Cognitive Processing
https://doi.org/10.1007/s10339-019-00933-y

RESEARCH ARTICLE

Mapping semantic space: property norms and semantic richness

Emiko J. Muraki[1] · David M. Sidhu[1] · Penny M. Pexman[1]

Received: 28 May 2019 / Accepted: 10 September 2019

Abstract

In semantic property listing tasks, participants list many features for some concepts and fewer for others. This variability in *number of features* (NoF) has been used in previous research as a measure of a concept's semantic richness, and such studies have shown that in lexical-semantic tasks responses tend to be facilitated for words with high NoF compared to those for words with low NoF, even when many other relevant factors are controlled (Pexman et al. in Psychon Bull Rev 9:542–549, 2002; Mem Cogn 31:842–855, 2003; Psychon Bull Rev 15:161–167, 2008; Goh et al. in Front Psychol, 2016. https://doi.org/10.3389/fpsyg.2016.00976). Furthermore, shared features (i.e., features that are shared by multiple words) appear to facilitate responses in lexical-semantic tasks to a greater degree than distinctive features (Devereux et al. in Cogn Sci 40:325–350, 2016; Grondin et al. in J Mem Lang 60:1–19, 2009). This previous work was limited, however, to relatively small sets of words, typically those extracted from the McRae norms (McRae et al. in Behav Res Methods 37(4):547–559, 2005). New property listing norms provide the opportunity to extract NoF values for many more items (Buchanan et al. in Behav Res Methods 51:1849–1863, 2019). The purpose of the present study was to test whether NoF effects generalize to this larger item set, and to explore how NoF is related to other measures of semantic richness, including subjective ratings of concreteness, imageability, body–object interaction, sensory experience, valence, arousal, and age of acquisition, as well as more objective measures like semantic diversity, number of associates, and lexical centrality. Using the new Buchanan norms, we found significant NoF effects in lexical decision (*is it a word or a nonword*?) and semantic decision (*is it concrete or abstract*?) tasks. We also found significant effects of words' number of shared (less distinctive) features in each task. Further, factor analyses of all semantic richness measures showed a distinct factor structure, suggesting that there are clusters of semantic richness dimensions that seem to correspond to more embodied semantic dimensions and more distributional semantic dimensions. Results are interpreted as evidence that semantic representation is multimodal and multidimensional, and provide new insights about the structure of semantic space.

Keywords Semantic decision · Lexical decision · Word recognition · Semantic richness

Introduction

Semantic property listing tasks have been used to explore a number of important questions about semantic representation. For instance, the semantic property norms collected by McRae and colleagues (e.g., McRae et al. 2005) were used to conduct some of the early work on *semantic richness effects* and helped to provide evidence for distributed accounts of semantic representation (Pexman et al. 2003; Pexman et al. 2002). Semantic richness refers to the phenomenon whereby words associated with relatively more semantic information tend to be responded to faster and/or more accurately in lexical and semantic tasks (for a review, see Pexman 2012). In the context of property listing, Pexman et al. (2002, 2003) reported that visual lexical

Handling editor: Barry Devereux (Queen's University Belfast).

Reviewers: Blair Armstrong (University of Toronto), Gabriel Recchia (University of Cambridge).

This article is part of the special topic 'Eliciting Semantic Properties: Methods and Applications'.

✉ Emiko J. Muraki
ejmuraki@ucalgary.ca

[1] University of Calgary, Calgary, Canada

decision and semantic decision responses were facilitated for words that had high numbers of features (NoF), compared to responses for words with low numbers of features. The explanation for these NoF effects was that words with many features generate more semantic activation than words with fewer features. Greater semantic activation then facilitates lexical decision responses via stronger feedback from the semantic units to orthographic units (Hino and Lupker 1996) and facilitates semantic decision responses via faster semantic settling (Pexman et al. 2003).

In more recent years, researchers have used the McRae norms to test for NoF effects in several different tasks, with results showing facilitory NOF effects in sentence reading (Cook et al. 2013; Pexman et al. 2003), auditory lexical decision (Devereux et al. 2016; Goh et al. 2016; Sajin and Connine 2014), auditory semantic decision (Goh et al. 2016), free recall (Hargreaves et al. 2012), picture naming (Taylor et al. 2012), and semantic decisions for picture stimuli (Taylor et al. 2012). These findings suggest that, at least when assessed with the McRae items, semantic richness effects are pervasive, and features capture an important dimension of semantic representation. These inferences are tempered somewhat by the fact that attempts to examine NoF effects with stimuli beyond those in the McRae norms have produced mixed results. While the McRae norms include only concrete words, Recchia and Jones (2012) collected property listing responses for both concrete and abstract words and found that in lexical decision and naming tasks NoF effects were only significant for concrete words. Using a set of French property listing norms, Robert and Rico Duarte (2016) found NoF effects in visual lexical decision for young adults but not for older adults.

While many of the previous studies examined effects of words' sheer numbers of features, several others have investigated whether some types of features might be particularly influential. In so doing, they have tested additional claims about the nature of semantic representation. For instance, Grondin et al. (2009) replicated the NoF effect in lexical and semantic decision tasks, using the McRae et al. norms, and argued, further, that the effect was largely driven by shared features. Grondin et al. defined shared features as those that were listed for more than two concepts in the norms (e.g., *has legs*), while distinctive features are those that were listed for two or fewer concepts in the norms (e.g., *moos*). Responses in lexical and semantic decision tasks were faster for words with many shared features than for words with few shared features; words with many distinctive features also enjoyed a processing advantage over words with few distinctive features, but this benefit was smaller than that for shared features. Grondin et al. argued that the number of shared features (NoSF) facilitated lexical-semantic processing because shared features are more readily activated; they are more familiar and are better cues to lexical or semantic status.

While Grondin et al. (2009) defined the distinctiveness of features categorically, others have done so continuously. Devereux et al. (2016) calculated the distinctiveness of a feature by taking the multiplicative inverse of the number of concepts for which it appears (e.g., a feature appearing for three concepts would have a distinctiveness of 1/3, or 0.33). They then calculated the average distinctiveness of features (DoF) for all features associated with a given concept. Their results were consistent with Grondin et al.: concepts with a lower DoF (i.e., having more features shared across multiple concepts) were responded to more quickly in an auditory lexical decision task. However, it is unclear whether the benefit observed for shared features (whether measured categorically or continuously) extends to larger, more heterogeneous word sets than the 500 or so concrete items in the McRae norms, which all belong to particular well-defined categories (e.g., tools, musical instruments).

Findings that feature statistics like NoSF and DoF influence lexical-semantic processing have been taken as evidence for feature-based accounts of conceptual knowledge (e.g., Taylor et al. 2012). In contrast, other findings have challenged the notion that any single dimension of representation will provide a comprehensive account of lexical-semantic knowledge. In recent years, many new dimensions of semantic richness have been identified, and several studies have shown that there can be simultaneous effects of several different richness dimensions in the same task context. For instance, Pexman et al. (2008) examined the joint effects of three richness dimensions—NoF, number of semantic neighbors (the number of words used in similar lexical contexts, Durda et al. 2006), and contextual dispersion (the number of content areas in which a word occurs, Zeno et al. 1995)—in lexical decision and semantic decision tasks. Results showed that all three richness dimensions had significant facilitory effects in lexical decision, and that both NoF and contextual dispersion had significant facilitory effects in the semantic decision task. Thus, Pexman et al. concluded that there are multiple dimensions of semantic richness that may each tap different aspects of semantic knowledge, and their influence varies as a function of task demands.

Similarly, Yap et al. (2012) showed that NoF effects were observed in lexical decision, naming, progressive demasking and semantic decision tasks and were found simultaneously with effects of other richness dimensions, including the sensorimotor dimensions of imageability (Cortese and Fugett 2004) and body–object interaction (BOI; Siakaluk et al. 2008). Finally, in semantic decision tasks involving words and, separately, picture stimuli, NoF effects were observed alongside effects of other richness dimensions, including survival-based (Amsel et al. 2012), sensory, and motor dimensions (Taikh et al. 2015). Thus, a number of studies suggest that semantic representation may be multidimensional, and many theories can accommodate some degree of multidimensionality (Andrews et al. 2009; Barsalou et al. 2008; Borghi and Cimatti 2010; Buchanan et al. 2001; Dove 2009; Louwerse 2010; Paivio 1971; Patterson et al. 2007). The previous studies that have examined the simultaneous effects of multiple semantic richness dimensions have, however, largely used the same set of concrete items from the McRae norms; the structure of semantic space beyond this item set has not been tested.

Recently, a new set of property listing norms was developed by Buchanan and colleagues for a larger, more varied set of over 4400 words (Buchanan et al. 2019). These new norms presented us with the opportunity to address two main research questions. The first question was whether NoF effects generalize beyond the McRae et al. word set. We used NoF values derived from the new Buchanan norms to examine NoF effects in lexical and semantic tasks. The second question was how NoF effects are related to other semantic richness effects. As noted, there is evidence that semantic space is multidimensional, and so we expected NoF to be only modestly related to other semantic richness dimensions. Further, based on the previous literature, there are several possibilities in terms of how NoF might cluster with other semantic richness effects, detailed next.

First, Hargreaves and Pexman (2014) distinguished language-based semantic richness effects from object-based semantic richness effects. They classified as language-based the semantic richness effects, like the number of semantic neighbors, that are derived from the statistics of language use. In contrast, they assumed that NoF, imageability, and BOI were all object-based richness effects since they capture attributes of words' actual object referents. Hargreaves and Pexman (2014) investigated the timecourse of semantic richness effects and found that in the semantic decision task language-based richness effects (average neighbor similarity, ANS; Shaoul and Westbury 2010) tended to emerge before object-based richness effects (NoF, imageability). Based on their findings, one would expect NoF to cluster with other object-based richness effects.

Second, predictions could be derived from the work of Santos et al. (2011). Santos et al. argued that property listing responses tap into two processes: word association (a linguistic process) and situated simulation (an imagery process). Thus, NoF may show similarity to linguistic-based richness effects *and* to sensory dimensions (e.g., imageability, sensory experience ratings; SER; Juhasz and Yap 2013) that capture simulation. In contrast, Barsalou (2003) and McRae et al. (2005) emphasized simulation as the basis for property listing responses. Hence, NoF may show more similarity to other semantic dimensions that depend on the processes of simulation and embodiment.

Third, the methods by which semantic richness dimensions are derived might influence their relationships. For instance, some dimensions, like NoF, are derived more objectively, from counts or other quantifications of the information on which they are based. These more objective dimensions include semantic diversity (SemD; the extent to which words appear in more diverse contexts; Hoffman et al. 2013), number of associates (NoA; the number of unique words generated as associates of a target word in a free association task; De Deyne et al. 2018), and lexical centrality (LexC; the frequency with which a given word appears as an associate to a target word in a free association task; De Deyne et al. 2018). Other dimensions are derived more subjectively, from Likert and other scale ratings. These subjective dimensions include BOI, SER, imageability, concreteness (Brysbaert et al. 2013), age of acquisition (AoA; estimated age at which word was acquired; Kuperman et al. 2012), valence (Warriner et al. 2013), and arousal (Warriner et al. 2013). If method of derivation matters, then the objectively derived measures may tend to cluster together and the subjectively derived measures may tend to cluster together.

Fourth, some semantic richness dimensions capture aspects of meaning that are more closely tied to our bodily experience of word referents and mental simulations engaged during word processing, including BOI, imageability, arousal, concreteness, and SER. If embodiment is an organizing principle for semantic space, then these dimensions may cluster together. Other semantic richness dimensions capture more distributional aspects of word meanings, like SemD (distribution across contexts) and AoA (distribution across time), and thus may be related to each other and may predict the same variance in lexical-semantic tasks. In the present study, we tested these possibilities by examining the factor structure of semantic richness and investigated how those factors are related to responses in lexical and semantic decision tasks.

Number of features as a predictor of word processing

Method

All analyses reported here were based on secondary data sources. Analyses were conducted using the statistical software R (R Core Team 2018). The effect of NoF on lexical-semantic task responses was investigated using feature production norms from an expanded feature production database produced by Buchanan et al. (2019), in which participants were asked to generate lists of features for a presented word. This new database includes feature norms for over 4400 words. For each cue word, we considered features with the same root as being instances of the same feature (e.g., *leave*, *leaving*, and *left* were all considered instances of the same feature). As in McRae et al. (2005), we only included features that were provided by at least 16% of the respondents. After these considerations, we counted the number of unique features provided for each word (NoF). As an illustration of the resulting features, the top three unique features for the word *apple* were: *fruit*, *red*, and *grow*; the top three features for the word *peace* were: *love*, *war*, and *calm*.

In addition, we calculated each feature's distinctiveness by taking the multiplicative inverse of the number of words for which it was listed (Devereux et al. 2016). We then obtained each word's distinctiveness of features (DoF) by averaging the distinctiveness of its features. Our predictors of interest were NoF and DoF; analyses also included length, frequency (log subtitle frequency; Brysbaert and New 2009) and orthographic Levenshtein distance (OLD; Yarkoni et al. 2008), to control for the standard lexical factors in lexical decision and semantic decision tasks. All predictors were standardized.

The dependent variables were obtained from three mega-studies: lexical decision response times and accuracy from the English Lexicon Project (*ELP*; Balota et al. 2007), response times and accuracy from the Calgary Semantic Decision Project (*CSDP*; Pexman et al. 2017), and response times and accuracy from the English Crowdsourcing Project (*ECP*; Mandera et al. 2019a). The full methods for each mega-study are provided in their respective papers, thus only brief descriptions are provided below. The ELP includes lexical decision (*is it a word or a nonword*?) response time and accuracy data for 40,841 words from 816 participants. The CSDP includes semantic decision (*is it concrete or abstract*?) response time and accuracy data for 10,000 words collected from 321 participants. Finally, the ECP includes response time data and accuracy from nearly 700,000 participants on word knowledge (*is this a word you know*?). We analyzed z-scored response times. Each analysis used the maximum number of available items in each mega-study dataset (e.g., an item not being present in the CSDP did not preclude its inclusion in the LDT analyses).

Results

ELP

Responses were significantly faster to words with a greater number of features ($b = -0.011$, $p < .001$) and to words with features that were less distinct (more shared; $b = 0.010$, $p < .001$). Responses were also significantly more accurate to words with a greater number of features ($b = 0.004$, $p < .001$). There was not a significant effect of feature distinctiveness on response accuracy ($b = -0.001$, $p = .40$) (Table 1).

CSDP

Responses were significantly faster to words with a greater number of features ($b = -0.154$, $p < .001$) and to words with features that were less distinct ($b = 0.034$, $p < .001$). Responses were also more accurate to words with a greater number of features ($b = 0.027$, $p < .001$). There was not a significant effect of feature distinctiveness on response accuracy ($b = -0.004$, $p = .21$) (Table 1).

ECP

Responses were significantly faster to words with a greater number of features ($b = -0.011$, $p = .002$) and to words with features that were less distinct ($b = 0.008$, $p = .02$). Responses were also more accurate to words with a greater number of features ($b = 0.001$, $p < .001$). There was not a significant effect of feature distinctiveness on response accuracy ($b = -0.000$, $p = .58$) (Table 1).

Table 1 Regression coefficients, standard errors, and semi-partial correlations for models predicting response time and accuracy in the ELP, CSDP, and ECP, using NoF and DoF

	ELP			CSDP			ECP		
	b	SE	sr^2	b	SE	sr^2	b	SE	sr^2
Response time									
NoF	−0.011	0.003	0.002***	−0.154	0.010	0.15***	−0.011	0.004	0.002**
DoF	0.010	0.003	0.002***	0.034	0.009	0.008***	0.008	0.004	0.001*
Accuracy									
NoF	0.004	0.001	0.003***	0.027	0.003	0.049***	0.001	0.000	0.005***
DoF	−0.001	0.001	0.000	−0.004	0.003	0.001	−0.000	0.000	0.000

Note that models also included length, frequency, and orthographic Levenshtein distance

We used the "SDSRegressionR" package in R (Mahometa 2018) to generate sr^2 values

ELP English Lexicon Project, *CSDP* Calgary Semantic Decision Project, *ECP* English Crowdsourcing Project, *NoF* number of features, *DoF* distinctiveness of features

*$p < .05$; **$p < .01$; ***$p < .001$

Factor analysis of semantic richness variables

Method

Lexical and semantic variables were obtained for the purposes of conducting an exploratory factor analysis. Table 2 presents the number of shared items among the three mega-study datasets from the analyses above (ELP, CSDP, and ECP) and the lexical and semantic variables. The lexical variables obtained for these analyses included letter length, OLD (Yarkoni et al. 2008), frequency (log subtitle frequency; Brysbaert and New 2009), and AoA (Kuperman et al. 2012). The semantic richness variables included NoF (Buchanan et al. 2019), imageability (Cortese and Fugett 2004; Schock et al. 2012), concreteness (Brysbaert et al. 2013), BOI (Pexman et al. 2019), SER (Juhasz and Yap 2013), ANS (Shaoul and Westbury 2010) LexC and NoA (De Deyne et al. 2018), SemD (Hoffman et al. 2013), valence (Warriner et al. 2013), and arousal (Warriner et al. 2013). NoF was extracted using the same process as described in the methods for the analyses above. NoA was extracted as the number of associates for a given cue word that were produced by at least two participants in De Deyne et al. as associates generated by only one individual can be idiosyncratic and less reliable (Nelson et al. 2004; Nelson and Schreiber 1992). As may be expected, the lexical and semantic variables show a high degree of correlation among one another, with only fifteen variable pairings having non-significant correlations, or a p value > .001, as presented in Fig. 1.

Table 2 Number of words available by mega-study dataset

Variable	ELP	CSDP	ECP
NoF	4012	957	4048
Length	21,256	4455	21,241
OLD	21,256	4455	21,241
Frequency	21,159	4455	21,146
AoA	21,255	4709	30,793
Imageability	4555	1050	4570
Concreteness	18,846	4709	23,858
BOI	8397	4709	8662
SER	5547	1031	5678
ANS	21,248	4617	23,643
LexC	19,076	4121	23,325
NoA	6814	1544	7095
SemD	14,895	2823	16,133
Valence	12,660	2980	13,669
Arousal	12,660	2980	13,669

ELP English Lexicon Project, *CSDP* Calgary Semantic Decision Project, *ECP* English Crowdsourcing Project, *NoF* number of features, *OLD* orthographic Levenshtein distance, *AoA* age of acquisition, *BOI* body–object interaction, *SER* sensory experience rating, *ANS* average neighborhood similarity, *LexC* lexical centrality, *NoA* number of associates, *SemD* semantic distance

Results

Exploratory factor analysis

An exploratory factor analysis was conducted to determine whether latent constructs among the 15 lexical and semantic variables could be identified. In total 1520 words had values for each of the 15 lexical and semantic variables and were included in the analysis. Bartlett's test of sphericity was significant, $\chi^2(105) = 10{,}715.87$, $p < .001$, and the Kaiser–Meyer–Olkin measure of sampling adequacy was 0.76, indicating the correlations observed among the data are adequate for a factor analysis using all items and variables. A scree plot (Fig. 2) of the observed data and a parallel analysis indicated that it was appropriate to extract a five-factor solution. The solution was extracted using principal axis factoring due to the variables being not normally distributed (Costello and Osborne 2005). Given the high degree of correlation among the variables entered into the factor analysis, and that the extracted factors were correlated at greater than $r = 0.35$, an oblimin oblique rotation was applied rather than an orthogonal rotation, in order to accurately represent the relationships among the factors (Costello and Osborne 2005). The final solution yielded five factors that accounted for 57.2% of the total variance, with the first two factors accounting for 34.7% of the variance alone. The structure coefficients for the factors are presented in Table 3.

Overall the analyses indicated five distinct latent constructs that explain variance among the lexical and semantic variables. Factor 1 appears to represent a visuomotor richness construct, with high loadings from variables that relate to bodily experience such as concreteness, BOI, and imageability. Notably, NoF loaded onto this factor as well. Factor 2 represents a distributional construct related to word meaning and associations, with high loadings from frequency, ANS, SemD, lexical centrality, NoA, and a negative loading from AoA. Factor 3 represents a purely orthographic construct, with high loadings from length and OLD. Factor 4 represents a sensory richness construct, comprised of high loadings from SER and arousal and was correlated with the Factor 1 visuomotor richness construct ($r = 0.40$), likely due to the overlap in sensory experience reflected in SER ratings and ratings such as concreteness, BOI and imageability. Finally, Factor 5 appeared to exclusively represent a valence construct, as its sole variable that loaded onto it was valence. The relationships between visuomotor richness, distributional, sensory richness, and valence factor scores are depicted in Fig. 3. Consistent

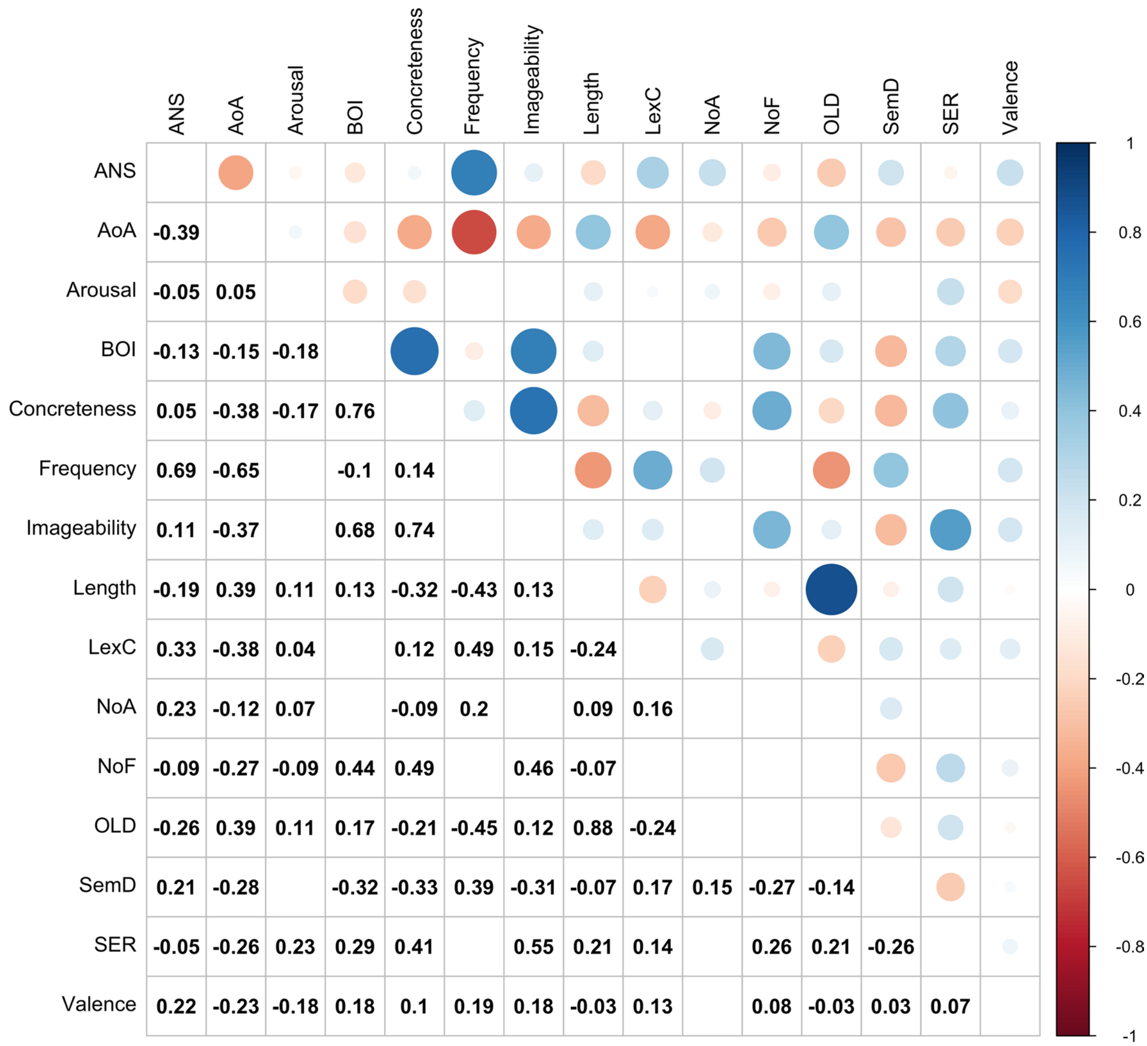

Fig. 1 Correlation matrix of the relationships between the lexical and semantic variables used in the factor analysis. The strengths and directions of coefficients are indicated by size and color of circles above the diagonal, and correlation coefficients are reported below the diagonal. Correlations with a *p* value greater than .001 have been suppressed. *ANS* average neighborhood similarity, *AoA* age of acquisition, *BOI* body–object interaction, *LexC* lexical centrality, *NoA* number of associates, *NoF* number of features, *OLD* orthographic Levenshtein distance, *SemD* semantic distance, *SER* sensory experience rating

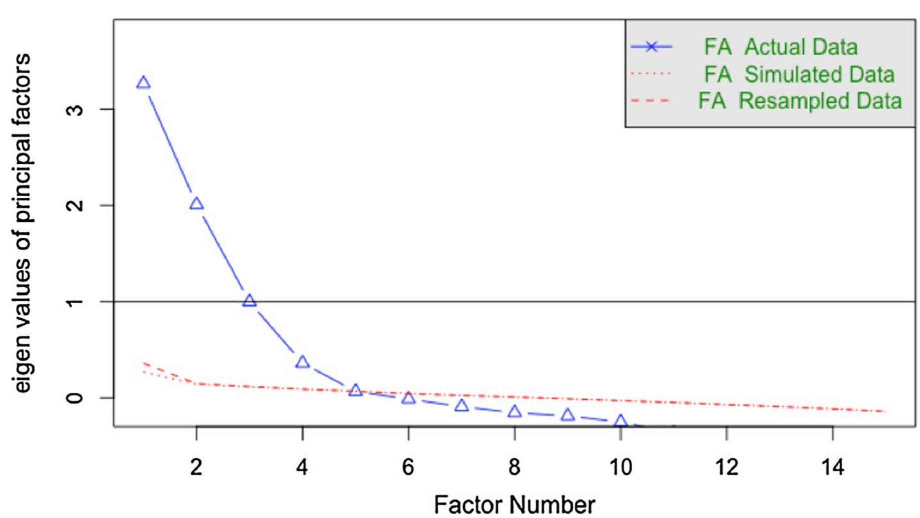

Fig. 2 Parallel analysis scree plot showing eigenvalues by number of factors based on actual lexical and semantic variable data, simulated data, and resampled data

Table 3 Factor loadings and communalities based on principal axis factoring with oblimin rotation for 15 word properties ($n = 1520$)

Property	Visuomotor richness	Distributional	Orthography	Sensory richness	Valence	Communality
Concreteness	0.924					0.870
Imageability	0.846					0.836
BOI	0.886					0.736
NoF	0.454					0.260
Frequency		0.918				0.884
ANS		0.716				0.539
SemD		0.573				0.539
LexC		0.528				0.390
AoA		−0.488				0.459
NoA		0.410				0.146
Length			0.962			0.908
OLD			0.848			0.768
SER				0.677		0.629
Arousal				0.465		0.320
Valence					0.737	0.541

Factor loadings greater than 0.3 are reported

OLD orthographic Levenshtein distance, *AoA* age of acquisition, *BOI* body–object interaction, *SER* sensory experience rating, *ANS* average neighborhood similarity, *LexC* lexical centrality, *NoA* number of associates, *SemD* semantic distance

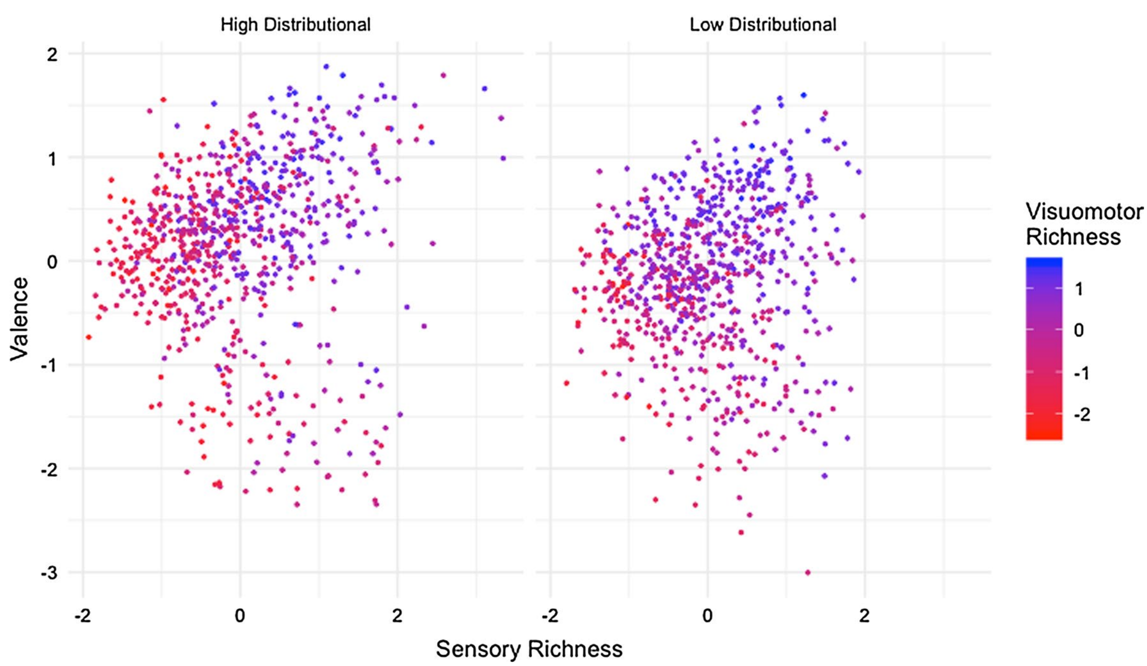

Fig. 3 Scatterplot of sensory richness and valence factor scores, as a function of visuomotor richness scores, plotted with a median split of distributional factor scores ($n = 1520$). A 3D interactive version of this chart is available at: https://plot.ly/~ejmuraki/1/

across high and low distribution words, high factor scores for visuomotor richness cluster with higher scores for sensory richness and valence. An interactive, three-dimensional mapping of this semantic space is available at: https://plot.ly/~ejmuraki/1/.

Factor scores for each of the 1520 words were extracted and entered into regression analyses to examine their relationships to lexical and semantic task response times and accuracy from the ELP, CSDP, and ECP. Descriptive statistics for the factor scores are presented in Table 4.

Kernel density plots show the distribution of the factor scores in Fig. 4. All factors show a normal distribution.

Regression analysis

Each analysis once again used the maximum number of available items in each mega-study dataset. The results of the regression analyses are presented in Table 5.

Table 4 Descriptive statistics of factor scores ($n = 1520$)

Descriptive statistic	Visuomotor richness	Distributional	Orthography	Sensory richness	Valence
Mean	0	0	0	0	0
Median	0.128	−0.030	−0.087	−0.066	0.079
SD	0.968	0.958	0.965	0.840	0.786
Skewness	−0.383	0.341	0.436	0.350	−0.604
Kurtosis	−0.964	−0.034	−0.114	−0.110	0.336

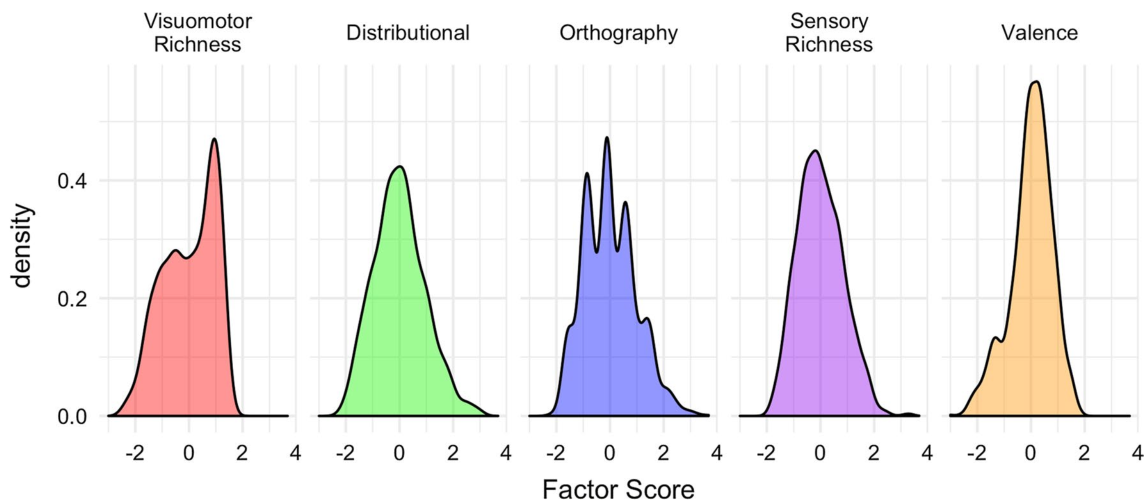

Fig. 4 Kernel density plots depicting distribution of latent factor scores ($n = 1520$)

Table 5 Regression coefficients, standard errors, and semi-partial correlations for ELP, CSDP, and ECP regression models

Factor	ELP ($n = 1519$)			CSDP ($n = 343$)			ECP ($n = 1517$)		
	b	SE	sr²	*b*	SE	sr²	*b*	SE	sr²
Response time									
Visuomotor richness	−0.014	0.005	0.003**	−0.205	0.020	0.167***	−0.003	0.008	0.000
Orthography	0.048	0.004	0.044***	0.060	0.018	0.018***	0.017	0.007	0.004*
Distributional	−0.121	0.005	0.240***	−0.092	0.018	0.041***	−0.019	0.007	0.004*
Valence	−0.007	0.006	0.000	−0.070	0.022	0.018***	−0.017	0.009	0.002
Sensory richness	−0.005	0.005	0.000	−0.038	0.023	0.004	−0.015	0.009	0.002
Accuracy									
Visuomotor richness	0.002	0.002	0.001	0.050	0.007	0.113***	0.001	0.000	0.002
Orthography	0.002	0.001	0.001	0.003	0.006	0.001	0.001	0.000	0.005**
Distributional	0.021	0.001	0.118***	0.012	0.006	0.008	0.004	0.000	0.144***
Valence	0.000	0.002	0.000	0.011	0.008	0.004	−0.000	0.000	0.000
Sensory richness	0.002	0.002	0.000	−0.002	0.008	0.000	0.001	0.000	0.004**

We used the "SDSRegressionR" package in R (Mahometa 2018) to generate sr² values

*$p < .05$; **$p < .01$; ***$p < .001$

ELP

Response time

The regression analysis for ELP indicated that the factor scores explained 47.6% of the variance in response time, $F(5, 1514) = 244.40$, $p < .001$. The visuomotor richness and orthographic and distributional factors were significant predictors of ELP response time, with higher visuomotor richness scores ($b = -0.014$, $p = .007$), lower orthography scores ($b = 0.048$, $p < .001$), and higher distributional scores ($b = -0.121$, $p < .001$) contributing to faster responses.

Accuracy

The regression analysis for ELP accuracy indicated that the factor scores explained 15.9% of the variance in response accuracy, $F(5, 1514) = 57.26$, $p < .001$. The distributional factor was the only significant predictor of ELP accuracy, with higher distributional scores ($b = 0.021$, $p < .001$) contributing to greater accuracy.

CSDP

Response time

The regression analysis for CSDP indicated that the factor scores explained 46% of the variance in response time, $F(5, 338) = 61.64$, $p < .001$. The visuomotor richness, orthographic, distributional, and valence factors were all significant predictors of CSDP response time, with higher visuomotor richness scores ($b = -0.207$, $p < .001$), lower orthographic scores ($b = 0.052$, $p = .004$), higher distributional scores ($b = -0.084$, $p < .001$), and higher valence scores ($b = -0.070$, $p = .002$) contributing to faster responses. These relationships are depicted in Fig. 5. (An interactive version of this chart is available at: https://plot.ly/~ejmuraki/3/.)

Accuracy

The regression analysis for CSDP accuracy indicated that the factor scores explained 22.2% of the variance in response accuracy, $F(5, 338) = 19.26$, $p < 0.001$. The visuomotor richness factor was a significant predictor of CSDP accuracy, with higher visuomotor richness scores ($b = 0.050$, $p < 0.001$) related to greater accuracy.

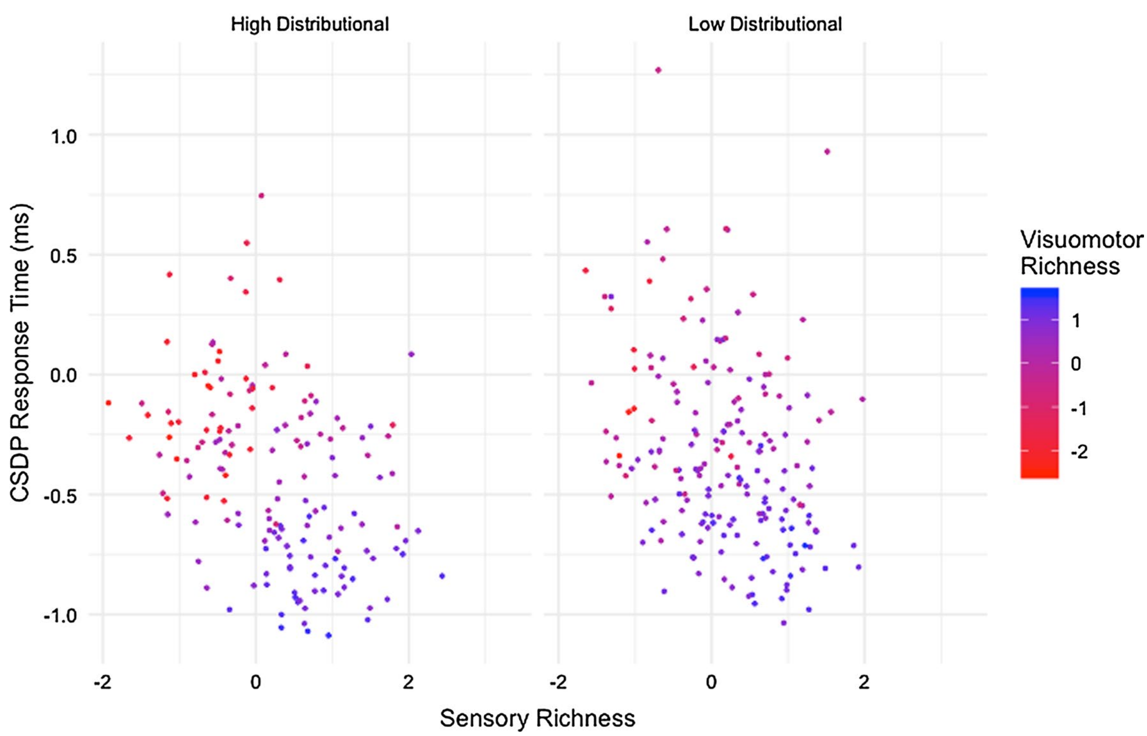

Fig. 5 Scatterplot of CSDP response time and sensory richness factor scores, as a function of visuomotor richness scores, plotted with a median split of distributional factor scores ($n = 343$). A 3D interactive version of this chart is available at: https://plot.ly/~ejmuraki/3/

ECP

Response time

The regression analysis for ECP response time indicated that the factor scores explained only 2.04% of the variance in response time, $F(5, 1512) = 6.28$, $p < .001$. The orthographic and distributional factors were significant predictors of ECP response time, with lower orthographic scores ($b = 0.017$, $p = .018$) and higher distributional scores ($b = -0.019$, $p = .011$) contributing to faster responses.

Accuracy

The regression analysis for ECP accuracy indicated that the factor scores explained 19.3% of the variance in response accuracy, $F(5, 1512) = 72.43$, $p < .001$. The orthographic, distributional, and valence factors were significant predictors of ECP accuracy, with higher orthographic scores ($b = 0.001$, $p = .002$), higher distributional scores ($b = 0.004$, $p < .001$), and higher valence scores ($b = 0.001$, $p = .008$) contributing to greater accuracy.

Discussion

The purpose of the present study was to address two main research questions. The first question was whether NoF effects in lexical and semantic tasks would generalize beyond the McRae et al. (2005) word set to the larger and more heterogeneous word set from Buchanan et al. (2019). Our results suggest that these effects do generalize to the new word set. NoF and DoF were both related to response times in the ELP, CSDP, and ECP. In addition, NoF was related to response accuracy in all three datasets. A comparison of NoF and DoF effects provides additional insights: NoF had stronger relationships to performance in the lexical and semantic tasks than did DoF, suggesting that while shared features may be influential, sheer NoF is the better predictor of lexical-semantic processing. As such, our conclusions are somewhat different than those of Grondin et al. (2009), who argued that shared features drive processing. We think a key difference may be the homogeneity of concepts in the McRae norms (used in the Grondin et al. analyses) compared to the more varied concepts in the Buchanan et al. (2019) norms. Grondin et al. proposed that shared features may facilitate decision-making for concrete words, as shared features can activate representations not only for the target cue, but for a supraordinate category to which the cue belongs. It is conceivable that using the items from the McRae et al. norms showed a processing advantage for these shared features because the items were from well-defined, concrete categories that could benefit from this type of category activation. The Buchanan et al. norms, on the other hand, are a much more heterogeneous set that include words from several different parts of speech with varying degrees of concreteness. As such, items may be less likely to be members of a common category that can benefit from category activation. These results are consistent with theories of semantic memory that emphasize a central role of feature information in semantic representation (Caramazza and Mahon 2003; McClelland and Rogers 2003); however, these theories emphasize that shared features should be most resistant to degradation and our findings suggest that this may not be true of concepts that do not fit into clearly defined categories.

As illustrated in Table 1, NoF explained considerably more variance in semantic decision (CDSP) than in the more lexical tasks (ELP, ECP). This is consistent with the findings of other studies, where semantic variables tend to explain a larger proportion of variance in semantic decision than in lexical decision tasks (e.g., Pexman et al. 2017; Taikh et al. 2015). The explanation is that lexical decisions rely more on orthographic familiarity (Balota et al. 1991; Hino and Lupker 1996), whereas semantic decisions rely more on semantic activation (Pexman et al. 2017).

The second research question was how NoF effects are related to the many semantic richness effects that have been identified in recent years. This included the object-based, bodily experience and situated simulation effects of BOI, concreteness, imageability, SER, valence and arousal, and the language- and distribution-based effects of SemD, NoA and LexC. We used an exploratory factor analysis to map these semantic dimensions onto distinct constructs. The results suggest that NoF is most related to variables reflecting situated simulation and embodied experience. The most prominent factor that emerged included high factor loadings from BOI, concreteness, imageability, and NoF, reflecting an overall visuomotor richness construct. This factor was correlated with a factor representing a sensory richness construct. In addition, variables that relate to distributional aspects of word meaning and association loaded together onto a distributional factor and orthographic variables were clearly reflected in another factor.

Thus, the latent factors extracted for the word set support an interpretation that NoF represents processes of situated simulation or bodily experience, rather than distributional semantic or lexical associations, consistent with the predictions of Barsalou (2003) and McRae et al. (2005) and only partially supporting the proposal put forward by Santos et al. (2011). Furthermore, the majority of the objectively measured variables appear to load together onto the distributional factor, whereas the subjectively measured variables are divided among three different factors. The only exception is NoF, which, although an objective measure, clusters with other subjective measures of visual, motor, and embodied experience. This inclusion of NoF in a factor alongside subjective measures suggests that the factors are not primarily defined as a function of how these variables have been operationalized or derived, but rather reflects that the variables may share some underlying representational processes. The resulting factor structure best supports the object-based, embodiment account of how NoF relates to other semantic richness effects.

It is important to note that the current set of feature norms do not distinguish between different types of features, such as those outlined in McRae et al. (2005). In the McRae et al. norms, features are categorized into types such as taxonomic, visual form, and surface. Though the current measure of NoF loads onto a visuomotor richness construct in the factor analysis, it may be that more fine-grained categorization and recounting of the features could result in a different factor structure that emphasizes differences between embodied or perceptual features and those features that relate to linguistic relationships or factual knowledge. It may also be the case that the feature listing task employed by Buchanan et al. (2019) emphasized or encouraged feature generation based on the physical or object-oriented properties of given items. In future research, more fine-grained classifications of the features could be useful to determine to what extent the feature listing task contributes to the relationships we have observed between NoF and other embodied or situated simulation variables.

The factors performed as expected in regression models predicting ELP, CSDP, and ECP response times and accuracy. Faster response times in the ELP lexical decision task were related to the visuomotor richness, orthographic and distributional factor scores, whereas the CSDP semantic decision task response times showed significant relationships to those factors that tap into richness of bodily or emotional experience, in addition to the relationships with the orthographic and distributional factors. The results of these analyses are consistent with aforementioned findings that recruitment of semantic variables is more important for semantic tasks, although visuomotor richness was related to response times in the lexical decision task. This may be a function of the fact that words with highly salient visuomotor referents also tend to be acquired earlier in life, which would be represented in the distributional factor score. Response times in the ECP were faster for words that had lower orthographic factor scores and higher distributional factor scores, in line with the original analyses presented by Mandera et al. (2019a).

The present study extends our current understanding of semantic richness effects, moving beyond a simply multidimensional representational model to show how these lexical and semantic dimensions are recruited in concert with one another, and as a function of task demands. When the task demands encourage focus on lexical processes, orthographic aspects of stimuli are more influential, whereas when the task involves more extensive semantic processing there is strong recruitment of bodily and sensory simulated experiences. Across all tasks, the distributional factor was a strong predictor of processing speed, aligning with a recent validation of distributional models (Mandera et al. 2019b) and providing strong support for distributional semantic theories such as latent semantic analysis (LSA) theory or word2vec that propose words with similar meanings tend to cluster together in corpus and that this clustering can be measured in order to count or predict semantic relationships (Landauer and Dumais 1997; Mikolov et al. 2013).

The fact that the visuomotor richness factor accounted for the greatest amount of variance among all the lexical and semantic dimensions included in the exploratory factor analysis suggests that it is an important component of lexical and semantic knowledge. While this by itself might suggest support for a strongly embodied account of word meaning (e.g., Glenberg 2015), other aspects of the results point to a different conclusion. The clear differentiation between factors representing bodily experience or sensory simulation and factors representing more linguistic dimensions such as frequency, AoA, word length, and OLD supports a weak embodiment view of semantic representation, which purports that semantic content is based in both linguistic and situated simulation representations (Meteyard et al. 2012). It is interesting to note that variables generally considered to capture embodied representations such as BOI, concreteness, imageability, SER, and arousal actually loaded onto two different factors. This suggests that even within embodied representations there are important distinctions that differentially influence language processing. Motor and visual experience appear to cluster together in the visuomotor richness factor and show a significant relationship to CSDP response times, whereas the more general sensory experience contributing to the construct of sensory richness shows no significant relationship. This could indicate that motor or visually situated simulations are more important to semantic representations than the other sensory experiences that are included in SER beyond sight and touch, such as taste, sound, and smell, particularly when the task demands require making an abstract versus concrete distinction. Certainly, these two factors are correlated, perhaps demonstrating the partial overlap of motor and visual sensory experience contributing to each.

Previous research had documented the multidimensional nature of lexical-semantic representations (e.g., Pexman et al. 2014). The present results extend those findings to suggest that there are clusters of related semantic dimensions in multidimensional semantic space. These clusters correspond to different types of semantic knowledge: bodily and interaction-based, sensorimotor experience, linguistic context, and emotion. These types of information will all need to be incorporated in theories of semantic representation. Indeed, several existing theories can do so, because they assume multiple representation systems for word meaning (e.g., Andrews et al. 2009; Barsalou et al. 2008; Dove 2009). Our results provide those theories with important information about how the multiple dimensions of semantic information might be both related and distinct. Our results are more difficult to reconcile with theories that assume a single dimension of semantic representation, whether it is derived from lexical co-occurrence information (e.g., Lund and Burgess 1996) or from embodiment (Glenberg 2015; Glenberg and Gallese 2012). Overall, the present findings demonstrate the utility of an expanded set of NoF norms to further our understanding of the structure of semantic space and will be useful to continued investigations of how multiple representation systems function in tandem to support word meaning.

Acknowledgements This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC), in the form of a Discovery Grant to PMP.

Funding This study was funded by the Natural Sciences and Engineering Research Council of Canada (RGPIN/03860-2018).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical standards All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee (Conjoint Faculties Research Ethics Board, REB13-0098) and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

Amsel BD, Urbach TP, Kutas M (2012) Perceptual and motor attribute ratings for 559 object concepts. Behav Res Methods 44:1028–1041

Andrews M, Vigliocco G, Vinson D (2009) Integrating experiential and distributional data to learn semantic representations. Psychol Rev 116:463–498

Balota DA, Ferraro FR, Connor LT (1991) On the early influence of meaning in word recognition: a review of the literature. In: Schwanenflugel P (ed) The psychology of word meanings. Erlbaum, Hillsdale, pp 187–218

Balota DA, Yap MJ, Hutchison KA, Cortese MJ, Kessler B, Loftis B, Neely JH, Nelson DL, Simpson GB, Treiman R (2007) The English Lexicon project. Behav Res Methods 39(3):445–459

Barsalou LW (2003) Abstraction in perceptual symbol systems. Philos Trans R Soc B 358:1177–1187

Barsalou LW, Santos A, Simmons WK, Wilson CD (2008) Language and simulation in conceptual processing. In: De Vega M, Glenberg AM, Graesser AC (eds) Symbols, embodiment, and meaning. Oxford University Press, Oxford, pp 245–283

Borghi A, Cimatti F (2010) Embodied cognition and beyond: acting and sensing the body. Neuropsychology 48:763–773

Brysbaert M, New B (2009) Moving beyond Kučera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. Behav Res Methods 41:977–990

Brysbaert M, Warriner AB, Kuperman V (2013) Concreteness ratings for 40 thousand generally known English word lemmas. Behav Res Methods 46(3):904–911

Buchanan L, Westbury C, Burgess C (2001) Characterizing semantic space: neighborhood effects in word recognition. Psychon Bull Rev 8:531–544

Buchanan EM, Valentine KD, Maxwell NP (2019) English semantic feature production norms: an extended database of 4436 concepts. Behav Res Methods 51:1849–1863

Caramazza A, Mahon BZ (2003) The organization of conceptual knowledge: the evidence from category-specific semantic deficits. Trends Cogn Sci 7(8):354–361

Cook AE, Colbert-Getz J, Kircher JC (2013) Number-of-features effects during reading: evidence from eye movements. Discourse Process 50(3):210–225

Cortese MJ, Fugett A (2004) Imageability ratings for 3000 monosyllabic words. Behav Res Methods Instrum Comput 36(3):384–387

Costello AB, Osborne JW (2005) Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. Pract Assess Res Eval 10(7):1–9

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

Devereux BJ, Taylor KI, Randall B, Geertzen J, Tyler LK (2016) Feature statistics modulate the activation of meaning during spoken word processing. Cogn Sci 40:325–350

Dove G (2009) Beyond perceptual symbols: a call for representational pluralism. Cognition 110:412–431

Durda K, Buchanan L, Caron R (2006). WordMine2 (online). www.wordmine2.org

Glenberg AM (2015) Few believe the world is flat: how embodiment is changing the scientific understanding of cognition. Can J Exp Psychol 69:165–171

Glenberg AM, Gallese V (2012) Action-based language: a theory of language acquisition, comprehension, and production. Cortex 48:905–922

Goh WD, Yap MJ, Lau MC, Ng MMR, Tan LC (2016) Semantic richness effects in spoken word recognition: a lexical decision and semantic categorization megastudy. Front Psychol. https://doi.org/10.3389/fpsyg.2016.00976

Grondin R, Lupker SJ, McRae K (2009) Shared features dominate semantic richness effects for concrete concepts. J Mem Lang 60:1–19

Hargreaves IS, Pexman PM (2014) Get rich quick: the signal to respond procedure reveals the time course of semantic richness effects during visual word recognition. Cognition 131:216–242

Hargreaves IS, Pexman PM, Johnson JS, Zdrazilova L (2012) Richer concepts are better remembered: number of features effects in free recall. Front Hum Neurosci. https://doi.org/10.3389/fnhum.2012.00073

Hino Y, Lupker SJ (1996) Effects of polysemy in lexical decision and naming: an alternative to lexical access accounts. J Exp Psychol Hum Percept Perform 22(6):1331–1356

Hoffman P, Lambon Ralph MA, Rogers TT (2013) Semantic diversity: a measure of semantic ambiguity based on variability in the contextual usage of words. Behav Res Methods 45:718–730. https://doi.org/10.3758/s13428-012-0278-x

Juhasz BJ, Yap MJ (2013) Sensory experience ratings for over 5000 mono- and disyllabic words. Behav Res Methods 45:160–168

Kuperman V, Stadthagen-Gonzalez H, Brysbaert M (2012) Age-of-acquisition ratings for 30,000 English words. Behav Res Methods 44(4):978–990

Landauer TK, Dumais ST (1997) A solution to Plato's problem: the latent semantic analysis theory of acquisition, induction and representation of knowledge. Psycol Rev 104(2):211

Louwerse M (2010) Symbol interdependency in symbolic and embodied cognition. Top Cogn Sci 3(2):1–30

Lund K, Burgess C (1996) Producing high-dimensional semantic space from lexical co-occurrence. Behav Res Methods Instrum Comput 28:203–208

Mahometa MJ (2018) The SDSRegression R package. R package version 0.8.6. https://github.com/MichaelJMahometa/SDSRegressionR

Mandera P, Keuleers E, Brysbaert M (2019a) Recognition times for 62 thousand English words: data from the English Crowdsourcing Project. Manuscript in review

Mandera P, Keuleers E, Brysbaert M (2019b) Explaining human performance in psycholinguistic tasks with models of semantic similarity based on prediction and counting: a review and empirical validation. J Mem Lang 92:57–78

McClelland JL, Rogers TT (2003) The parallel distributed processing approach to semantic cognition. Nature Reviews Neurosci 4(4):310–322

McRae K, Cree GS, Seidenberg MS, McNorgan C (2005) Semantic feature production norms for a large set of living and nonliving things. Behav Res Methods 37(4):547–559

Meteyard L, Cuadrado SR, Bahrami B, Vigliocco G (2012) Coming of age: a review of embodiment and the neuroscience of semantics. Cortex 48:788–804

Mikolov T, Sutskever I, Chen K, Corrado GS, Dean J (2013) Distributed representations of words and phrases and their compositionality. In: Burges CJC, Bottou L, Welling M (eds) Advances in neural information processing systems, pp 3111–3119. Retrieved from http://papers.nips.cc/book/advances-in-neural-information-processing-systems-26-2013

Nelson DL, Schreiber TA (1992) Word concreteness and word structure as independent determinants of recall. J Mem Lang 31(2):237–260. https://doi.org/10.1016/0749-596X(92)90013-N

Nelson DL, McEvoy CL, Schreiber TA (2004) The University of South Florida free association, rhyme, and word fragment norms. Behav Res Methods Instrum Comput 36(3):402–407

Paivio A (1971) Imagery and verbal processes. Holt, Rinehart, and Winston, New York

Patterson K, Nestor PJ, Rogers TT (2007) Where do you know what you know? The representation of semantic knowledge in the human brain. Nat Rev Neurosci 8:976–987

Pexman PM (2012) Meaning-level influences on visual word recognition. In: Adelman JS (ed) visual word recognition, vol 2. Meaning and context, individuals and development. Psychology Press, Hove, pp 24–43

Pexman PM, Lupker SJ, Hino Y (2002) The impact of feedback semantics in visual word recognition: number of features effects in lexical decision and naming tasks. Psychon Bull Rev 9:542–549

Pexman PM, Holyk GG, Monfils MH (2003) Number of features effects and semantic processing. Mem Cogn 31:842–855

Pexman PM, Hargreaves IS, Siakaluk PD, Bodner GE, Pope J (2008) There are many ways to be rich: effects of three measures of semantic richness on visual word recognition. Psychon Bull Rev 15:161–167

Pexman PM, Siakaluk PD, Yap MJ (eds) (2014) Meaning in mind: semantic richness effects in language processing. Frontiers E-Books, Lausanne

Pexman PM, Heard A, Lloyd E, Yap MJ (2017) The Calgary semantic decision project: concrete/abstract decision data for 10,000 English words. Behav Res Methods 49:407–417

Pexman PM, Muraki EJ, Sidhu DM, Siakaluk PD, Yap MJ (2019) Quantifying sensorimotor experience: body-object interaction ratings for more than 9000 English words. Res Methods, Behav. https://doi.org/10.3758/s13428-018-1171-z

R Core Team (2018) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. https://www.R-project.org/

Recchia G, Jones MN (2012) The semantic richness of abstract concepts. Front Hum Neurosci 6:1–16. https://doi.org/10.3389/fnhum.2012.00315

Robert C, Rico Duarte L (2016) Semantic richness and aging: the effect of number of features in the lexical decision task. J Psycholinguist Res 45(2):359–365. https://doi.org/10.1007/s10936-015-9352-8

Sajin SM, Connine CM (2014) Semantic richness: the role of semantic features in processing spoken words. J Mem Lang 70:13–35. https://doi.org/10.1016/j.jml.2013.09.006

Santos A, Chaigneau SE, Simmons WK, Barsalou LW (2011) Property generation reflects word association and situated simulation. Lang Cogn 3(1):83–119. https://doi.org/10.1515/LANGCOG.2011.004

Schock J, Cortese MJ, Khanna MM (2012) Imageability estimates for 3000 disyllabic words. Behav Res Methods 44(2):374–379. https://doi.org/10.3758/s13428-011-0162-0

Shaoul C, Westbury C (2010) Exploring lexical co-occurrence space using HiDEx. Behav Res Methods 42:393–413

Siakaluk PD, Pexman PM, Aguilera L, Owen WJ, Sears CR (2008) Evidence for the activation of sensorimotor information during visual word recognition: the body-object interaction effect. Cognition 106:433–443

Taikh A, Hargreaves IS, Yap M, Pexman PM (2015) Semantic classification of pictures and words. Q J Exp Psychol 68(8):1502–1518. https://doi.org/10.1080/17470218.2014.975728

Taylor KI, Devereux BJ, Acres K, Randall B, Tyler LK (2012) Contrasting effects of feature-based statistics on the categorisation and basic-level identification of visual objects. Cognition 122(3):363–374. https://doi.org/10.1016/j.cognition.2011.11.001

Warriner AB, Kuperman V, Brysbaert M (2013) Norms of valence, arousal, and dominance for 13,915 English lemmas. Behav Res Methods 45(4):1191–1207

Yap MJ, Pexman PM, Wellsby M, Hargreaves IS, Huff M (2012) An abundance of riches: cross-task comparisons of semantic richness effects in visual word recognition. Front Hum Neurosci. https://doi.org/10.3389/fnhum.2012.00053

Yarkoni T, Balota D, Yap M (2008) Moving beyond Coltheart's N: a new measure of orthographic similarity. Psychon Bull Rev 15:971–979

Zeno SM, Ivens SH, Millard RT, Duvvuri R (eds) (1995) The educator's word frequency guide. Touchstone Applied Science Associates, Brewster