On the Geometry of Concreteness

Christian Wartena
Hochschule Hannover
Expo Plaza 12
30539 Hannover, Germany
`christian.wartena@hs-hannover.de`

Abstract

In this paper we investigate how concreteness and abstractness are represented in word embedding spaces. We use data for English and German, and show that concreteness and abstractness can be determined independently and turn out to be completely opposite directions in the embedding space. Various methods can be used to determine the direction of concreteness, always resulting in roughly the same vector. Though concreteness is a central aspect of the meaning of words and can be detected clearly in embedding spaces, it seems not as easy to subtract or add concreteness to words to obtain other words or word senses like e.g. can be done with a semantic property like gender.

1 Introduction

In the current paper we aim to shed some light on the way concreteness and abstractness are represented in word embeddings. This might help to better understand the concept of concreteness that seems to be an important semantic property used to explain various phenomena in language and language understanding. Ultimately, it might also contribute a little bit to the understanding of the semantic spaces in which we embed words for many tasks.

1.1 Research Questions

The first question we want to address is to what degree concreteness can be represented as a vector in the embedding space. Such a vector v_{concr} should have the property that either the cosine between a vector v_w for a word w and v_{concr} or the length of the projection of v_w on v_{concr} corresponds to the concreteness value w.

The second question concerns the relation between concreteness and abstractness. Is abstractness characterized by a direction in the embeddings space in a similar way as concreteness is, or is abstractness just the absence of concreteness? The studies of Hill and Korhonen (2014) and Naumann et al. (2018) suggest that abstractness could go into many different directions and is quite different from concreteness since abstract words occur in more diverse contexts. On the other hand side, in all psycholinguistic studies concreteness and abstractness are treated as the two extremes on one scale.

Finally, we want to know, whether concreteness is a property that can be added to or removed from words, like e.g. gender can be separated and used to explicitly relate words like *king* to *queen*. There seem to be many cases of regular polysemy in which one reading of the word is more concrete than the other one. Examples are the polysemy between buildings and institutions for words like *school*, *church*, *parliament*, *theater*, etc. or between process and the result of the process (like e.g. *creation*, that can either denote the process of creating something or the thing that is created) or between a function and the person holding that function. In all cases it is clear that one reading is more concrete than the other one. What we want to know is to which extend the difference between the two meanings is determined by concreteness.

1.2 Concreteness

Concreteness is a core semantic property of words that has received a lot of attention in psycholinguistic research. Friendly et al. (1982) define concrete words as words that "refer to tangible objects, materials or persons which can be easily perceived with the senses". Brysbaert et al. (2014) define concreteness as the degree to which the concept denoted by a word refers to a perceptible entity. Theijssen et al. (2011) point out that in general two concepts of concreteness are used that do not completely overlap, namely *sensory perceivability* and *specificity*. However, they also note that most subjects in tests interpret concreteness as *sensory perceivability*. Also in a corpus study they could show that in cases where concreteness plays a role

204

in the choice of a syntactic construction, *sensory perceivability* is the best predictor.

Various studies suggest that concrete and abstract words are represented and processed differently by the human brain (see a.o. Binder et al. (2005); Kousta et al. (2011); Borghi et al. (2017)). E.g. it is assumed that concreteness influences learning, recognition memory and the speed of visual recognition, reading and spelling (Spreen and Schulz, 1966; Hargis and Gickling, 1978; Sadoski et al., 2004; Palmer et al., 2013; Neath and Surprenant, 2020). Moreover, studies conducted on abstract and concrete words also found that the participants remembered concrete words better than the abstract words (for an overview of various studies see e.g. Yui et al., 2017). This difference is explained by the Dual Coding Theory (Paivio, 1970) according to which concrete concepts are stored verbally and visually in mind while abstract concepts are only stored verbally. A difference in recognition ease and speed is explained by the Context Availability Hypothesis (Schwanenflugel and Shoben, 1983; Schwanenflugel, 2013). This hypothesis states that it is crucial to evoke the context of a word to access its meaning and that it is easier to construct the appropriate context for concrete than for abstract words.

Among others Hill et al. (2014) and Naumann et al. (2018) have shown that abstract words occur in more broad and diverse contexts than concrete words. Furthermore, it was noted in several studies (see e.g. (Tanaka et al., 2013)) and investigated in detail by Frassinelli et al. (2017) and Naumann et al. (2018) that concrete words tend to occur in the context of other concrete words and abstract words in the context of other abstract words.

Most studies that collected or predicted concreteness values for words either ignored the fact that many words have several senses or excluded ambiguous words. The statement of Gilhooly and Logie (1980) still seems to be valid: "The problem of word ambiguity has generally been overlooked in compiling lists of words measured on various attributes." Only a few mostly smaller studies collected concreteness judgments for different word senses. These are, as far as we know, (Gilhooly and Logie, 1980) for English, (Hager, 1994) for German, and more recently (Đurđević et al., 2017) for Serbian and both (Reijnierse et al., 2019) and (Scott et al., 2019) for English words.

1.3 Organization of this paper

The remainder of the paper is organized as follows. In section 2 start with an overview of the few studies that try to identify concreteness in embedding spaces. In section 3 we describe the data we have used. In the following sections we present a series of experiments to get a better understanding of the representation of concreteness in embedding spaces: in section 4 we compare several possibilities to determine the direction of concreteness and abstractness in an embedding space, in section 5 we compare the mutual similarity between concrete and abstract words and finally in section 6 we have a short look at the possibilities to represent the meaning of ambiguous words with a concrete and an abstract sense.

2 Related Work

Word embeddings are widely used as a proxy for the meaning of words but in fact word embeddings are chiefly compact representations of the contexts in which they occur. Since concrete words occur preferably in the context of other concrete words and since concrete words are used as object to sensory verbs we expect that concreteness can be found in word embeddings. Indeed a number studies have shown the presence of concreteness in word embeddings: Rothe et al. (2016) try to find low-dimensional feature representations of words in which at least some dimensions correspond to interpretable properties of words. One of these dimensions is concreteness. For training and testing they use Google News embeddings and two subsets of frequent words from the norms of Brysbaert et al. (2014). For their test set of 8,694 frequent words they found a moderate correlation with the human judgments (Kendall's $\tau = 0.623$). Similarly, Hollis and Westbury (2016) investigated which dimensions of word embeddings correlate to one of the classical word norms. They found no direct correlations, but after reducing the number of dimensions for a set of words by applying Singular Value Decomposition, they found a strong correlation between one of the dimensions and concreteness. Charbonnier and Wartena (2019, 2020) train regression models on word embeddings to predict concreteness values, thus showing that concreteness information is present in the embeddings.

3   Materials

The answers to the research questions might depend on the embeddings we use. Nevertheless, we will restrict the experiments to just two embeddings, one for English and one for German, and for the moment being assume that results for other embeddings will be similar.

For English we use the 300 dimensional fastText embeddings without subword information trained on the Common Crawl with 600 billion tokens. For German we also use 300 dimensional fastText embeddings trained on the Common Crawl and Wikipedia. Both embeddings are available at the fastText site (https://fasttext.cc/).

The concreteness values are taken from Brysbaert et al. (2014) for English and from the merged dataset from Charbonnier and Wartena (2020) for German. Since concreteness is most clearly defined for nouns, from both datasets we use only nouns for which we also have embeddings. In the data from Brysbaert et al. (2014) the words are rated between 1.0 and 5.0. The ratings for the German data range from 1.0 to 7.0. As examples of concrete nouns we take for the English data all nouns rated above 4.0 and for German all nouns rated above 6.0. As clearly abstract nouns we use nouns rated below 2.7 for English and rated below 4 for German. This results in the numbers given in Table 1.

Table 1: Number of abstract and concrete nouns used.

	English	German
nouns	18,307	3,281
concrete nouns	6,345	753
abstract nouns	5,713	1,072

4   Concreteness vectors

In this section we will compare different methods to build prototypical vectors for concreteness and abstractness.

4.1   Methods

A straightforward method to obtain a vector for concreteness is to take the average embedding of all concrete words and subtract the average embedding of all words. As a second method we can take embeddings of concrete and abstract words, apply principal component analysis (PCA) and hope that the most important component represents concreteness. Finally, we can use use linear regression to find a vector that fits best to the concreteness values in the data set.

Since concreteness is most clearly defined for nouns, we take the average of all embeddings of concrete nouns and subtract the average of all noun vectors. The same can be done for abstract nouns and if, hopefully, the vectors for concreteness and abstractness roughly point in opposite directions, we can compute the average of the concrete and the opposite of the abstract vector, to get one vector representing concreteness and abstractness. Formally, let v_n be the average of the word embeddings of all nouns, v_{cn} the average of all embeddings of all concrete nouns and v_{an} the average of all embeddings of all abstract nouns, for the sets of abstract and concrete nouns as defined in section 3. Now let

$$v_{concr} = v_{cn} - v_n, \quad (1)$$
$$v_{abstr} = v_{an} - v_n. \quad (2)$$

For convenience we will use unit vectors defined as usual by setting $\hat{v}_{concr} = \frac{v_{concr}}{|v_{concr}|}$ and $\hat{v}_{abstr} = \frac{v_{abstr}}{|v_{abstr}|}$.

A vector based both on concrete and on abstract words can be defined as

$$v_{ca} = \frac{\hat{v}_{concr} - \hat{v}_{abstr}}{2}, \quad (3)$$
$$\hat{v}_{ca} = \frac{v_{ca}}{|v_{ca}|}. \quad (4)$$

For the principal component analysis we take the same sets of concrete and abstract words and put their embeddings in one matrix on which we perform PCA with 12 components. We take the first component as concreteness vector that we will call v_{pca} in the following.

For the regression we use all nouns, not just the most concrete and abstract ones. For all words we use their length normalized embeddings. As first option we use standard multiple linear regression, minimizing the sum of squared errors between real and predicted concreteness value. We let v_{regr}^2 be the vector of the regression coefficients. Since we use the squared errors, the linear regression is quite sensitive to outliers. As an alternative we use linear regression with Huber loss function that is defined as:

$$L_H(\alpha) = \begin{cases} \frac{1}{2}(\alpha)^2 & \text{for } |\alpha| \leq \delta, \\ \delta\left(|\alpha| - \frac{1}{2}\delta\right), & \text{otherwise,} \end{cases} \quad (5)$$

where $\alpha = y - f(x)$ is the residual or prediction error. For both the German and the English data

we set $\delta = 0.25$. Finally, we add $\gamma \, ||w||_1$ as a regularization term, where w is the vector of regression coefficients. We set $\gamma = 1 \cdot 10^{-4}$. We call the resulting vector of coefficients v_{regr}^1.

4.2 Results

We do not have any method to access the quality of the vectors obtained by the different methods, but we can at least compare them. Furthermore, we can compute the correlation between real concreteness values of a word and the length of the projections of word embeddings on the concreteness vectors. The later value cannot be seen as a real concreteness prediction, but gives some indication how well the concreteness vector fits to the actual data.

For the English data we find $|v_{concr}| = 0.152$ and $|v_{abstr}| = 0.156$, for German $|v_{concr}| = 0.222$ and $|v_{abstr}| = 0.160$. Here we do not see a noticeable difference between concrete and abstract words.

Tables 2 and 3 give the cosine similarities between the various concreteness vectors for English and German respectively.

The first remarkable observation is that, both for the English and German data, the angle between v_{concr} and v_{abstr} is almost 180 degrees. This is maybe the most remarkable result of the present study: the vectors computed independently for distinct sets of concrete and abstract words are almost perfectly diametrically opposed! This suggests that abstractness and concreteness are indeed to extremes on the scale of the same property.

Furthermore, we see that all vectors are very much alike, except v_{regr}^2, the vector of coefficients of a classical linear regression model. Here indeed extreme values seem to dominate and specify a direction different from those obtained by all other methods.

A second indication for the quality of the concreteness vectors is the degree to which they can be used to predict the concreteness of individual words. Ideally, the length of the projection of a embedding vector on the concreteness vector would correspond to the empirically determined concreteness values. As it is not clear whether the length of an embedding value has any meaning or just the direction is important, we also could assume that the cosine between a word vector and the concreteness vector should be used. In Table 4 we therefore give the correlation (Pearsons's r and Kendalls's τ) for cosine and projection length. We should not interpret these numbers as an attempt to predict the concreteness. In the first place it would be easy to design a better (non linear) prediction model and in the second place we did not split into training and test data to make a sound prediction experiment (However, the vectors were, dependent on the method, computed using only a small part of the data, e.g. only nouns and the results might moreover not change, when one or a few words would be left out from the data).

Again we see that the values for English and German are almost the same. In both cases we see that v_{regr}^2 gives the best correlation, which is not very surprising since this vector was optimized for Pearson correlation. More remarkable is the fact that, especially for the English data, the correlation of v_{regr}^1 with the concreteness judgements is not much worse. Furthermore, we see that the cosine is a much better predictor for the concreteness values than the projection length. Given that the cosine is just the projection length of the unit vector of the word embedding, this suggests that vector length in word embeddings is not relevant and only the direction matters. Finally, the correlation is in the same order of magnitude as the correlation found by Rothe et al. (2016) but much behind the results from from Charbonnier and Wartena (2019), who use a non-linear classifier and additional morphological information.

5 Diversity of concrete and abstract words

As discussed above it has been observed that abstract words occur in more diverse contexts than concrete words. Does this also mean that abstract words are more diverse? I.e., can words be concrete just in one way but abstract in many different ways? To answer this question we selected randomly 100 words from our set of concrete and 100 from the set of abstract words. We compute the average cosine similarity for all pairs of words within each set and within the union of both sets. The results are given in Table 5. We see here no large differences between the abstract and concrete nouns. The abstract nouns even seem to be slightly more similar to each other than the concrete nouns. This again suggests that abstractness and concreteness are quite symmetric properties. The average similarity within each set (4,950 pairs for each set) is clearly larger that within the entire set of 200 nouns (i.e. 19,900 pairs), showing the importance of concreteness for

Table 2: Cosine similarities between concreteness vectors computed using different methods for English data.

	v_{concr}	v_{abstr}	v_{ca}	v_{pca}	v_{regr}^1	v_{regr}^2
v_{concr}	1.000	-	-	-	-	-
v_{abstr}	-0.945	1.000	-	-	-	-
v_{ca}	0.986	-0.986	1.000	-	-	-
v_{pca}	0.916	-0.882	0.912	1.000	-	-
v_{regr}^1	0.955	-0.960	0.971	0.812	1.000	-
v_{regr}^2	0.630	-0.589	0.618	0.447	0.695	1.000

Table 3: Cosine similarities between concreteness vectors computed using different methods for German data.

	v_{concr}	v_{abstr}	v_{ca}	v_{pca}	v_{regr}^1	v_{regr}^2
v_{concr}	1.000	-	-	-	-	-
v_{abstr}	-0.917	1.000	-	-	-	-
v_{ca}	0.979	-0.979	1.000	-	-	-
v_{pca}	0.937	-0.888	0.932	1.000	-	-
v_{regr}^1	0.945	-0.990	0.988	0.914	1.000	-
v_{regr}^2	0.572	-0.585	0.591	0.457	0.593	1.000

Table 4: Correlation (Pearsons's r) and rank correlation (Kendalls's τ) of concreteness values with the lengths of the projection of each word vector on a concreteness vector and the correlation with the cosines between each word vectors and a concreteness vector for different concreteness vectors.

		projection		cosine	
		P's r	K's τ	P's r	K's τ
English	v_{ca}	0.74	0.61	0.85	0.65
	v_{pca}	0.63	0.55	0.78	0.59
	v_{regr}^1	0.78	0.64	0.86	0.67
	v_{regr}^2	0.81	0.67	0.89	0.71
German	v_{ca}	0.71	0.56	0.80	0.59
	v_{pca}	0.64	0.49	0.74	0.54
	v_{regr}^1	0.72	0.56	0.80	0.60
	v_{regr}^2	0.78	0.62	0.84	0.65

Table 5: Average cosine similarity between 100 abstract and 100 concrete nouns

	English	German
concr	0.13	0.22
abstr	0.15	0.23
concr $\cup$ abstr	0.11	0.18

similarity in the embedding space.

6 Concreteness and regular polysemy

A word like *school* can refer to the schoolhouse or to the educational institution. In our sets of word embeddings there is only one embedding for all meanings of the word school, even including the sense of a group (as in a school of fish), a group of artists or thinkers and even the verb *to school*. We would hope that if we add a little bit of con-

creteness to the embedding of school, we get an embedding that is a bit closer to the embedding of schoolhouse and if we add some abstractness, the embedding becomes more similar to other abstract concepts from education. As a first indication to see whether this is indeed the case, we visualize the distances between a few ambiguous words (*school*, *university* and *hospital* for English and *Schule* (school), *Universität* (university) and *Fabrik* (factory) for German) along with some related concrete and abstract words. For each ambiguous word w we use the original embedding v_w as well as $v_w + 0.2\hat{v}_{ca}$ and $v_w - 0.2\hat{v}_{ca}$. We add $0.2\ \hat{v}_{ca}$ since 0.2 is roughly the length of the projection of the most abstract and the most concrete words on $\hat{v}_{ca}$. In the visualization the variants are labeled with the original word and either an a or c. The projection in a two-dimensional space is done with tSNE (Van der Maaten and Hinton, 2008). The results are shown in Figure 1 [1].

---

[1] The translation of all German words used in this figure and the subsequent tables is given in the appendix.
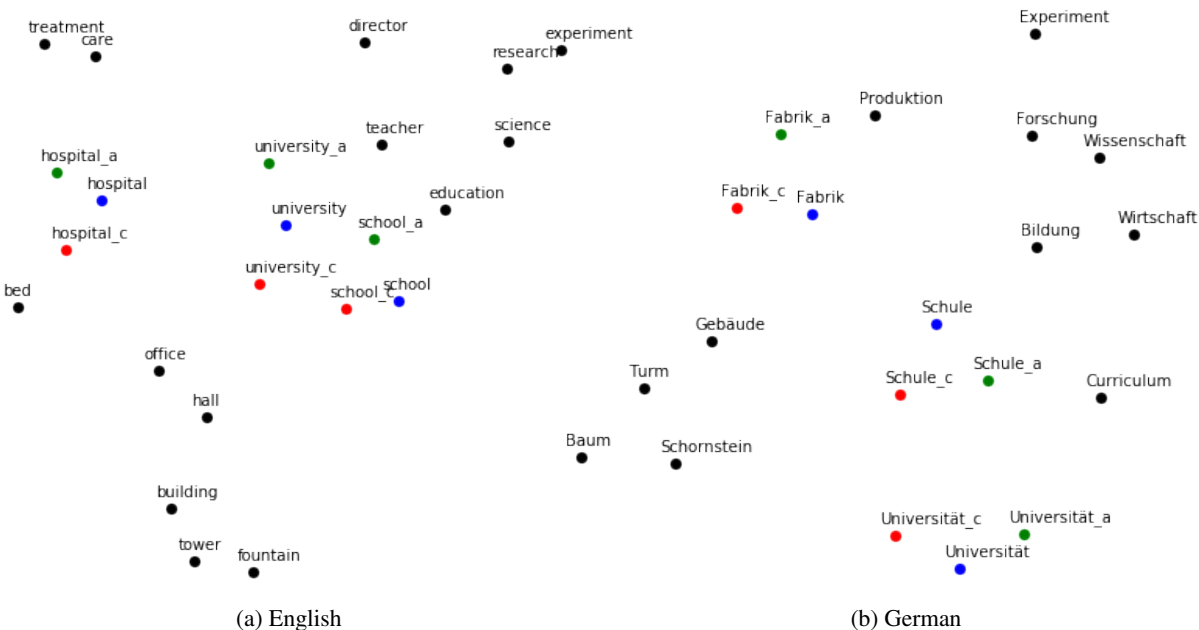
(a) English   (b) German

Figure 1: Three ambiguous English (left) and German (right) words with concrete and abstract variants and some related words in a two dimensional projection of the embedding space.

Since we added clearly abstract and concrete words concreteness becomes a clear dimension in the visualization. For all words we see that the concrete variants indeed moved into the direction of the related concrete words and similar for the abstract variants.

In order to know whether the concrete and abstract variants of the embeddings really become more similar to synonyms of the respective senses we selected 10 ambiguous words for English and German along with a closely related word for the abstract and for the concrete sense. For German we selected 10 words that are ambiguous between a building (or location) and an institution. For English we selected words that either denote a process or an actor involved in that process. Here we tried to select words that do not have too many senses, are predominantly used as noun and for the related words we tried to find synonyms that do not have the same ambiguity. Most of the related words were taken from the synsets in WordNet (Miller, 1995) to make the choices somewhat more objective. Now for each word we add (subtract) $0.2\ \hat{v}_{ca}$ and determine how much the resulting vector is closer to the embedding of the related concrete (abstract) word than the original vector. The results for are give in Tables 6 and 7.

In all cases we see that the improvement is very small or even negative. For some of the word pairs, like *Parlament – Parlamentsgebäude* (Parliament -

Parliament's house) or *Schule – Schulgebäude*, it seems that the second word is a real synonym of the building sense of the first word and we would expect that the cosine similarity would be much larger when adding the concreteness to the general vector. Thus we have to conclude that though the senses of these ambiguous German words clearly have different degrees of concreteness, the difference between the senses is much more than just the concreteness.

For the English words that are ambiguous between a process and an entity, adding concreteness in two cases even makes the pairs more dissimilar and adding abstractness only in one case makes the word more similar to a synonym of the process reading.

7 Conclusion

We have seen that concreteness can be identified as a direction in the word embedding space. Various methods, based on many words with concreteness values or just on a view highly concrete words give almost the same vector for concreteness. Moreover, the cosine of these vectors with the embeddings of words correlates strongly with the concreteness judgments of human subjects of these words. Thus our first research question can be answered positively.

Furthermore, we see that concreteness and abstractness are quite symmetric properties. We can

Table 6: Ten ambiguous English words with each time one word related to the concrete sense (person or artifact) and one word related to the abstract sense (process). The column after the related word gives the cosine between the embeddings of the word and the related word; the column labeled δ gives the improvement if cosine similarity when adding (resp. subtracting) $0.2\,\hat{v}_{ca}$ to the embedding of the original word.

word	concr. related	cos	δ	abstr. related	cos	δ
passage	passageway	0.52	0.05	transition	0.30	0.00
entry	entranceway	0.32	0.03	debut	0.16	-0.01
creation	world	0.25	-0.00	founding	0.31	-0.02
shot	scene	0.37	0.00	stroke	0.23	-0.02
opposition	opponent	0.51	0.01	resistance	0.38	-0.03
help	assistant	0.19	0.03	assistance	0.58	-0.00
opening	gap	0.36	-0.02	initiative	0.24	0.02
replacement	successor	0.38	-0.07	replacing	0.60	-0.04
storage	storehouse	0.39	0.01	warehousing	0.50	-0.01
shipment	freight	0.50	0.02	dispatch	0.48	-0.02

Table 7: Ten ambiguous German words with each time one word related to the concrete sense (building or location) and one word related to the abstract sense (institution). The column after the related word gives the cosine between the embeddings of the word and the related word; the column labeled δ gives the improvement if cosine similarity when adding (resp. subtracting) $0.2\,\hat{v}_{ca}$ to the embedding of the original word.

word	concr. related	cos	δ	abstr. related	cos	δ
Parlament	Parlamentsgebäude	0.70	0.01	Politik	0.47	0.01
Laden	Schuppen	0.28	0.06	Einzelhandel	0.45	-0.01
Gericht	Gerichtsgebäude	0.57	0.03	Urteil	0.55	0.01
Schule	Schulgebäude	0.64	0.02	Lernen	0.45	0.01
Büro	Bürohaus	0.55	0.01	Arbeit	0.43	0.02
Polizei	Polizeiwache	0.65	0.03	Ordnung	0.27	0.01
Kirche	Kirchturm	0.60	0.05	Religion	0.51	0.01
Universität	Hörsaal	0.42	0.04	Forschung	0.39	0.01
Theater	Schauspielhaus	0.72	-0.00	Kultur	0.46	0.01
Fabrik	Schornstein	0.31	0.07	Produktion	0.57	0.01

compute vectors for concreteness and abstractness independently and found both for English and German that the angle between these vectors is almost 180 degrees. Moreover, we do not see any indication that all concrete form one cluster while abstract words are distributed more uniformly through the embedding space or the other way around. Thus, we also can give a positive answer to the second research question.

Finally, we hoped that we would find pairs of words that just differ w.r.t. the concreteness dimension, like the words king and queen only differ w.r.t. the gender dimension. At least we would like to find words with different senses, where the degree of concreteness is the main difference between the senses. Though there are many polysemous words, that seem to be good candidates and though we can make suggestive visualizations for selected examples, our last experiment is not very encouraging in this respect. In the first place it has to be noted that the evaluation is quite problematic since we do not know what the embedding of the specific senses of a word should be. Nevertheless, at least in the case of the building/institution ambiguity the senses the senses are clearly distinguished by concreteness, but there are many more differences between the senses than just this aspect. The last result does not mean that it is not possible to learn the relation between vectors for different senses of a word in the case of regular polysemy, but the relation is more complex than just linearly adding concreteness to the embedding.

References

Jeffrey R Binder, Chris F Westbury, Kristen A McKiernan, Edward T Possing, and David A Medler. 2005. Distinct brain systems for processing concrete and abstract concepts. *Journal of cognitive neuroscience*, 17(6):905–917.

Anna M Borghi, Ferdinand Binkofski, Cristiano Castelfranchi, Felice Cimatti, Claudia Scorolli, and Luca Tummolini. 2017. The challenge of abstract concepts. *Psychological Bulletin*, 143(3):263.

Marc Brysbaert, Amy Beth Warriner, and Victor Kuperman. 2014. Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, 46(3):904–911.

Jean Charbonnier and Christian Wartena. 2019. Predicting word concreteness and imagery. In *Proceedings of the 13th International Conference on Computational Semantics-Long Papers*, pages 176–187.

Jean Charbonnier and Christian Wartena. 2020. Predicting the Concreteness of German Words. In *Proceedings of Konvens / SwissText*.

Diego Frassinelli, Daniela Naumann, Jason Utt, and Sabine Schulte m Walde. 2017. Contextual characteristics of concrete and abstract words. In *IWCS 2017 — 12th International Conference on Computational Semantics — Short papers*.

Michael Friendly, Patricia E. Franklin, David Hoffman, and David C. Rubin. 1982. The Toronto Word Pool: Norms for imagery, concreteness, orthographic variables, and grammatical usage for 1,080 words. *Behavior Research Methods & Instrumentation*, 14(4):375–399.

K. J. Gilhooly and R. H. Logie. 1980. Meaning-dependent ratings of imagery, age of acquisition, familiarity, and concreteness for 387 ambiguous words. *Behavior Research Methods & Instrumentation*, 12(4):428–450.

Willy Hager. 1994. Bildhaftigkeit, Konkretheit-Abstraktheit und Bedeutungshaltigkeit von 63 mehrdeutigen Substantiven. In Willi Hager and Marcus Hasselhorn, editors, *Handbuch deutschsprachiger Wortnormen*, chapter 3.6, pages 212–217. Hogrefe Verlag für Psychologie, Göttingen.

Charles H Hargis and Edward E Gickling. 1978. The function of imagery in word recognition development. *The Reading Teacher*, 31(8):870–874.

Felix Hill and Anna Korhonen. 2014. Concreteness and subjectivity as dimensions of lexical meaning. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 725–731.

Felix Hill, Anna Korhonen, and Christian Bentz. 2014. A quantitative empirical analysis of the abstract/concrete distinction. *Cognitive science*, 38(1):162–177.

Geoff Hollis and Chris Westbury. 2016. The principals of meaning: Extracting semantic dimensions from co-occurrence models of semantics. *Psychonomic bulletin & review*, 23(6):1744–1756.

Stavroula-Thaleia Kousta, Gabriella Vigliocco, David P Vinson, Mark Andrews, and Elena Del Campo. 2011. The representation of abstract words: why emotion matters. *Journal of Experimental Psychology: General*, 140(1):14.

George A Miller. 1995. Wordnet: a lexical database for english. *Communications of the ACM*, 38(11):39–41.

Daniela Naumann, Diego Frassinelli, and Sabine Schulte im Walde. 2018. Quantitative Semantic Variation in the Contexts of Concrete and Abstract Words. In *Proceedings of the 7th Joint Conference on Lexical and Computational Semantics*, pages 76–85, New Orleans, LA, USA.

Ian Neath and Aimée M Surprenant. 2020. Concreteness and disagreement: Comment on Pollock (2018). *Memory & cognition*, 48(4):683–690.

Allan Paivio. 1970. On the functional significance of imagery. *Psychological Bulletin*, 73(6):385.

Shekeila D Palmer, Lucy J MacGregor, and Jelena Havelka. 2013. Concreteness effects in single-meaning, multi-meaning and newly acquired words. *Brain research*, 1538:135–150.

W. Gudrun Reijnierse, Christian Burgers, Marianna Bolognesi, and Tina Krennmayr. 2019. How polysemy affects concreteness ratings: The case of metaphor. *Cognitive Science*, 43(8):e12779.

Sascha Rothe, Sebastian Ebert, and Hinrich Schütze. 2016. Ultradense word embeddings by orthogonal transformation. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 767–777. Association for Computational Linguistics.

Mark Sadoski, Victor L Willson, Angelia Holcomb, and Regina Boulware-Gooden. 2004. Verbal and nonverbal predictors of spelling performance. *Journal of Literacy Research*, 36(4):461–478.

Paula J Schwanenflugel. 2013. Why are abstract concepts hard to understand? In *The psychology of word meanings*, pages 235–262. Psychology Press.

Paula J Schwanenflugel and Edward J Shoben. 1983. Differential context effects in the comprehension of abstract and concrete verbal materials. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9(1):82.

Graham G. Scott, Anne Keitel, Marc Becirspahic, Bo Yao, and Sara C. Sereno. 2019. The glasgow norms: Ratings of 5,500 words on nine scales. *Behavior Research Methods*, 51(3):1258–1270.

Otfried Spreen and Rudolph W. Schulz. 1966. Parameters of abstraction, meaningfulness, and pronunciability for 329 nouns. *Journal of Verbal Learning & Verbal Behavior*, 5(5):459–468.

Shinya Tanaka, Adam Jatowt, Makoto P. Kato, and Katsumi Tanaka. 2013. Estimating content concreteness for finding comprehensible documents. In *Proceedings of the Sixth ACM International Conference on Web Search and Data Mining*, WSDM '13, pages 475–484, New York, NY, USA. ACM.

Daphne Theijssen, Hans van Halteren, Lou Boves, and Nelleke Oostdijk. 2011. On the difficulty of making concreteness concrete. *Computational Linguistics in the Netherlands Journal*, 1:61–77.

Dušica Filipović Đurđević, Aleksandar Kostić, and Zorana Đinđića. 2017. Number, relative frequency, entropy, redundancy, familiarity, and concreteness of word senses: Ratings for 150 serbian polysemous nouns. In *Selected Papers From the 4th and 5th Workshop on Psycholinguistic, Neurolinguistic and Clinical Linguistic Research*, volume 2 of *Studies in Language and Mind*, pages 13–50. Filozofski fakultet u Novom Sadu.

Laurens Van der Maaten and Geoffrey Hinton. 2008. Visualizing data using t-SNE. *Journal of machine learning research*, 9(11).

Lin Yui, Roslin Ng, and Hiran Perera-WA. 2017. Concrete vs abstract words–what do you recall better? a study on dual coding theory. Technical report, PeerJ Preprints.

Appendix: Translation of German words used in the figures and tables.

Word	Translation
Arbeit	work, labor
Baum	tree
Bildung	education
Büro	office
Bürohaus	office building
Curriculum	curriculum
Einzelhandel	retail
Experiment	experiment
Fabrik	factory
Forschung	research
Gebäude	building
Gericht	court
Gerichtsgebäude	court building
Hörsaal	lecture hall
Kirche	church
Kirchturm	church tower
Kultur	culture
Laden	shop
Lernen	to learn
Ordnung	order
Parlament	parliament
Parlamentsgebäude	parliament's house
Politik	politics
Polizei	police
Polizeiwache	Police station
Produktion	production
Religion	religion
Schauspielhaus	playhouse, theater
Schornstein	chimney
Schule	school
Schulgebäude	school building
Schuppen	shed
Theater	theater
Turm	tower
Universität	university
Urteil	verdict, judgment
Wirtschaft	economy
Wissenschaft	science