

## Online processing of Verb–Argument Constructions: lexical decision and meaningfulness\*

NICK C. ELLIS

*Department of Psychology, University of Michigan*

*(Received 01 June 2015 – Revised 01 January 2016 – Accepted 04 March 2016 –  
First published online 15 July 2016)*

### ABSTRACT

This paper discusses how patterns of construction usage, implicitly learned over lifelong experience, tune the language processing system for fluent interactive lexical, syntactic, and semantic access. It reports on three experiments that investigate online processing of Verb–Argument Constructions (VACs) and the degree to which this is effected by (i) verb frequency in the language, (ii) verb frequency in the VAC, (iii) VAC–verb contingency, and (iv) verb prototypicality in terms of centrality within the VAC semantic network. Experiment 1 tested lexical decision of VAC exemplars presented as successive verb–preposition pairs. Experiment 2 tested lexical decision of VAC exemplars presented as arbitrarily interrupted verb–adverb–preposition pairs. Experiment 3 had participants judge whether two-word utterances were meaningful or not. All of the experiments show effects of Verb Frequency and Verb–VAC frequency: learners have rich implicit statistical knowledge of verb–VAC type–token frequency that guides processing. Lexical decision is additionally driven by semantic prototypicality (but not VAC–verb contingency  $\Delta P_{cw}$ ), whereas meaning judgment is affected by VAC–verb contingency  $\Delta P_{cw}$  (but not semantic prototypicality). These findings, I argue, index the spreading activation of unconscious meaning representation in lexical decision in comparison to the election of a unitary interpretation in conscious comprehension. I conclude that speeded automatic VAC processing involves rich associations, tuned by verb type and token frequencies, their contingencies of usage, and their histories of prototypical and specific interpretations which interface syntax, lexis, and semantics.

---

[\*] University of Michigan, Department of Psychology, 530 Church St., Ann Arbor, MI 48109, USA. e-mail: ncellis@umich.edu

**KEYWORDS:** Construction Grammar, usage-based acquisition and processing, Verb–Argument Constructions, lexical decision, frequency, semantic networks, contingency, conscious and unconscious meaning, psycholinguistic processing.

## 1. Background

Usage-based approaches to language hold that we learn the patterns of language from language usage and that knowledge of these patterns underlies fluent language processing. Related inquiries involve cross-disciplinary investigations of usage (What are the patterns of language?), acquisition (How do we acquire them?), and processing (How is online processing tuned to the regularities of usage?). No single approach is enough to understand the complex adaptive system of language (Beckner et al., 2009; Ellis & Larsen-Freeman, 2006; Ellis, Römer, & O’Donnell, 2016). Language and its learning are adapted to human cognition. Human language cognition adapts to the regularities of language usage.

Relevant research progresses broadly as follows:

1. Cognitive and corpus linguistics focus upon usage. They show how language is highly structured and pervaded by collocations and phraseological patterns, that every word has its own local grammar, and that language constructions are motivated by semantics and communicative functions: lexis, syntax, and semantics are inseparable (Biber & Reppen, 2015; Trousdale & Hoffmann, 2013).
2. Psychological research into symbolic and statistical language learning investigates the range of human abilities for implicit associative and statistical learning, concept learning and categorization, and explicit declarative learning and analogy-making – abilities which have the potential to learn the symbols, sequences, and patterns of language and which imbue our every waking moment (Rebuschat & Williams, 2012).
3. Child and second language acquisition research charts the stages of learners coming to know their language, using longitudinal corpora of development and supporting these with experiments focused upon process (Ambridge & Lieven, 2011; Ambridge & Rowland, 2013; Gass & Mackey, 2011; Granger, Gilquin, & Meunier, 2015).
4. Psycholinguistics catalogues the many ways in which our language processing is sensitive to the statistical regularities of language experience at every level of structure (Ellis, 2002; Gaskell, 2007; Traxler & Gernsbacher, 2011).

Usage-based approaches to language acquisition hold that schematic constructions emerge as prototypes from the conspiracy of memories of

particular exemplars that language users have experienced (Ellis, O'Donnell, & Römer, 2012; Goldberg, 1995, 2006; Trousdale & Hoffmann, 2013) over their lifetime of language processing. The three experiments reported in this paper, therefore, investigate online processing of abstract Verb–Argument Constructions (VACs) and its sensitivity to the statistics of usage in terms of verb exemplar type–token frequency distribution, VAC–verb contingency, and VAC–verb semantic prototypicality.

Consider the novel utterance “it mandools across the ground”. You know that *mandool* is a verb of motion and have some idea of its action semantics even though you have never come across this nonsense word before. Theories of construction grammar hold that VACs inherit their schematic meaning from the conspiracy of all of the examples you have experienced. *Mandool* gets its interpretation from the echoes of the verbs that you have heard occupying this VAC. Your language processing system parses “it mandools across the ground” as a Verb Locative (VL) construction, then the paradigmatic associations of the types of verb that you have experienced occupying this VL ‘V across N’ VAC – *come, walk, move, ... , scud, skitter, and flit* – come to mind to guide your interpretation.

If constructions are indeed learned like this, as schematic signs, as form–meaning pairings, then the general principles of associative learning and categorization should be evident in their processing (Ellis & Ogden, 2015). The learning and processing of cue–outcome contingencies should be affected by: (i) form frequency in the input, (ii) contingency of form–function mapping, and (iii) function (prototypicality of meaning).

## 1.1. PRINCIPLES OF VAC COGNITION

### 1.1.1. *Construction frequency*

Learning, memory, and perception are all affected by frequency of usage: the more times we experience something, the stronger our memory for it, and the more fluently it is accessed. Language processing is sensitive to usage frequency at all levels of language representation: phonology and phonotactics, reading, spelling, lexis, morphosyntax, formulaic language, language comprehension, grammaticality, sentence production, and syntax – high-frequency forms are learned more easily and processed more fluently (Ellis, 2002). So high-frequency verbs in the language should be processed faster than low-frequency verbs. I will refer to this as Verb Frequency. Likewise, the more times we experience conjunctions of features, the more they become associated in our minds and the more these conditional frequencies subsequently affect perception and categorization (Harnad, 1987; Lakoff, 1987). So, in particular, verbs which appear more often in particular VACs should be more associated with those frames, and processed faster. I will refer to this as Verb-VAC frequency.

1.1.2. *Contingency of form–function mapping*

Psychological research into associative learning has long recognized that while frequency of form is important, more so is contingency of mapping (Shanks, 1995). Consider how, in the learning of the category of birds, while eyes and wings are equally frequently experienced features in the exemplars, it is wings which are distinctive in differentiating birds from other animals. Wings are important features to learning the category of birds because they are reliably associated with class membership; eyes are not. Raw frequency of occurrence is less important in categorization than is the contingency between cue and interpretation (Rescorla, 1968).

So, in VAC processing, lexical cues which are more faithful to a VAC should be more telling. In my research with long-time collaborators, we use the one-way dependency statistic  $\Delta P$  (Allan, 1980) to measure contingency. This has been shown to predict cue–outcome learning in the associative learning literature (Shanks, 1995) as well as in psycholinguistic studies of form–function contingency in construction usage, knowledge, and processing (Ellis, 2006; Ellis & Ferreira-Junior, 2009; Gries & Ellis, 2015).

Consider the contingency table showing the four possible combinations of the presence or absence of a VAC and a verb illustrated in Table 1 where *a, b, c, d* represent frequencies, so, for example, *a* is the number of times the cue and the outcome co-occurred; *c* is the number of times the outcome occurred without the cue; etc.

TABLE 1. *A contingency table showing the four possible combinations of events showing the presence or absence of a target cue and an outcome*

	Outcome		No outcome	
Cue	a	b	c	d
No cue	c	d		

The effects of conjoint frequency, verb frequency, and VAC frequency are illustrated for three cases below:

	$\Delta P$ Construction $\rightarrow$ Word				$\Delta P$ Word $\rightarrow$ Construction			
	Conjoint Frequency	VAC Frequency	Verb Frequency		Conjoint Frequency	Verb Frequency	VAC Frequency	
lie <i>across</i>	a	a+b	a+c	$\Delta P_{cw}$	a	a+b	a+c	$\Delta P_{wc}$
stride	44	5,261	13,190	0.0076	44	13,190	5,261	0.0030
<i>across</i>	44	5,261	1,049	0.0083	44	1,049	5,261	0.0416
crowd	44	50,070	749	0.0008	44	749	50,070	0.0559
<i>into</i>								

$\Delta P$  is the probability of the outcome given the cue minus the probability of the outcome in the absence of the cue. When these are the same, when the outcome is just as likely when the cue is present as when it is not, there is no covariation between the two events and  $\Delta P = 0$ .  $\Delta P$  approaches 1.0 as the presence of the cue increases the likelihood of the outcome and approaches  $-1.0$  as the cue decreases the chance of the outcome – a negative association.

$$\Delta P = P(O|C) - P(O|\neg C) = \frac{a}{a+b} - \frac{c}{c+d}$$

$\Delta P$  is a directional measure. We can consider the association between a VAC as cue and a particular verb type as the outcome (I will call this  $\Delta P_{cw}$  for construction  $\rightarrow$  word). Alternately we can consider the association between a verb as cue and a particular VAC as the outcome ( $\Delta P_{wc}$ ).

$\Delta P$  is affected by the conjoint frequency of construction and verb in the corpus ( $a$ ), but also by the frequency of the verb in the corpus (Verb Frequency), the frequency of the VAC in the corpus, and the number of verbs in the corpus. For illustration, the lower part of Table 1 considers three exemplars, *LIE across*, *STRIDE across*, and *CROWD into*, which all have the same conjoint frequency of 44 in a corpus of 17,408,901 VAC instances. This is the value of Verb-VAC frequency described in 1.1.1. However, while  $\Delta P_{Construction \rightarrow Word}$  ( $\Delta P_{cw}$ ) for *LIE across* and *STRIDE across* are approximately the same, that for *CROWD into* is an order of magnitude less.  $\Delta P_{wc}$  shows a different pattern – the values for *STRIDE across* and *CROWD into* are over ten times greater than for *LIE across*.

### 1.1.3. *Function (prototypicality of meaning)*

Categories have graded structure, with some members being better exemplars than others. In the prototype theory of concepts (Rosch & Mervis, 1975; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976), the prototype as an idealized central description is the best example of the category, appropriately summarizing the most representative attributes of a category. As the typical instance of a category, a prototype serves as the benchmark against which surrounding, less representative instances are classified. In semantic network theories of meaning, related concepts are more closely and strongly connected, and when one concept is activated, so activation spreads to neighboring nodes (Anderson, 1983). In these views, the prototype has two advantages: The first is a frequency factor: the greater the token frequency of an exemplar, the more it contributes to defining the category, and the greater the likelihood it will be considered the prototype (Rosch & Mervis, 1975; Rosch et al., 1976). Thus it is the response that is most associated with the concept in its own right.

But beyond that, it gets the network centrality advantage. When any response is made, it spreads activation and reminds other members in the set. The prototype is most connected at the center of the network and, like Rome, all roads lead to it. Thus it receives the most spreading activation. Ellis, O'Donnell, and Römer (2014) consider spreading activation as it might apply to VACs. As symbolic form–function mappings, the VAC lexico-syntactic frame is associated by usage experience with a network of meanings. When the VAC is activated, prototypical verb meanings are more readily awakened.

#### 1.1.4. *Investigating effects of frequency, contingency, and prototypicality in VAC processing*

In order to investigate these factors in VAC processing, the first step is an analysis of the relevant statistics in a large corpus of representative usage, the second is a psycholinguistic analysis of the processing of VACs selected to vary on these dimensions.

#### 1.2. CORPUS ANALYSIS OF VACs IN USAGE

Ellis and O'Donnell (2011, 2012) investigated the type–token distributions of 20 Verb–Locative (VL) VACs such as 'V(erb) *across* n(oun phrase)' in the British National Corpus (BNC, 2007), a 100-million-word corpus of English usage. The other locatives sampled were *about, after, against, among, around, as, at, between, for, in, into, like, of, off, over, through, towards, under, and with*. They searched a dependency-parsed version of the BNC for specific VACs previously identified in the Grammar Patterns volume resulting from the COBUILD corpus-based dictionary project (Francis, Hunston, & Manning, 1996). The details of the linguistic analyses, as well as subsequently modified search specifications in order to improve precision and recall, are fully described in Ellis and O'Donnell (2011, 2012) and Römer, O'Donnell, and Ellis (2015). This corpus linguistic research demonstrated:

1. The frequency profile of the verbs in each VAC follows a Zipfian profile (Zipf, 1935) whereby a few verbs take the lion's share: the highest-frequency types account for the most linguistic tokens. Zipf's law states that, in human language, the frequency of words decreases as a power function of their rank: the most frequent verb occurs roughly twice as often as the second most frequent, roughly three times as often as the third most frequent, etc.
2. VACs are selective in their verb form family occupancy: individual verbs select particular constructions; particular constructions select particular verbs; there is high contingency ( $\Delta P$  as described above in Section 1.1.2) between verb types and constructions. This means that

the Zipfian profiles seen in point 1 above are not those of the verbs in English as a whole – instead their constituency and rank ordering are special to each VAC.

3. The most frequent verb in each VAC is prototypical of that construction's functional interpretation, albeit generic in its action semantics.
4. VACs are coherent in their semantics. This was assessed using WordNet (Miller, 2009), a distribution-free semantic database based upon psycholinguistic theory, as an initial resource to investigate the similarity/distance between verbs. Then networks science algorithms (de Nooy, Mrvar, & Batagelj, 2010) were used to build semantic networks in which the nodes represent verb types and the edges strong semantic similarity for each VAC. Standard measures of network density, average clustering, degree centrality, transitivity, etc. were then used to assess the cohesion of these semantic networks and verb type connectivity within the network. Betweenness centrality was used as a measure of a verb node's centrality in the VAC network (McDonough & De Vleeschauwer, 2012). Betweenness centrality was developed to quantify the brokerage role played by an individual between other humans in a social network (Freeman, 1977). It is defined as the number of shortest paths from all nodes to all others that pass through that node. In semantic networks, central nodes are those which are prototypical of the network as a whole.

These corpus analyses thus demonstrated that the cognitive principles of categorization reviewed in Section 1.1 applied in usage. But what about in human cognition?

### 1.3. ANALYSIS OF KNOWLEDGE OF VACs

Ellis et al. (2014) used free association and verbal fluency tasks to investigate verb–argument constructions (VACs) and the ways in which their processing is sensitive to these statistical patterns of usage (verb type–token frequency distribution, VAC–verb contingency, verb–VAC semantic prototypicality). In Experiment 1, 285 native speakers of English generated the first word that came to mind to fill the V slot in 40 sparse VAC frames such as 'he \_\_ across the ...', 'it \_\_ of the ...', etc. In Experiment 2, 40 English speakers generated as many verbs that fit each frame as they could think of in a minute. For each VAC, Ellis et al. compared the results from the experiments with the corpus analyses of usage described above in Section 1.2. For both experiments, multiple regression analyses predicting the frequencies of verb types generated for each VAC showed independent contributions of (i) verb frequency in the VAC, (ii) VAC–verb contingency, and (iii) verb prototypicality in terms of centrality within the VAC semantic network. Ellis et al. (2014) contend that

the fact that native-speaker VACs implicitly represent the statistics of language usage implies that they are learned from usage.

#### 1.4. MOTIVATIONS FOR THE CURRENT EXPERIMENTS

These findings show that lexis, syntax, and semantics are richly associated in VAC processing. However, free-association tasks can involve conscious rather than automatic processing, especially those achieved over the time span of a minute. Various deliberate search strategies can come to play. It is difficult to conclude, therefore, that these results imply that VACs are ‘mentally represented’ as part of the constructicon. Although the findings are compatible with that idea, they are not conclusive. For example, the native speakers in the one-minute tasks might be building ad hoc categories (Barsalou, 2010) based on information (such as frequency information, contingencies, etc.) in order to engage in the association task. An ad hoc category is a novel category constructed spontaneously to achieve a goal relevant in the current situation (e.g., constructing *WAYS OF CATCHING MOLES* while seeing their destruction of the lawn). These categories are novel – they have not been entertained previously. They are constructed spontaneously and do not reside as knowledge structures in long-term memory waiting to be retrieved. They help achieve a task-relevant goal by organizing knowledge relevant to the current situation in ways that support effective goal pursuit.

Therefore, the data provided in the free-association data do not force the conclusion that frequency, contingency, and prototypicality of verb–frame pairings are mentally represented as separate VACs. Instead, online processing experiments are needed to explore the generality of these findings and their implications for representation. In the remainder of this paper I report on three experiments which focus on construction access, paradigmatic associations, and processing for meaning. Ellis (2016) reports a parallel line of investigations which focuses upon the statistical binding of syntagmatic VAC forms, firstly for recognition, then for naming.

There is a rich psycholinguistic tradition investigating effects of such factors as frequency, contingency, prototypicality, imagery, semantics, neighborhood density, word length, spelling regularity, morphological transparency, and orthographic depth in *LEXICAL* access and processing (e.g., Cortese & Balota, 2012; Gaskell, 2007; Gernsbacher, 1994; Meyer & Schvaneveldt, 1971; Seidenberg & McClelland, 1989; Traxler & Gernsbacher, 2011). Demonstrations of the effects of these factors upon automatic processing are taken as indications that lexical constructions are stored in long-term memory rather than constructed ad hoc. The current experiments therefore adapt relevant lexical processing paradigms and apply them to construction access and processing. If these reveal the same sorts of effects, it encourages the



conception of a unified construction where words and VACs alike are symbolic representations, acquired from usage, statistics and all, with their subsequent processing tuned probabilistically to usage experience.

## 2. Experiment 1: lexical decision

The lexical decision task is one of the most commonly used psycholinguistic techniques to study word recognition, lexical access, and the organization of semantic memory. The procedure involves measuring how quickly people classify letter strings as words or nonwords. Lexical decision latencies are faster for words than nonwords and are sensitive to word frequency, increasing by a constant number of milliseconds for each log unit of frequency (Rubenstein & Pollack, 1963; Scarborough, Cortese, & Scarborough, 1977). In a study of the lexical decision and naming of 2,428 monosyllabic words, Balota, Cortese, Sergent-Marshall, Spieler, and Yap (2004) found that semantic factors such as imageability and the semantic connectivity between a word and other words had effects above and beyond other lexical and sublexical factors such as frequency and neighborhood density.

Meyer and Schvaneveldt (1971) adapted the paradigm to provide a simple, but powerful, empirical and theoretical approach to studying subconscious mental structures and processes whereby people represent and retrieve information in long-term memory. In their classic experiment they measured response times as people made lexical decisions when a pair of letter strings, presented simultaneously, were both words. In conditions in which both stimuli were words, some of the pairs were related (e.g., DOCTOR and NURSE) and others were unrelated (e.g., CHAIR and FLOWER). The key finding was that response time was faster for related words than for unrelated words, consistent with the concept of spreading activation: as a consequence of prior information processing, related concepts in long-term memory become prepared or 'primed' for later use, speeding recognition when people subsequently encounter other words associated with them. The semantic-priming phenomenon revealed by such facilitation opened new windows onto lexical and long-term memory organization and subconscious mental processes arising from prior exposure to successive related concepts. The work rapidly became a 'citation classic' in the Scientific Citation Index. Measures of priming effects have become some of the most widely published dependent variables in cognitive science (McNamara, 2005) as well as in other fields, for example in the Implicit Association Test that is much used in social psychology (Greenwald & Banaji, 1995).

Neely (1991) reviewed twenty years of research of findings and theories of semantic priming in visual word recognition. Various mechanisms for priming have been implicated, including both semantic priming when the prime and

the target are from the same semantic category (e.g., DOCTOR and NURSE), and associative priming where the prime and the target are associated but are from different semantic categories (e.g., RAKE and LEAF), and whether the effects might be mediated by lexical and/or conceptual relations.

This technique is therefore ripe for application to study VAC processing. Furthermore, since from our prior research (see Ellis et al., 2016) we have measures of Verb Frequency, Verb-VAC association frequency, VAC-verb contingency, and Verb-VAC semantic prototypicality, we might be able to separately identify the degree to which these different lexical, syntagmatic, and semantic dimensions are represented so to affect lexical decision latency.

## 2.1. PARTICIPANTS

The participants were forty-nine university students at a large mid-western university taking an introductory course in psychology and participating in the subject pool for course requirement. The age range was 18–22 ( $M = 18.42$ ,  $SD = 0.76$ ). Sixteen were male, thirty-three were female. Thirty reported knowing one, sixteen knowing two, and three knowing three languages. Forty-one reported that English was their first language.

## 2.2. METHOD

### 2.2.1. Stimulus materials

Ellis et al. (2014) identified the verb lemmas which together covered the top 95% of verb token uses in the BNC. They counted their token frequencies in the BNC (Verb Frequency), along with the frequency with which they occupied Verb–Locative (VL) VACs such as ‘V(erb) *across* n(oun phrase)’ (Verb-VAC frequency), the contingency between construction and word ( $\Delta P_{cw}$ ), and the semantic prototypicality of the verb in the construction (betweenness centrality). The range of VL VACs included *about*, *across*, *against*, *among*, *around*, *between*, *for*, *into*, *like*, *of*, *off*, *over*, *through*, *towards*, *under*, *with*. The current experiment required a subset of stimuli which as far as possible factorially manipulated these dimensions, keeping them as independent as possible. The first step, therefore, was to regress each of the factors against the others. So, for example,  $\log_{10}VACfrequency$  was regressed against  $\log_{10}corpusfrequency$ ,  $\log_{10}\Delta P_{cw}$ , and  $\log_{10}centrality$ , and the  $\log_{10}VACfrequency$  residuals were saved for each verb. In similar fashion,  $\log_{10}\Delta P_{cw}$  was regressed against  $\log_{10}corpusfrequency$ ,  $\log_{10}VACfrequency$ , and  $\log_{10}centrality$ , and the  $\log_{10}\Delta P_{cw}$  residuals were saved for each verb. And so on. Thus, for a verb-VAC pairing, we knew whether a verb was particularly high (or low) on one of these dimensions against the background of what might be expected from the levels of the other predictors. For each VAC,

we then chose example verbs which reflected high, medium, and low semantic prototypicality, high, medium, and low VACfrequency, and high, medium, and low  $\Delta P_{cw}$ . We also selected high (+), medium (0), and low (-) corpus frequency verbs which never appear in the construction. We stripped the VACs down from 'V(erb) preposition n(oun phrase)' to their bare minimum, i.e., the verb preposition collocation. Examples for the case of 'V about n' are sem+ *move about*; sem0 *float about*; sem- *lie about*; vacfreq+ *chat about*; vacfreq0 *jump about*; vacfreq- *point about*;  $\Delta P+$  *talk about*;  $\Delta P0$  *understand about*;  $\Delta P-$  *tell about*; never *reduce about*; never *catch about*; never *appoint about*. The complete set of 192 stimuli so constructed is shown in Appendix A, shown in the supplementary material online (available at <<http://dx.doi.org/10.1017/langcog.2016.18>>), alongside their Verb Frequency, Verb-VAC frequency, VAC-verb contingency, and Verb-VAC semantic prototypicality. These steps did not achieve complete orthogonality, but they did reduce the association of these predictors from the higher levels typically found in natural language to those correlations shown in Table B1 in the online supplementary material.

We used the complete set of 192 stimuli in Appendix A (see online supplementary material) as well as, for each, a yoked stimulus which had a nonword substituted in either the verb or preposition slot: half were randomly chosen for each. The nonwords were generated using the ARC Nonword Database (Rastle, Harrington, & Coltheart, 2002), selecting nonwords between 2 and 8 letters in length which had orthographically existing onsets, orthographically existing bodies, and legal bigrams in English. These steps resulted in there being 384 stimuli in all.

### 2.2.2. Procedure

The experiment was scripted in *PsychoPy* v1.80.03 (Peirce, 2007) and run on iMac computers. Participants were instructed that they would be shown two letter strings, side by side. First they would see a fixation point, then the strings would appear. Their task was to judge whether both of these are words or not as quickly as possible after they appeared, pressing 'm' if both are words, or 'z' if one of them is not a word. The experiment began with sixteen practice items which paired verbs not used in the experiment proper with the VAC prepositions (e.g., *bring about*, *meet across*, *set against*); again, half had nonword substitutions. Trial order was randomized individually for each participant. We recorded the reaction time from stimulus onset, as well as the correctness of the response. The experiment as a whole took between 30 and 45 minutes.

The data files for all participants were concatenated. We analyzed only the trials where the stimuli were both words. In order to remove outliers,

RT data were Winsorized within each participant, trimming 5% of responses: for each participant, this set RTs below the 2.5th percentile to the value of the 2.5th percentile, and RTs above the 97.5th percentile to the 97.5th percentile. Over all participants and items, 96.8% were judged correctly to be valid lexical strings with a mean judgment RT of 0.78 sec. We log transformed the Winsorized RTs.

### 2.3. RESULTS

#### 2.3.1. *Lexical decision RT*

Figure 1 shows the means of the judgment RTs for each stimulus string as a function of Verb Frequency, Verb-VAC frequency, VAC-verb contingency, Verb-VAC semantic prototypicality, and VAC length. To assess their independent effects, we performed a glmm of  $\log_{10}RT$  against the five predictors with participant and VAC as independent random intercepts using the R package lme4 (Bates, Maechler, Bolker, & Walker, 2015). The summary results are shown in Table B2 in the supplementary material online, where it can be seen that there are separate independent effects of Verb Frequency ( $t = -8.45$ ), Verb-VAC frequency ( $t = -2.73$ ), and Verb-VAC semantic prototypicality ( $t = -2.37$ ). Increasing frequency of the verb in the language as a whole, frequency of the verb in the VAC, and verb semantic prototypicality all increase the speed of making a lexical decision. In order to graph these separate effects, we used the *R* library by Fox (2003). To obtain a model without random effects, we ran a glm of the L10RTs against our five predictors and plotted the independent effects to the same scale in Figure 1.

#### 2.3.2. *Lexical decision judgments*

We ran the same glmm model (though with family=binomial) on the judgments of lexicality. The summary results are shown in Table B3 in the supplementary material online, where it can be seen that there are separate independent effects of each of the predictors: Verb Frequency ( $z = 6.38$ ), Verb-VAC frequency ( $z = 1.75$ ), semantic prototypicality ( $z = 2.96$ ), and stimulus length ( $z = 5.59$ ). As with the RTs, increasing frequency of the verb in the language as a whole, frequency of the verb in the VAC, and verb semantic prototypicality all increase the likelihood of an all-word string to be judged as such.

### 2.4. INTERIM DISCUSSION

Both the latency and judgment data show that VAC lexical decisions are a function of word frequency in the language, of syntagmatic factors (Verb-VAC frequency), and of semantic factors (Verb-VAC semantic prototypicality).

Lexical Decision Adjacent Words RT

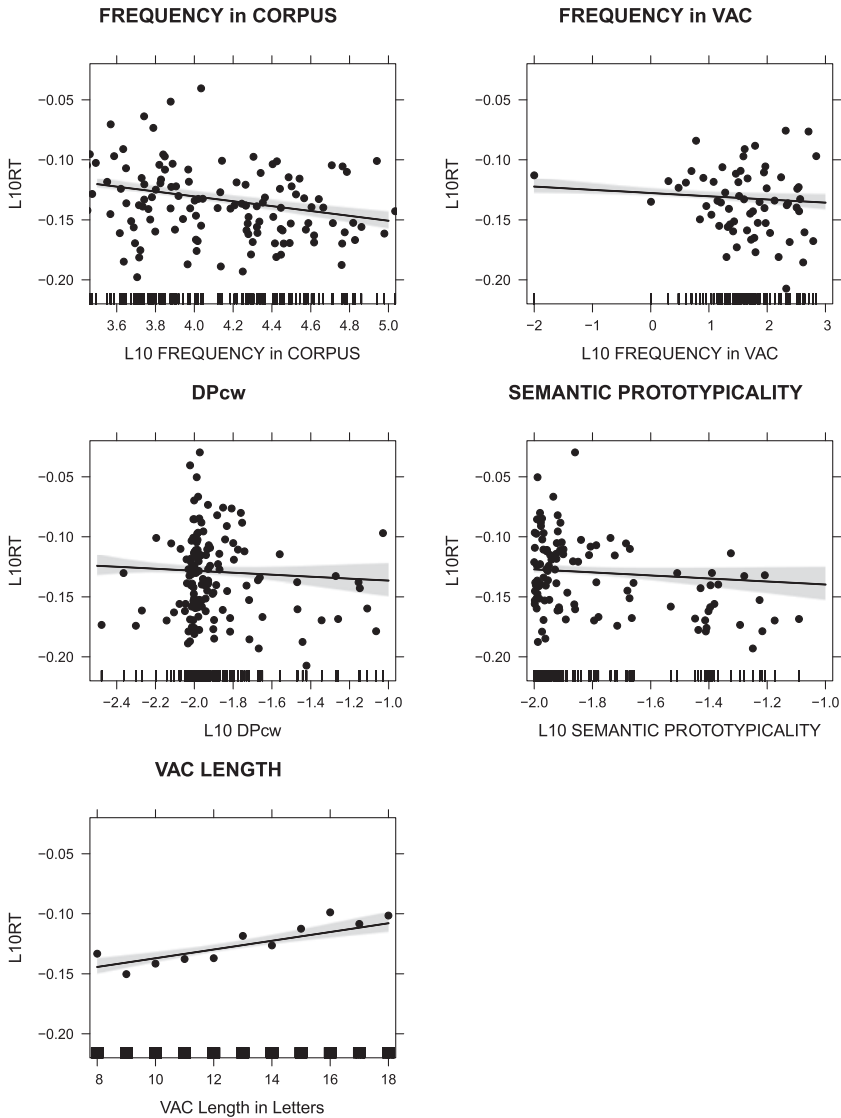


Fig. 1. Independent effect sizes of (a) frequency of the verb in the corpus, (b) frequency of the verb in the VAC, (c) VAC-Verb contingency ( $\Delta P_{cw}$ ), (d) verb semantic prototypicality (betweenness centrality), and (e) stimulus length upon adjacent lexical decision RT.

These findings reinforce an interpretation of VACs as symbolic bindings involving both form and meaning according to the strengths of their association in language usage.

### 3. Experiment 2: lexical decision of interposed constituents

One possible criticism is that much of these effects might be largely due to mere collocation, i.e., to transitional probabilities among immediately sequential words. McDonald and Shillcock (2003) demonstrated how statistical information latent in the linguistic environment can contribute to reading behavior. Using eye-tracking they demonstrated that the transitional probabilities between words had a measurable influence on fixation durations, and using a simple Bayesian statistical model they showed that lexical probabilities derived by combining transitional probability with the prior probability of a word's occurrence provided the most parsimonious account of the eye-movement data. They suggested that the brain is able to draw upon statistical information in order to rapidly estimate the lexical probabilities of upcoming words: a computationally inexpensive mechanism that may contribute to proficient reading. Such exploitation of transitional probabilities in reading could be an essential part of the rationality of usage-based processing. But it would be good to know if additionally VAC processing is more structurally abstract than this, i.e., if it reflected the binding of a verb and its VAC preposition even if these are non-contiguous.

In this experiment, therefore, we repeated the lexical decision study but now with three words, interrupting the verb–preposition exemplars of Experiment 1 by inserting a randomly chosen, unassociated, intervening item (an adverb from the set of *quickly, happily, sadly, easily, gladly, wildly, calmly, always, often, carefully, quietly*) with the result that, while the verb–preposition associations were of the same order as in Experiment 1, the trigram stimuli were themselves of very low transitional probability.

This procedure is similar to that of Schvaneveldt and Meyer (1973), who interposed an unrelated category between the related prime and target (e.g., DOCTOR PAPER NURSE) and had subjects make a 'yes' response if all three stimuli were words and a 'no' response otherwise. With this procedure, priming still occurred, even though the subjects were sequentially processing the three simultaneously presented items (see reviews by Masson, 1991; Neely, 1991).

#### 3.1. PARTICIPANTS

The participants were forty-three university students at a large mid-western university taking an introductory course in psychology and participating in

the subject pool for course requirement. The age range was 18–23 ( $M = 18.56$ ,  $SD = 0.88$ ). Twelve were male, thirty-one were female. Twenty-eight reported knowing one, eleven knowing two, and four knowing three languages. Forty reported that English was their first language.

### 3.2. METHOD

#### 3.2.1. Stimulus materials

We used the same 192 VAC stimuli as in Experiment 1, except that for the twelve exemplars for each VAC preposition, we randomly inserted one adverb from the set of *quickly*, *happily*, *sadly*, *easily*, *gladly*, *wildly*, *calmly*, *always*, *often*, *carefully*, *quietly* between the verb and the preposition. To get a feel for the effects of this manipulation on overall meaning, consider how these additions affected the exemplars for the *between* and *for* VACs shown in Appendix A in the supplementary material online, which became: *ran slowly between*; *paused quickly between*; *opened happily between*; *remembered sadly between*; *switched easily between*; *transferred gladly between*; *checked wildly between*; *granted calmly between*; *distinguish always between*; *spills often between*; *worked carefully between*; *coincide quietly between*; *holds slowly for*; *proceeds quickly for*; *sat happily for*; *protects sadly for*; *opted easily for*; *flowed gladly for*; *advised wildly for*; *reminds calmly for*; *asked always for*; *display often for*; *departed carefully for*; *deem quietly for*; etc.

As in Experiment 1, for each of the 192 all-word responses there was a yoked control item where one of the three elements was replaced by a nonword that was generated using the ARC Nonword Database (Rastle et al., 2002), selecting nonwords between 2 and 8 letters in length which had orthographically existing onsets, orthographically existing bodies, and legal bigrams in English. The verb, the adverb, and the preposition were replaced equally often (64 times), so nonword status was independent of serial position in the VAC. These steps resulted in there being 384 stimuli in all.

#### 3.2.2. Procedure

The experiment was scripted in *PsychoPy* v1.80.03 (Peirce, 2007) and run on iMac computers. Participants were instructed that they would be shown three letter strings, side by side. First they would see a fixation point, then the strings would appear. Their task was to judge whether all three of these are words or not as quickly as possible after they appeared, pressing ‘m’ if all are words, or ‘z’ if one of them is not a word. The experiment began with sixteen practice items which paired verbs not used in the experiment proper with the random adverbs and then the VAC prepositions; again, half had nonword substitutions. Trial order was randomized individually for each participant.

We recorded the reaction time from stimulus onset, as well as the correctness of the response. The experiment as a whole took between 30 minutes and 45 minutes.

As in Experiment 1, the data files for all participants were concatenated. We analyzed only the trials where the stimuli were all words. In order to remove outliers, RT data were Winsorized within each participant, trimming 5% of responses: For each participant, this set RTs below the 2.5th percentile to the value of the 2.5th percentile, and RTs above the 97.5th percentile to the 97.5th percentile. Over all participants and items, 95.4% were judged correctly to be valid lexical strings with a mean judgment RT of 0.96 sec. We log transformed the Winsorized RTs.

### 3.3. RESULTS

#### 3.3.1. *Lexical decision RT*

Figure 2 shows the means of the judgment RTs for each stimulus string as a function of Verb Frequency, Verb-VAC frequency, VAC-verb contingency, Verb-VAC semantic prototypicality, and VAC length. To assess their independent effects, we performed a glmm of  $\log_{10}RT$  against the five predictors with participant and VAC as independent random intercepts using the R package *lme4* (Bates et al., 2015). The summary results are shown in Table B4 in the supplementary online material, where it can be seen that there are separate independent effects of Verb Frequency ( $t = -8.18$ ), Verb-VAC frequency ( $t = -2.89$ ), and Verb-VAC semantic prototypicality ( $t = -2.78$ ). Increasing frequency of the verb in the language as a whole, frequency of the verb in the VAC, and verb semantic prototypicality all increase the speed of making a lexical decision. In order to graph these separate effects, we used the *R* library by Fox (2003). To obtain a model without random effects, we ran a glm of the L10RTs against our five predictors and plotted the independent effects to the same scale in Figure 2.

#### 3.3.2. *Lexical decision judgments*

We ran the same glmm model (though with `family=binomial`) on the judgments of lexicality. The summary results are shown in Table B5 in the online supplementary material, where it can be seen that there are separate independent effects of each of the predictors: Verb Frequency ( $z = 7.08$ ), Verb-VAC frequency ( $z = 2.13$ ), semantic prototypicality ( $z = 2.72$ ), and stimulus length ( $z = 7.90$ ). As with the RTs, increasing frequency of the verb in the language as a whole, frequency of the verb in the VAC, and verb semantic prototypicality all increase the likelihood of an all-word string to be judged as such.



Lexical Decision Interrupted Words RT

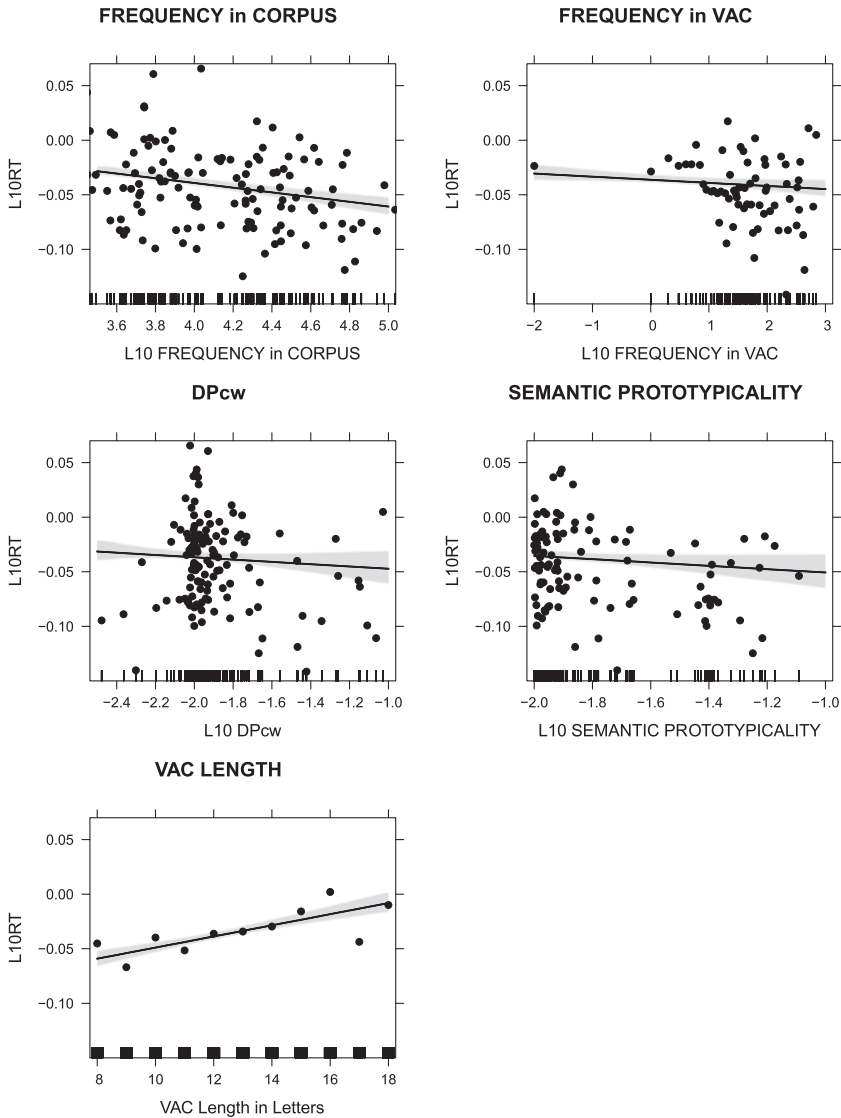


Fig. 2. Independent effect sizes of (a) frequency of the verb in the corpus, (b) frequency of the verb in the VAC, (c) VAC-Verb contingency ( $\Delta P_{cw}$ ), (d) verb semantic prototypicality (betweenness centrality), and (e) stimulus length upon interrupted lexical decision RT.

### 3.4. INTERIM DISCUSSION

These results are remarkably like those of Experiment 1. Adding a random intervening adverb, which breaks up the verb–preposition collocation but which does not systematically affect the verb-VAC meaning, simply increases the overall processing time (from  $M = 0.78$  sec to  $M = 0.96$  sec) and marginally drops the hit rate from 96.8% to 95.4%, but it does not change the pattern of variables which affect processing: effects of Verb-VAC frequency ( $t = -2.89$ ) and Verb-VAC semantic prototypicality ( $t = -2.78$ ) are maintained. We conclude that VAC associations are between the constituents themselves rather than between merely sequential elements.

## 4. Experiment 3: judging the meaningfulness of VACs

Experiments 1 and 2 show automatic effects of spread of semantic association in lexical decision tasks which in themselves do not explicitly call for semantic processing. What happens when we do direct participants to process the stimuli for meaning? In order to study respondents' processing of the meaning of word sequences that is as fast and automatic as possible, we designed a very simple task where participants were asked to consciously judge, as quickly as possible, whether the two-word sequences of Experiment 1 made sense to them or not. As in Experiment 1, they were given two possible responses, 'yes' and 'no', but were otherwise left to their own devices – no feedback was given.

### 4.1. PARTICIPANTS

The participants were forty-five university students at a large mid-western university taking an introductory course in psychology and participating in the subject pool for course requirement. The age range was 17–23 ( $M = 18.64$ ,  $SD = 1.30$ ). Eighteen were male, twenty-seven were female. Thirty-one reported knowing one, eleven knowing two, and three knowing three languages. Forty reported that English was their first language.

### 4.2. METHOD

#### 4.2.1. Stimulus materials

We generated a control item for each of these 192 VAC exemplars by randomly yoking other particles (e.g., *since*, *round*, *during*) to the originals. For example, for the VAC collocations *break against*, *crash against*, *stand against*, there was a matching pair, *break during*, *crash during*, *stand during*, etc. The arbitrary pairing were designed to give less meaningful foils than the authentic VACs, so as to allow some baseline against which meaning could be judged. We did

not analyze the meaningfulness as judged by our respondents for these arbitrary pairings. Instead we assessed the effects of Stimulus Length, Verb Frequency, Verb-VAC frequency, VAC-verb contingency, and Verb-VAC semantic prototypicality upon meaningfulness rating and meaningfulness rating RT (ms.) for only the authentic exemplars.

#### 4.2.2. Procedure

The experiment was scripted in *PsychoPy* v1.80.03 (Peirce, 2007) and run on iMac computers. Participants were instructed that there would be two words, one after another, and that their task was to judge whether these make sense to them or not as quickly as possible after they appear by pressing the relevant key on the keyboard (*m* for yes, *z* for no). Participants pressed the space bar when they were ready for the next trial. One second later, the first word of the pair appeared. Two hundred and fifty ms later, the second word appeared underneath it. After 16 practice trials, there followed the 384 trials presented in a different random order for each participant. The experiment as a whole took between thirty and forty minutes.

### 4.3. RESULTS

The data files for all participants were concatenated. We analyzed RTs and judgments for the 192 VAC exemplars rather than their control items. Over these items, the mean judgment RT was 1.086 sec, and 60% were judged to be meaningful. Remember that 25% of pairings were never found in the corpus (e.g., never *reduce about*; never *catch about*; never *appoint about*) and so were effectively meaningless, so the baseline is 75%. See Appendix A in the supplementary material online for the items used. We log transformed the RTs. We separately analyzed judgment RT and meaningfulness.

#### 4.3.1. Meaning judgment RTs

To assess the independent effects of VAC Stimulus Length (in letters), Verb Frequency, Verb-VAC frequency, VAC-verb contingency, and Verb-VAC semantic prototypicality, we performed a glm of  $\log_{10}RT$  against the five predictors with participant and VAC as independent random intercepts using the R package *lme4* (Bates et al., 2015). The summary results are shown in Table B6 in the online supplementary material, where it can be seen that there are separate independent effects of Verb Frequency ( $t = 2.24$ ), Verb-VAC frequency ( $t = -4.90$ ), and VAC-verb contingency  $\Delta P_{cw}$  ( $t = -5.87$ ). The overall frequency of the verb in the language slows meaningfulness

judgment, whereas increased frequency of the verb in the VAC and  $\Delta P_{cw}$  both lead to faster judgments.

In order to graph separate effects, we used the *R* library by Fox (2003), again from a model without random intercepts – a glm of the L10RTs against our five predictors – plotting the independent effects to the same scale in Figure 3.

#### 4.3.2. *Meaning judgments*

We ran the same model (though with family=binomial) on the judgments of meaningfulness. The summary results are shown in Table B7 in the online supplementary material, where it can be seen that there are separate independent effects of each of the predictors: Verb Frequency ( $z = -8.30$ ), Verb-VAC frequency ( $z = 26.25$ ), and VAC-verb contingency ( $z = 9.26$ ). As with the RTs, the overall frequency of the verb in the language detracts from meaningfulness, while the frequency of the verb in the VAC and  $\Delta P_{cw}$  both have significant positive effects.

#### 4.3.3. *Interim discussion*

Clearly the meaning judgment task encourages participants to do something beyond mere lexical access – while the mean RT in Experiment 1 was  $M = 0.78$  sec, the mean judgment RT in Experiment 3 was 1.086 sec. The RT and judgment data together show the same effects. Higher-frequency verbs are somewhat less meaningful in VAC judgment. In contrast, the higher the conditional frequency of the verb in the VAC, and the higher the contingency, the more meaningful the VAC is judged to be, and the faster the judgment is made. Against our initial expectations, and contra the results for the Lexical Decision Experiments 1 and 2, there is no effect of semantic prototypicality. I shall return to this below.

## 5. **General discussion**

Note that the same statistical models were used for analysis across the three experiments. Experiment 1, involving adjacent constituents, demonstrated effects of Verb Frequency, Verb-VAC frequency, and Verb-VAC semantic prototypicality upon lexical decision RT, and of Verb Frequency, Verb-VAC frequency, semantic prototypicality, and stimulus length upon judgments of lexicality itself. Experiment 2, involving interposed constituents, gave very similar results: effects of Verb Frequency, Verb-VAC frequency, and Verb Frequency, Verb-VAC frequency, upon lexical decision RT, and Verb Frequency, Verb-VAC frequency, semantic prototypicality, and stimulus length upon

Semantic Decision RT

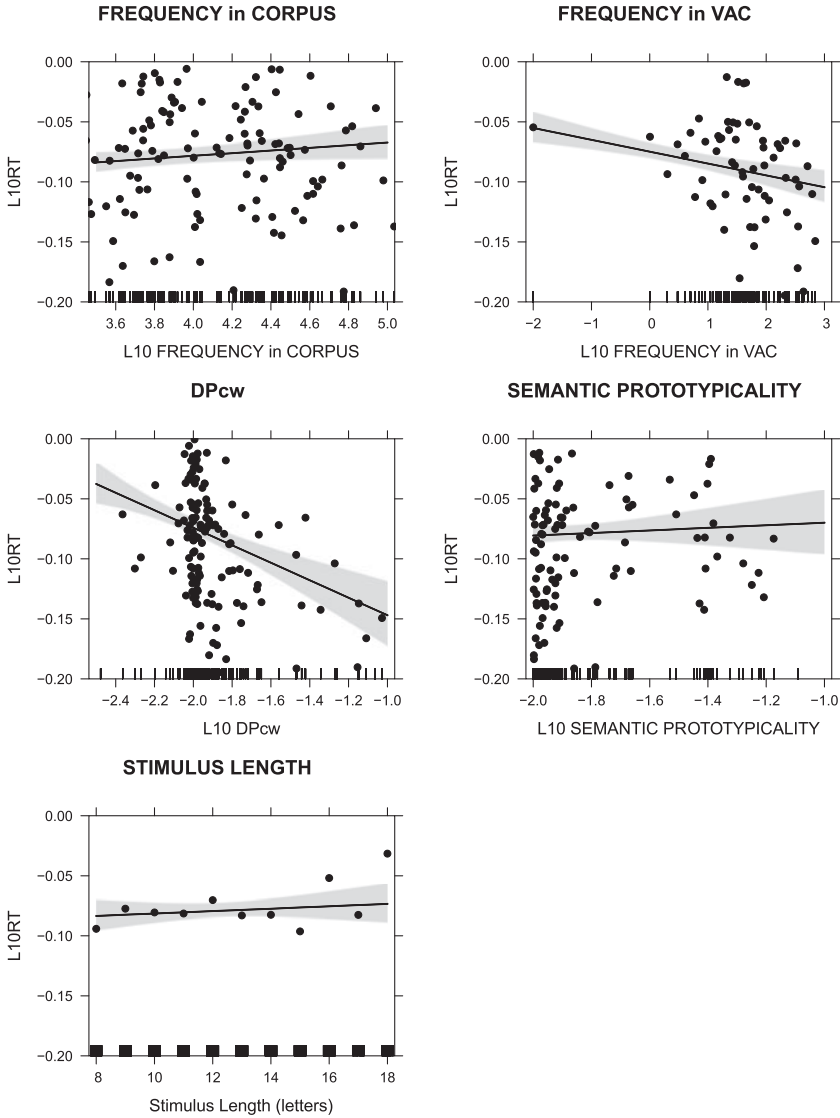


Fig. 3. Independent effect sizes of (a) frequency of the verb in the corpus, (b) frequency of the verb in the VAC, (c) VAC-Verb contingency ( $\Delta P_{cw}$ ), (d) verb semantic prototypicality (betweenness centrality), and (e) stimulus length upon meaningfulness judgment RT.

judgments of lexicality itself. Experiment 3 demonstrated independent effects of Verb Frequency, Verb-VAC frequency, and VAC-verb contingency  $\Delta P_{cw}$  upon meaningfulness judgment RT, and of Verb Frequency, Verb-VAC frequency, and VAC-verb contingency upon judged meaningfulness itself. Thus, all of the experiments and outcome measures here show effects of Verb Frequency and Verb-VAC frequency. These statistics of verb-VAC type token frequency are clearly represented. However, lexical decision is additionally driven by semantic prototypicality (but not VAC-verb contingency  $\Delta P_{cw}$ ), whereas meaning judgment is affected VAC-verb contingency  $\Delta P_{cw}$  (but not semantic prototypicality). Let us consider each in turn.

### 5.1. LEXICAL DECISION

It is standard that the recognition of individual words is a function of their prior experience as indexed by word frequency in the language (Balota, Yap, & Cortese, 2006). The finding that lexical decisions of VACs is affected by the frequency of the verb is thus no surprise. The effect of Verb-VAC frequency is more potent: perception is sensitive to the pairing of the verb and the VAC. This could reflect sensitivity to syntagmatic sequence, i.e., their collocation, or it could reflect sensitivity to the binding of the verb to the VAC as a whole, meaning and all.

There are many other demonstrations that language users have implicit knowledge of sequences of language (for review see Ellis, 2012). For example, reading time is affected by collocational and sequential probabilities. Bod (2001), using a lexical-decision task, showed that high-frequency three-word sentences such as “I like it” were reacted to faster than low-frequency sentences such as “I keep it”, by native speakers. Ellis, Frey, and Jalkanen (2009) used lexical decision to demonstrate that native speakers preferentially process frequent verb–argument and booster/maximizer–adjective two-word collocations. Durrant and Doherty (2010) used lexical decision to assess the degree to which the first word of low- (e.g., *famous saying*), middle- (*recent figures*), high-frequency (*foreign debt*), and high-frequency and psychologically associated (*estate agent*) collocations primed the processing of the second word in native speakers. The highly frequent and high-frequency associated collocations evidenced significant priming. Arnon and Snider (2010) used a phrasal decision task (“Is this phrase possible in English or not?”) to show that comprehenders are also sensitive to the frequencies of compositional four-word phrases: more frequent phrases (e.g., *don’t have to worry*) were processed faster than less-frequent phrases (*don’t have to wait*), even though these were matched for the frequency of the individual words or substrings.

The replication of the eye-tracking demonstrations of anticipation of words based on transitional probabilities by McDonald and Shillcock (2003) for

VACs rather than words in our results in Experiment 2, where the verb–preposition pairings were discontinuous as a result of interposed adverbs, shows that the relevant statistics are not simple bigram collocations, but instead it is the probabilistic association between the verb and the VAC that is represented.

The additional independent effects of verb prototypicality in Experiments 1 and 2 show that these are not mere syntagmatic effects, but rather that VAC meaning is represented as well, and that VACs containing verbs which are more semantically central are processed faster. These results hark back to the semantic priming effects first observed in two-word and in three-word interrupted lexical decision (Meyer & Schvaneveldt, 1971), which they took as evidence of spreading semantic activation rather than facilitated lexical access.

A relevant conceptualization is that of interactive-activation in connectionist models of lexical processing (Balota et al., 2006; McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982; Seidenberg & McClelland, 1989). Models with multiple independent layers of detectors (features, letters, words, meanings), with mutual inhibition of units within levels, but activation cascading both upwards and downwards between these levels, allow partial activation of meaning-level activations to in turn partially activate the representations that produced those representations (Balota, Ferraro, & Connor, 1991, p. 213; Balota et al., 2006; Steyvers & Tenenbaum, 2005). Seeing *Jump* activates the VL VACs with which *jump* is associated, which activates *VL-down* semantic space, which in turn sends activation downwards to the logogen for *down*, making it more likely to fire. It is not just statistical association between word forms (that's the effect of verb-VAC frequency). It really involves semantics, because additionally, verbs more prototypical of the VAC semantic meaning cause greater activation.

## 5.2. ASSESSING MEANING

Experiment 3 demonstrated independent effects of Verb Frequency, Verb-VAC frequency, and VAC-verb contingency  $\Delta P_{cw}$  upon meaningfulness judgment RT, and of Verb Frequency, Verb-VAC frequency, and VAC-verb contingency upon judged meaningfulness itself. Both outcome measures show effects of Verb Frequency and Verb-VAC frequency. These statistics of verb-VAC type token frequency clearly affect processing.

However, while lexical decision (Experiments 1 and 2) was strongly driven by semantic prototypicality (but not VAC-verb contingency  $\Delta P_{cw}$ ), meaning judgment (Experiment 3) was affected by VAC-verb contingency  $\Delta P_{cw}$ , but not semantic prototypicality. This was not expected. We included meaning judgment as a task because we expected semantic prototypicality to have a greater

effect here. This speeded and open meaning judgment task has not been widely used in past research, and so there are few direct leads in the literature.

One finding that seems relevant is from Ellis and Ferreira-Junior (2009), who asked people to rate the prototypicality of verbs in VL VACs. The verb *go* was rated as 7.4 out of 9 in terms of the degree to which it matched the prototypical schematic meaning (even though in network analyses it is quite central). However, a number of other more peripheral verbs received a higher rating: *walk* (9.0), *move* (8.8), *run* (8.8), *travel* (8.8), *come* (8.4), *drive* (8.2), *arrive* (8.0), *jump* (8.0), *return* (8.0), and *fall* (7.8). The same occurred for the VOL VAC, where *put* was rated 8.0 in terms of how well it described the construction schema, yet it was surpassed in the ratings by *bring* (8.6), *move* (8.6), *send* (8.6), *take* (8.6), *carry* (8.4), *drive* (8.4), *drop* (8.4), *pass* (8.4), *push* (8.4), *hit* (8.2), and *pull* (8.2), which are all more specific in their action semantics.

Prototypical verbs, by dint of their wide usage, have less-specific meanings and are less imageable. Toglia and Battig (1978) report information derived from college students' ratings of a large number and variety of individual words (and some nonwords) for seven basic semantic characteristics (concreteness, imagery, familiarity, pleasantness, number of attributes or features, categorizability, and meaningfulness). They do not include all of our verbs, and we should remember that they presented verbs alone, out of VAC context. Their ratings for the VL verbs used in the present experiments were: for imageability *go* (364), *walk* (470), *move* (428), *travel* (520), *come* (408), *jump* (506); for meaningfulness *go* (430), *walk* (505), *move* (413), *travel* (506), *come* (322), *jump* (466). For our VOL verbs, the imageability ratings were *put* (263), *move* (413), *send* (423), *take* (337), *carry* (393), *drop* (417), *pass* (479), *pull* (446); the meaningfulness ratings *put* (297), *move* (428), *send* (384), *take* (360), *carry* (436), *drop* (400), *pass* (440), *pull* (410). Prototypical verbs are semantically general and rather less imageable, hence their often being called LIGHT VERBS (Clark, 1978; Ninio, 1999; Pinker, 1989). Theakston, Lieven, Pine, and Rowland (2004) list the range of light verbs defined according to the criteria as applied by Clark, Ninio, and Pinker (semantic generality, frequency, and tendency to grammaticalize cross-linguistically) as *bring*, *come*, *do*, *get*, *give*, *go*, *make*, *put*, and *take*.

This semantic lightness and lack of imageability parallels the lack of semantic prototypicality effect in the present experiments when our participants are asked to consciously judge the VAC exemplars for meaning. The verb-VAC combinations that are judged to be more meaningful and judged so faster are those with a high verb-VAC contingency. It is the contingency of verb and VAC which gives special, specific, readily accessible meanings. So the high  $\Delta P_{cw}$  *talk about*, *distinguish between*, *fall into*, and *look around* are judged more meaningful than the high semantically prototypical *move about*, *run between*, *travel into*, and *go around*.



## 5.3. CONSCIOUS AND UNCONSCIOUS MEANING

My collaborators and I believe that these differences whereby lexical decision is driven by semantic prototypicality (but not VAC-verb contingency  $\Delta P_{cw}$ ), whereas meaning judgment is affected VAC-verb contingency  $\Delta P_{cw}$  (but not semantic prototypicality), might reflect differences between unconscious semantic access in lexical decision, and conscious comprehension in the meaning judgment experiment. Unconscious semantic access involves spreading access. In discussing the results of the masked lexical priming experiments of Marcel (1980, 1983), where the prime word was masked down to a subliminal level, Dehaene summarizes: “after flashing the word *bank*, both *money* and *water* were primed ... Thus our unconscious mind is clear enough to store and retrieve, in parallel, all the possible semantic associations of a word ... The unconscious mind proposes, while the conscious mind selects” (Dehaene, 2014, p. 66).

While you are conscious of words in your visual focus, you definitely did not just now consciously label the word ‘focus’ as a noun (Baars, 1997). On reading it, you were surely unaware of its nine alternative meanings, though in a different sentence you would instantly have brought a different meaning to mind. What happens to the other meanings? Psycholinguistic evidence from experiments like those of Marcel demonstrates that some of them exist unconsciously for a few tenths of a second before your brain decides on the right one. Most words (more than 80% in English) have multiple meanings, but only one at a time can become conscious. Comprehension (etymologically, ‘together catching’) requires the assemblage of fragments of meaning, and this is done faster when the pieces go together well. High verb-VAC contingency, as reflected by  $\Delta P_{cw}$ , speeds the dynamic competition among the massively parallel constituency of the unconscious mind to elect (Koch, 2004, pp. 24, 173) a current oneness to the fleeting stream of conscious experience (Dehaene & Changeux, 2004; Dehaene, Sergent, & Changeux, 2003).

## 5.4. LIMITATIONS AND CONCLUSION

While there is advantage in using the same stimuli and statistical analysis models in the various experiments here, it would be sensible to replicate this research with different samples of stimuli. However hard we tried, it was impossible to achieve a sample of stimulus items where the predictor variables were completely orthogonal. Furthermore, as can be seen from the scatterplots overlaid upon the effects plots, some of our variables, particularly contingency, are patchily distributed.

Stripping down the VAC to the verb–preposition collocation adds problematic confounds to our interpretation. Consider, for example, the verb–preposition

collocation *throw up*. If this were presented to subjects, then whatever reaction they had could be due to *throw up* as an intransitive prepositional verb (e.g., ‘He threw up because he had too much to eat’), or as an idiomatic transitive phrasal verb (e.g., ‘He threw up his hands in despair’), or as a compositional transitive phrasal verb (e.g., ‘He threw up his car keys to her’). Thus, there is an as yet unidentified amount of variability on the data that may create, amplify, or weaken the correlations found here. There is much scope for other online processing measures too, e.g., reading rate as measured by moving window self-paced reading or eye-tracking. Additionally, there are much more sophisticated methods for investigating online semantic processing. There is good scope for using visual-world paradigms here.

Finally, the stimuli we used were the end of a long series of operationalizations of measures including NLP searches of a 100 million word corpus, statistical and definitional decisions regarding semantic analysis and network building. Each step has its own associated error. Starting again from scratch would be the best triangulation.

These three experiments were designed to address the concern that the effects of usage characteristics upon VAC processing shown previously in free association tasks might reflect conscious processing and the use of ad hoc categories. We have replicated their generality in speeded automatic online processing tasks. All of the experiments and outcome measures here show effects of Verb Frequency and Verb-VAC frequency. These statistics of verb-VAC type token frequency are clearly represented and guide processing. Lexical decision is additionally driven by semantic prototypicality (but not VAC-verb contingency  $\Delta P_{cw}$ ), whereas meaning judgment is affected VAC-verb contingency  $\Delta P_{cw}$  (but not semantic prototypicality). We believe these findings index the spreading activation of unconscious meaning representation in comparison to the election of a unitary interpretation in conscious comprehension.

I conclude therefore that speeded automatic online VAC processing involves rich associations, tuned by verb type and token frequencies, their contingencies of usage, and their histories of interpretations, both specific and prototypical, which interface syntax, lexis, and semantics. The results encourage the conception of a unified construct where words and VACs alike are symbolic representations, acquired from usage, statistics and all, with their subsequent processing tuned probabilistically to usage experience.

### Supplementary material online

For Supplementary Appendices online please go to: <<http://dx.doi.org/10.1017/langcog.2016.18>>.

## REFERENCES

- Allan, L. G. (1980). A note on measurement of contingency between two binary variables in judgment tasks. *Bulletin of the Psychonomic Society*, **15**, 147–149.
- Ambridge, B., & Lieven, E. (2011). *Child language acquisition: contrasting theoretical approaches*. Cambridge: Cambridge University Press.
- Ambridge, B., & Rowland, C. F. (2013). Experimental methods in studying child language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science*, **4**(2), 149–168.
- Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning & Verbal Behavior*, **22**(3), 261–295.
- Arnon, I., & Snider, N. (2010). More than words: frequency effects for multi-word phrases. *Journal of Memory and Language*, **62**, 67–82.
- Baars, B. J. (1997). In the theatre of consciousness: Global Workspace Theory, a rigorous scientific theory of consciousness. *Journal of Consciousness Studies*, **4**, 292–309.
- Balota, D., Cortese, M., Sergent-Marshall, S. D., Spieler, D., & Yap, M. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, **133**(2), 283–316.
- Balota, D., Ferraro, F., & Connor, L. (1991). On the early influence of meaning in word recognition: a review of the literature. In P. J. Schwanenflugel (Ed.), *The psychology of word meanings* (pp. 187–222). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Balota, D. A., Yap, M. J., & Cortese, M. J. (2006). Visual word recognition: the journey from features to meaning (a travel update). In M. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics*, 2nd ed. (pp. 285–375). Amsterdam: Elsevier.
- Barsalou, L. W. (2010). Ad hoc categories. In P. C. Hogan (Ed.), *The Cambridge encyclopedia of the language sciences* (pp. 87–88). New York: Cambridge University Press.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, **67**(1), 1–48.
- Beckner, C., Blythe, R., Bybee, J., Christiansen, M. H., Croft, W., Ellis, N. C., ... Schoenemann, T. (2009). Language is a complex adaptive system. Position paper. *Language Learning*, **59** Supplement 1, 1–26.
- Biber, D., & Reppen, D. (Eds.) (2015). *The Cambridge handbook of English corpus linguistics*. Cambridge: Cambridge University Press.
- BNC. (2007). BNC XML Edition. Online: <<http://www.natcorp.ox.ac.uk/corpus/>>.
- Bod, R. (2001). *Sentence memory: storage vs. computation of frequent sentences*. Paper presented at the CUNY-2001, Philadelphia, Pennsylvania.
- Clark, E. V. (1978). Discovering what words can do. In D. Farkas, W. M. Jacobsen, & K. W. Todrys (Eds.), *Papers from the parasession on the lexicon, Chicago Linguistics Society April 14–15, 1978* (pp. 34–57). Chicago: Chicago Linguistics Society.
- Cortese, M., & Balota, D. (2012). Visual word recognition in skilled adult readers. In Michael J. Spivey, Ken McRae, & Marc F. Joanisse (Eds.), *Cambridge handbook of psycholinguistics* (pp. 159–185). Cambridge: Cambridge University Press.
- de Nooy, W., Mrvar, A., & Batagelj, V. (2010). *Exploratory social network analysis with Pajek*. Cambridge: Cambridge University Press.
- Dehaene, S. (2014). *Consciousness and the brain: deciphering how the brain codes our thoughts*. New York: Penguin/Random House.
- Dehaene, S., & Changeux, J.-P. (2004). Neural mechanisms for access to consciousness. In M. Gazzaniga (Ed.), *The cognitive neurosciences*, 3rd ed. (pp. 1145–1158). Cambridge, MA: MIT Press.
- Dehaene, S., Sergent, C., & Changeaux, J.-P. (2003). A neuronal network model linking subjective reports and objective physiological data during conscious perception. *Proceedings of the National Academy of Science USA*, **100**, 8520–8525.
- Durrant, P., & Doherty, A. (2010). Are high-frequency collocations psychologically real? Investigating the thesis of collocational priming. *Corpus Linguistics and Linguistic Theory*, **6**, 125–155.
- Ellis, N. C. (2002). Frequency effects in language processing: a review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, **24**(2), 143–188.

- Ellis, N. C. (2006). Language acquisition as rational contingency learning. *Applied Linguistics*, *27*(1), 1–24.
- Ellis, N. C. (2012). Formulaic language and second language acquisition: Zipf and the phrasal teddy bear. *Annual Review of Applied Linguistics*, *32*, 17–44.
- Ellis, N. C. (2016). On-line processing of Verb–Argument Constructions: visual recognition thresholds and naming. *Review of Cognitive Linguistics*, *14*(1), 105–135.
- Ellis, N. C., & Ferreira-Junior, F. (2009). Constructions and their acquisition: islands and the distinctiveness of their occupancy. *Annual Review of Cognitive Linguistics*, *7*, 111–139.
- Ellis, N. C., Frey, E., & Jalkanen, I. (2009). The psycholinguistic reality of collocation and semantic prosody (1): lexical access. In U. Römer & R. Schulze (Eds.), *Exploring the lexis–grammar interface* (pp. 89–114). Amsterdam: John Benjamins.
- Ellis, N. C., & Larsen-Freeman, D. (Eds.) (2006). Language emergence: implications for applied linguistics. *Special Issue of Applied Linguistics*, *27*(4).
- Ellis, N. C., & O'Donnell, M. B. (2011). Robust language acquisition – an emergent consequence of language as a complex adaptive system. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 3512–3517). Austin, TX: Cognitive Science Society.
- Ellis, N. C., & O'Donnell, M. B. (2012). Statistical construction learning: Does a Zipfian problem space ensure robust language learning? In J. Rebuschat & J. Williams (Eds.), *Statistical learning and language acquisition* (pp. 265–304). Berlin: Mouton de Gruyter.
- Ellis, N. C., O'Donnell, M. B., & Römer, U. (2012). Usage-based language: investigating the latent structures that underpin acquisition. *Currents in Language Learning*, *1*, 25–51.
- Ellis, N. C., O'Donnell, M. B., & Römer, U. (2014). The processing of Verb–Argument constructions is sensitive to form, function, frequency, contingency, and prototypicality. *Cognitive Linguistics*, *25*, 55–98.
- Ellis, N. C., & Ogden, D. C. (2015). Language cognition. *Journal of Child Language*, *42*, 182–186.
- Ellis, N. C., Römer, U., & O'Donnell, M. B. (2016). *Usage-based approaches to language acquisition and processing: cognitive and corpus investigations of construction grammar*. Malden, MA: Wiley-Blackwell.
- Fox, J. (2003). Effect displays in R for generalised linear models. *Journal of Statistical Software*, *8*(15), 1–27.
- Francis, G., Hunston, S., & Manning, E. (Eds.) (1996). *Grammar patterns 1: verbs* (The COBUILD Series). London: Harper Collins.
- Freeman, L. (1977). A set of measures of centrality based upon betweenness. *Sociometry*, *40*, 35–41.
- Gaskell, G. (Ed.) (2007). *Oxford handbook of psycholinguistics*. Oxford: Oxford University Press.
- Gass, S., & Mackey, A. (2011). *Handbook of second language acquisition*. London: Routledge / Taylor Francis.
- Gernsbacher, M. A. (1994). *A handbook of psycholinguistics*. San Diego, CA: Academic Press.
- Goldberg, A. E. (1995). *Constructions: a construction grammar approach to argument structure*. Chicago: University of Chicago Press.
- Goldberg, A. E. (2006). *Constructions at work: the nature of generalization in language*. Oxford: Oxford University Press.
- Granger, S., Gilquin, G., & Meunier, F. (Eds.) (2015). *The Cambridge handbook of learner corpus research*. Cambridge: Cambridge University Press.
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*(1), 4–27.
- Gries, S. Th., & Ellis, N. C. (2015). Statistical measures for usage-based linguistics. *Currents in Language Learning*, *2*, 228–255.
- Harnad, S. (Ed.) (1987). *Categorical perception: the groundwork of cognition*. New York: Cambridge University Press.
- Koch, C. (2004). *The quest for consciousness: a neurobiological approach*. Englewood, CO: Roberts and Company.

- Lakoff, G. (1987). *Women, fire, and dangerous things: what categories reveal about the mind*. Chicago, IL: University of Chicago Press.
- Marcel, A. J. (1980). Conscious and preconscious recognition of polysemous words: locating the selective effects of prior verbal context. In Raymond S. Nickerson (Ed.), *Attention and performance VIII* (pp. 435–457). Hillsdale, NJ: Lawrence Erlbaum.
- Marcel, A. J. (1983). Conscious and unconscious perception: experiments on visual masking and word recognition. *Cognitive Psychology*, **15**(2), 197–237.
- Masson, M. E. J. (1991). A distributed memory model of context effects in word identification. In D. Besner & G. W. Humphreus (Eds.), *Basic processes in reading: visual word recognition* (pp. 233–263). Hillsdale, NJ: Lawrence Erlbaum Associates.
- McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, **88**(5), 375–407.
- McDonald, S. A., & Shillcock, R. C. (2003). Eye movements reveal the on-line computation of lexical probabilities during reading. *Psychological Science*, **14**(6), 648–652.
- McDonough, K., & De Vleeschauwer, J. (2012). Prompt type frequency, auditory pattern discrimination, and EFL learners' production of wh-questions. *Studies in Second Language Acquisition*, **34**(3), 355–377.
- McNamara, T. P. (2005). *Semantic priming: perspectives from memory and word recognition*. Hove: Psychology Press.
- Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, **90**, 227–234.
- Miller, G. A. (2009). WordNet – about us. Online: <<http://wordnet.princeton.edu>> (last accessed 1 March 2010).
- Neely, J. H. (1991). Semantic priming effects in visual word recognition: a selective review of current findings and theories. In D. Besner & G. W. Humphreus (Eds.), *Basic processes in reading: visual word recognition* (pp. 264–336). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Ninio, A. (1999). Pathbreaking verbs in syntactic development and the question of prototypical transitivity. *Journal of Child Language*, **26**, 619–653.
- Peirce, J. W. (2007). PsychoPy – psychophysics software in Python. *Journal of Neuroscience Methods*, **162**(1), 8–13.
- Pinker, S. (1989). *Learnability and cognition: the acquisition of argument structure*. Cambridge, MA: Bradford Books.
- Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: the ARC Nonword Database. *Quarterly Journal of Experimental Psychology*, **55A**, 1339–1362.
- Rebuschat, P., & Williams, J. N. (Eds.) (2012). *Statistical learning and language acquisition*. Berlin: Mouton de Gruyter.
- Rescorla, R. A. (1968). Probability of shock in the presence and absence of CS in fear conditioning. *Journal of Comparative and Physiological Psychology*, **66**, 1–5.
- Römer, U., O'Donnell, M. B., & Ellis, N. C. (2015). Using COBUILD grammar patterns for a large-scale analysis of verb-arguments in syntactic constructions: exploring corpus data and speaker knowledge. In M. Charles, N. Groom, & S. John (Eds.), *Corpora, grammar, text and discourse: in honour of Susan Hunston* (pp. 43–71). Amsterdam: John Benjamins.
- Rosch, E., & Mervis, C. B. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, **104**, 192–233.
- Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, **8**, 382–439.
- Rubenstein, H., & Pollack, I. (1963). Word predictability and intelligibility. *Journal of Verbal Learning and Verbal Behavior*, **2**(2), 147–158.
- Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: II. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review*, **89**(1), 60–94.
- Scarborough, D. L., Cortese, C., & Scarborough, H. S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance*, **3**(1), 1–17.

- Schvaneveldt, R. W., & Meyer, D. E. (1973). Retrieval and comparison processes in semantic memory. *Attention and Performance*, **IV**, 395–409.
- Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, **96**(4), 523–568.
- Shanks, D. R. (1995). *The psychology of associative learning*. New York: Cambridge University Press.
- 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.
- Theakston, A. L., Lieven, E. V. M., Pine, J. M., & Rowland, C. F. (2004). Semantic generality, input frequency and the acquisition of syntax. *Journal of Child Language*, **31**, 61–99.
- Toglia, M. P., & Battig, W. F. (1978). *Handbook of semantic word norms*. Hillsdale, NJ: Lawrence Erlbaum.
- Traxler, M., & Gernsbacher, M. A. (Eds.) (2011). *Handbook of psycholinguistics*. New York: Academic Press.
- Trousdale, G., & Hoffmann, T. (Eds.) (2013). *Oxford handbook of Construction Grammar*. Oxford: Oxford University Press.
- Zipf, G. K. (1935). *The psycho-biology of language: an introduction to dynamic philology*. Cambridge, MA: MIT Press.