Title
From Words to Behaviour via Semantic Networks

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 38(0)

Authors
Rotaru, Armand
Vigliocco, Gabriella
Frank, Stefan

Publication Date
2016

Peer reviewed

From Words to Behaviour via Semantic Networks

Armand S. Rotaru (ucjt401@ucl.ac.uk) **Gabriella Vigliocco (g.vigliocco@ucl.ac.uk)**
Faculty of Brain Sciences, University College London,
WC1H 0DS, London, United Kingdom

Stefan L. Frank (s.frank@let.ru.nl)
Centre for Language Studies, Radboud University,
6525 HT Nijmegen, the Netherlands

Abstract

The contents and structure of semantic networks have been the focus of much recent research, with major advances in the development of distributional models. In parallel, connectionist modeling has extended our knowledge of the processes engaged in semantic activation. However, these two lines of investigation have rarely brought together. Here, starting from a standard textual model of semantics, we allow activation to spread throughout its associated semantic network, as dictated by the patterns of semantic similarity between words. We find that the activation profile of the network, measured at various time points, can successfully account for response times in the lexical decision task, as well as for subjective concreteness and imageability ratings.

Keywords: computational modelling; semantic networks; text corpora; lexical decision; concreteness; imageability

Introduction

In the last 15 years, a great deal of effort was invested in collecting extensive behavioural norms, for lexical semantic tasks such as free association (Nelson, McEvoy, & Schreiber, 2004), similarity judgement (Bruni, Tran, & Baroni, 2014; Silberer & Lapata, 2014), feature generation (McRae, Cree, Seidenberg, & McNorgan, 2005; Recchia & Jones, 2012; Vinson & Vigliocco, 2008). In addition, large norms have been obtained for tasks that rely primarily on orthographic and phonological processing, but also include a semantic component, such as lexical decision (Balota et al., 2007; Keuleers, Lacey, Rastle, & Brysbaert, 2012).

This wealth of data has allowed researchers to start exploring the ties that link language to perception and action, in a more methodical and in-depth manner than was previously possible. At a general level, especially within the fields of computational linguistics and natural language processing, representational similarity analysis has been employed in order to study verbal and visual semantic representations across domains of knowledge (Kriegeskorte, Mur, & Bandettini, 2008; for a recent review, see Kriegeskorte & Kievit, 2013). This approach is inspired by several embodied theories of cognition in which the semantic system is considered to rely on integrated modal (especially visual) and amodal representations (Barsalou, Santos, Simmons, & Wilson, 2008; Louwerse, 2007; Vigliocco, Meteyard, Andrews, & Kousta, 2009). The research following said approach has shown that unimodal (i.e., verbal or visual), but especially multimodal (i.e., verbal-visual) distributional models (for a detailed review, see Bruni, Tran, & Baroni, 2014) can provide a good account of human task performance in a number of semantic tasks. Such studies demonstrated that integrating information from two modalities provides a better account of behavioural data than that offered by the individual modalities, across a wide range of models and integration methods, even for abstract concepts, such as *peace* and *freedom* (Bruni, Tran, & Baroni, 2014; Hill & Korhonen, 2014; Hill, Reichart & Korhonen, 2014). The results are consistent with those of previous studies (Andrews, Vigliocco, & Vinson, 2009; Louwerse, 2011; Maki & Buchanan, 2008; Riordan & Jones, 2011; Sadeghi, McClelland, & Hoffman, 2015; Steyvers, 2010), indicating that language and perception can be seen as highly redundant, yet complementary, sources of semantic information.

Differences in the reliance upon one or the other modalities, as well as in degree and strength of association to other concepts, have been argued to underscore difference across domains of knowledge. For example, representational richness has been argued to underlie the distinction between concrete and abstract concepts, whereby concrete concepts are richer than abstract ones when it comes to perceptual and motor elements, but poorer with respect to introspective and linguistic elements (see Gee, Nelson, & Krawczyk, 1999; Hill, Korhonen, & Bentz, 2014; Pecher, Boot, & Van Dantzig, 2011; Vigliocco et al., 2009; Wiemer-Hastings & Xu, 2005). A large number of studies have used comprehensive behavioural norms and subjective ratings to evaluate the role of semantic richness, using different measures of richness such as number of features as well as contextual and semantic diversity, to name a few (for reviews, see Jones, Johns, & Recchia, 2012; Mirman & Magnuson, 2008; Yap, Pexman, Wellsby, Hargreaves, & Huff, 2012). Not surprisingly, the results paint a rather complex picture, where semantic richness effects are both task-general and task-specific, have both an early and a late impact on task behaviour (Hargreaves & Pexman, 2014), and either facilitate or hinder task performance (Mirman & Magnuson, 2008).

Here, we attempt to bring a fresh perspective in the study of how concepts (both concrete and abstract) are represented and, crucially, processed, by developing a computational model that accounts for previous findings by incorporating both structural and dynamical elements. In particular, we explore the extent to which we can predict response times and accuracies in visual word recognition (i.e., lexical decision), as well as both concreteness and imageability ratings, starting from distributional models of semantics (Mandera, Keuleers, & Brysbaert, 2015; Westbury et al., 2013) supplemented by simple assumptions concerning the dynamic spreading of activation during processing.

Method

Model

We derive semantic richness measures of words from a probabilistic model of semantic processing, in the following manner: (a) pre-process the written part of the British National Corpus (Leech, Garside, & Bryant, 1994), by converting all the words to lowercase, eliminating punctuation marks and removing words whose absolute frequencies are less than 5; (b) construct 300-dimensional vector representations for the words in the BNC, by employing the Skipgram[1] model (Mikolov, Chen, Corrado, & Dean, 2013); (c) compute a representational similarity matrix *DM* from said vectors, using vector cosine as a measure of similarity between the vectors (i.e., words); (d) set to zero all the negative values in *DM*, as a means of reducing the amount of noise present (i.e., vector cosines which carry very little semantic information); (e) normalize the rows of the matrix, such that each row sums to one, and that any value $DM(I,J)$ can be interpreted as the strength of the directional connection from word W_I to word W_J ; (f) consider the discrete Markov chain associated with *DM*, which we denote as *MARKOV*(*DM*), and compute the state of *MARKOV*(*DM*) at steps $K = 1$ through $K = 7$, namely $S_K(DM)$; (g) for each word W and each K between 1 and 7, count the number of close neighbours of W ($numNeigh_K$). A word V is considered a close neighbour of W if $P(V \mid S_K(DM)) > thresh_K$, where $thresh_K$ are lower thresholds.

In the end, we are left with seven free parameters (i.e., $thresh_{1\text{-}7}$) and seven semantic richness measures (i.e., $numNeigh_{1\text{-}7}$), as well as with a few fixed parameters for the underlying Skipgram model.[2] Although our richness measures are all derived in a very similar manner, they have rather different interpretations, at least from a graph-theoretical perspective (Koschützki, Lehmann, Peeters, Richter, Tenfelde-Podehl, & Zlotowski, 2005). The meaning of each measure is briefly described in Table 1.

Table 1. Semantic richness measures computed by our model, and their tentative interpretation. For clarity, only the distinguishing aspects of each measure are presented.

Semantic richness measure	Graph theoretical interpretation
$numNeigh_1$	# of close neighbours
$numNeigh_2$	# of connections between close neighbours
	# of distant neighbours
	# of connections between close and distant neighbours
$numNeigh_3$	# of connections between distant neighbours
	# of connections between distant and close neighbours
$numNeigh_{4\text{-}7}$	# of close direct and indirect neighbours
	# of very distant neighbours

Data Analysis

We focus on four dependent measures: concreteness ratings (Brysbaert, Warriner, & Kuperman, 2014), imageability ratings (Gilhooly & Logie, 1980; Stadthagen-Gonzalez & Davis, 2006), and both accuracies and response times from a lexical decision task, for a subset of 2,328 words from Keuleers, Lacey, Rastle, and Brysbaert (2012).

We include the following <u>baseline variables</u>: (log) contextual diversity, (log) frequency (van Heuven, Mandera, Keuleers, & Brysbaert, 2014), familiarity (Gilhooly & Logie, 1980; Stadthagen-Gonzalez & Davis, 2006), age of acquisition (Kuperman, Stadthagen-Gonzalez, & Brysbaert, 2012), (squared) hedonic valence (Warriner, Kuperman, & Brysbaert, 2013), number of letters, Coltheart's N (i.e., the number of words that can be produced by substituting one letter of a given word for any other, such that the result is a valid word; Coltheart, Davelaar, Jonasson, & Besner, 1977), orthographic Levenshtein distance (OLD20; the average orthographic editing distance between a word and its twenty closest neighbours in the lexicon; Yarkoni, Balota, & Yap, 2008), and phonological Levenshtein distance (PLD20; the average phonological distance between a word and its twenty closest neighbours in the lexicon; Suárez, Tan, Yap, & Goh, 2011). In addition we include semantic diversity (Hoffman, Lambon Ralph, & Rogers, 2013) as a baseline measure. This latter has been argued to capture basic semantic differences across concepts as represented in distributional semantic networks. Our <u>variables of interest</u> are the seven measures of semantic richness (i.e., $numNeigh_{1\text{-}7}$).

We run two multiple linear regressions, one for the baseline variables, and one for the complete set of predictors (i.e., the baseline variables and our semantic richness measures). Since our richness measures are very strongly correlated with one another, we partial out any variance shared with other predictors, such that $numNeighR_K$ = Res

[1] We prefer the Skipgram model for two main reasons: firstly, it is nearly state-of-the-art when it comes to accounting for behavioral data (Baroni, Dinu, & Kruszewski, 2014; Mandera, Keuleers, & Brysbaert, 2015); secondly, several freely available, computationally efficient and well documented implementations of the model exist (https://code.google.com/p/word2vec/).

[2] We use the Skipgram implementation provided by the *gensim* tool (Řehůřek & Sojka, 2010), with the following parameter values: *alpha* = 0.025 (initial learning rate), *window* = 5 (radius of sliding window), *sample* = 0 (amount of downsampling), *negative* = 0 (amount of negative sampling), and *iter* = 1 (number of iterations over the entire corpus).

($numNeigh_K$ ~ $Baseline$ + $numNeighR_1$ + … + $numNeighR_{K-1}$), for all values of K between 1 and 7. Therefore, our predictors consist of *Baseline* and $numNeighR_{1\text{-}7}$, whereas our dependent variables are *Log RT*, *Accuracy*, *Concreteness* and *Imageability*.

We employ one half of the words for model tuning, and the other half for model evaluation. We derive the optimal values for our predictors by using a variant of the simplex method (Lagarias, Reeds, Wright, & Wright, 1998), with (negative) total amount of variance explained serving as the objective function. In order to avoid local minima, we run 100 iterations of the optimisation process, and keep only the best result.

Results

The results are displayed in Tables 2 and 3. Our semantic richness measures can account for a significant amount of variance in concreteness and imageability ratings, as well as in response times in the lexical decision task. However, they do not explain variance in lexical decision accuracy over and above the baseline measures (Table 2). Table 3 shows the regression weights for all predictors and dependent variables.

Table 2. Percentage of variance accounted for by two models: a baseline model, and a combined one, consisting of all the predictors in the baseline model plus the semantic richness measures $numNeighR_1$ through $numNeighR7$ (all values are significant at .001 level, except for accuracy in the "combined – baseline" comparison)

Dependent variable	Baseline	Combined	Combined – baseline
Log RT	47.93	49.80	1.87
Accuracy	27.09	27.81	0.72
Concreteness	35.40	58.59	23.19
Imageability	31.90	53.24	21.34

Discussion and Conclusions

We develop a model that takes into account both the structural properties of semantics networks, as well as their dynamic aspects, by considering the flow of semantic activation generated by the automatic processing of individual words. An important result of looking at both structure and dynamics is that it allows us to assess the effects of both direct and indirect, mediated semantic relations between words, rather than limiting our analysis to strong, direct semantic links. Our results suggest that the explanatory power of text-based semantic representations is currently being underestimated, as a consequence of not taking into consideration the additional information provided by spreading activation mechanisms. By ignoring these simple processes, the extra information they generate would have to be integrated into the representations by

Table 3. Regression weights and their associated significance values, namely <0.1 (†), <0.05 (*), <0.01 (**), and <0.001 (***). Log RT = (log) response time; ACC = accuracy; CONC = concreteness; IMAG = imageability.

Predictor \ Outcome	Log RT	ACC	CONC	IMAG
(Intercept)	6.611 ***	.676 ***	7.143 ***	7.405 ***
Semantic diversity	.009	-.022 **	-1.219 ***	-1.401 ***
Log contextual diversity	-.025 ***	.026 ***	-.385 ***	-.539 ***
Log frequency	-8.07e-4	-.008 *	.190 ***	.288 ***
Familiarity	-.035 ***	.024 ***	.169 ***	.373 ***
Age of acquisition	.003	-.003	-.313 ***	-.477 ***
Squared hedonic valence	-.004 ***	.002 *	-.090 ***	.025 †
Number of letters	.007 **	.005 *	.040	.031
Coltheart's N	.001 †	-1.48e-4	.012 †	.025 **
OLD20	.002	-6.54e-5	-.148	.058
PLD20	.012 *	-8.50e-4	-.263 ***	-.342 ***
$NumNeighR_1$	-.006 **	.004 *	.181 ***	.401 ***
$NumNeighR_2$	.004 *	-.003 †	-173 ***	-.231 ***
$NumNeighR_3$	-.001	-3.41e-4	-.132 ***	-.271 ***
$NumNeighR_4$	-.001	-1.93e-4	-.250 ***	-.116 ***
$NumNeighR_5$	3.31e-4	-2.91e-4	-.263 ***	-.218 ***
$NumNeighR_6$	.002	-7.74e-4	-.057 **	.014
$NumNeighR_7$	-2.12e+6 **	-.002	.150 ***	.167 ***

design, which would lead to the conflation of representations and processes.

Based on the results presented in Table 2, it seems that our model is considerably more suitable for predicting concreteness and imageability ratings, than reaction time and accuracy in the word recognition task. We believe that this phenomenon might be due to differences between the requirements of the lexical decision task on the one hand, and those of the concreteness/imageability rating task, on the other. Since our model assumes that the string of letters received as input is already a word, it is not surprising that it fares rather poorly in predicting lexical decision response time and accuracy. In contrast, the rating task involves making a considerably more elaborate discrimination, one between concrete/imageable and abstract/non-imageable words, all of which are present in our model (Buchanan, Westbury, & Burgess, 2001).

Beyond the promising initial results, we believe that our model has a number of advantages, which recommend it as a potentially useful tool in the study of semantic processing. In our opinion, the main quality of our model is that it ties together a number of competing modelling approaches, and combines many of their strengths, while avoiding most of their limitations.

Firstly, our model has a pronounced connectionist and/or dynamical systems flavour to it (Anderson, 1983; for a review, see McClelland et al., 2010), whereby the dynamics of the model can be interpreted in terms of "spreading activation". In this case, activation flows from an initial concept to its neighbours, then to the neighbours of its neighbours, and so on, until the system reaches a global "attractor" state (i.e., an eigenstate). However, unlike other existing models (Chen & Mirman, 2012; Hoffman & Woollams, 2015; Rogers & McClelland, 2004), it has a large number of nodes and feedforward/feedback/recurrent connections, making it slightly more realistic and comprehensive. As a result, it might provide better insight into the distinct contribution of structural and task related aspects of semantic behaviour. One potentially promising approach in this regard comes from network science and the theory of stochastic processes, two methodologies which have attracted an increasing amount of attention in cognitive science (De Deyne & Storms, 2008; Ferrer i Cancho & Solé, 2001; Gruenenfelder, Recchia, Rubin, & Jones, in press; Steyvers & Tenenbaum, 2005; Utsumi, 2015; for a general review of network-based analyses of cognition, see Baronchelli, Ferrer i Cancho, Pastor-Satorras, Chater, & Christiansen, 2013). Another possibility might be to use a respond-to-signal paradigm (Ratcliff, 2006; Hargreaves & Pexman, 2014), which would provide additional quantitative insights on the accumulation of task-specific and task-independent information during task performance (e.g., in the word naming or the lexical decision tasks).

Secondly, our model can be seen as a probabilistic one (Griffiths, Chater, Kemp, Perfors, & Tenenbaum, 2010), such that at each step, the model makes use of its underlying Markov chain, namely *MARKOV*(*DM*), in order to perform multi-step inferences. In contrast to other probabilistic models, such as Topics (Griffiths, Steyvers, & Tenenbaum, 2007), our model is non-hierarchical and does not undergo any form of dimensionality reduction, which means that the inferences are easier to interpret and that less semantic information is lost. Said inferences allow us to assess the strength of both direct and indirect semantic relations between words (Steyvers, Shiffrin, & Nelson, 2004; Howard, Shankar, & Jagadisan, 2011), for instance by testing whether certain words and/or associations between words are crucial for successfully carrying out a semantic task. Moreover, we can also examine the manner in which semantic cues restrict and guide the inference process, as is the case in tasks such as semantic fluency (Hills, Jones, & Todd, 2012), continued free association (De Deyne & Storms, 2008), and extralist cued recall (Nelson, Kitto, Galea, McEvoy, & Bruza, 2013).

Finally, our model is relatively simple, from a structural point of view, and is completely transparent in terms of its parameters. Taken together, these features make our model easy to run, and facilitate comparisons across different subsets of participants, stimuli and tasks. Also, as a results of its simplicity, our current model is very much open to extensions, for instance in order to increase its neuropsychological plausibility.

References

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, *22*(3), 261-295.

Andrews, M., Vigliocco, G., & Vinson, D. (2009). Integrating experiential and distributional data to learn semantic representations. *Psychological Review*, *116*(3), 463-498.

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English lexicon project. *Behavior Research Methods*, *39*(3), 445-459.

Baronchelli, A., Ferrer i Cancho, R., Pastor-Satorras, R., Chater, N., & Christiansen, M. H. (2013). Networks in cognitive science. *Trends in Cognitive Sciences*, *17*(7), 348-360.

Baroni, M., Dinu, G., & Kruszewski, G. (2014). Don't count, predict! A systematic comparison of context-counting vs. context-predicting semantic vectors. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*.

Barsalou, L. W., Santos, A., Simmons, W. K., & Wilson, C. D. (2008). Language and simulation in conceptual processing. In *Symbols and Embodiment: Debates on meaning and cognition*.

Bruni, E., Tran, N. K., & Baroni, M. (2014). Multimodal Distributional Semantics. *Journal of Artificial Intelligence Research*, *49*(1), 1-47.

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known

English word lemmas. *Behavior Research Methods*, *46*(3), 904-911.
Buchanan, L., Westbury, C., & Burgess, C. (2001). Characterizing semantic space: Neighborhood effects in word recognition. *Psychonomic Bulletin & Review*, *8*(3), 531-544.
Chen, Q., & Mirman, D. (2012). Competition and cooperation among similar representations: toward a unified account of facilitative and inhibitory effects of lexical neighbors. *Psychological Review*, *119*(2), 417-430.
Coltheart, M., Davelaar, E., Jonasson, T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI*. Hillsdale, NJ: Erlbaum.
De Deyne, S., & Storms, G. (2008). Word associations: Network and semantic properties. *Behavior Research Methods*, *40*(1), 213-231.
Ferrer i Cancho, R., & Solé, R. V. (2001). The small world of human language. *Proceedings of the Royal Society of London B: Biological Sciences*, *268*(1482), 2261-2265.
Gee, N. R., Nelson, D. L., & Krawczyk, D. (1999). Is the concreteness effect a result of underlying network interconnectivity? *Journal of Memory and Language*, *40*(4), 479-497.
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*(4), 395-427.
Griffiths, T. L., Chater, N., Kemp, C., Perfors, A., & Tenenbaum, J. B. (2010). Probabilistic models of cognition: Exploring representations and inductive biases. *Trends in Cognitive Sciences*, *14*(8), 357-364.
Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. (2007). Topics in semantic representation. *Psychological Review*, *114*(2), 211-244.
Gruenenfelder, T. M., Recchia, G., Rubin, T., & Jones, M. N. (in press). Graph-theoretic properties of networks based on word association norms: implications for models of lexical semantic memory. *Cognitive Science*.
Hargreaves, I. S., & Pexman, P. M. (2014). Get rich quick: The signal to respond procedure reveals the time course of semantic richness effects during visual word recognition. *Cognition*, *131*(2), 216-242.
Hill, F., & Korhonen, A. (2014). Learning abstract concept embeddings from multi-modal data: Since you probably can't see what I mean. In *Proceedings of EMNLP*.
Hill, F., Korhonen, A., & Bentz, C. (2014). A quantitative empirical analysis of the abstract/concrete distinction. *Cognitive Science*, *38*(1), 162-177.
Hill, F., Reichart, R., & Korhonen, A. (2014). Multi-modal models for concrete and abstract concept meaning. *Transactions of the Association for Computational Linguistics*, *2*, 285-296.
Hills, T. T., Jones, M. N., & Todd, P. M. (2012). Optimal foraging in semantic memory. *Psychological Review*, *119*(2), 431-440.
Hoffman, P., Ralph, M. A. L., & Rogers, T. T. (2013). Semantic diversity: a measure of semantic ambiguity based on variability in the contextual usage of words. *Behavior Research Methods*, *45*(3), 718-730.
Hoffman, P., & Woollams, A. M. (2015). Opposing effects of semantic diversity in lexical and semantic relatedness decisions. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(2), 385-402.
Howard, M. W., Shankar, K. H., & Jagadisan, U. K. (2011). Constructing semantic representations from a gradually changing representation of temporal context. *Topics in Cognitive Science*, *3*(1), 48-73.
Jones, M. N., Johns, B. T., & Recchia, G. (2012). The role of semantic diversity in lexical organization. Canadian *Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *66*(2), 115-124.
Keuleers, E., Lacey, P., Rastle, K., & Brysbaert, M. (2012). The British Lexicon Project: Lexical decision data for 28,730 monosyllabic and disyllabic English words. *Behavior Research Methods*, *44*(1), 287-304.
Koschützki, D., Lehmann, K. A., Peeters, L., Richter, S., Tenfelde-Podehl, D., & Zlotowski, O. (2005). Centrality indices. In *Network analysis*. Springer: Berlin Heidelberg.
Kriegeskorte, N., & Kievit, R. A. (2013). Representational geometry: integrating cognition, computation, and the brain. *Trends in Cognitive Sciences*, *17*(8), 401-412.
Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis–connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *2*.
Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, *44*(4), 978-990.
Lagarias, J. C., Reeds, J. A., Wright, M. H., & Wright, P. E. (1998). Convergence properties of the Nelder--Mead simplex method in low dimensions. *SIAM Journal on Optimization*, *9*(1), 112-147.
Leech, G., Garside, R., & Bryant, M. (1994). CLAWS4: the tagging of the British National Corpus. In *Proceedings of the 15th Conference on Computational Linguistics* (vol. 1).
Louwerse, M. M. (2007). Symbolic or embodied representations: A case for symbol interdependency. In *Handbook of Latent Semantic Analysis*.
Louwerse, M. M. (2011). Symbol interdependency in symbolic and embodied cognition. *Topics in Cognitive Science*, *3*(2), 273-302.
Maki, W. S., & Buchanan, E. (2008). Latent structure in measures of associative, semantic, and thematic knowledge. *Psychonomic Bulletin & Review*, *15*(3), 598-603.
Mandera, P., Keuleers, E., & Brysbaert, M. (2015). How useful are corpus-based methods for extrapolating psycholinguistic variables? *The Quarterly Journal of Experimental Psychology*, *68*(8), 1623-1642.
McClelland, J. L., Botvinick, M. M., Noelle, D. C., Plaut, D. C., Rogers, T. T., Seidenberg, M. S., & Smith, L. B. (2010). Letting structure emerge: connectionist and

dynamical systems approaches to cognition. *Trends in Cognitive Sciences*, *14*(8), 348-356.
McDonald, S. A., & Shillcock, R. C. (2001). Rethinking the word frequency effect: The neglected role of distributional information in lexical processing. *Language and Speech*, *44*(3), 295-322.
McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods*, *37*(4), 547-559.
Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. In *Proceedings of Workshop at ICLR*.
Mirman, D., & Magnuson, J. S. (2008). Attractor dynamics and semantic neighborhood density: processing is slowed by near neighbors and speeded by distant neighbors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(1), 65-79.
Nelson, D. L., Kitto, K., Galea, D., McEvoy, C. L., & Bruza, P. D. (2013). How activation, entanglement, and searching a semantic network contribute to event memory. *Memory & Cognition*, *41*(6), 797-819.
Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 402-407.
Pecher, D., Boot, I., & van Dantzig, S. (2011). Abstract concepts: sensory-motor grounding, metaphors, and beyond. In *The Psychology of Learning and Motivation* (vol. 54).
Ratcliff, R. (2006). Modeling response signal and response time data. *Cognitive Psychology*, *53*(3), 195-237.
Recchia, G., & Jones, M. N. (2012). The semantic richness of abstract concepts. *Frontiers in Human Neuroscience*, *6*.
Řehůřek, R., & Sojka, P. (2010) Software Framework for Topic Modelling with Large Corpora. In *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*.
Riordan, B., & Jones, M. N. (2011). Redundancy in perceptual and linguistic experience: Comparing feature-based and distributional models of semantic representation. *Topics in Cognitive Science*, *3*(2), 303-345.
Rogers, T. T., & McClelland, J. L. (2004). *Semantic Cognition: A Parallel Distributed Processing Approach*. MIT Press.
Sadeghi, Z., McClelland, J. L., & Hoffman, P. (2015). You shall know an object by the company it keeps: An investigation of semantic representations derived from object co-occurrence in visual scenes. *Neuropsychologia*, *76*, 52-61.
Silberer, C., & Lapata, M. (2014). Learning grounded meaning representations with autoencoders. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*.
Stadthagen-Gonzalez, H., & Davis, C. J. (2006). The Bristol norms for age of acquisition, imageability, and familiarity. *Behavior Research Methods*, *38*(4), 598-605.
Steyvers, M. (2010). Combining feature norms and text data with topic models. *Acta Psychologica*, *133*(3), 234-243.
Steyvers, M., Shiffrin, R. M., & Nelson, D. L. (2004). Word association spaces for predicting semantic similarity effects in episodic memory. In *Experimental cognitive psychology and its applications: Festschrift in honor of Lyle Bourne, Walter Kintsch, and Thomas Landauer*.
Steyvers, M., & Tenenbaum, J. B. (2005). The Large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science*, *29*(1), 41-78.
Suárez, L., Tan, S. H., Yap, M. J., & Goh, W. D. (2011). Observing neighborhood effects without neighbors. *Psychonomic Bulletin & Review*, *18*(3), 605-611.
van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology*, *67*(6), 1176-1190.
Utsumi, A. (2015). A Complex Network Approach to Distributional Semantic Models. *PLoS ONE*, *10*(8), e0136277.
Vigliocco, G., Meteyard, L., Andrews, M., & Kousta, S. (2009). Toward a theory of semantic representation. *Language and Cognition*, *1*(2), 219-247.
Vinson, D. P., & Vigliocco, G. (2008). Semantic feature production norms for a large set of objects and events. *Behavior Research Methods*, *40*(1), 183-190.
Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, *45*(4), 1191-1207.
Westbury, C. F., Shaoul, C., Hollis, G., Smithson, L., Briesemeister, B. B., Hofmann, M. J., & Jacobs, A. M. (2013). Now you see it, now you don't: on emotion, context, and the algorithmic prediction of human imageability judgments. *Frontiers in Psychology*, *4*.
Wiemer-Hastings, K., & Xu, X. (2005). Content differences for abstract and concrete concepts. *Cognitive Science*, *29*(5), 719-736.
Yap, M. J., Pexman, P. M., Wellsby, M., Hargreaves, I. S., & Huff, M. (2012). An abundance of riches: cross-task comparisons of semantic richness effects in visual word recognition. *Frontiers in Human Neuroscience*, *6*(72).
Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's N: A new measure of orthographic similarity. *Psychonomic Bulletin & Review*, *15*(5), 971-979.