literal and figurative contexts in parallel data.

References

Berger, M. (2022). Transfer learning parallel metaphor using bilingual embeddings. In *Proceedings of the 3rd workshop on figurative language processing (flp)* (pp. 13–23). (https://aclanthology.org/2022.flp-1.3.pdf)

Berger, M., Kiwitt, N., & Reimann, S. (2024, May). Applying transfer learning to German metaphor prediction. In N. Calzolari, M.-Y. Kan, V. Hoste, A. Lenci, S. Sakti, & N. Xue (Eds.), *Proceedings of the 2024 joint international conference on computational linguistics, language resources and evaluation (lrec-coling 2024)* (pp. 1383–1392). Torino, Italia: ELRA and ICCL. Retrieved from https://aclanthology.org/2024.lrec-main.123/

Bisang, W., Hock, H. H., Winter, W., Stefanowitsch, A., & Gries, S. T. (2006). *Corpus-based approaches to metaphor and metonymy.* Mouton de Gruyter.

Bizzoni, Y., & Ghanimifard, M. (2018). Bigrams and bilstms two neural networks for sequential metaphor detection. In *Proceedings of the workshop on figurative language processing* (pp. 91–101). (https://aclanthology.org/W18-0911.pdf)

Borbely, A. F. (1998). A psychoanalytic concept of metaphor. *The International journal of psycho-analysis*, *79*(5), 923. (https://www.proquest.com/docview/1298185260)

Casasanto, D. (2009). When is a linguistic metaphor a conceptual metaphor. *New directions in cognitive linguistics*, *24*, 127–146. (https://www.degruyterbrill.com/document/doi/10.1075/hcp.24.11cas/html)

Cettolo, M., Girardi, C., & Federico, M. (2012). Wit3: Web inventory of transcribed and translated talks. In *Proceedings of the conference of european association for machine translation (eamt)* (pp. 261–268). (https://cris.fbk.eu/bitstream/11582/104409/1/WIT3-EAMT2012.pdf)

Choi, M., Lee, S., Choi, E., Park, H., Lee, J., Lee, D., & Lee, J. (2021). Melbert: Metaphor detection via contextualized late interaction using metaphorical identification theories. *arXiv preprint arXiv:2104.13615*. (https://arxiv.org/abs/2104.13615)

Clausen, Y., & Nastase, V. (2019). Metaphors in text simplification: To change or not to change, that is the question. In *Proceedings of the fourteenth workshop on innovative use of nlp for building educational applications* (pp. 423–434). (https://aclanthology.org/W19-4444.pdf)

Dorst, A. G. (2015). More or different metaphors in fiction? a quantitative cross-register comparison. *Language and Literature*, *24*(1), 3–22. (https://journals.sagepub.com/doi/pdf/10.1177/0963947014560486)

Even-Zohar, I. (1971). *Introduction to a theory of literary translation* (Unpublished doctoral dissertation). Ph. D. Thesis, Mimeo. Tel Aviv University (Hebrew, with extensive English summary).

Fellbaum, C. (2010). Wordnet. In *Theory and applications of ontology: computer applications* (pp. 231–243). Springer. (https://link.springer.com/book/10.1007/978-90-481-8847-5)

Gao, G., Choi, E., Choi, Y., & Zettlemoyer, L. (2018). Neural metaphor detection in context. *arXiv preprint arXiv:1808.09653*. (https://arxiv.org/abs/1808.09653)

Gordon, J., Hobbs, J. R., May, J., Mohler, M., Morbini, F., Rink, B., ... Wertheim, S. (2015). A corpus of rich metaphor annotation. In *Proceedings of the third workshop on metaphor in nlp* (pp. 56–66). (https://aclanthology.org/W15-1407.pdf)

Gray, R. M. (2011). *Entropy and information theory.* Springer Science & Business Media.

Holmes, J. S. (2021). *Translated!: Papers on literary translation and translation studies. with an introduction by raymond van den broeck* (Vol. 7). Brill. (https://brill.com/display/title/27760)

Holyoak, K. J., & Stamenković, D. (2018). Metaphor comprehension: A critical review of theories and evidence. *Psychological bulletin*, *144*(6), 641. (http://www.cogsci.bme.hu/~ktkuser/KURZUSOK/BMETE47MC15/2019_2020_1/Cikkek/HolyoakStamenkovic2018.pdf)

Jolliffe, I. (2005). Principal component analysis. *Encyclopedia of statistics in behavioral science.*

Kintsch, W. (2000). Metaphor comprehension: A computational theory. *Psychonomic bulletin & review*, *7*(2), 257–266. (https://link.springer.com/content/pdf/10.3758/bf03212981.pdf)

Koehn, P. (2005). Europarl: A parallel corpus for statistical machine translation. In *Proceedings of machine translation summit x: papers* (pp. 79–86). (https://aclanthology.org/2005.mtsummit-papers.11.pdf)

Kolmogorov, A. N. (1933). Sulla determinazione empirica di una legge didistribuzione. *Giorn Dell'inst Ital Degli Att*, *4*, 89–91.

Köper, M., & Schulte im Walde, S. (2017). Improving verb metaphor detection by propagating abstractness to words, phrases and individual senses. In *Proceedings of the 1st workshop on sense, concept and entity representations and their applications* (pp. 24–30). (https://aclanthology.org/W17-1903.pdf)

Kövecses, Z. (2018). Metaphor universals in literature. *Antares: letras e humanidades*, *10*(20), 154–168.

Kövecses, Z. (2000). The scope of metaphor. In A. Barcelona (Ed.), *Metaphor and metonymy at the crossroads* (pp. 79–92). Berlin, Germany: Mouton. (https://www.degruyter.com/document/doi/10.1515/9783110894677/pdf)

Kövecses, Z. (2016). Conceptual metaphor theory. In *The routledge handbook of metaphor and language.* Taylor & Francis. (https://www.jbe-platform.com/content/journals/10.1075/arcl.6.08kov)

Lakoff, G. (2009). The neural theory of metaphor. *Available at SSRN 1437794*. (https://escholarship.org/content/qt9n6745m6/qt9n6745m6.pdf)

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by.* University of Chicago Press. (https://www.academia.edu/download/50816110/Metaphor_we_live_by.pdf)

Landauer, T. K. (2007). Lsa as a theory of meaning. *Handbook of latent semantic analysis*, *3*, 32. (http://cognaction.org/cogs105/readings/LSA.pdf)

Langacker, R. W. (1987). *Foundations of cognitive grammar: Volume i: Theoretical prerequisites* (Vol. 1). Stanford university press.

Lin, Z., Ma, Q., Yan, J., & Chen, J. (2021). Cate: A contrastive pre-trained model for metaphor detection with semi-supervised learning. In *Proceedings of the 2021 conference on empirical methods in natural language processing* (pp. 3888–3898). (https://aclanthology.org/anthology-files/anthology-files/pdf/emnlp/2021.emnlp-main.316.pdf)

Liu, J., O'Hara, N., Rubin, A., Draelos, R., & Rudin, C. (2020). Metaphor detection using contextual word embeddings from transformers. In *Proceedings of the second workshop on figurative language processing* (pp. 250–255). (https://aclanthology.org/2020.figlang-1.34.pdf)

Liu, Y., Ye, H., Weissweiler, L., Wicke, P., Pei, R., Zangenfeind, R., & Schütze, H. (2023). A crosslingual investigation of conceptualization in 1335 languages. In *Proceedings of the 61st annual meeting of the association for computational linguistics (volume 1: Long papers)* (pp. 12969–13000). Toronto, Canada: Association for Computational Linguistics. (https://aclanthology.org/2023.acl-long.726/)

Luong, M.-T., Pham, H., & Manning, C. D. (2015). Bilingual word representations with monolingual quality in mind. In *Proceedings of the 1st workshop on vector space modeling for natural language processing* (pp. 151–159). (https://aclanthology.org/W15-1521.pdf)

Maudslay, R. H., Pimentel, T., Cotterell, R., & Teufel, S. (2020). Metaphor detection using context and concreteness. In *Proceedings of the second workshop on figurative language processing* (pp. 221–226). (https://aclanthology.org/2020.figlang-1.pdf)

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*, *26*. (https://proceedings.neurips.cc/paper/2013/file/9aa42b31882ec039965f3c4923ce901b-Paper.pdf)

Mohler, M., Tomlinson, M. T., & Rink, B. (2015). Cross-lingual semantic generalization for the detection of metaphor. *Int. J. Comput. Linguistics Appl.*, *6(2)*, 117-140. (https://www.gelbukh.com/ijcla/2015-2/IJCLA-2015-2-pp-117-140-Cross-lingual-Semantic-Generalization.pdf)

Papakyriakopoulos, O., Hegelich, S., Serrano, J. C. M., & Marco, F. (2020). Bias in word embeddings. In (p. 446–457). Association for Computing Machinery. (https://dl.acm.org/doi/pdf/10.1145/3351095.3372843)

Pellauer, D., & Dauenhauer, B. (2002–2022). *Paul Ricoeur.* online; Stanford Encyclopedia of Philosophy. (https://plato.stanford.edu/entries/ricoeur/)

Pennington, J., Socher, R., & Manning, C. D. (2014). GloVe: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)* (pp. 1532–1543). (https://aclanthology.org/D14-1162.pdf)

Piccioni, S. (2013). What can metaphor tell us about the language of translation? *Procedia - Social and Behavioral Sciences*, *95*, 354-362.

Pragglejaz Group. (2007). MIP: A method for identifying metaphorically used words in discourse. *Metaphor and symbol*, *22*(1), 1–39. (https://www.tandfonline.com/doi/abs/10.1080/10926480709336752)

Ricoeur, P. (1978). *The rule of metaphor: Multi-disciplinary studies of the creation of meaning in language.* Routledge and Kegan Paul. (trans. Robert Czerny with Kathleen McLaughlin and John Costello, S. J., https://muse.jhu.edu/pub/50/monograph/book/105427)

Sanchez-Bayona, E., & Agerri, R. (2022). Leveraging a new spanish corpus for multilingual and crosslingual metaphor detection. *arXiv e-prints*, arXiv–2210.

Schäffner, C. (2004). Metaphor and translation: some implications of a cognitive approach. *Journal of pragmatics*, *36*(7), 1253–1269. (https://www.sciencedirect.com/science/article/pii/S0378216604000244?)

Schuster, J., & Markert, K. (2023). Nut-cracking sledgehammers: Prioritizing target language data over bigger language models for cross-lingual metaphor detection. In E. Breitholtz, S. Lappin, S. Loaiciga, N. Ilinykh, & S. Dobnik (Eds.), *Proceedings of the 2023 clasp conference on learning with small data (lsd)* (pp. 98–106). Gothenburg, Sweden: Association for Computational Linguistics. Retrieved from https://aclanthology.org/2023.clasp-1.12

Semino, E., & Demjén, Z. (2016). *The routledge handbook of metaphor and language.* Taylor & Francis. (https://www.academia.edu/download/79134201/9781315672953.pdf)

Shutova, E. (2010). Automatic metaphor interpretation as a paraphrasing task. In *Human language technologies: The 2010 annual conference of the north american chapter of the association for computational linguistics* (pp. 1029–1037). (https://aclanthology.org/N10-1147.pdf)

Shutova, E., & Teufel, S. (2010). Metaphor corpus annotated for source-target domain mappings. In *Lrec* (Vol. 2, pp. 2–2). (https://openreview.net/pdf?id=8TVq60hx1bB)

Song, W., Zhou, S., Fu, R., Liu, T., & Liu, L. (2021). Verb metaphor detection via contextual relation learning. In C. Zong, F. Xia, W. Li, & R. Navigli (Eds.), *Proceedings of the 59th annual meeting of the association for computational linguistics and the 11th international joint conference on natural language processing (volume 1: Long papers)* (pp. 4240–4251). Online: Association for Computational Linguistics. Retrieved from https://aclanthology.org/2021.acl-long.327 doi: 10.18653/v1/2021.acl-long.327

Steen, G. J., Dorst, A. G., Herrmann, J. B., Kaal, A. A., & Krennmayr, T. (2010). Metaphor in usage. *Cognitive Linguistics, vol. 21, no. 4*. (https://www.degruyter.com/document/doi/10.1515/cogl.2010.024/html)

Teich, E. (2003). *Cross-linguistic variation in system and text: A methodology for the investigation of translations and comparable texts.* de Gruyter.

Tsvetkov, Y., Boytsov, L., Gershman, A., Nyberg, E., & Dyer, C. (2014). Metaphor detection with cross-lingual model transfer. In *Proceedings of the 52nd annual meeting of the association for computational linguistics (volume 1: Long papers)* (pp. 248–258). (https://aclanthology.org/P14-1024.pdf)

Tsvetkov, Y., Mukomel, E., & Gershman, A. (2013). Cross-lingual metaphor detection using common semantic features. In *Proceedings of the first workshop on metaphor in nlp* (pp. 45–51). (https://aclanthology.org/W13-0906.pdf)

Upadhyay, S., Faruqui, M., Dyer, C., & Roth, D. (2016). Cross-lingual models of word embeddings: An empirical comparison. *arXiv preprint arXiv:1604.00425*. (https://arxiv.org/pdf/1604.00425)

Van den Broeck, R. (1981). The limits of translatability exemplified by metaphor translation. *Poetics today*, *2*(4), 73–87. (https://www.jstor.org/stable/pdf/1772487.pdf)

Vaswani, A. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*.

Veale, T., Shutova, E., & Klebanov, B. B. (2016). Metaphor: A computational perspective. *Synthesis Lectures on Human Language Technologies*, *9*(1), 1–160. (https://openreview.net/pdf?id=8dwgVhag_wk)

Vukovic, N. T. (2007). Corpus-based approaches to metaphor and metonymy. *Suvremena Lingvistika*, *33*(63), 133–134. (https://www.researchgate.net/publication/292744446_Corpus-based_approaches_to_metaphor_and_metonymy)

Vulić, I., & Moens, M.-F. (2015). Monolingual and cross-lingual information retrieval models based on (bilingual) word embeddings. In *Proceedings of the 38th international acm sigir conference on research and development in information retrieval* (pp. 363–372). (https://dl.acm.org/doi/pdf/10.1145/2766462.2767752)

Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., Macherey, W., … others (2016). Google's neural machine translation system: Bridging the gap between human and machine translation. *arXiv preprint arXiv:1609.08144*. (https://arxiv.org/abs/1609.08144)

Yan, D., Noël, D., & Wolf, H.-G. (2010). Patterns in metaphor translation: a corpus-based case study of the translation of fear metaphors between english and chinese. *Using Corpora in Contrastive and Translation Studies. Newcastle: Cambridge Scholars Publishing*, 40–61. (http://www.lancaster.ac.uk/fass/projects/corpus/UCCTS2008Proceedings/papers/Ding_et_al.pdf)

Zipf, G. K. (2016). *Human behavior and the principle of least effort: An introduction to human ecology.* Ravenio books.

Appendix A: Translation Graphs (Exp. 1)

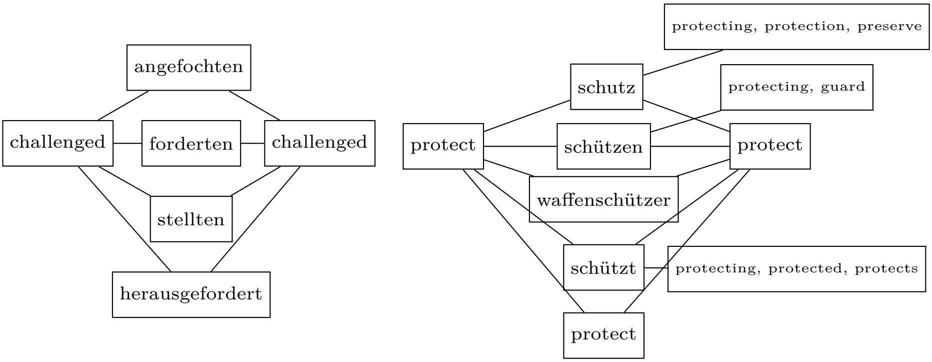


Figure 8: Translation graphs of "challenged" and "protect"

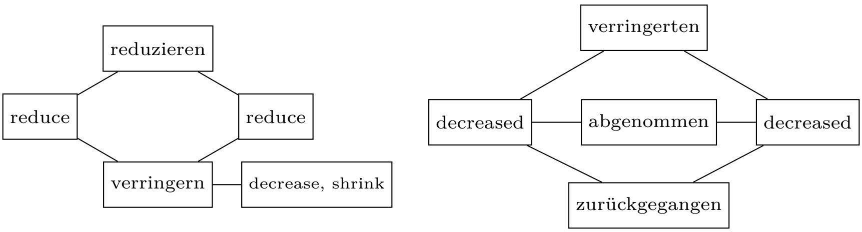


Figure 9: Translation graphs of "reduce" and "decreased"

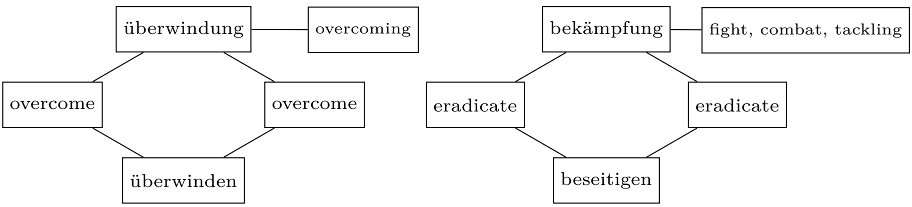


Figure 10: Translation graphs of "overcome" and "eradicate"

Appendix B: Significance Statistics (Exp. 3)

Table 6: Cos(ine) sim(ilarity), Z-test, significance level and p-value of the KS-test, stability and entropy; strong saturation indicates significant difference

lexelt	cos sim	Z-val	α	p-val
feeds	0.61	0.05	0.10	0.001
relief	0.68	0.07	0.16	0.000
decrease	0.56	0.06	0.10	0.002
bulwark	0.74	0.01	0.08	0.025
protected	0.78	0.01	0.04	0.478
escape	0.59	0.01	0.03	0.771
assault	0.69	0.02	0.06	0.180
reduce	0.73	0.03	0.04	0.730
evil	0.86	0.01	0.04	0.518
damage	0.86	0.01	0.05	0.224
killed	0.86	0.02	0.03	0.848
threatens	0.76	0.02	0.09	0.003
feeding	0.61	0.02	0.02	0.992
challenge	0.71	0.02	0.04	0.644
death	0.85	0.01	0.05	0.304
impact	0.67	0.01	0.06	0.100
defend	0.81	0.03	0.04	0.559
defended	0.49	0.02	0.10	0.001
guard	0.61	0.01	0.05	0.275
preserve	0.67	0.03	0.12	0.000
protects	0.63	0.05	0.11	0.000
safeguard	0.51	0.02	0.05	0.368
growth	0.91	0.01	0.05	0.224
advance	0.65	0.04	0.11	0.000
hole	0.79	0.01	0.08	0.013
block	0.56	0.04	0.07	0.052
hurdles	0.86	0.01	0.05	0.335
burden	0.75	0.05	0.04	0.730
progress	0.53	0.00	0.09	0.005
path	0.56	0.12	0.06	0.143
win	0.71	0.03	0.11	0.000
compete	0.76	0.02	0.06	0.113
competition	0.68	0.03	0.11	0.000
competitive	0.73	0.02	0.10	0.002
race	0.77	0.04	0.08	0.011
season	0.65	0.06	0.12	0.000
debate	0.77	0.00	0.04	0.602
revolution	0.83	0.02	0.05	0.403
oppose	0.80	0.04	0.02	0.973
winning	0.50	0.05	0.06	0.113
won	0.73	0.02	0.04	0.687
play	0.61	0.01	0.03	0.882
played	0.58	0.03	0.03	0.957
overcome	0.60	0.06	0.07	0.068
overcoming	0.68	0.04	0.04	0.439

Table 7: Cos(ine) sim(ilarity), Z-test result, statistical significance for given p-value of the Kolmogorov–Smirnov test, strong saturation indicates significant difference

lexelt	cos sim	Z-val	α	p-val
Erleichterung	0.62	0.05	0.05	0.368
verringert	0.57	0.01	0.06	0.180
Bollwerk	0.45	0.06	0.08	0.016
Sucht	0.67	0.01	0.11	0.000
behindert	0.66	0.04	0.08	0.016
Verringerung	0.49	0.04	0.06	0.100
böse	0.86	0.02	0.04	0.730
zerstört	0.86	0.02	0.07	0.039
Steuererleichterungen	0.56	0.06	0.02	0.973
Opfer	0.65	0.02	0.08	0.010
töten	0.75	0.02	0.11	0.000
getötet	0.78	0.01	0.10	0.001
droht	0.77	0.00	0.04	0.518
Opfern	0.61	0.07	0.03	0.848
ernährt	0.61	0.01	0.04	0.730
Zusammenbruch	0.68	0.05	0.03	0.771
Verlust	0.68	0.02	0.04	0.478
besiegen	0.69	0.05	0.09	0.006
Tod	0.70	0.08	0.14	0.000
Ende	0.70	0.01	0.11	0.000
schützt	0.71	0.03	0.07	0.052
verletzt	0.53	0.05	0.04	0.559
verteidigen	0.83	0.02	0.07	0.068
Schützen	0.83	0.00	0.04	0.602
kämpfen	0.87	0.03	0.12	0.000
greift	0.57	0.01	0.06	0.180
Entwicklung	0.66	0.04	0.11	0.000
Wachstum	0.80	0.02	0.08	0.010
fördert	0.53	0.01	0.08	0.016
bringt	0.77	0.05	0.11	0.000
Fortschritt	0.69	0.02	0.10	0.002
Weg	0.79	0.02	0.03	0.811
Suche	0.52	0.04	0.04	0.518
Loch	0.66	0.03	0.07	0.052
Hindernis	0.58	0.04	0.04	0.559
Hürden	0.73	0.03	0.02	0.984
Belastung	0.69	0.04	0.03	0.811
Förderung	0.65	0.03	0.06	0.100
konkurrieren	0.72	0.00	0.03	0.882
Steuerwettbewerb	0.65	0.06	0.06	0.113
Kampf	0.84	0.00	0.04	0.518
Revolution	0.85	0.01	0.04	0.730
Krieg	0.77	0.01	0.07	0.045
Gegner	0.61	0.03	0.10	0.001
widersetzt	0.61	0.03	0.05	0.224
verlieren	0.54	0.03	0.07	0.077
erzielte	0.55	0.00	0.03	0.848
Waffenrechten	0.77	0.01	0.10	0.000
Tragödie	0.74	0.01	0.09	0.007
gewinnen	0.55	0.05	0.10	0.001
gewonnen	0.67	0.03	0.10	0.001
ablehnt	0.68	0.01	0.03	0.811
Überwindung	0.66	0.07	0.05	0.275
konfrontiert	0.52	0.01	0.03	0.771
stärken	0.63	0.03	0.04	0.559
Schwierigkeiten	0.54	0.06	0.03	0.911

Correspondence

Maria Berger

Ruhr University Bochum
German Department
Bochum, Germany
maria.berger-a2l@rub.de

Tatjana Scheffler

Ruhr University Bochum
German Department
Bochum, Germany
tatjana.scheffler@rub.de