Cognition 169 (2017) 46–53

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Brief article

Effects of semantic neighborhood density in abstract and concrete words

Megan Reilly, Rutvik H. Desai *

University of South Carolina, 220 Discovery I, 915 Greene St., Columbia, SC 29208, United States

ARTICLE INFO

Article history:
Received 28 November 2016
Revised 3 August 2017
Accepted 8 August 2017
Available online 14 August 2017

Keywords:
Concreteness
Abstract semantics
Semantic neighborhood density

ABSTRACT

Concrete and abstract words are thought to differ along several psycholinguistic variables, such as frequency and emotional content. Here, we consider another variable, semantic neighborhood density, which has received much less attention, likely because semantic neighborhoods of abstract words are difficult to measure. Using a corpus-based method that creates representations of words that emphasize featural information, the current investigation explores the relationship between neighborhood density and concreteness in a large set of English nouns. Two important observations emerge. First, semantic neighborhood density is higher for concrete than for abstract words, even when other variables are accounted for, especially for smaller neighborhood sizes. Second, the effects of semantic neighborhood density on behavior are different for concrete and abstract words. Lexical decision reaction times are fastest for words with sparse neighborhoods; however, this effect is stronger for concrete words than for abstract words. These results suggest that semantic neighborhood density plays a role in the cognitive and psycholinguistic differences between concrete and abstract words, and should be taken into account in studies involving lexical semantics. Furthermore, the pattern of results with the current feature-based neighborhood measure is very different from that with associatively defined neighborhoods, suggesting that these two methods should be treated as separate measures rather than two interchangeable measures of semantic neighborhoods.

1. Introduction

Traditionally, the study of conceptual processing has focused mainly on concrete nouns, whose perceptual properties are relatively easy to articulate and whose similarities tend to fall into well-defined categories. However, abstract concepts often fail to fit the same models as concrete entities, and they tend to differ from concrete words along several psycholinguistic variables. For example, abstract words tend to have higher emotional arousal (Newcombe, Campbell, Siakaluk, & Pexman, 2012; Vigliocco, Meteyard, Andrews, & Kousta, 2014; Zdrazilova & Pexman, 2013). In concrete words, there is evidence that semantic neighborhood density plays an important role in behavior (Buchanan, Westbury, & Burgess, 2001; Mirman & Magnuson, 2008); however, measures of neighborhood density rely on calculating semantic similarity between concepts, which can be difficult to compare between abstract and concrete concepts because of their fundamentally different semantic organization (Dalla Volta, Fabbri-Destro, Gentilucci, & Avanzini, 2014).

Most concrete concepts can be readily organized into categories, based on the perceptual and functional characteristics that overlap within a category (Cree & McRae, 2003). In such a system, semantic similarity and neighborhood density can be measured by comparing manually collected lists of feature norms, e.g., "cat" and "dog" share the feature "has a tail" (McRae, Cree, Seidenberg, & McNorgan, 2005; Pexman, Hargreaves, Siakaluk, Bodner, & Pope, 2008). Abstract concepts, on the other hand, are more difficult to describe based on their semantic neighbors and features, leading to the assumption that abstract concepts are semantically "impoverished" (Paivio, 2010; Plaut & Shallice, 1993). However, abstract words arguably have rich meaning, even if their individual properties are difficult to describe using predicates such as 'is a', 'has', 'contains', 'is made of'; perhaps the variables of importance to abstract and concrete words are different. Abstract words are known to differ from concrete words in terms of the modality of semantic content associated with them. Concrete nouns tend to possess visual and/or motor characteristics, while abstract words tend to have more emotional content (Crutch, Troche, Reilly, & Ridgway, 2013). Thus, the features of abstract words are not easily described using manual norm collection.

An alternative approach that uses corpus-based models, such as the Hyperspace Analogue to Language model (HAL; Burgess &

* Corresponding author.
E-mail address: rutvik@sc.edu (R.H. Desai).

http://dx.doi.org/10.1016/j.cognition.2017.08.004

Lund, 1997) or Latent Semantic Analysis (LSA; Landauer & Dumais, 1997), can be used instead of manually collected norms. These models measure similarity between concepts based on the lexical contexts in which they tend to co-occur. This type of method is automated and therefore can easily be applied to calculate vectors for thousands of words. A limitation of corpus-based methods has been that they tend to emphasize associative or thematic relationships at the expense of taxonomic or featural relationships. For example, a term-to-term comparison of associates "cow" and "milk" using LSA (lsa.colorado.edu) has a similarity index of 0.60 along a scale of −1 to 1, while "cow" and "bull" (taxonomically related) have a similarity of only 0.21.

A few attempts have been made to use vector spaces to compare concrete and abstract words. For example, the WINDSORS model of semantic space (Durda & Buchanan, 2008) is a variant of HAL that aims to prevent artificially dense neighborhoods for high-frequency words. Danguecan and Buchanan (2016) used the WINDSORS model to estimate semantic neighborhood densities for both concrete and abstract words and compared high- and low-density words for several behavioral tasks. The authors found that semantic neighborhood density effects were task-specific, such that high neighborhood density influenced reaction times (RTs) during a lexical decision (LD) task and go/no-go LD task, but not more semantic tasks such as sentence relatedness. In these LD tasks, they showed that RTs for high-density words tended to be slower than for low-density words, but that this effect was specific to abstract concepts.

These results suggest that semantic neighborhood density is an important measure to the study of concrete and abstract concepts. However, the stimuli used by Danguecan and Buchanan (2016) were not balanced for any quantitative measure of concreteness across high and low neighborhood density levels. An analysis of their stimuli using a comprehensive set of concreteness ratings (Brysbaert, Warriner, & Kuperman, 2014) reveals a significant interaction between the Concrete/Abstract and High/Low Density conditions ($F[1] = 6.54$, $p = 0.012$) such that although high- and low-density concrete words were balanced for concreteness, abstract words were not. High-density abstract words were in fact significantly more concrete than low-density abstract words (t [2.69], $p = 0.010$). Thus, it is difficult to draw conclusions about whether unique effects within abstract words are due to differences in semantic neighborhood density or the inclusion of very highly abstract words for the high-density group. Additionally, this method does not eliminate a basic problem facing most vector spaces, namely the conflation of associative and featural similarity.

The development of a semantic space that includes both concrete and abstract words has great potential for investigating the cognitive and neural processes that underlie language processing. For example, comprehending abstract words selectively activates the anterior temporal lobe (ATL; Sabsevitz, Medler, Seidenberg, & Binder, 2005; Wang, Conder, Blitzer, & Shinkareva, 2010), an area which is also activated during detection of "unique entities", including famous and familiar people and landmarks (Gorno-Tempini & Price, 2001; Grabowski et al., 2001; Olson, McCoy, Klobusicky, & Ross, 2013; Ross & Olson, 2010; Tranel, 2009). Unique entities by definition have few semantic neighbors. If abstract words have more sparse neighborhoods than concrete words, then the ATL's role in accessing abstract concepts may reflect a more general role in accessing sparsely encoded semantic information.

Additionally, a feature-based measure of neighborhood density in concrete and abstract words may help explain behavioral differences between them. Recchia and Jones (2012) used a traditional measure of neighborhood density (HAL) that emphasizes associative information. The authors found that for abstract words, lexical decisions were faster for words that had more neighbors. For concrete words this measure had no significant effect. Since concrete words are thought to be organized according to semantic features, a feature-based neighborhood density measure may reveal an effect that is unique to concrete words, in contrast to Recchia and Jones' unique effect for abstract words. Additionally, although dense associative semantic neighborhoods may elicit faster responses, dense taxonomically related neighborhoods are more likely to create semantic competition and slow responses.

Here, our goals were (1) to compare semantic neighborhood densities in a large set of abstract and concrete words, to examine whether they differ systematically using a vector space that emphasizes featural, and not associative, similarity; and (2) to examine if semantic neighborhood characteristics can account for some of the behavioral differences between abstract and concrete words.

To this end, the current study measures neighborhood density using a semantic space that accommodates both concrete and abstract words. Roller and Erk (2016) propose a semantic space based on contextual similarity that restricts contextual search in a corpus to a small window of neighboring words. This semantic space was calculated by counting pairwise co-occurrences of words that appear near each other in a corpus, then using dimensionality reduction to calculate a 300-value vector describing each word. The specific procedure is based on existing models that aim to automate detection of hypernymy (superordinate-subordinate relationships such as animal-cat). As a result, the vector space captures taxonomic relationships successfully but not other types of similarity (Erk, 2016). Thus, it incorporates the conceptual advantages of feature-based models with the more flexible and scalable method of collecting semantic information from a corpus. Recall that LSA estimates "cow" and "milk" as closer semantic neighbors than "cow" and "bull". In the current measure of similarity, the Euclidean distance between the vectors for "cow" and "milk" is 318.8, and the distance between "cow" and "bull" as 161.3 (meaning that the taxonomic neighbor "bull" is a closer neighbor to "cow" than its associative neighbor "milk"). The current investigation is a direct comparison of neighborhood density and behavior in 3500 English lexical items available from the Roller and Erk (2016) semantic space.

2. Methods

Stimuli were taken from a set of calculated semantic vectors using the methods of Roller and Erk (2016). This corpus included 232,585 English lexical items, including nouns, verbs, adjectives, proper names, function words, hyphenated phrases, and other lexical items. To narrow the scope to nouns and ensure that we had access to psycholinguistic variables for all items, we limited the dataset of interest to words that fit the following a priori criteria:

- Words are labeled as nouns according to the Roller & Erk norms, and are listed as "Noun" in the Princeton Wordnet Norms (Fellbaum, 1998).
- Words are included in the Warriner, Kuperman, and Brysbaert (2013) concreteness norms.
- Words are included in the Brysbaert et al. (2014) emotional valence and arousal norms.

To calculate a given word's semantic neighborhood, its vector was compared with the vector for every word in the full (232,585-word) Roller & Erk dataset, and the Euclidean distance between the two words was calculated. Thus, each word had a semantic 'neighborhood' consisting of 232,585 semantic distances. For each word, the average distance of its nearest 10 neighbors was calculated; we refer to this measure as semantic neighborhood distance (SND). Note that a high SND value, here, is equivalent to a low neighborhood density (sparse neighborhood). Neighborhood density was also calculated for 3-, 25-, and 50-word neighborhoods. Generally, the smaller the neighborhood, the greater the difference between concrete and abstract words (see below). Here, we mainly focus on the 10-neighborhood measure.

Because frequency and SND are highly correlated in this dataset (see Results), an additional preprocessing step was taken to decorrelate frequency and SND. For each word, the product of its log frequency $*$ log semantic distance for 10 neighbors was calculated; words with very high or very low frequency-SND products were gradually removed until the remaining words were not correlated on these measures ($p > 0.10$). Concreteness ratings were split into thirds, such that the highest third of all words were classified as "Concrete", and the lowest third were classified as "Abstract". This left 3489 words that were at one concreteness extreme or the other (5198 including intermediate concreteness words). Removing the middle concreteness values is important in light of recent criticism of concreteness norms. Pollock (2017) showed that concreteness values near the mean of the ratings are often much higher in variability across ratings than words at either extreme of the scale. This suggests that these middle values are not, in fact, of medium concreteness, but rather that raters have different salient meanings for those words and therefore there is disagreement as to their concreteness. A word with a high standard deviation across ratings is not a reliable measure of concreteness, and Pollock suggested that these words tend to elicit behavior more similar to concrete words than abstract words, instead of being truly in the middle between the two. Therefore, in categorical analyses, only the highest and lowest thirds of the data were included. For analyses with continuous variables (i.e., regression), both the full and the trimmed datasets were included.

In addition, the following variables were collected for every word:

- Log HAL Frequency (Brysbaert & New, 2009).
- Concreteness (Brysbaert et al., 2014).
- Percent Known (a measure of Familiarity) (Brysbaert et al., 2014).
- Valence (Warriner et al., 2013).
- Emotional Arousal (Warriner et al., 2013).
- Lexical decision and naming reaction times (RTs; Balota et al., 2007).

Reaction times were collected by the English Lexicon Project (ELP), which included minor preprocessing. RTs were listed for correct responses only, averaged across several hundred participants after outliers were removed (any responses less than 200 ms, greater than 3000 ms, or above or below 3 standard deviations from that participant's mean RT).

To examine the difference in SND between abstract and concrete words, a Pearson correlation compared SND and concreteness. A *t*-test also compared SND in categorically sorted concrete and abstract words. Linear regressions were performed to examine whether there is a relationship between concreteness and SND even when other semantic variables (frequency, emotional arousal) are taken into account.

To address the second question regarding the effect of SND on LD and naming RT, median splits were performed on emotional arousal and SND, and an omnibus analysis of variance (ANOVA) evaluated the effects of each variable on reaction times. Additional multiple linear regressions were performed which included several lexical (frequency, length, familiarity) and semantic (emotional arousal, valence) variables as regressors, in order to examine how a word's SND and concreteness interact when other variables are accounted for. Similar regressions were also performed separately for concrete and abstract words. Here, the term SND refers to the average distance between a word and its nearest 10 neighbors, unless otherwise specified.

Finally, a second set of LD RTs were collected from the British Lexicon Project (Keuleers, Lacey, Rastle, & Brysbaert, 2012) to attempt a replication of the behavioral results. This included a set of 5910 words which were present in the BLP and the lexical and semantic norms listed above. The same analyses were applied to the ELP and BLP data except for the frequency-correlation correction ($r = 0.69$, $t[5908] = 73.5$, $p < 0.0001$). It should also be noted that a very low proportion of BLP items had low concreteness relative to the ELP data: 10.3% of words in the BLP had a concreteness rating less than 2.5 on a 1–5 scale (as opposed to 17.8% in the ELP dataset). Thus, the distribution of concreteness values is concentrated on a narrower range with higher concreteness than the American dataset, which exacerbates the issue described above in Pollock (2017) that words with middle-of-the-range concreteness ratings do not really represent a midway point between concrete and abstract. When abstract and concrete words were determined using the top and bottom thirds of the concreteness ratings, there were 1970 words of each abstract and concrete in the BLP dataset.

A complete dataset including the densities for neighborhood sizes of 3, 10, 25, and 50 for all 200K+ lexical items in Roller and Erk (2016) can be found here: http://www.mccauslandcenter.sc.edu/delab/sites/sc.edu.delab/files/attachments/SND_reilly_desai.txt. This dataset also includes several thousand entries consisting entirely of punctuation and numerals which were removed prior to the current analysis.

3. Results

In the full ELP dataset, frequency and SND were highly correlated ($r = 0.655$, $t[9317] = 83.7$, $p < 0.0001$). Therefore, the ELP analysis reported here will use a decorrelated dataset with fewer words, in which the frequency/SND correlation is absent ($r = 0.022$, $t[5196] = 1.59$, $p > 0.10$). The BLP dataset is used as is, due to its smaller size and bias towards more concrete words. Both sets contained many words with "intermediate" concreteness values. To examine SND effects in words that can be more clearly classified as concrete or abstract, we examine the top and bottom third of the concreteness ratings (mean [SD] for ELP: abstract, 2.42 [0.46]; concrete, 4.69 [0.20]; BLP: abstract, 2.70 [0.53]; concrete, 4.75 [0.16]).

Table 1 shows the full correlation matrix for the semantic variables. Concreteness and SND were significantly correlated ($r = -0.049$, $t[5196] = -3.57$, $p = 0.0004$) such that increasing concreteness corresponds to a denser neighborhood. The smaller the neighborhood, the greater the difference between concrete and abstract words. For 3- and 10-word neighborhoods, abstract words have a significantly higher SND (i.e., more sparse neighborhood) than concrete words (3 neighbors: $t[3448] = 3.59$, $p = 0.0003$; 10 neighbors: $t[3452] = 2.73$, $p = 0.006$; see Fig. 1). Using a 25-word or 50-word neighborhood, abstract and concrete words did not differ, although the average neighborhood for abstract words was qualitatively more sparse. Consistent with previous findings, abstract words also had higher average emotional arousal ratings ($t[3486] = 14.3$, $p < 0.0001$) than concrete words.

A linear regression on SND that included all variables as regressors also showed a significant relationship between concreteness and SND even when the other variables are accounted for (concreteness extremes only: $t[3434] = -2.81$, $p = 0.005$; full dataset: $t[5191] = -2.50$, $p = 0.013$). When an interaction term is included between emotional arousal and concreteness, this interaction is significant (concreteness extremes only: $t[3433] = -3.32$,

Table 1
Correlation coefficients between each pair of semantic variables of interest.

	SND	Valence	Arousal	Concreteness
Frequency	0.022	0.067***	0.013	−0.058***
Concreteness	−0.049**	0.142***	−0.192***	
Arousal	−0.043*	−0.217***		
Valence	0.066***			

* p < 0.05.
** p < 0.001.
*** p < 0.0001.

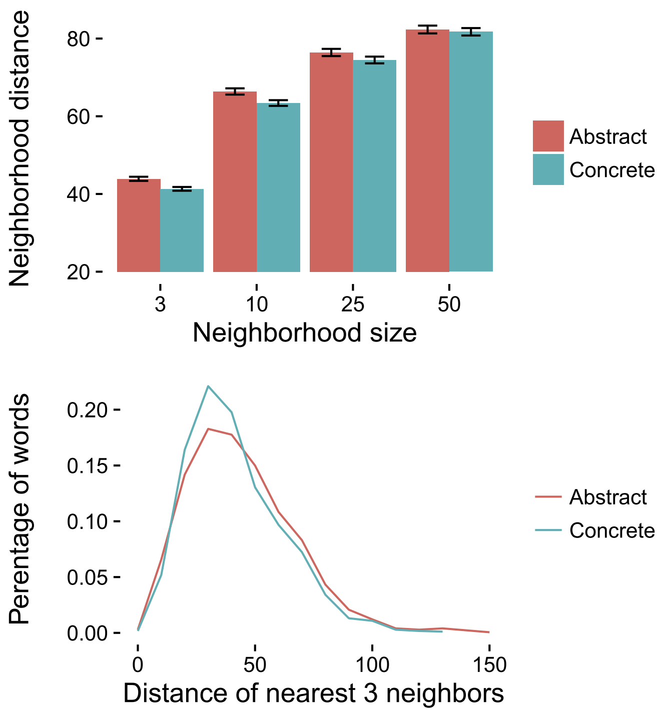

Fig. 1. Above: average SND for semantic neighborhoods of 3-, 10-, 25-, and 50-word neighborhoods. Error bars represent standard error. Below: distribution of neighborhood densities in concrete and abstract words for a 3-word neighborhood.

$p = 0.0009$; full dataset: $t[5190] = -3.85$, $p = 0.0001$), suggesting that the relationship between emotional arousal and SND differs between concrete and abstract words. Fig. 2 visualizes the relationships between SND, concreteness and emotional arousal.

Turning to the behavioral effects, words were split categorically along arousal and SND measures using median splits (Fig. 3, right). An omnibus ANOVA on LD RTs showed that in both ELP and BLP datasets, there were significant main effects of concreteness and SND, such that RTs for concrete and sparse words were faster than for abstract and dense words, respectively (ELP: concreteness, $F[1,3433] = 223$, $p < 0.0001$; SND, $F[1,3433] = 27.6$, $p < 0.0001$; BLP: concreteness, $F[1,3990] = 9.57$, $p = 0.002$; SND, $F[1,3990] = 1164$, $p < 0.0001$). The main effect of emotional arousal was only significant in the ELP dataset ($F[1,3433] = 5.77$, $p = 0.016$). The only interaction in either dataset was the interaction between concreteness and SND, which was significant in the ELP and marginal in the BLP. In both datasets, RTs for concrete words with sparse neighborhoods were particularly fast (ELP: $F[1,3433] = 31.7$, $p < 0.0001$; BLP: $F[1,3990] = 2.81$, $p = 0.094$). Within each SND/arousal category, concrete words had faster RTs than abstract words in the ELP (Low/Sparse: $t[697] = 11.90$, $p < 0.0001$; High/Sparse: $t[614] = 9.06$, $p < 0.0001$; Low/Dense: $t[656] = 4.16$, $p < 0.0001$; High/Dense: $t[828] = 4.13$, $p < 0.0001$). The BLP results showed faster responses for concrete words only within the sparse categories (Low/Sparse: $t[863] = 3.30$, $p = 0.001$; High/Sparse: $t[852] = 3.39$, $p = 0.0007$).

Linear regressions were also performed treating emotional arousal and SND as continuous variables, in addition to length, frequency, emotional valence, and percent known (familiarity). An interaction term between concreteness and SND was also included in the model. Reaction time values were log transformed because they were not normally distributed. The slope and t-value of each variable is listed in Table 2. Consistent with the categorical analysis for LD, the interaction term between concreteness and SND was significant, such that lexical decision of concrete words is selectively faster for words with sparse neighborhoods relative to abstract words in the ELP data (concreteness extremes only: $t[3432] = -2.96$, $p = 0.0031$; full dataset: $t[5111] = -2.40$, $p = 0.017$) and BLP data, but only when the middle concreteness values were not included (concreteness extremes only: $t[3989] = -2.34$, $p = 0.0193$; full dataset: $t[5901] = -0.55$, $p > 0.1$). The lack of effect in the full BLP dataset likely reflects the fact that "middle"

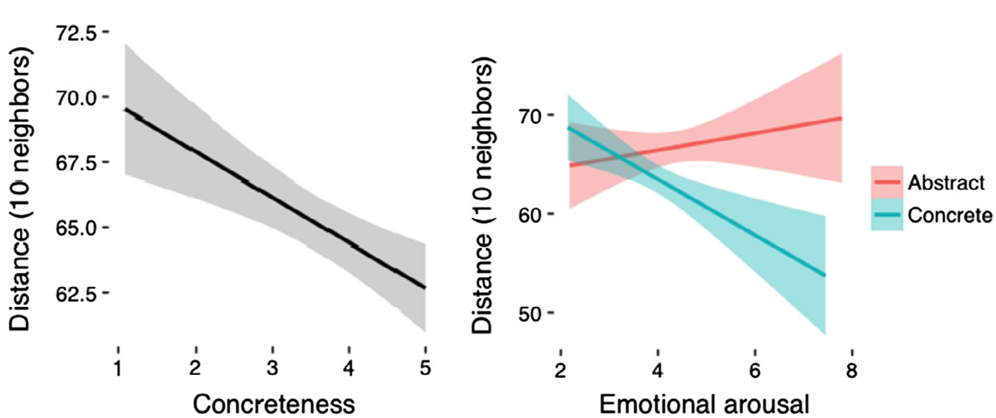

Fig. 2. Relationships between SND and concreteness (left) and emotional arousal (right). Shaded areas represent standard error.

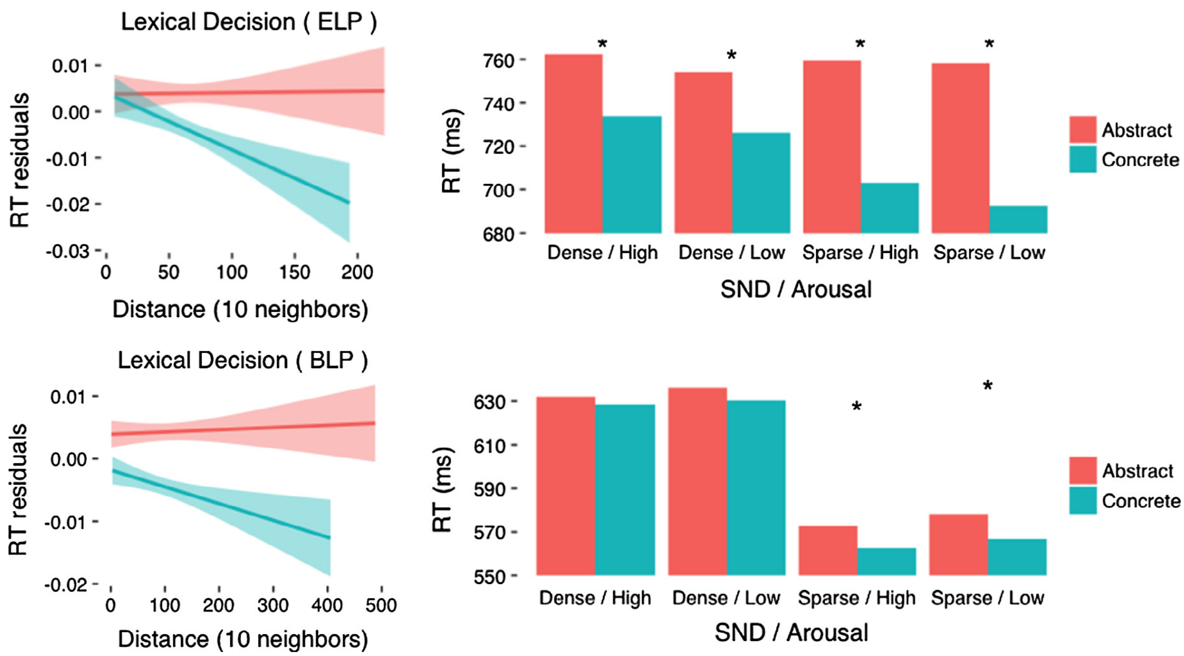

Fig. 3. Effects of semantic variables of interest on ELP (top) and BLP (bottom), lexical decision reaction times. The left subfigures use continuous measures of semantic distance regressing out the effects of all other lexical and semantic variables. Right subfigures use median splits across SND and Emotional Arousal. '*' represents a p-value less than 0.05 in a *t*-test between abstract and concrete words.

Table 2
Results of linear regressions on lexical decision (LD) and naming reaction times. Because log RTs were used and estimates were very small, all estimates and standard errors are multiplied by 1000 for increased readability. For the ELP results, the intercept estimate is 3344 and the multiple R^2 for the entire model is 0.387. For the BLP, the intercept estimate is 3242 and the multiple R^2 for the entire model is 0.543.

	Beta	Std. error	*t*	*p*
ELP LD				
Concreteness	−1.44	3.35	−0.43	>0.1
Emotional arousal	−0.589	0.863	−0.683	>0.1
Frequency***	−12.6	0.72	−17.5	<0.0001
Length***	9.37	0.344	27.3	<0.0001
Percent known***	−440	27.9	−15.8	<0.0001
SND	0.0077	0.031	0.248	>0.1
Valence***	−3.84	0.612	−6.28	<0.0001
SND * concreteness**	−0.135	0.0456	−2.96	0.00314
BLP LD				
Concreteness***	−6.97	1.65	−4.21	<0.0001
Emotional arousal***	−4.61	0.575	−8.03	<0.0001
Frequency***	−31.9	0.889	−36.3	<0.0001
Length***	3.22	0.319	10.1	<0.0001
Percent known***	−368	20.3	−18.2	<0.0001
SND	0.0051	0.0098	0.52	>0.1
Valence***	−2.03	0.433	−4.69	<0.0001
SND * concreteness*	−0.03	0.0128	−2.34	0.0193

^ $p < 0.10$.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.0001$.

concreteness values are actually high in concreteness relative to the abstract words, and likely do not reflect true intermediate values (Pollock, 2017).

Fig. 3 visualizes the influence of SND and concreteness on LD RTs. In order to visualize the continuous effects of SND while accounting for other variables, a linear regression was performed on the RTs which included every variable except SND. The residuals from this regression (that is, the RTs adjusted for non-SND variables) are the dependent variable in the lefthand figures of Fig. 3.

Separate regressions were also performed within concrete and abstract words. Within concrete words, SND had a significant influence on LD RTs ($t[1733] = -3.13$, $p = 0.0018$) and naming RTs ($t[1733] = -2.76$, $p = 0.0059$). Abstract words did not show any effect of SND on RT. The BLP data did not show any significant effects within a particular category.

Naming RTs (available only in the ELP; Fig. 4) showed a similar pattern of results to the LD data. Like in the LD analysis, the categorical ANOVA revealed significant main effects of Concreteness (F[1,3433] = 200, $p < 0.0001$), SND (F[1,3433] = 6.31, $p = 0.012$), and Arousal (F[1,3433] = 6.61, $p = 0.010$) as well as an interaction between SND and Concreteness (F[1,3433] = 32.1, $p < 0.0001$). Concrete word naming was faster than abstract words within every

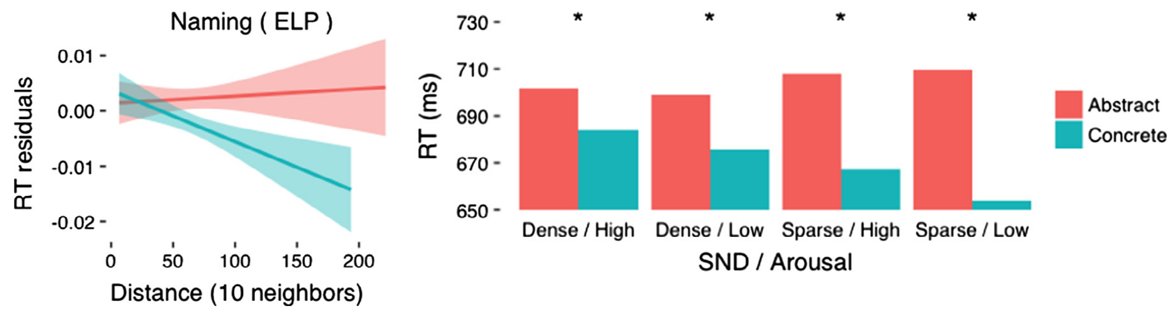

Fig. 4. Effects of SND on naming RTs in ELP. The left figure depicts the residual RTs after regressing out non-SND lexical and semantic variables. The right figure uses a median split across Arousal and SND. '*' represents a p-value less than 0.05 in a *t*-test between abstract and concrete words.

SND/arousal category (Low/Sparse: $t[662] = 11.1$, $p < 0.0001$; High/Sparse: $t[638] = 8.01$, $p < 0.0001$; Low/Dense: $t[644] = 4.16$, $p < 0.0001$; High/Dense: $t[825] = 3.37$, $p = 0.0008$). Also like the LD analysis, a linear regression including all variables of interest yielded a significant interaction between SND and concreteness (concreteness extremes only: $t[3432] = -2.70$, $p = 0.007$; full dataset: $t[5111] = -2.157$, $p = 0.031$).

4. Discussion

The current study used a vector-space measure of semantic similarity to calculate semantic neighborhood density for both concrete and abstract words, using a measure that emphasizes taxonomic, as opposed to thematic or associative aspects of the concept. An analysis of 3500 words showed that abstract words tend to have more sparse neighborhoods than concrete words, and that the faciliatory effect of neighborhood sparseness on lexical decision is stronger for concrete words than for abstract words.

The difference in neighborhood density between concrete and abstract words also appears to be strongest for very small neighborhoods; for example, the density of a word's three nearest neighbors is much higher for concrete than abstract words, but the density of a word's fifty nearest neighbors is similar for concrete and abstract words. In other words, abstract words tend to have a few very close neighbors, unlike concrete words. As the definition of a near neighborhood is relaxed, and more distant neighbors are included, the difference between the two reduces. This can be understood in terms of their origins. Concrete entities can be divided into natural kinds and artifacts. Natural kinds are either living things that are products of evolution (e.g., animals, plants) or inanimate objects (such as mountains or lakes) that follow the principles of physics. Due to their origins, they tend to have many similar features that exhibit "coherent covariation" (McClelland & Rogers, 2003). For example, any given pair of mammals very likely share the features "has two eyes", "can see", "has fur", "has legs", and "can walk"; as a result, members of the mammal family are easy to access, and are often preserved in patients even when other semantic information has been lost (Devlin, Gonnerman, Andersen, & Seidenberg, 1998; Moss, Tyler, Durrant-peatfield, & Bunn, 1998). Artifacts are construed to serve particular functions and to be used by humans (or other animals), and hence are constrained by anatomy. For example, many tools have handles of a certain size so they can be grasped by the human hand, have sharp edges if they are meant for cutting, have a weight range such that they can be lifted by people, and so on, again resulting in coherent covariation. Such classes have members that contain many overlapping features, and differ by a single or a few features, resulting in many close neighbors. Hence, close neighbors of a concrete concept also tend overlap with each other. Abstract entities, on the other hand, are not directly constrained by physical or biological principles. While they can have many neighbors for a sufficiently generous definition of neighborhood, high overlap in conceptual content is not common. For example, in the current dataset, the word "justice" has a dense semantic neighborhood, but its nearest neighbors are not close neighbors with each other ("accountability", "autonomy", "commerce"). Conversely, "camel", which has a much more sparse neighborhood overall, maintains close neighbors that share some features ("bison", "boar", "donkey"). In terms of semantic access, the pattern of consistency across many categories of concrete nouns can drive the SND-specific effects, which are weaker for abstract than concrete words.

The finding that concrete words have closer near neighbors than abstract words has potential implications for the functional roles of neural areas that are sensitive to abstract words, for example, the anterior temporal lobe (ATL) (Noppeney & Price, 2004; Sabsevitz et al., 2005; Wang et al., 2010). The ATL is known to have a role in processing unique entities. A word with a sparse neighborhood (i.e., many abstract words) is, in a sense, "unique" relative to other words in the lexicon. Similarly, Rogers, Ralph, Hodges, and Patterson (2004) showed that patients with semantic dementia, who tend to have ATL damage, have selectively impaired performance for two other types of "unique" information: words with low bigram/trigram frequency, and object images with atypical visual features for their category. Together, these findings suggest the possibility that the ATL plays a general role in processing sparsely encoded, unique semantic information.

Additionally, the neighborhoods of concrete and abstract words pattern differently for words with high emotional arousal. Abstract words appear to be relatively sparse regardless of emotional content, while concrete words have particularly dense neighborhoods if they are high in emotional arousal. Emotionally charged concrete words seem to have many close semantic neighbors: for example, "bully" has a dense neighborhood, including neighbors such as "bastard", "brat", and "madman". Conversely, "treason", which is among the most emotionally charged abstract words, has a more distant set of closest neighbors in this dataset ("adultery", "bribery", "evasion").

During lexical decision, there is a selective effect of neighborhood density on concrete words: sparse, concrete words are accessed more quickly than others. This effect emerges even when emotional valence and arousal are accounted for. These results are an interesting reversal of Recchia and Jones (2012), who used corpus norms that load more heavily on associative information to measure the impact of semantic neighborhood density on lexical decision RTs. They showed that increasing associative-based neighborhood density has a selective faciliatory effect on abstract words. Here, we find that a decreasing feature-based neighborhood density has a selective faciliatory effect on concrete words. Two conclusions can be drawn from these differences. First, as predicted, concrete concepts are selectively affected by feature-based measures, while Recchia & Jones found that abstract concepts are selectively affected by associative-based measures. Second, higher associative-based density speeded RTs in Recchia & Jones' results, but here the fastest RTs were elicited for low-density words. Naming RTs also showed a similar pattern of results. Selective effects of SND on these tasks can also be understood in terms of patterns of feature overlap mentioned above. More competition for the selection of the correct response can occur for concrete words with high neighborhood density, because concepts in the neighborhood tend to be densely interconnected and activate each other. Concrete words with sparse neighborhoods do not face this competition, and hence are relatively fast. Because abstract words do not exhibit similar featural overlap, there is no relative advantage of having few neighbors. Dell's interactive two-step model of lexical access implements similar mechanisms (Dell, 1986; Dell & O'Seaghdha, 1992; Schwartz, Dell, Martin, Gahl, & Sobel, 2006). In this model, when feature units are activated (e.g., from a picture in a picture naming task), they activate localist word units in the word layer. Due to feature overlap, not just the word unit itself but its taxonomic neighbors are also activated to some extent (e.g., picture of a lion activates 'lion', and also 'tiger', 'leopard' units). With noise in the system, the model makes taxonomic errors, similar to those made by patients. Similar mechanisms can operate in lexical decision and naming tasks, due to bi-directional or interactive connections between the layers. Orthographic forms activate word units, which activate corresponding semantic feature units, which in turn activate word units of the semantic neighbors of the target word. In healthy subjects, many taxonomic neighbors for concrete words are reflected in increased RTs due to this competition. The current results, combined with those of Recchia and Jones, point to a fundamental difference in representation and processing of concrete and abstract words; featural overlap is important for concrete words, with dense featural neighborhoods resulting in relative slowing in lexical decision and naming tasks, while associative information is more important for abstract words, with stronger associations resulting in relative facilitation.

The difference between the current results and those of Recchia and Jones (2012) suggests that semantic neighborhood density is not a single measure. Indeed, the two methods for calculating SND – associative and feature-based – appear to have very different and independent effects on behavior. Moving forward, we recommend that semantic neighborhood density should not be used as a blanket term to describe both feature-based and associative density, and the two should be treated as different psycholinguistic measures.

Finally, the current results differ somewhat from the Danguecan and Buchanan (2016) findings, which can be expected given the differences in vector spaces between the two studies. Danguecan & Buchanan found selectively slower RTs for abstract words with dense neighborhoods, although this condition was also more highly abstract than the other conditions, which may have also slowed RTs. Using a much larger set of words, the current results suggest that LD RTs are particularly fast for concrete words with sparse neighborhoods.

The current analysis is a preliminary investigation of the influence of semantic neighborhood density on the study of concrete and abstract words. Additional methods of measuring semantic neighborhood density would be useful, e.g., differently modeled corpus-based semantic spaces, or a measure of semantic feature ratings that can accommodate the variables of importance to abstract words (Crutch et al., 2013). As an initial result, however, we find clear evidence that abstract and concrete words differ on at least one measure of semantic neighborhood density, and both SND and emotional arousal affect behavior differently for concrete and abstract words, even when other lexical variables are accounted for. This highlights the importance of taking semantic neighborhood density into account in studies of the cognitive and neural basis of semantics.

Acknowledgments

This research was supported in part by NIH/NIDCD grant R01 DC010783 (R. H. D.).

References

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchinson, K. A., Kessler, B., Loftus, B., ... Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods, 39*(3), 445–459.

Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41*(4), 977–990. http://dx.doi.org/10.3758/BRM.41.4.977.

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

Buchanan, L., Westbury, C. F., & Burgess, C. (2001). Characterizing semantic space: Neighborhood effects in word recognition. *Psychonomic Bulletin & Review, 8*(3), 531–544.

Burgess, C., & Lund, K. (1997). Representing abstract words and emotional connotation in a high-dimensional memory space. *Cognitive Science Proceedings*.

Cree, G. S., & McRae, K. (2003). Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns). *Journal of Experimental Psychology: General, 132*(2), 163–201. http://dx.doi.org/10.1037/0096-3445.132.2.163.

Crutch, S. J., Troche, J., Reilly, J., & Ridgway, G. R. (2013). Abstract conceptual feature ratings: The role of emotion, magnitude, and other cognitive domains in the organization of abstract conceptual knowledge. *Frontiers in Human Neuroscience, 7*, 186. http://dx.doi.org/10.3389/fnhum.2013.00186.

Dalla Volta, R., Fabbri-Destro, M., Gentilucci, M., & Avanzini, P. (2014). Spatiotemporal dynamics during processing of abstract and concrete verbs: An ERP study. *Neuropsychologia, 61*, 163–174. http://dx.doi.org/10.1016/j.neuropsychologia.2014.06.019.

Danguecan, A. N., & Buchanan, L. (2016). Semantic neighborhood effects for abstract versus concrete words. *Frontiers in Psychology, 7*, 1034. http://dx.doi.org/10.3389/fpsyg.2016.01034.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93*(3), 283–321.

Dell, G. S., & O'Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition, 42*, 287–314.

Devlin, J. T., Gonnerman, L. M., Andersen, E. S., & Seidenberg, M. S. (1998). Category-specific semantic deficits in focal and widespread brain damage: A computational account. *Journal of Cognitive Neuroscience, 10*(1), 77–94.

Durda, K., & Buchanan, L. (2008). WINDSOR: Windsor improved norms of distance and similarity of representations of semantics. *Behavior Research Methods, 40*(3), 705–712. http://dx.doi.org/10.3758/brm.40.3.705.

Erk, K. (2016). What do you know about an alligator when you know the company it keeps? *Semantics and Pragmatics, 9*. http://dx.doi.org/10.3765/sp.9.17.

Fellbaum, C. (1998). *WordNet: An electronic lexical database*. Cambridge, MA: MIT Press.

Gorno-Tempini, M. L., & Price, A. R. (2001). Identification of famous faces and buildings. *Brain, 124*, 2087–2097.

Grabowski, T. J., Damasio, A., Tranel, D., Ponto, L. L. B., Hichwa, R. D., & Damasio, A. (2001). A role for left temporal pole in the retrieval of words for unique entities. *Human Brain Mapping, 13*, 199–212.

Keuleers, E., Lacey, P., Rastle, K., & Brysbaert, M. (2012). The British Lexicon Project: Lexical decision data for 28,730 monosyllabic and disyllabic English words. *Behavior Research Methods, 44*(1), 287–304. http://dx.doi.org/10.3758/s13428-011-0118-4.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 194*(2), 211–240.

McClelland, J. L., & Rogers, T. T. (2003). The parallel distributed processing approach to semantic cognition. *Nature Reviews Neuroscience, 4*(4), 310–322. http://dx.doi.org/10.1038/nrn1076.

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

Mirman, D., & Magnuson, J. S. (2008). Attractor dynamics and semantic neighborhood density processing is slowed by near neighbors and speeded by distant neighbors. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(1), 65–79.

Moss, H. E., Tyler, L. K., Durrant-peatfield, M., & Bunn, E. M. (1998). 'Two Eyes of a See-through': Impaired and intact semantic knowledge in a case of selective deficit for living things. *Neurocase, 4*(4–5), 291–310. http://dx.doi.org/10.1080/13554799808410629.

Newcombe, P. I., Campbell, C., Siakaluk, P. D., & Pexman, P. M. (2012). Effects of emotional and sensorimotor knowledge in semantic processing of concrete and

abstract nouns. *Frontiers in Human Neuroscience, 6*(275), 1–15. http://dx.doi.org/10.3389/fnhum.2012.00275.
Noppeney, U., & Price, C. J. (2004). Retrieval of abstract semantics. *Neuroimage, 22*(1), 164–170. http://dx.doi.org/10.1016/j.neuroimage.2003.12.010.
Olson, I. R., McCoy, D., Klobusicky, E., & Ross, L. A. (2013). Social cognition and the anterior temporal lobes: A review and theoretical framework. *Social Cognitive and Affective Neuroscience, 8*(2), 123–133. http://dx.doi.org/10.1093/scan/nss119.
Paivio, A. (2010). Dual coding theory and the mental lexicon. *The Mental Lexicon, 5*(2), 205–230. http://dx.doi.org/10.1075/ml.5.2.04pai.
Pexman, P. M., Hargreaves, I. S., Siakaluk, P. D., Bodner, G. E., & Pope, J. (2008). There are many ways to be rich: Effects of three measures of semantic richness on visual word recognition. *Psychonomic Bulletin & Review, 15*(1), 161–167. http://dx.doi.org/10.3758/pbr.15.1.161.
Plaut, D. C., & Shallice, T. (1993). Deep dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology, 10*(5).
Pollock, L. (2017). Statistical and methodological problems with concreteness and other semantic variables: A list memory experiment case study. *Behavioral Research Methods.* http://dx.doi.org/10.3758/s13428-017-0938-y. Advance online publication.
Recchia, G., & Jones, M. N. (2012). The semantic richness of abstract concepts. *Frontiers in Human Neuroscience, 6*, 315. http://dx.doi.org/10.3389/fnhum.2012.00315.
Rogers, T. T., Ralph, M. A., Hodges, J. R., & Patterson, K. (2004). Natural selection: The impact of semantic impairment on lexical and object decision. *Cognitive Neuropsychology, 21*(2), 331–352. http://dx.doi.org/10.1080/02643290342000366.
Roller, S., & Erk, K. (2016). Relations such as hypernymy: Identifying and exploiting Hearst patterns in distributional vectors for lexical entailment. *arXiv:1605.05433 [cs.CL].*
Ross, L. A., & Olson, I. R. (2010). Social cognition and the anterior temporal lobes. *Neuroimage, 49*(4), 3452–3462. http://dx.doi.org/10.1016/j.neuroimage.2009.11.012.
Sabsevitz, D. S., Medler, D. A., Seidenberg, M., & Binder, J. R. (2005). Modulation of the semantic system by word imageability. *Neuroimage, 27*(1), 188–200. http://dx.doi.org/10.1016/j.neuroimage.2005.04.012.
Schwartz, M., Dell, G., Martin, N., Gahl, S., & Sobel, P. (2006). A case-series test of the interactive two-step model of lexical access: Evidence from picture naming. *Journal of Memory and Language, 54*(2), 228–264. http://dx.doi.org/10.1016/j.jml.2005.10.001.
Tranel, D. (2009). The left temporal pole is important for retrieving words for unique concrete entities. *Aphasiology, 23*(7–8), 867–884. http://dx.doi.org/10.1080/02687030802586498.
Vigliocco, G., Meteyard, L., Andrews, M., & Kousta, S. (2014). Toward a theory of semantic representation. *Language and Cognition, 1*(02), 219–247. http://dx.doi.org/10.1515/langcog.2009.011.
Wang, J., Conder, J. A., Blitzer, D. N., & Shinkareva, S. V. (2010). Neural representation of abstract and concrete concepts: A meta-analysis of neuroimaging studies. *Human Brain Mapping, 31*(10), 1459–1468. http://dx.doi.org/10.1002/hbm.20950.
Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods, 45*(4), 1191–1207. http://dx.doi.org/10.3758/s13428-012-0314-x.
Zdrazilova, L., & Pexman, P. M. (2013). Grasping the invisible: Semantic processing of abstract words. *Psychonomic Bulletin & Review, 20*(6), 1312–1318. http://dx.doi.org/10.3758/s13423-013-0452-x.