1

Lexicosemantic, affective, and distributional norms for 1,000 Dutch adjectives

Steven Verheyen[1]*, Simon De Deyne[2], Sarah Linsen[1], & Gert Storms[1]

[1]KU Leuven, Leuven, Belgium

[2]University of Melbourne, Melbourne, Australia

*Address for correspondence:

Steven Verheyen

Laboratory for Experimental Psychology

Tiensestraat 102 box 3711

3000 Leuven, Belgium

Tel: +32 16 37 30 12

E-mail: steven.verheyen@kuleuven.be

Abstract

The research of the word is still very much the research of the noun. Adjectives have been largely overlooked, despite being the second largest word class in many languages and serving an important communicative function because of the rich, nuanced qualifications they afford. Adjectives are also ideally suited to study the interface between cognition and emotion, as they naturally cover the entire range of lexicosemantic variables like imageability (*infinite-green*) and affective variables like valence (*sad-happy*). We illustrate this by showing how the centrality of words in the mental lexicon varies as a function of the words' affective dimensions using newly collected norms for 1,000 Dutch adjectives. The norms include the lexicosemantic variables age of acquisition, familiarity, concreteness, and imageability; the affective variables valence, arousal, and dominance; and a variety of distributional variables, including network statistics resulting from a large-scale word association study. The norms are freely available from https://osf.io/nyg8v/ for researchers studying adjectives proper or for whom adjectives constitute convenient stimuli to study other topics such as vagueness, inference, spatial cognition, or affective word processing.

Keywords: age of acquisition, arousal, concreteness, dominance, familiarity, imageability, valence.

Lexicosemantic, affective, and distributional norms for 1,000 Dutch adjectives

1. Introduction

The research of the word is predominantly the research of the noun (Clark & Paivio, 2004; Medin, Lynch, & Solomon, 2000). Adjectives in particular have been overlooked, perhaps because not all languages have this word class and because they are often defined negatively, as a set of lexical items that are distinct from the universal noun and verb classes on morphological and syntactic grounds (Dirven & Taylor, 1988; Dixon, 1982; Vogel, 2004). In many languages adjectives are, however, undeniably important parts of speech, both because of their number and because of their semantic role. The CELEX count for Dutch words, for instance, contains 95,657 nouns, 13,912 adjectives, and 11,837 verbs (Baayen, Piepenbrock, & van Rijn, 1993). In the "Small World of Words" Dutch word association norms (SWOW-NL, De Deyne & Storms, 2008b) 18% of the produced associations are adjectives, making them the second most frequently produced word class (after nouns with 72%). Semantically, nouns refer to concepts, whereas adjectives refer to properties (Gärdenfors, 2000). In communication adjectives thus allow one to distinguish or identify instances that are referred to by the same noun (as in: hand me the *tall* glass; Dixon, 1982). Adjectives also constitute the majority of the hubs in semantic networks, connecting distinct parts of the mental lexicon (e.g., the adjective *white* connects the semantically remote words North Pole and sink; De Deyne & Storms, 2008b). As such, they contribute considerably to the richness and flexibility of our language.

The predominance of nouns is also reflected in available norm studies (Bird, Franklin, & Howard, 2001). Whenever adjectives are included in norm studies, they are either few in

number (Altarriba, Bauer, & Benvenuto, 1999; Berrian, Metzler, Kroll, & Clark-Meyers, 1979; Grühn & Smith, 2008; Võ, Conrad, Kuchinke, Urton, Hofmann, & Jacobs, 2009) or data for a small number of variables is obtained (Anderson, 1968; Bird, Franklin, & Howard, 2001; Brysbaert, Stevens, De Deyne, Voorspoels, & Storms, 2014; Kuperman, Stadthagen-Gonzalez, & Brysbaert, 2012; Lynott & Connell, 2009; van Loon-Vervoorn, 1985). In order to avoid sampling bias, to allow for generalization, and to ensure that the critical lexical variables can be controlled for, norming data for a variety of words and variables are required, however (Clark & Paivio, 2004; Kousta, Vinson, & Vigliocco, 2009). In line with other recent studies (e.g., Quadflieg, Michel, Bukowski, & Samson, 2014), the current study aimed to accommodate the paucity of normative data for adjectives by having 1,300 students provide ratings on seven variables for 1,000 Dutch adjectives.

We purposefully include both distributional and experiential variables. We follow Andrews, Vigliocco, and Vinson (2009) in defining distributional variables as specifying how words are statistically distributed across different spoken or written texts, and experiential variables as capturing perceived attributes or properties associated with the referents of the words. Distributional variables thus pertain to linguistic information (e.g., frequency, orthographic neighborhood size, summated bigram frequency), while experiential variables result from experience with the physical world (Andrews, Vigliocco, & Vinson, 2009; Barsalou, Santos, Simmons, & Wilson, 2008; Santos, Chaigneau, Simmons, & Barsalou, 2011). Among the experiential variables one sometimes distinguishes between lexicosemantic ones (e.g., age of acquisition, familiarity, concreteness, imageability) and affective ones (e.g., valence, arousal, dominance), but the distinction is hard to make in any case (Andrew, Vigliocco, & Vinson, 2009; Citron, Weekes, & Ferstl, 2014).

There is a growing consensus that integration of distributional and experiential information is an asset for research domains at both sides of the spectrum (Kousta, Vigliocco, Vinson, Andrews, & Del Campo, 2011; Langacker, 1987; Talmy, 2000) and empirical evidence to this point is accumulating (e.g., Dolan, 2002; Kousta, Vinson, & Vigliocco, 2009; Kuperman, Estes, Brysbaert, & Warriner, 2014). In light of the roles adjectives play in our language (see above) we feel it is especially important to include experiential variables in addition to linguistic ones to adequately describe the stimulus domain (Clark & Paivio, 2004; De Deyne, Voorspoels, Verheyen, Navarro, & Storms, 2014). Lexicosemantic variables like concreteness and imageability, and affective variables like valence and arousal are important to include since the adjectival domain covers words representing the very abstract (*ideal*, *infinite*) to the very concrete (*blue*, *green*) and includes both emotion words (*happy*, *sad*) and neutral words (*level*, *similar*).

In the following sections we describe the details of the data collection and demonstrate how the norms might be applied through an investigation of the structure of a word association network. In line with the arguments provided above, this demonstration will show that the consideration of both experiential and distributional information proves fruitful for understanding the organization of the mental lexicon.

2. Collection of norms

    2.1. Participants

    Our participants were university students because this is the population typically tested in the studies for which the ratings are intended. A total of 1,300 Dutch speaking students were recruited at KU Leuven, Belgium. They participated in exchange for course credit.

Each participant provided ratings for only one variable so that the ratings for one variable could not influence or contaminate the ratings for another variable.

2.2. Materials

The 488 Dutch adjectives from van Loon-Vervoorn (1985) were included. The 328 English adjectives from Berrian et al. (1979) were translated into Dutch and were included as well. Other research purposes (reported in De Deyne, Voorspoels, Verheyen, Navarro, & Storms, 2014) prompted us to add additional adjectives for a total of 1,000.

Attention was drawn in the instructions to the fact that all the stimuli were to be regarded as adjectives, because the classification of some words may be arguable without context (e.g., *abstract, bitter, fine, firm, flat, gold, human, ideal, light, minute, musical, patient, piercing, safe, stiff, sweet*).

2.3. Procedure

Experiential variables

In collecting the ratings, we employed a modular approach, in which the full sample of words was divided into five separate blocks, except for age of acquisition (AoA), for which we employed 35 blocks. This decision was made because for AoA participants needed to indicate the age at which they first acquired a word, which was deemed more effortful than providing a judgment on a scale, like for the other variables (see below). For each variable two (AoA, concreteness, imageability) or four (arousal, dominance, familiarity, valence) permutations of the adjectives in each block were randomly distributed across 20

participants. Participants completed the ratings on paper (AoA, familiarity, connotation) or on a computer (concreteness, imageability, arousal, dominance).

All ratings were performed on 7-point scales, except for AoA, for which participants were asked to enter the age (in years) at which they thought they had learned the word. Participants were also given the option to indicate that they did not know a word. Where possible, the instructions were taken from other Dutch norming studies.

For *AoA* we used the same instructions Moors et al. (2013) did. For each word, the participants were asked to enter the age (in years) at which they thought they had learned the word.

The instructions for *familiarity* were taken from De Deyne et al. (2008). Participants rated how familiar they were with each word in the list by indicating how often they had encountered or used the word. Ratings were indicated on a 7-point scale with the ends labeled never (1) and very often (7).

The instructions for *concreteness* were taken from Van der Goten, De Vooght, and Kemps (1999). Participants rated each word's level of abstraction on a 7-point scale ranging from very abstract (1) to very concrete (7).

The instructions for *imageability* were taken from Paivio, Yuille, and Madigan (1968). Participants rated each word's capacity to arouse a mental image by indicating how easily they could form a mental image of the word on a 7-point scale with the ends labeled difficult (1) and easy (7).

The instructions for *valence* were taken from Van der Goten, De Vooght, and Kemps (1999) and invited participants to rate the valence of each word by indicating on a 7-point scale the extent to which it evoked a very bad (1) or a very good (7) feeling.

Following Moors et al. (2013), participants were asked to indicate the degree to which a word refers to something very passive/calm (1) or very active/arousing (7), and to something very weak/submissive (1) or very strong/dominant (7) for *arousal* and *dominance*, respectively.

Distributional variables

The experiential norms were supplemented with distributional norms. For each adjective, we included the number of characters (*nchar*) and the number of syllables (*nsyl*), as well as the orthographic neighborhood size (*neighb*), the summated position-nonspecific bigram frequency (*bigram*), and the frequency (*celex_freq*) based on the Dutch version of the CELEX database (Baayen, Piepenbrock, & van Rijn, 1993). The variable *neighb* indicates the number of CELEX words that are obtained by changing one letter of the adjective, whereas *bigram* indicates the frequency of the adjacent letter pairs of the adjective across all the words included in CELEX. The variable *subtlex_freq* lists the lemma frequency of the adjectives in the SUBTLEX-NL database (Keuleers, Brysbaert, & New, 2010). It indicates the frequency with which any form of the 1,000 stimulus words occurs in a corpus of Dutch subtitles as an adjective (based on part-of-speech tagging).

The variables *in-strength* and *betweenness* were taken from the Small World of Words project (https://smallworldofwords.org/), a large-scale study that aims to build a map of the human lexicon in the major languages of the world (including Dutch) using word association data (De Deyne, Navarro, Perfors, Brysbaert, & Storms, 2018). To this end,

participants are shown 15 cue words and are invited to provide the three words that first come to mind in response to each cue (see De Deyne & Storms, 2008a, for details). These responses are used to construct a word association network from which the above measures can be derived (see De Deyne, Navarro, & Storms, 2013, for details). The in-strength s_i of a word i reflects the number of links that go into the node representing the word in the word association network and as such provides an indication of the word's importance in the mental lexicon. It is operationalized as the weighted sum of the incoming links:

$$s_i = \sum_{j \in N} w_{ij}$$

with w_{ij} the weight of the link connecting nodes i and j and N the number of nodes in the network. Betweenness is another indicator of the centrality of a node in the network. It takes the global structure of the word association network more in account than in-strength does in that it captures how often a node is located on the shortest path between other nodes in the network. Betweenness was calculated with the iGraph R package (Csardi & Nepusz, 2006).The betweenness b_i of a node i in the network is operationalized as:

$$b_i = \frac{2}{(n * n - 3 * n + 2)} \sum_{\substack{h,j \in N, \\ h \neq j, h \neq i, j \neq i}} \frac{\rho_{hj}(i)}{\rho_{hj}}$$

with $\rho_{hj}(i)$ corresponding to the weighted number of shortest paths from h to j passing through i and ρ_{hj} the total number of shortest paths from h to j. The summed proportions are divided by the number of all possible pairs h,j. The first term of the equation is used to normalize the score to take into account networks of different sizes.

2.4. Results

The full set of norms may be downloaded from https://osf.io/nyg8v/. In addition to the distributional variables, the norms include the original responses, mean values, standard deviations, and % unknown for the rated attributes of each word. The file also contains English translations of the Dutch materials.

Reliability

Table 1

Estimates of the Reliability of the Experiential Variables for each Block of Adjectives

	Block				
Measure	1	2	3	4	5
Familiarity	0.96	0.95	0.96	0.95	0.94
Concreteness	0.80	0.85	0.82	0.88	0.85
Imageability	0.86	0.93	0.90	0.88	0.95
Valence	0.97	0.97	0.97	0.96	0.96
Arousal	0.91	0.93	0.93	0.94	0.90
Dominance	0.90	0.90	0.92	0.91	0.94

The reliability of the ratings within each block was evaluated by applying the Spearman-Brown formula to the split-half correlations (Spearman, 1904). Table 1 shows the reliabilities of the variables familiarity, concreteness, imageability, valence, arousal, and dominance for each of the five blocks. The reliabilities for AoA, for which there were 35 blocks, ranged between .95 and .99. The reliability values indicate a substantial degree of

agreement between the participants. Do note, however, that because of homonymy the consistency may be less for some words included in the modules. The rating variability of each word, indicated by the standard deviation, is arguably a good indication of this ambiguity.

Validity

All the collected ratings can be validated by correlating the mean ratings obtained in our study with those obtained in other studies. In the section below, *N* will always refer to the number of overlapping stimuli. The values reported between square brackets pertain to the subset of the stimuli that only have an adjectival reading in Dutch. We made it clear in our instructions that all words should be considered adjectives, but since other studies included words from other word classes as well, participants in these studies might have responded toward the noun reading, which in turn might have affected the rating.

Our mean AoA ratings correlate .93 [.92] with those in Brysbaert, Stevens, De Deyne, Voorspoels, and Storms (2014) for the *N*=473 [416] stimuli included in both studies. When the ratings of four Dutch AoA norming studies are combined (Ghyselinck, Custers, & Brysbaert, 2003; Ghyselinck, De Moor, & Brysbaert, 2000; Moors et al., 2013; Brysbaert et al., 2014) the overlap increases to 959 [707] and the correlation becomes .94 [.94]. Our mean valence, arousal, and dominance ratings correlate .96 [.97], .92 [.93], and .90 [.92], respectively with those in Moors et al. (2013) for the 494 [298] stimuli included in both studies. Our mean valence ratings correlate .96 [.96] with those for the 154 adjectives that were also included in Hermans and De Houwer (1994). Our mean familiarity ratings correlate .82 [.90] with the familiarity ratings taken from the same paper. These substantial correlations with external norms indicate that the ratings for AoA, familiarity,

valence, arousal, and dominance are well in line with those of previous studies. The magnitude of the correlations approaches the theoretical maximum, considering that the reliability of the ratings (see Table 1) constitutes a ceiling for the correlations of the ratings with external variables (Spearman, 1904). Mismatches due to noun readings of particular adjectives in earlier studies appear to be limited. Restricting the correlation to stimuli that only have an adjectival reading does not impact the correlation heavily.

The mean imageability ratings for the *N*=488 [305] adjectives in van Loon-Vervoorn (1985) correlate .76 [.74] with our mean imageability ratings and .45 [.40] with our mean concreteness ratings. The magnitude of the latter correlation coefficient is very similar to the .46 [.41] correlation observed between imageability and concreteness in our own study for the subset of 488 [305] adjectives. The mean concreteness ratings for the 957 [707] adjectives in Brysbaert et al. (2014) correlate .68 [.69] with our mean imageability ratings and .49 [.47] with our mean concreteness ratings. Although the latter correlation does not seem to be strongly affected by the potentially different interpretation of the word class in Brysbaert et al. (2014), it is considerably lower than one would expect a priori. We believe that the composition of the stimulus set and differences in the wording of the rating scale might be responsible for this discrepancy. Compared to the participants in Brysbaert et al., our participants appeared to use a restricted range of the concreteness scale, which might have reduced the correlation. The 5th and 95th percentile of the mean concreteness ratings for the overlapping adjectives on the 5-point rating scale used by Brysbaert et al. are 1.40 and 4.13, while the corresponding values for our 7-point rating scale are 3.25 and 6.00. It thus appears that our participants rarely used the lower points on the concreteness scale, presumably because unlike the participants in Brysbaert et al. the stimuli they were asked to rate did not include nouns and verbs, many

of which are arguably relatively concrete compared to adjectives. This observation has ramifications for the way in which concreteness norms collected for specific word classes – be it nouns, adjectives, or verbs – are used. These data cannot be straightforwardly combined because the relative magnitudes might not reflect comparable levels of abstractness. We believe that the lack of agreement between our concreteness ratings and those of Brysbaert and colleagues is also due to the use of different operationalizations of the concreteness scale, which might carry different meanings. In their instructions for the concreteness rating task, Brysbaert et al. equated 'abstract' with language-based and 'concrete' with experiential. This might explain why the Brysbaert concreteness ratings correlate better with our imageability ratings than they do with our concreteness ratings. Following Van Der Goten, De Vooght, and Kemps (1999), we employed rather general concreteness instructions, in which the interpretation of 'abstract' and 'concrete' were left to the participants, while in the imageability instructions a rather explicit reference is made to the (visual) senses. This seems to agree better with Brysbaert et al.'s concreteness instructions, which also refer to experience. The observation that scales intended to measure the same construct can yield different results due to what appear to be minor phrasing differences is a cause of concern and warrants further investigation.

Correlations

The collected ratings can be further validated by verifying whether the pattern of inter-correlations between the variables is similar to that of other norming studies. Table 2 shows the correlations between the seven experiential variables in our study.

Table 2

Correlations between the Experiential Variables collected for 1,000 Dutch Adjectives

Variable	1	2	3	4	5	6	7
1. AoA	-						
2. Familiarity	-.70	-					
3. Concreteness	-.31	.20	-				
4. Imageability	-.47	.23	.50	-			
5. Valence	-.07	.18	.04	-.02	-		
6. Arousal	.00	.12	-.01	.02	.43	-	
7. Dominance	.05	.08	-.00	-.01	.45	.79	-

A natural comparison is one with the AoA, valence, arousal, and dominance norms of

Moors et al. (2013) for 4,300 Dutch words; 494 of which are included in our study. We

found that AoA shares hardly any variance with the affective variables valence, arousal,

and dominance (see Table 2). This is also the case in the Moors et al. data, both for the

overlapping stimuli (*r*= -.06, *r*=.05, *r*=-.01, respectively) and the entire stimulus set (*r*=-.17,

r=.03, *r*=.08). The correlations among the affective variables in Table 2 are also mirrored

in the Moors et al. data for the overlapping stimuli, where the correlation between

arousal and dominance is found to be the strongest (*r*=.71), and the correlation between

valence and dominance (*r*=.56) is found to be somewhat stronger than that between

valence and arousal (*r*=.27). This was to be expected given the strong correlation between

our data and those of Moors et al. (all *r*>.90; see above). The correlations among the

affective variables are, however, more pronounced for the overlapping stimuli than for

the stimulus set as a whole ($r= -.01$, $r=.27$, $r=.59$, respectively), indicating that the observed structure of the affective variables might be specific to adjectives, and might not generalize across word classes.

Another natural comparison is that with Brysbaert et al. (2014) where AoA and concreteness were found to correlate -.34 across 25,882 Dutch words from different word classes. When the stimulus set is restricted to the 473 adjectives that are also included in our stimulus set, the correlation becomes -.35. Both correlations are comparable to the -.31 correlation across our 1,000 adjectives (see Table 2).

Finally, we established a less pronounced relationship between familiarity and valence than Hermans and De Houwer (1994) did. While we observed a correlation of .18 (see Table 2), they found that familiarity and valence correlated .25 across 370 adjectives and .34 across the 154 adjectives included in both studies.

There are no Dutch norming studies available that include both AoA, familiarity, concreteness, and imageability. For comparison, we therefore turn to two English studies that incorporate all four lexicosemantic variables. Gilhooly and Logie (1980) contains ratings for 1,944 nouns. Bird, Franklin, and Howard (2001) report ratings of AoA ($N=2,694$), familiarity ($N=1,217$), concreteness ($N=1,070$), and imageability ($N=2,019$) for nouns, verbs, adjectives, numerals, adverbs, and function words. For ease of comparison, the correlations between the four lexicosemantic variables are shown next to one another in Table 3.

Table 3

Correlations between Lexicosemantic Variables collected in the Current and Two Previous Studies

Variable	This study				Gilhooly & Logie (1980)				Bird et al. (2001)			
	1	2	3	4	1	2	3	4	1	2	3	4
1. AoA	-	-	-	-	-	-	-	-	-	-	-	-
2. Familiarity	-.70	-	_	_	-.72	-	_	_	-.52	-	_	_
3. Concreteness	-.31	.20	-	_	-.50	.11	-	_	-.05	-.11	-	_
4. Imageability	-.47	.23	.50	-	-.72	.43	.78	-	-.50	-.03	.66	-

The inter-correlations between the lexicosemantic variables AoA, familiarity, concreteness, and imageability appear comparable across languages and word classes. The correlations we observed for the 1,000 Dutch adjectives tended to be in between the correlations reported in Gilhooly and Logie (1980) and Bird et al. (2001) for English stimuli that were mostly comprised of nouns. This was the case for the correlation between AoA and familiarity, between AoA and concreteness, and between familiarity and imageability. The -.47 correlation we found between AoA and imageability was somewhat smaller than the correlations reported in the literature but very close to the -.50 correlation reported in Bird et al. (2001). The .20 correlation between familiarity and concreteness was somewhat higher than the correlations in the literature but close to the .11 correlation in Gilhooly and Logie (1980). The largest discrepancy was found for the correlation between concreteness and imageability. We found a moderate correlation between the two variables, while they were found to correlate strongly in the English norming studies. As

discussed in the previous section, we believe the wording of the concreteness scale and the composition of the stimulus set to be responsible for this discrepancy (see above).

3. Application

The idea that the mental lexicon can be thought of as an organized network based on meaningful word associations, dates back to at least Deese (1966). It is currently regaining popularity due to the ability to compile large, contemporary word association datasets through crowdsourcing (De Deyne, Navarro, Perfors, Brysbaert, & Storms, 2018) and the observation that these data are particularly apt at capturing semantic relations across words of varying levels of abstraction (De Deyne, Verheyen, & Storms, 2015) and relatedness (De Deyne, Navarro, Perfors, & Storms, 2016). Word association data display assortativity for valence, arousal, and dominance: cues of a particular affective quality tend to elicit responses with a similar affective quality (Pollio, 1964; Staats & Staats, 1959; Van Rensbergen, Storms, De Deyne, 2015). Accurate predictions of words' standings on all three affective dimensions can also be obtained from word association data (Vankrunkelsven, Verheyen, Storms, & De Deyne, 2018; Van Rensbergen, De Deyne, & Storms, 2015). Therefore, word association data have the potential to uncover the extent to which there are systematic relationships between the manner in which words are organized in the mental lexicon and the words' affective dimensions, which have been claimed to be an integral part of the stored word meaning (Osgood, Suci, & Tannenbaum, 1957; Samsonovich & Ascoli, 2010).

Investigations of the extent to which distributional variables like word frequency (Steyvers & Tenenbaum, 2005) and contextual diversity (Hills, Maouene, Riordan, & Smith, 2010), and lexicosemantic variables like AoA (Steyvers & Tenenbaum, 2005) affect the inter-connectivity

of words in a word association network have already been undertaken. However, no study has systematically looked at the effects of the affective variables valence, arousal, and dominance on the organisation of the mental lexicon. The current norms allow us to undertake this investigation for adjectives, which not only can be assumed to cover the entire range of the affective variables under investigation, but also constitute an interesting class of words to investigate because of the role they have been shown to play in establishing the small-world structure of the word association network. As we already mentioned in the introduction, the majority of the hubs in semantic networks – those words connecting remote parts of the mental lexicon – are adjectives (De Deyne & Storms, 2008b).

3.1 Method

We chose to evaluate the effect of the affective variables on the organisation of the mental lexicon by regressing the variables in-strength and betweenness on the distributional and experiential variables included in the norms. Whereas in-strength represents a local measure of the organisation of the word association network, taking just the number of links between a word and its directly connected neighbors into account, betweenness provides an indication of the broader context of the word in the network through the proportion of times it features along indirect paths that connect it with all other words (see also section 2: *Distributional measures*). Both measures indicate the centrality of words in the network and have been shown to explain a host of findings in the word processing and memory literature (e.g., Hutchison, 2003; Nelson & McEvoy, 2000).

Following the investigation of the non-linear effects that valence and arousal have on lexical decision by Kousta, Vinson, and Vigliocco (2009), we carried out an ordinary least squares

regression analysis of in-strength and betweenness with the same predictors included in their study, with the addition of dominance and with the position non-specific mean bigram frequency instead of the mean positional bigram frequency. The model thus incorporated all linear and nonlinear effects of our experiential variables (AoA, familiarity, concreteness, imageability, valence, arousal, dominance) and the distributional variables orthographic neighborhood size*,* mean bigram frequency, number of characters, and SUBTLEX word frequency.[1] We took the square root of neighborhood size and logarithmically transformed in-strength, betweenness, and word frequency before entering them in the regression.

All statistical analyses were carried out in R version 3.5.1 (R Development Core Team, 2007). We used restricted cubic splines (Harrell, 2001) to model nonlinear relationships between the predictors and the dependent variables (in all instances, using three knots at quantiles {.1, .5, .9} on a given variable). Alpha was set to .05 to establish significance.

3.2 Results

The adjusted R^2 was .70 for in-strength and .65 for betweenness, indicating that the predictors in our model captured a considerable amount of variability in both dependent variables. Adjectives tended to score higher on in-strength and betweenness the shorter, more frequent, more familiar, more imageable, and earlier acquired they are (see the complete overview of the regression results in the Appendix). Here, we will focus on the effects of valence, arousal, and dominance on in-strength and betweenness. The top row of Figure 1 shows the results for in-strength. The bottom row contains the results for

---

[1] We obtained similar results when using the CELEX word frequency instead of the SUBTLEX word frequency.

betweenness. From left to right, the panels show the partial effects of valence, arousal, and dominance, respectively, when all other predictors in the model are set to their median values.
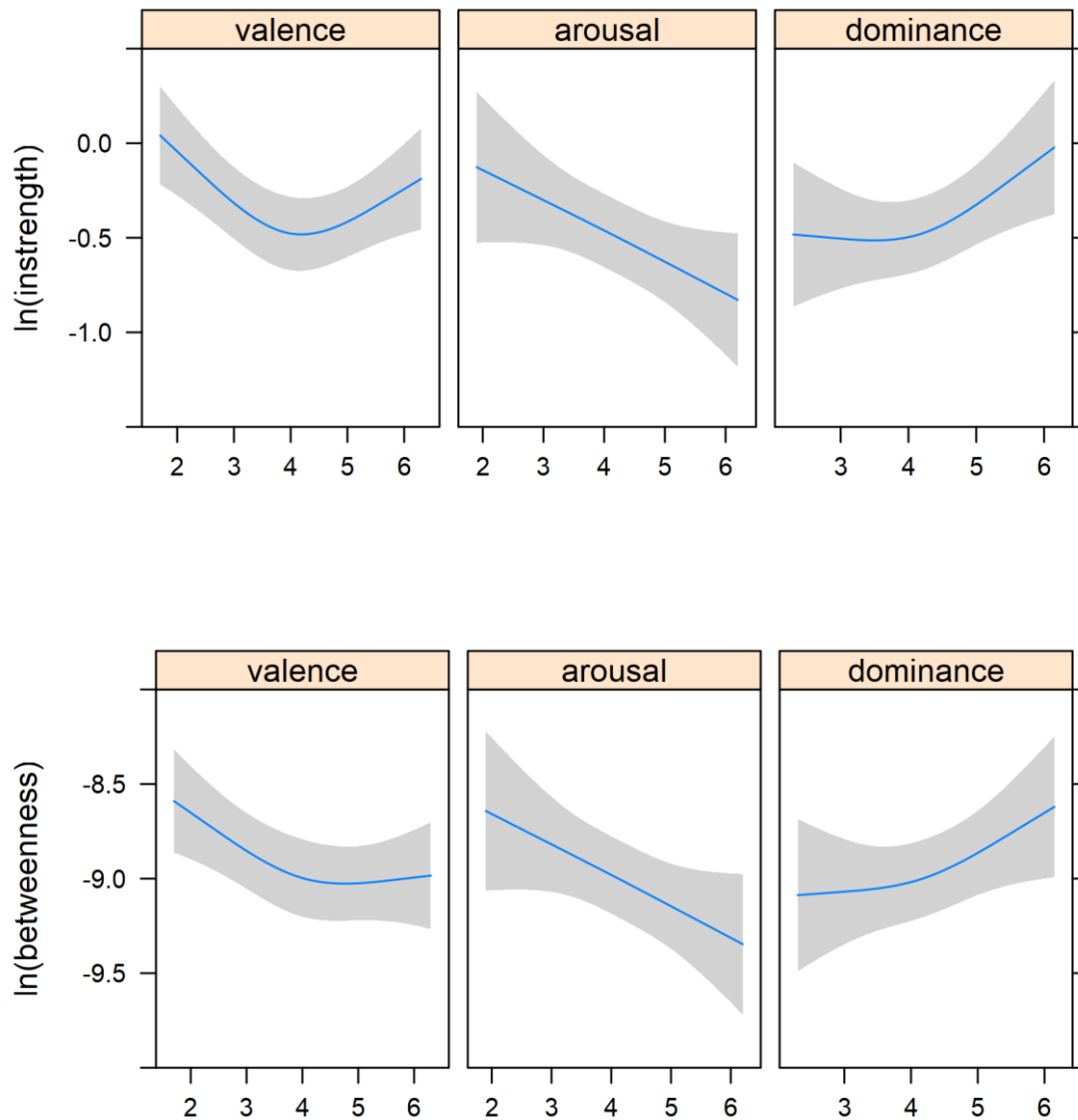

Figure 1. Plots of the partial effects of the affective variables valence (left), arousal (middle), and dominance (right) on *In*(in-strength) (top) and *In*(betweenness) (bottom) for median values of all the other predictors. The grey zone indicates a 95% confidence interval.

Valence and arousal were found to be significant predictors of both in-strength and betweenness. Valence displays a U-shaped relationship with in-strength (see top left panel in Figure 1) indicating that positive and negative words receive more incoming links (F(2,930)=8.73, p<.001; nonlinear: F(1,930)=13.18, p<.001). The results for betweenness indicate that negative words occupy a more central position in the word association network than neutral and positive words do (F(2,930)=6.00, p<.01; nonlinear: F(1,930)=3.20, p=.07; bottom left panel in Figure 1).

Words with an active/arousing character tended to receive fewer incoming links and feature less frequently in the connection between other words: Arousal had a linear effect on both in-strength (F(2,889)=5.10, p=.006; nonlinear: F(1,889)=.00, p=.95) and betweenness (F(2,889)=4.61, p=.01; nonlinear: F(1,889)=.00, p=.97).

Dominance significantly predicted in-strength (F(2,930)=4.16, p=.02; nonlinear: F(1,930)=2.75, p=.10), indicating that words tend to receive more incoming links the stronger/more dominant they tend to be. Dominance was not a significant predictor of betweenness (F(2,930)=2.94, p=.05; nonlinear: F(1,930)=.99, p=.32).

3.3 Discussion

The study of the mental lexicon has largely neglected affective variables, perhaps because previous studies were biased towards concrete neutral words. Recently, however, researchers have begun to acknowledge that affect accounts for considerable variability in the meaning of words (De Deyne, Navarro, Collell, & Perfors, 2018), particularly of abstract

words (Lenci, Lebani, & Passaro, 2018; Wang et al., 2017) and adjectives (De Deyne, Voorspoels, Verheyen, Navarro, & Storms, 2014). By regressing indices of the centrality of 1,000 adjectives in a word association network on a variety of distributional, lexicosemantic, and affective variables, we found that valence, arousal, and dominance are among the organizing principles of the mental lexicon. That is, the centrality of adjectives in the mental lexicon varies as a function of the words' affective dimensions: Negative adjectives and adjectives low in arousal take a more central position in the word association network than neutral adjectives do, and as such connect remote parts of the word association network. Highly affective adjectives (be it positive or negative), adjectives low in arousal, and adjectives that express strength take up a prominent position in the network as well. They are provided as an associate more often than other adjectives. The early research on word association has already established that words elicit more variable responses the more positive they are (Johnson & Lim, 1964; Koen, 1962). Our findings go beyond these early results in that they pertain to a much larger set of words and take the structure of the entire word association network into account. We are currently investigating the extent to which our findings generalize across word classes and languages. This investigation also intends to explain any discrepancies in the findings for in-strength vs. betweenness. One of the working hypotheses for the observation of a U-shaped relationship between valence and in-strength but not betweenness, is the disproportional distribution of positive and negative words across the lexicon (Boucher & Osgood, 1969; Dodds et al., 2015).

We have already completed a comparable analysis involving 2,042 Dutch nouns, which shows significant linear and non-linear effects of valence, arousal, and dominance on both

in-strength and betweenness[2]. The nature of the relationship between the centrality

measures and the affective dimensions for these nouns appears similar to that for the

adjectives. In-strength and betweenness display a U-shaped relationship with valence, a

decreasing relationship with arousal, and an increasing relationship with dominance. This

result thus confirms our general finding that there are systematic relationships between the

manner in which words are organized in the mental lexicon and the words' affective

dimensions.

Word associations arguably involve more semantic elaboration than the lexical decision and

word naming tasks for which the effect of emotional content on language processing has

been investigated so far (Estes & Adelman, 2008ab; Kousta, Vinson, & Vigliocco, 2009;

Kuperman, Estes, Brysbaert, & Warriner, 2014; Larsen, Mercer, Balota, & Strube, 2008;

Rodríguez-Ferreiro & Davies, 2018; Vinson, Ponari, & Vigliocco, 2014; Yap & Seow, 2014).

The involvement of valence, arousal, and also dominance – a variable not considered in

previous studies on language processing – in the organization of the mental lexicon, suggests

that all three of these affective variables should be used as controls both in studies without

and with semantic elaboration (contrary to the current practices; but see Moffat, Siakaluk,

Sidhu, & Pexman, 2015, for a notable exception).

---

[2] We employed the same regression model as for the adjectives, with the exception that (i) we reflected and then logarithmically transformed concreteness and imageability because of their negative skew for nouns, and (ii) we used Keuleers et al.'s (2015) z score measure of word prevalence (the percentage of a population knowing a word) instead of rated familiarity to ascertain a sufficiently high number of overlapping words across norming studies. We used the valence, arousal, dominance, and AoA norms from Moors et al. (2013), Brysbaert et al.'s (2014) concreteness norms, and van Loon-Vervoorn's (1985) imageability norms.

4. General Discussion

We presented lexicosemantic (age of acquisition, familiarity, concreteness, imageability), affective (valence, arousal, dominance), and distributional variables (number of characters, number of syllables, summated position-nonspecific bigram frequency, orthographic neighborhood size, and word frequency) for 1,000 Dutch adjectives. The ratings of the lexicosemantic and affective variables proved very reliable. Wherever possible, we compared our ratings to ratings obtained in other studies, which tended to include fewer adjectives. In the case of age of acquisition, familiarity, imageability, valence, arousal, and dominance, the resulting correlations were considerable, validating our measurements. In the case of concreteness, the correlation was less pronounced, suggesting that different constructs are being measured in the various norming studies and indicating that claims regarding the comparability of experiential norms should dwell on the phrasing of the accompanying instructions and the composition of the stimulus set.

The pattern of inter-correlations between the included variables also mirrored that observed in previous norming studies, again validating our measurements, except for the correlations involving concreteness due to the use of different instructions in different studies. We found that the affective variables were stronger correlated among adjectives than across words from different word classes. The observation that valence, arousal, and dominance might correlate differently depending on the word class adds to the ongoing methodological discussion about affective variable ratings. Moors et al. (2013) noted that different patterns of inter-correlations between affective variables have been reported. They attributed these differences to (i) differences in the make-up of the stimulus set, (ii) the instruction to rate the stimuli vs. the participants' feelings in response to the stimuli, and (iii) the use of a

between vs. a within-subjects design. In the latter case, participants might be more inclined to emphasize differences between the different variables. No gold standard is currently available for the measurement of lexicosemantic and affective variables, and individual researchers seem to choose the phrasing that best suits their reading of the underlying construct. To establish a gold rating scale standard it would appear that one would have to turn to external variables for validation. Especially for the affective variables this seems possible in that these experiential variables should supposedly correlate with physiological or behavioral measures. However, there is now a growing consensus that the latter measures constitute distinct aspects of emotional experience, suggesting that a gold standard is out of reach (Mauss & Robinson, 2009). For concreteness, the instruction issue might be resolved by turning to ratings of how stimuli are perceived through each of the perceptual modalities (Lynott & Connell, 2009, 2013). Whereas concreteness rating instructions have tended to overemphasize the visual modality or have left the modality unspecified (Brysbaert, Warriner, & Kuperman, 2014; Connell & Lynott, 2012), having participants provide separate ratings for each modality (visual, haptic, auditory, olfactory, gustatory) might make the ensuing data more straightforward to interpret.

Using the available norms, we investigated how affective variables provide insight into the organisation of the mental lexicon. We found that adjectives with a pronounced negative or positive valence, receive more incoming links, giving rise to a U-shaped relationship between valence and in-strength. Adjectives low in arousal and adjectives high in dominance were found to have a higher number of incoming links. The adjectives that tend to take the most central positions in a word association network were found to be negatively valenced and low in arousal. Although word associations are often referred to as constituting a corpus, suggesting that the information they contain is language-based distributional in nature, they

have been shown to contain both semantic and lexical information (De Deyne, Verheyen, & Storms, 2015, 2016; see also Collins & Loftus, 1975; Szalay & Deese, 1978). The current findings indicate that they also yield affective information about the words for which associations have been gathered (see also De Deyne, Navarro, Collell, & Perfors, 2018).

A multimodal distributional view on word meaning, according to which meaning is both embodied in modal representations and informed by word usage in context, currently prevails (e.g., Barsalou, Santos, Simmons, & Wilson, 2008; Louwerse, 2011). The Affective Embodiment Account (AEA; Vigliocco, Meteyard, Andrews, & Kousta, 2009) is an extension of this view, arguing that affect should be considered another element of meaning that is grounded in experience based on internal states, affective experiences, or implicit evaluative appraisals. Moreover, the AEA claims that affective grounding is not limited to emotion words, but extends to most words (Kousta, Vinson, & Vigliocco, 2009). The involvement of valence, arousal, and dominance in the organization of the mental lexicon supports the AEA.

Our norms add a substantive number of adjectives to the growing set of Dutch experiential and distributional norms, and can easily be connected to the behavioural norm data that are amassing for Dutch, pertaining to lexical decision (Brysbaert, Stevens, Mandera, & Keuleers, 2016; Keuleers, Diependaele, & Brysbaert, 2010), word prevalence (Keuleers, Stevens, Mandera, & Brysbaert, 2015), text reading (Cop, Dirix, Drieghe, & Duyck, 2017), and word fragment completion (Heyman, Van Akeren, Hutchison, & Storms, 2016). We believe the norms would benefit both research that studies adjectives proper (for instance, to establish a typology; Dixon, 1982; Raskin & Nirenburg, 1998) or in which adjectives constitute the preferred stimulus material such as vagueness (Hampton, 2011; Kennedy, 2007; Van Rooij, 2011; Verheyen & Egré, 2018), spatial cognition (Bianchi, Savardi, & Burro, 2011; Bianchi,

Savardi, & Kubovy, 2011), affective word processing (Bernat, Bunce, & Shevrin, 2001; Herbert, Kissler, Junghofer, Peyk, & Rockstroh, 2006), and inference (Gotzner, Solt, & Benz, 2018; Ruytenbeek, Verheyen, & Spector, 2017). The norms can be used both as explanatory variables (Gilet & Jallais, 2011; Kuperman, Estes, Brysbaert, & Warriner, 2014) and control variables (Estes & Adelman, 2008a; Larsen, Mercer, & Balota, 2006). Our own application demonstrates that the study of adjectives can also constitute an avenue into phenomena of words more general.

Open Practices Statement

All the data and materials are available at https://osf.io/nyg8v/. R scripts for conducting the regression analyses are available at https://osf.io/8zhpw/. None of the procedures were preregistered. The data, materials, and scripts used in this article are licensed under a Creative Commons Attribution 4.0 International License (CC-BY), which permits use, sharing, adaptation, distribution, and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. Any other third-party material in this article is included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

Author note

References

Altarriba, J., Bauer, L. M., & Benvenuto, C. (1999). Concreteness, context availability, and

imageability ratings and word associations for abstract, concrete, and emotion words.

Behavior Research Methods, Instruments, & Computers, 31, 578-602.

https://doi.org/10.3758/bf03200738

Anderson, N. H. (1968). Likableness ratings of 555 personality-trait words. *Journal of

Personality and Social Psychology, 9*, 272-279. https://doi.org/10.1037/h0025907

Andrews, M., Vigliocco, G., & Vinson, D. P. (2009). Integrating experiential and distributional

data to learn semantic representations. *Psychological Review, 116*, 463-498.

https://doi.org/10.1037/a0016261

Baayen, R., Piepenbrock, R., & van Rijn, H. (1993). The CELEX lexical database [CD-ROM].

Philadelphia, PA: University of Pennsylvania, Linguistic Data Consortium.

Barsalou, L. W., Santos, A., Simmons, W. K., & Wilson, C. D. (2008). Language and simulation in conceptual processing. In M. De Vega, A. M. Glenberg, & A. C. Graesser (Eds.), *Symbols, embodiment, and meaning* (pp. 245-283). Oxford, UK: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199217274.003.0013

Bernat, E., Bunce, S., & Shevrin, H. (2001). Event-related brain potentials differentiate positive and negative mood adjectives during both supraliminal and subliminal visual processing. *International Journal of Psychophysiology, 42*, 11-34. https://doi.org/10.1016/s0167-8760(01)00133-7

Berrian, R. W., Metzler, D. P., Kroll, N. E. A., & Clark-Meyers, G. M. (1979). Estimates of imagery, ease of definition, and animateness for 328 adjectives. *Journal of Experimental Psychology: Human Learning and Memory, 5*, 435-447. https://doi.org/10.1037//0278-7393.5.4.435

Bianchi, I., Savardi, U., & Burro, R. (2011). Perceptual ratings of opposite spatial properties: Do they lie on the same dimension? *Acta Psychologica, 138*, 405-418. https://doi.org/10.1016/j.actpsy.2011.08.003

Bianchi, I., Savardi, U., & Kubovy, M. (2011). Dimensions and their poles: A metric and topological theory of opposites. *Language and Cognitive Processes, 26*, 1232-1265. https://doi.org/10.1080/01690965.2010.520943

Bird, H., Franklin, S., & Howard, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior Research Methods, Instruments, & Computers, 33*, 73-79. https://doi.org/10.3758/bf03195349

Boucher, J., & Osgood, C. E. (1969). The Pollyanna hypothesis. *Journal of Verbal Learning and Verbal Behaviour, 8*, 1-8. https://doi.org/10.1016/s0022-5371(69)80002-2

Brysbaert, M., Stevens, M., De Deyne, S., Voorspoels, W., & Storms, G. (2014). Norms of age of acquisition and concreteness for 30,000 Dutch words. *Acta Psychologica, 150*, 80-84. https://doi.org/10.1016/j.actpsy.2014.04.010

Brysbaert, M., Stevens, M., Mandera, P., & Keuleers, E. (2016). The impact of word prevalence on lexical decision times: Evidence from the Dutch Lexicon Project 2. *Journal of Experimental Psychology: Human Perception & Performance, 42*, 441-458. https://doi.org/10.1037/xhp0000159

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

Citron, F. M. M., Weekes, B. S., & Ferstl, E. C. (2014). How are affective word ratings related to lexicosemantic properties? Evidence from the Sussex Affective Word List. *Applied Psycholinguistics, 35*, 313-331. https://doi.org/10.1017/s0142716412000409

Clark, J. M., & Paivio, A. (2004). Extensions of the Paivio, Yuille, and Madigan (1968) norms. *Behavior Research Methods, Instruments, & Computers, 36*, 371-383. https://doi.org/10.3758/bf03195584

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*, 407–428. https://doi.org/10.1037//0033-295x.82.6.407

Connell, L., & Lynott, D. (2012). Strength of perceptual experience predicts word processing performance better than concreteness or imageability. *Cognition, 125*, 452-465. https://doi.org/10.1016/j.cognition.2012.07.010

Cop, U., Dirix, N., Drieghe, D., & Duyck, W. (2017). Presenting GECO: An eyetracking corpus of monolingual and bilingual sentence reading. *Behavior Research Methods, 49*, 602-615. https://doi.org/10.3758/s13428-016-0734-0

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems,* 1695, 1-9*.*

De Deyne, S., Navarro, D., Collell, G., & Perfors, A. (2018, November 28). *Visual and affective grounding in language and mind.* https://doi.org/10.31234/osf.io/q97f8

De Deyne, S., Navarro, D. J., Perfors, A., Brysbaert, M., & Storms, G. (2018). The "Small World of Words" English word association norms for over 12,000 cue words. *Behavior Research Methods.* https://doi.org/10.3758/s13428-018-1115-7

De Deyne, S., Navarro, D. J., Perfors, A., & Storms, G. (2016). Structure at every scale: A semantic network account of the similarities between unrelated concepts. *Journal of Experimental Psychology: General, 145*, 1228-1254. https://doi.org/10.1037/xge0000192

De Deyne, S., Navarro, D. J., & Storms, G. (2013). Better explanations of lexical and semantic cognition using networks derived from continued rather than single word associations. *Behavior Research Methods, 45*, 480–498. https://doi.org/10.3758/s13428-012-0260-7

De Deyne, S., & Storms, G. (2008a). Word associations: Norms for 1,424 Dutch words in a continuous task. *Behavior Research Methods, 40*, 198-205. https://doi.org/10.3758/brm.40.1.198

De Deyne, S., & Storms, G. (2008b). Word associations: Network and semantic properties. *Behavior Research Methods, 40*, 213-231. https://doi.org/10.3758/brm.40.1.213

De Deyne, S., Verheyen, S., Ameel, E., Vanpaemel, W., Dry, M. J., Voorspoels, W., & Storms, G. (2008). Exemplar by feature applicability matrices and other Dutch normative data for semantic concepts. *Behavior Research Methods, 40*, 1030-1048. https://doi.org/10.3758/brm.40.4.1030

De Deyne, S., Verheyen, S., & Storms, G. (2015). The role of corpus size and syntax in deriving lexico-semantic representations for a wide range of concepts. *The Quarterly Journal of Experimental Psychology, 68*, 1643-1664. https://doi.org/10.1080/17470218.2014.994098

De Deyne, S., Verheyen, S., & Storms, G. (2016). Structure and organization of the mental lexicon: A network approach derived from syntactic dependency relations and word associations. In A. Mehler, A. Lücking, S. Banish, P. Blanchard, & B. Frank-Job (Eds.), *Towards a theoretical framework for analyzing complex linguistic networks* (pp. 47-79). Berlin, Germany: Springer. https://doi.org/10.1007/978-3-662-47238-5_3

De Deyne, S., Voorspoels, W., Verheyen, S., Navarro, D. J., & Storms, G. (2014). Accounting for graded structure in adjective categories with valence-based opposition relationships. *Language and Cognitive Processes*, *29*, 568-583. https://doi.org/10.1080/01690965.2013.794294

Deese, J. (1966). *The structure of associations in language and thought*. Baltimore: Johns Hopkins Press.

Dirven, R., & Taylor, J. (1988). The conceptualisation of vertical space in English: The case of tall. In B. Rudzka-Ostyn (Ed.), *Topics in Cognitive Linguistics* (pp. 379-402). Amsterdam, The Netherlands: John Benjamins. https://doi.org/10.1075/cilt.50.15dir

Dixon, R. (1982). *Where have all the adjectives gone?* The Hague, The Netherlands: Mouton. https://doi.org/10.1515/9783110822939

Dodds, P. S., Clark, E. M., Desu, S., Frank, M. R., Reagan, A. J., Williams, J. R., et al. (2015). Human language reveals a universal positivity bias. *PNAS, 112*, 2389-2394. https://doi.org/10.1073/pnas.1411678112

Dolan, R. J. (2002). Emotion, cognition, and behavior. *Science, 298*, 1191-1194. https://doi.org/10.1126/science.1076358

Estes, Z., & Adelman, J. S. (2008a). Automatic vigilance for negative words in lexical decision and naming: Comment on Larsen, Mercer, and Balota (2006). *Emotion, 8*, 441-444. https://doi.org/10.1037/1528-3542.8.4.441

Estes, Z., & Adelman, J. S. (2008b). Automatic vigilance for negative words is categorical and general. *Emotion, 8*, 453–457. https://doi.org/10.1037/a0012887

Gärdenfors, P. (2000). *Conceptual spaces: The geometry of thought*. Cambridge, MA: The MIT Press. https://doi.org/10.7551/mitpress/2076.001.0001

Ghyselinck, M., Custers, R., & Brysbaert, M. (2003). Age-of-acquisition rations for 2332 Dutch words from 49 different semantic categories. *Psychologica Belgica, 43*, 181–214.

Ghyselinck, M., De Moor, W., & Brysbaert, M. (2000). Age-of-acquisition ratings for 2,816

Dutch four- and five-letter nouns. *Psychologica Belgica, 40*, 77–98.

Gilet, A. L., & Jallais, C. (2011). Valence, arousal, and word associations. *Cognition & Emotion,*

25, 740-746. https://doi.org/10.1080/02699931.2010.500480

Gilhooly, K. J., & Logie, R. H. (1980). Age-of-acquisition, imagery, concreteness, familiarity,

and ambiguity measures for 1,944 words. *Behavior Research Methods & Instrumentation,*

12, 395-427. https://doi.org/10.3758/bf03201693

Gotzner, N., Solt, S., & Benz, A. (2018). Scalar diversity, negative strengthening, and

adjectival semantics. *Frontiers in Psychology, 9*: 1659.

https://doi.org/10.3389/fpsyg.2018.01659

Grühn, D., & Smith, J. (2008). Characteristics for 200 words rated by young and older adults:

Age-dependent evaluations of German adjectives (AGE). *Behavior Research Methods, 40*,

1088-1097. https://doi.org/10.3758/brm.40.4.1088

Hampton, J. A. (2011). Concepts and Natural Language. In R. Belohlavek, & G. J. Klir (Eds.),

Concepts and Fuzzy Logic (pp. 233-258). Cambridge, MA: The MIT Press.

https://doi.org/10.7551/mitpress/8842.003.0011

Harrell, F. E. (2001). *Regression modeling strategies*. New York, NY: Springer-Verlag.

https://doi.org/10.1007/978-1-4757-3462-1

Herbert, C., Kissler, J., Junghofer, M., Peyk, P., & Rockstroh, B. (2006). Processing of

emotional adjectives: Evidence from startle EMG and ERPs. *Psychophysiology, 43*, 197-206.

https://doi.org/10.1111/j.1469-8986.2006.00385.x

Hermans, D., & De Houwer, J. (1994). Affective and subjective familiarity ratings of 740 Dutch words. *Psychologica Belgica, 34*, 115-139.

Heyman, T., Van Akeren, L., Hutchison, K. A., & Storms, G. (2016). Filling the gaps: A speeded word fragment completion megastudy. *Behavior Research Methods, 48*, 1508-1527. https://doi.org/10.3758/s13428-015-0663-3

Hills, T., Maouene, J., Riordan, B., & Smith, L. (2010). The associative structure of language: Contextual diversity in early word learning. *Journal of Memory and Language, 63*, 259–273. https://doi.org/10.1016/j.jml.2010.06.002

Hutchison, K. A. (2003). Is semantic priming due to association strength or feature overlap? A microanalytic review. *Psychonomic Bulletin & Review, 10*, 785-813. https://doi.org/10.3758/bf03196544

Johnson, R. C., & Lim, D. (1964). Personality variables in associative production. *Journal of General Psychology, 71*, 349–350.

Kennedy, C. (2007). Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy, 30*, 1-45. https://doi.org/10.1007/s10988-006-9008-0

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new frequency measure for Dutch words based on film subtitles. *Behavior Research Methods, 42*, 643-650. https://doi.org/10.3758/brm.42.3.643

Keuleers, E., Diependaele, K., & Brysbaert, M. (2010). Practice effects in large-scale visual word recognition studies: A lexical decision study on 14,000 Dutch mono- and disyllabicwords and nonwords. *Frontiers in Psychology, 1*, 1-15.

https://doi.org/10.3389/fpsyg.2010.00174

Keuleers, E., Stevens, M., Mandera, P., & Brysbaert, M. (2015). Word knowledge in the crowd: Measuring vocabulary size and word prevalence in a massive online experiment. *The Quarterly Journal of Experimental Psychology, 68*, 1665-1692. https://doi.org/10.1080/17470218.2015.1022560

Koen, F. (1962). Polarization, *m*, and emotionality in words. *Journal of Verbal Learning and Verbal Behavior, 1*, 183-187. https://doi.org/10.1016/s0022-5371(62)80027-9

Kousta, S.-T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The representation of abstract words: Why emotion matters. *Journal of Experimental Psychology: General, 140*, 14-34. https://doi.org/10.1037/a0021446

Kousta, S.-T., Vinson, D. P., & Vigliocco, G. (2009). Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition, 112*, 473-481. https://doi.org/10.1016/j.cognition.2009.06.007

Kuperman, V., Estes, Z., Brysbaert, M., & Warriner, A. B. (2014). Emotion and language: Valence and arousal affect word recognition. *Journal of Experimental Psychology: General*, *143*, 1065-1081. https://doi.org/10.1037/a0035669

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods, 44*, 978-990. https://doi.org/10.3758/s13428-012-0210-4

Langacker, R. W. (1987). *Foundations of cognitive grammar*. *Vol.1*: *Theoretical prerequisites*. Stanford, CA: Stanford University Press.

Larsen, R. J., Mercer, K. A., & Balota, D. A. (2006). Lexical characteristics of words used in emotional Stroop experiments. *Emotion, 6*, 62-72. https://doi.org/10.1037/1528-3542.6.1.62

Larsen, R. J., Mercer, K. A., Balota, D. A., & Strube, M. J. (2008). Not all negative words slow down lexical decision and naming speed: Importance of word arousal. *Emotion, 8*, 445-452. https://doi.org/10.1037/1528-3542.8.4.445

Lenci, A., Lebani, G. E., & Passaro, L. C. (2018). The emotions of abstract words: A distributional semantic analysis. *Topics in Cognitive Science, 10*, 550-572. https://doi.org/10.1111/tops.12335

Louwerse, M. M. (2011). Symbol interdependency in symbolic and embodied cognition. *Topics in Cognitive Science, 3*, 273–302. https://doi.org/10.1111/j.1756-8765.2010.01106.x

Lynott, D., & Connell, L. (2009). Modality exclusivity norms for 423 object properties. *Behavior Research Methods, 41*, 558-564. https://doi.org/10.3758/brm.41.2.558

Lynott, D., & Connell, L. (2013). Modality exclusivity norms for 400 nouns: The relationship between perceptual experience and surface word form. *Behavior Research Methods, 45*, 516-526. https://doi.org/10.3758/s13428-012-0267-0

Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion : A review. *Cognition and Emotion, 23*, 209-237. https://doi.org/10.1080/02699930802204677

Medin, D. L., Lynch, E.B., & Solomon, K. O. (2000). Are there kinds of concepts? *Annual Review of Psychology, 51*, 121-147. https://doi.org/10.1146/annurev.psych.51.1.121

Moffat, M., Siakaluk, P. D., Sidhu, D. M., & Pexman, P. M. (2015). Situated conceptualization and semantic processing: effects of emotional experience and context availability in

semantic categorization and naming tasks. *Psychonomic Bulletin & Review, 22*, 408-419.

https://doi.org/10.3758/s13423-014-0696-0

Moors, A., De Houwer, J., Hermans, D., Wanmaker, S., van Schie, K., Van Harmelen, A., et al.

(2013). Norms of valence, arousal, dominance, and age of acquisition for 4300 Dutch words.

Behavior Research Methods, 45, 169–177. https://doi.org/10.3758/s13428-012-0243-8

Nelson, D. L., & McEvoy, C. L. (2000). What is this thing called frequency? *Memory &*

Cognition, 28, 509-522. https://doi.org/10.3758/bf03201241

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*.

Urbana, IL: University of Illinois Press.

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness

values for 925 nouns. *Journal of Experimental Psychology, 76*, 1-25.

https://doi.org/10.1037/h0025327

Pollio, H. (1964). Some semantic relations among word-associates. *The American Journal of*

Psychology, 77, 249–256. https://doi.org/10.2307/1420131

Quadflieg, S., Michel, C., Bukowski, H. & Samson, D. (2014). A database of psycholinguistic

and lexical properties for French adjectives referring to human and/or nonhuman attributes.

Canadian Journal of Experimental Psychology, 68, 67-76.

https://doi.org/10.1037/cep0000001

R Development Core Team (2007). *R: A language and environment for statistical computing*.

R Foundation for Statistical Computing, Vienna, Austria.

Raskin, V., & Nirenburg, S. (1998). An applied ontological semantic microtheory of adjective

meaning for natural language processing. *Machine Translation, 13*, 135-227.

https://doi.org/10.1023/A:1008039100776

Rodríguez-Ferreiro, J., & Davies, R. (2018). The graded effect of valence on word recognition

in Spanish. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

https://doi.org/10.1037%2Fxlm0000616

Ruytenbeek, N., Verheyen, S., & Spector, B. (2017). Asymmetric inference towards the

antonym: Experiments into the polarity and morphology of negated adjectives. *Glossa: A*

Journal of General Linguistics, 2(1): 92 1-27. https://doi.org/10.5334/gjgl.151

Samsonovich, A. V., & Ascoli, G. A. (2010). Principal semantic components of language and

the measurement of meaning. *PLoS ONE, 5*, e10921.

https://doi.org/10.1371/journal.pone.0010921

Santos, A., Chaigneau, S. E., Simmons, W. K., & Barsalou, L. W. (2011). Property generation

reflects word association and situated simulation. *Language and Cognition, 3*, 83-119.

https://doi.org/10.1515/langcog.2011.004

Spearman, C. (1904). The proof and measurement of association between two things. *The*

American Journal of Psychology, 15, 72–101. https://doi.org/10.2307/1412159

Staats, A. W., & Staats, C. K. (1959). Meaning and m: correlated but separate. *Psychological*

Review, 66, 136–44. https://doi.org/10.1037/h0047970

Steyvers, M., & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks:

statistical analyses and a model of semantic growth. *Cognitive Science, 29*, 41–78.

https://doi.org/10.1207/s15516709cog2901_3

Szalay, L. B., & Deese, J. (1978). *Subjective meaning and culture: An assessment through word associations*. Hillsdale, NJ: Lawrence Erlbaum.

Talmy, L. (2000). *Toward a cognitive semantics: Concept structuring systems.* Cambridge, MA: MIT Press. https://doi.org/10.7551/mitpress/6847.001.0001

Van Der Goten, K., De Vooght, G., & Kemps, E. (1999). Concreteness and emotional valence ratings of 399 Dutch nouns. *Psychologica Belgica, 39*, 49-70.

Vankrunkelsven, H., Verheyen, S., Storms, G., & De Deyne, S. (2018). Predicting lexical norms: A comparison between a word association model and text-based word co-occurrence models. *Journal of Cognition, 1*(1), 45. http://doi.org/10.5334/joc.50

Van Loon-Vervoorn, W. A. (1985). *Voorstelbaarheidswaarden van Nederlandse woorden* [Imageability ratings for Dutch words] Lisse, The Netherlands: Swets & Zeitlinger.

Van Rensbergen B., De Deyne S., Storms G. (2015). Estimating affective word covariates using word association data. *Behavior Research Methods, 48*, 1644-1652. https://doi.org/10.3758/s13428-015-0680-2

Van Rensbergen B., Storms G., De Deyne S. (2015). Examining assortativity in the mental lexicon: Evidence from word association data. *Psychonomic Bulletin & Review, 22*, 1717-1724. https://doi.org/10.3758/s13423-015-0832-5

Van Rooij, R. (2011). Vagueness and linguistics. In G. Ronzitti (Ed.), *Vagueness: A guide* (pp. 123-170). Dordrecht, The Netherlands: Springer. https://doi.org/10.1007/978-94-007-0375-9_6

Vigliocco, G., Meteyard, L., Andrews, M., & Kousta, S. (2009). Toward a theory of semantic representation. *Language and Cognition, 1*, 219–247. https://doi.org/10.1515/langcog.2009.011

Vinson, D., Ponari, M., & Vigliocco, G. (2014). How does emotional content affect lexical processing? *Cognition & Emotion, 28*, 737-746. https://doi.org/10.1080/02699931.2013.851068

Verheyen, S., & Egré, P. (2018). Typicality and graded membership in dimensional adjectives. *Cognitive Science, 42*, 2250-2286. https://doi.org/10.1111/cogs.12649

Võ, M. L.-H., Conrad, M., Kuchinke, L., Urton, K., Hofmann, M. J., & Jacobs, A. M. (2009). The Berlin affective word list reloaded (BAWL-R). *Behavior Research Methods, 41*, 534-538. https://doi.org/10.3758/brm.41.2.534

Vogel, A. (2004). *Swedish Dimensional Adjectives*. Stockholm, Sweden: Almqvist & Wiksell.

Wang, X., Wu, W., Ling, Z., Xu, Y., Fang, Y., Wang, X., Binder, J. R., Men, W., Gao, J.H., & Bi, Y. (2017). Organizational principles of abstract words in the human brain. *Cerebral Cortex, 28*, 4305–4318. https://doi.org/10.1093/cercor/bhx283

Yap, M. J., & Seow, C. S. (2014). The influence of emotion on lexical processing: Insights from RT distributional analysis. *Psychonomic Bulletin & Review, 21*, 526-533. https://doi.org/10.3758/s13423-013-0525-x

Appendix

Table A1.

Summary of the Regression Results for Linear and Non-Linear Model Predictions of ln(in-strength) (N=953). Adjusted R^2: .70.

Predictor	Coeff.	Std.Error	*t*	Pr(>|*t*|)	df	Partial SS	MS	*F*	*p*
Intercept	-1.198	.719	-1.67	.096					
valence	-.281	.067	-4.17	<.0001	2	15.448	7.724	8.73	.000
Nonlinear	.308	.085	3.63	.000	1	11.654	11.654	13.18	.000
arousal	-.159	.100	-1.59	.112	2	9.020	4.510	5.10	.006
Nonlinear	-.007	.111	-.06	.953	1	.003	.003	.00	.953
dominance	-.031	.111	-.28	.779	2	7.354	3.677	4.16	.016
Nonlinear	.198	.120	1.66	.098	1	2.434	2.434	2.75	.098
concreteness	.144	.105	1.37	.172	2	5.475	2.738	3.10	.046
Nonlinear	-.275	.129	-2.14	.033	1	4.040	4.040	4.57	.033
imageability	.172	.072	2.39	.017	2	13.165	6.582	7.44	.001
Nonlinear	-.047	.088	-.53	.599	1	.245	.245	.28	.599
AoA	-.175	.044	-3.99	<.0001	2	17.076	8.538	9.66	.000
Nonlinear	.096	.041	2.37	.018	1	4.950	4.950	5.60	.018
familiarity	.471	.073	6.41	<.0001	2	170.085	85.042	96.17	<.0001
Nonlinear	.147	.083	1.77	.077	1	2.776	2.776	3.14	.077
ln(subtlex_freq)	.332	.031	10.85	<.0001	2	156.832	78.416	88.68	<.0001
Nonlinear	-.172	.041	-4.23	<.0001	1	15.822	15.822	17.89	<.0001
nchar	-.244	.055	-4.45	<.0001	2	31.662	15.831	17.90	<.0001
Nonlinear	.154	.055	2.80	.005	1	6.918	6.918	7.82	.005

sqrt(neighb)	.074	.081	.92	.360	2	1.012	.506	.57	.565
Nonlinear	-.134	.126	-1.07	.286	1	1.010	1.010	1.14	.286
meanbigram	.000	.000	1.22	.221	2	1.648	.824	.93	.394
Nonlinear	.000	.000	-.92	.360	1	0.743	.743	.84	.360
Total nonlinear					11	47.691	4.336	4.90	<.0001
Regression					22	2006.166	91.189	103.12	<.0001
Error					930	822.390	.884		

Table A2.

Summary of the Regression Results for Linear and Non-Linear Model Predictions of

ln(betweenness) (N=953). Adjusted R^2: .65.

Predictor	Coeff.	Std.Error	*t*	Pr(>|*t*|)	df	Partial SS	MS	*F*	*p*
Intercept	-10.162	.759	-13.40	<.0001					
valence	-.206	.071	-2.90	.004	2	11.795	5.898	6.00	.003
Nonlinear	.160	.089	1.79	.074	1	3.149	3.149	3.20	.074
arousal	-.161	.105	-1.53	.127	2	9.065	4.532	4.61	.010
Nonlinear	-.005	.117	-.04	.969	1	.001	.001	.00	.969
dominance	.026	.117	.22	.822	2	5.773	2.887	2.94	.054
Nonlinear	.125	.126	.99	.321	1	.970	.970	.99	.321
concreteness	.086	.111	.77	.439	2	5.851	2.925	2.97	.052
Nonlinear	-.232	.136	-1.71	.089	1	2.860	2.860	2.91	.089
imageability	.170	.076	2.24	.025	2	7.059	3.530	3.59	.028
Nonlinear	-.111	.093	-1.19	.234	1	1.396	1.396	1.42	.234
AoA	-.142	.046	-3.09	.002	2	12.730	6.365	6.47	.002
Nonlinear	.069	.043	1.60	.110	1	2.512	2.512	2.55	.110
familiarity	.502	.077	6.49	<.0001	2	149.012	74.506	75.75	<.0001
Nonlinear	.058	.088	.66	.512	1	.423	.423	.43	.512
ln(subtlex_freq)	.321	.032	9.94	<.0001	2	141.360	70.680	71.86	<.0001
Nonlinear	-.176	.043	-4.09	<.0001	1	16.491	16.491	16.77	<.0001
nchar	-.262	.058	-4.54	<.0001	2	40.245	20.122	20.46	<.0001
Nonlinear	.157	.058	2.70	.007	1	7.191	7.191	7.31	.007
sqrt(neighb)	.084	.085	.98	.326	2	.970	.485	.49	.611
Nonlinear	-.097	.132	-.73	.463	1	.531	.531	.54	.463

meanbigram	.000	.000	1.81	.071	2	3.355	1.678	1.71	.182
Nonlinear	.000	.000	-1.55	.122	1	2.363	2.363	2.40	.122
Total nonlinear					11	36.693	3.336	3.39	.000
Regression					22	1759.545	79.979	81.32	<.0001
Error					930	914.685	.984		