Topics in Cognitive Science (2017) 1–17

ISSN:1756-8757 print / 1756-8765 online
DOI: 10.1111/tops.12256

This article is part of the topic "More Than Words: The Role of Multiword Sequences in Language Learning and Use," Morten H. Chrisiansen and Inbal Arnon (Topic Editors). For a full listing of topic papers, see http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1756-8765/earlyview

Thinking About Multiword Constructions: Usage-Based Approaches to Acquisition and Processing

Nick C. Ellis,[a,b] Dave C. Ogden[b]

[a]*Department of Psychology, University of Michigan*
[b]*Department of Linguistics, University of Michigan*

Received 11 September 2015; received in revised form 10 August 2016; accepted 18 August 2016

Abstract

Usage-based approaches to language hold that we learn multiword expressions as patterns of language from language usage, and that knowledge of these patterns underlies fluent language processing. This paper explores these claims by focusing upon verb–argument constructions (VACs) such as "V(erb) *about* n(oun phrase)." These are productive constructions that bind syntax, lexis, and semantics. It presents (a) analyses of usage patterns of English VACs in terms of their grammatical form, semantics, lexical constituency, and distribution patterns in large corpora; (b) patterns of VAC usage in child-directed speech and child language acquisition; and (c) investigations of VAC free-association and psycholinguistic studies of online processing. We conclude that VACs are highly patterned in usage, that this patterning drives language acquisition, and that language processing is sensitive to the forms of the syntagmatic construction and their distributional statistics, the contingency of their association with meaning, and spreading activation and prototypicality effects in semantic reference. Language users have rich implicit knowledge of the statistics of multiword sequences.

Keywords: Verb–argument constructions; Corpus analysis; First language acquisition; Online processing; Second language acquisition; Type-token frequency; Form-function contingency; Semantic prototypicality

Correspondence should be sent to Nick C. Ellis, Department of Psychology, University of Michigan, 3215 E. Hall 1109, 530 Church Street, Ann Arbor, MI 48109-1043. E-mail: ncellis@umich.edu

1. Introduction

I've got a thing about multiword constructions, how bout you? Analyses of large collections of text demonstrate that natural language makes considerable use of recurrent patterns of words. Sinclair (1991, p. 110) summarized the results of the pioneering COBUILD lexicographic analysis of English corpora in his *Principle of Idiom:* "a language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analyzable into segments. To some extent this may reflect the recurrence of similar situations in human affairs; it may illustrate a natural tendency to economy of effort; or it may be motivated in part by the exigencies of real-time conversation."

A collocation is a sequence of words that co-occur more often than would be expected by chance. We talk about *strong tea* and *powerful computers*, but not *strong computers* or *powerful tea*. We prefer *blazing row*, *heated dispute*, *closely attached*, *deeply affected*, and *severely punished* to *heated row*, *blazing dispute*, *very attached*, *very affected*, *very punished*. Some of these pairs of words go better together than others, despite their being equally grammatical (see Culicover, Jackendoff, & Audring, 2017, this issue). "In the store of familiar collocations there are expressions for a wide range of familiar concepts and speech acts, and the speaker is able to retrieve these as wholes or as automatic chains from the long-term memory; by doing this he minimizes the amount of clause-internal encoding work to be done and frees himself to attend to other tasks in talk-exchange, including the planning of larger units of discourse" (Pawley & Syder, 1983, p. 192). There are many psycholinguistic studies demonstrating that language users have rich knowledge of the statistics of multiword sequences (see Ellis, 2002, 2012; MacDonald & Seidenberg, 2006, for reviews).

Frequency of use leads to automatization and fluency of production (Anderson, 1992; Segalowitz, 2010). Forms that are used highly frequently become phonologically eroded, and "words used together fuse together" (Bybee, 2003) (after Hebb's [1949] "Cells that fire together, wire together"). The phenomenon is graded—the degree of reduction is a continuous function of the frequency of the target word and its conditional probability given the previous and following words (Bybee & Hopper, 2001; Jurafsky, Bell, Gregory, & Raymond, 2001) and also upon the frequency of the multiword string (Arnon & Cohen Priva, 2013; Arnon & Snider, 2010; Reali & Christiansen, 2007). *How bout you?*, above, is a nice example—it is such a frequent sequence in everyday speech that the natural abbreviation of *about* has become formalized in the orthography. *I've* is another.

Collocations play key roles both in usage-based theories of human language learning and in the practice of computational linguistics. Usage-based approaches to first language (L1) acquisition hold that grammar emerges as children create linguistic constructions from their analysis of recurring sequences of language (Ambridge & Lieven, 2011; Theakston & Lieven, 2017, this issue; Tomasello, 2003). Distributional data have long been recognized as a cue to syntactic word class (Kiss, 1973; Redington, Chater, & Finch, 1998). In second (L2) and foreign language learning, collocational knowledge is central to the attainment of nativelike fluency and nativelike idiomaticity (Pawley & Syder, 1983; Wray, 2002). Crossley, Salsbury, and Mcnamara (2015) analyzed spoken and written texts produced by beginning, intermediate, and advanced L2 learners, and native speakers which had been scored for both analytic and holistic features of lexical proficiency by trained language proficiency testing raters. Multiple regression analysis showed that while collocation accuracy, lexical diversity, and word frequency were all significant predictors of human evaluations of lexical proficiency, collocation accuracy explained the greatest amount of variance in the proficiency scores (84% in the written samples and 89% in the spoken samples). Arnon and Christiansen (2017, this issue) and McCauley and Christiansen (2017, this issue) explore the degree to which multiword sequences are similarly potent in first- and second-language acquisition.

As a proficient user of English, you know many thousands of collocations—you know many thousands of words, and each word has its particular selection preferences for other words. For example, you know a lot about the collocations of *about*: *there's a lot of it about*, *a thing about*, *about time*, *an article about*, *a lot of fuss about nothing*, *about 5 percent*, *about as useful as* ... (a chocolate parasol, or something equally ironic), etc.

You also know about verb–argument constructions (VACs) that include prepositions like *about*. VACs are semi-abstract patterns that comprise verbs and the arguments they occur with, such as "V *about* N," "V *into* N," "V *of* N," "V N *in* N," "V N N," etc. Construction grammar argues that VACs encode as their prototypical senses event types that are basic to human experience—those of something moving, something being in a state, someone causing something, someone possessing something, something causing a change of state or location, someone causing a change of possession, something undergoing a change of state or location, something having an effect on someone, etc. (Croft, 2012; Goldberg, 1995, 2006; Levin, 1993). Each word of the construction contributes individual meaning, and the verb meanings in VACs are usually central. But the larger configuration of words as a whole carries meaning too. The VAC as a category inherits its schematic meaning from the conspiracy of all of the examples experienced in usage.

Thus our experience of language allows us to converge upon similar interpretations of novel utterances like "the ball mandools across the ground" and "the teacher spugged the boy the book." You know that *mandool* is a verb of motion and have some idea of how mandooling works—its action semantics. You know that *spugging* involves some sort of gifting, that the teacher is the donor, the boy the recipient, and that the book is the transferred object. How is this possible, given that you have never heard these verbs before? There is a close relationship between the types of verb that typically appear within constructions, hence their meaning as a whole is inducible from the lexical items experienced within them. So your reading of "the ball mandools across the ground" is driven by an abstract "V *across* noun" VAC which has inherited its schematic meaning from all of the relevant examples you have heard, and your interpretation of *mandool* emerges from the echoes of the verbs that occupy this VAC—words like *come*, *walk*, *move*, ... , *scud*, *skitter*, and *flit*, each weighted according to its frequency of experience.

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. There are three parts of this claim relating to usage, acquisition, and processing:

1. Corpus linguistic demonstration that language usage is highly structured and pervaded by collocations and phraseological patterns, that every word has its own local grammar, and that language forms are motivated by semantics and communicative functions: that is, that lexis, syntax, and semantics are inseparable.
2. Psychological demonstration that general learning mechanisms (implicit associative and statistical learning, concept learning and categorization, and explicit declarative learning and analogy-making) are sufficient to acquire knowledge of these patterns.
3. Psycholinguistic demonstration that our language processing is sensitive to the statistical regularities of these patterns. A VAC inherits its schematic meaning from the constituency of the verb exemplars experienced within it, weighted according to the frequency of their experience and the reliability of their association to that construction (their contingency), and their degree of prototypicality in the semantics of the VAC.

This paper summarizes relevant evidence for VACs. It presents (a) analyses of usage patterns of English VACs in terms of their grammatical form, semantics, lexical constituency, and distribution patterns in large corpora; (b) patterns of VAC usage in child-directed and child language acquisition; and (c) psycholinguistic investigations of both considered and automatic online processing. The research program investigates a range of VACs and we will summarize general findings, but when particular illustrations are called for, we will write about *about*.

2. Corpus analyses of language usage

Ellis and O'Donnell (2011, 2012) investigated the type-token distributions of 20 VACs such as "V(erb) *into* n(oun phrase)" in a 100-million-word corpus of English usage. The VAC prepositions sampled included *about*, *across*, *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 British National Corpus (BNC, 2007) 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 described in Römer, O'Donnell, and Ellis (2013). The steps were, for each VAC, such as the pattern "V *about* n":

1. Generate a list of verb types that occupy the construction (e.g., *talk*, *think*, *be*, *know*, *worry* . . ., *dream*, . . .).

2. Produce a frequency-ranked type-token profile of these verbs (e.g., *talk* 3,832, … *complain* 427, … *dream* 65, … *rattle* 5, …). Determine whether this is Zipfian (Goldberg, 2006; Zipf, 1949). Zipfian distributions exhibit a characteristic long-tail in a plot of rank against frequency. Zipf's law, like other power-law distributions, is most easily observed when plotted on doubly logarithmic axes, where the relationship between log (rank order) and log (frequency) is linear. The advised method to do this is via the (complementary) cumulative distribution (Adamic, 2002). We generated logarithmic plots and linear regressions to examine the extent of this trend using logarithmic binning of frequency against log cumulative frequency. The binning allows us to select and illustrate an example verb type from each frequency band. The plot for "V *about* n" is shown in Fig. 1.
3. Because some verbs are faithful to one construction while others are more promiscuous, calculate measures of contingency which reflect the statistical association between verb and VAC. We adopted various measures of contingency in usage, including ΔP (Ellis & Ferreira-Junior, 2009; Shanks, 1995): the association of *give* with the ditransitive (ΔP Word → Construction) is 0.025, that for *leave* is 0.001, the association of the ditransitive with *give* (ΔP Construction → Word) is *give* 0.314, that for *leave* is 0.003.

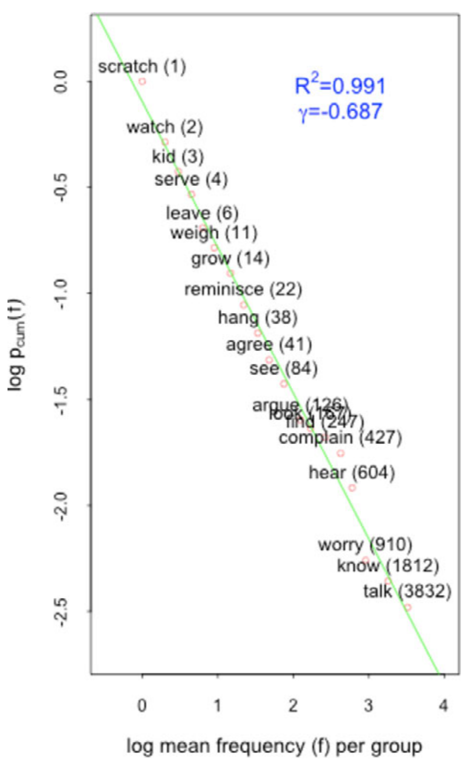

Fig. 1. British National Corpus verb type distribution for "V *about* n."

4. Using WordNet, a distribution-free semantic database based upon psycholinguistic theory (Miller, 2009), measure the semantic similarity of the meanings of the verbs occupying each construction and apply networks science algorithms (de Nooy, Mrvar, & Batagelj, 2010) 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. are then used to assess the cohesion of these semantic networks. We also apply algorithms for the detection of communities within the networks representing different semantic sets (Clauset, Newman, & Moore, 2004). The network for "V *about* n" is shown in Fig. 2. The network is fairly dense. The hubs, shown here as larger nodes, are those that are most connected, that is, have the highest degree. They are *go*, *move*, *run*, *get*, *feel*, and *see*—the prototypical Verb Locative "V *about* n" senses. However, there are also subcommunities, shown in different colors, for example one relating to physical and visual survey (*look*, *glance*). Other subcommunities include more metaphorical explorations of problem spaces, with hubs and attached communities for mentation (*think*, *know*, *remember*, *reminisce*), concern (*care*, *mind*, *bother*, *fuss*, *fret*), communication (*talk*, *speak*, *rave*, *gossip*, *chat*, *mutter*, *complain*, *grumble*, *protest*), and inquiry (*ask*, *question*, *consult*, *enquire*, *inquire*, *wonder*). Note that however sensible and intuitive these communities might seem, they are the objective result of statistical means for the detection of communities, somewhat akin to statistical cluster analysis. Note also that both degree and centrality in the network is unrelated to token frequency in the corpus; it simply reflects verb type connectivity within WordNet. Betweenness centrality is a measure of a node's centrality in a network equal to the number of shortest paths from all vertices to all others that pass through that node (McDonough & De Vleeschauwer, 2012). In semantic networks, central nodes are those which are prototypical of the network as a whole.

This research demonstrated: (a) The frequency distribution for the types occupying the verb slot of each VAC is Zipfian, with the most frequent verb taking the lion's share of the distribution. (b) The most frequent verbs in each VAC are prototypical of the VAC meaning—they are the hubs in the construction's semantic network, being well-connected while also fairly generic in their action semantics. (c) VACs are selective in their verb form family occupancy: individual verbs select particular constructions; particular constructions select particular verbs; there is high contingency between verb types and constructions. (d) VACs are coherent in their semantics. Psychology theory relating to the statistical learning of categories suggests that high frequencies of experience of labels which are highly contingent with prototypical exemplars are factors which make concepts rapidly and robustly learnable (Elio & Anderson, 1984; Goldberg, Casenhiser, & Sethuraman, 2004; Posner & Keele, 1968). So these aspects of usage could potentially make linguistic constructions robustly learnable by similar means.

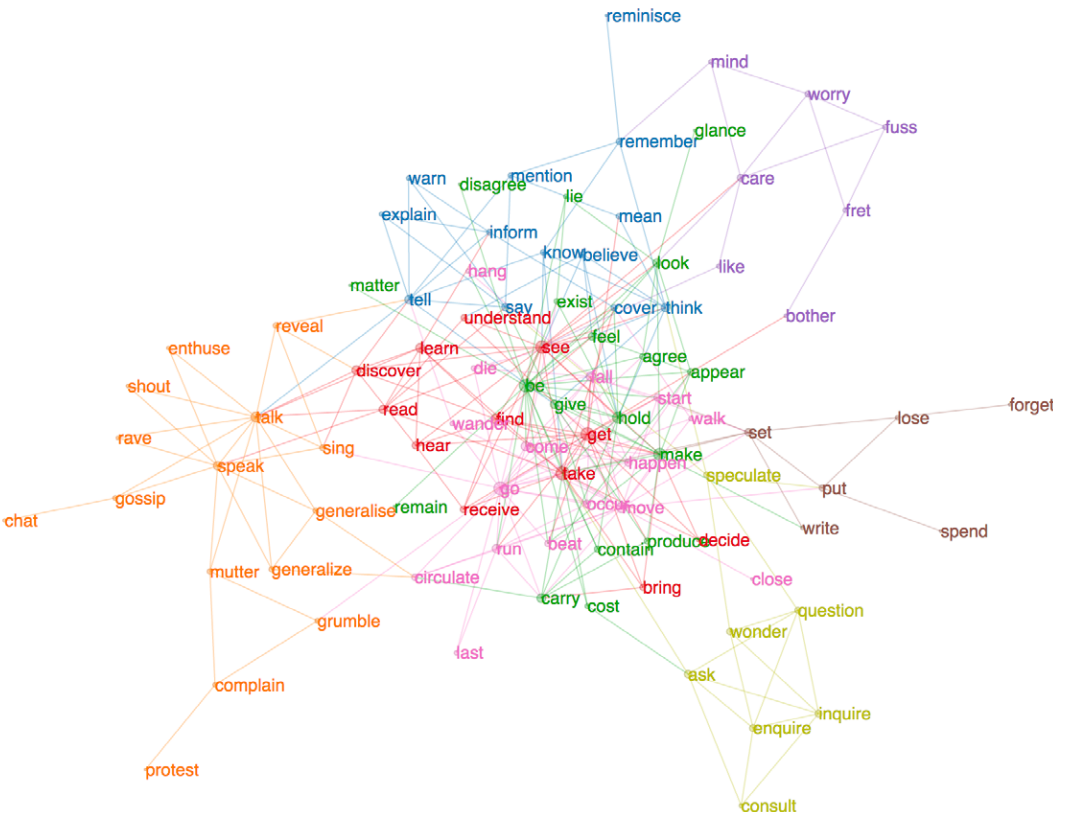

Fig. 2. Semantic network for verbs occupying the “V *about* n” verb–argument construction.

3. Corpus analyses of language acquisition

Goldberg et al. (2004) examined several hundred hand-coded tokens of the intransitive motion (V + $OBL_{path/location}$, henceforth “VL”), caused motion (V + OBJ + $OBL_{path/location}$, henceforth “VOL”), and ditransitive (V + OBJ_1 + OBJ_2, henceforth “VOO”) VACs in mother and child speech in the Bates, Brown, and MacWhinney corpora in the CHILDES database (MacWhinney, 2000) to show that in both mother and child speech, the distribution of verb type constituents in these constructions is heavily skewed, with one very high-frequency type and many low-frequency types:

- The verb object locative (VOL) [Subj V Obj $Obl_{path/loc}$] construction was exemplified in children’s speech by *put* 31% of the time, *get* 16%, *take* 10%, and *do/pick* 6%, a profile mirroring that of the mothers’ speech to these children (with *put* appearing 38% of the time in this construction that was otherwise exemplified by 43 different verbs).
- The verb locative (VL) [Subj V $Obl_{path/loc}$] construction was used in children’s speech with *go* 51% of the time, matching the mothers’ 39%.

- The ditransitive (VOO) [Subj V Obj Obj2] was filled by *give* between 53% and 29% of the time in five different children, with mothers' speech filling the verb slot in this frame by *give* 20% of the time.

They argued that this promotes acquisition: Tokens of one particular verb account for the lion's share of instances of each particular argument frame, and this pathbreaking verb is also the one with the prototypical meaning from which that construction is derived.

To test the generality of these patterns of parental usage and child language acquisition to scale, Ogden and Ellis (in Ellis, Römer & O'Donnell, 2016, chapter 7) report a large-scale analysis of these schematic VACs along with their more specific variants (e.g., for VL, "V *in* n," "V *on* n," "V *into* n," "V *about* n," etc.) in the morpho-syntactically parsed US and UK English data available from the CHILDES database (MacWhinney, 2000) to determine the distribution of verbs in VACs in child-directed speech (CDS; 4,809,299 words) and child language acquisition (2,559,260 words). Searches used the part-of-speech and dependency tagging provided in the CHILDES database (Sagae, Davis, Lavie, MacWhinney, & Wintner, 2007) and included both linear and dependency criteria that were refined based on recursively spot-checking results. The study reports a wide variety of analyses; here we concentrate upon the VL constructions involving the 20 prepositions reported in Section 2 plus others that were at least as frequent in the CHILDES corpus in the young child stratum Child I, ages 1;6 to 3;0 and the Parent Young stratum containing utterances from parents speaking with children younger than 3;0.

To test that the distribution of verb constituents in a VAC is near-Zipfian, plots of log frequency against log cumulative frequency were produced, again using binning to select and illustrate an example verb type from each frequency band. Fig. 3 plots on the left log10 cumulative frequency against log10 mean frequency per group for "V *about* n" in the CDS that these children were hearing, and on the right the corresponding plot for the verbs which the children produced in this construction. The distributions are Zipfian, and the rank order of verbs in child language parallels that in the input that they hear. The cumulative frequency plot of verb frequency against MLU in the child language shows how *talk* leads this VL VAC from the very start and takes the lion's share throughout development, being joined later by *run*, *think*, *know*, *laugh*, etc.

The correlation between log child verb frequency in the VAC and that in the CDS is 0.75 for *about*. The way child language follows the input for this VACs is graphed in Fig. 4.

This pattern is followed for all of the VACs studied. Child verb–VAC frequency of usage follows adult verb–VAC frequency of usage with r values $\approx$ 0.8. Child verb–VAC frequency of usage follows adult semantic prototypicality as indexed by betweenness centrality of the verb in the adult semantic network with r values $\approx$ 0.4–0.5. Together, the 30 or so separate, more specific "V *prep* n" VACs sum and resonate into a schematic "V locative construction," which communicates "someone or something moving to a new place or in a new direction" (Goldberg, 2006). Fig. 5 shows the Zipfian distribution ($R^2 > .98$) of verb types in this generic construction in the CDS (left) and the young

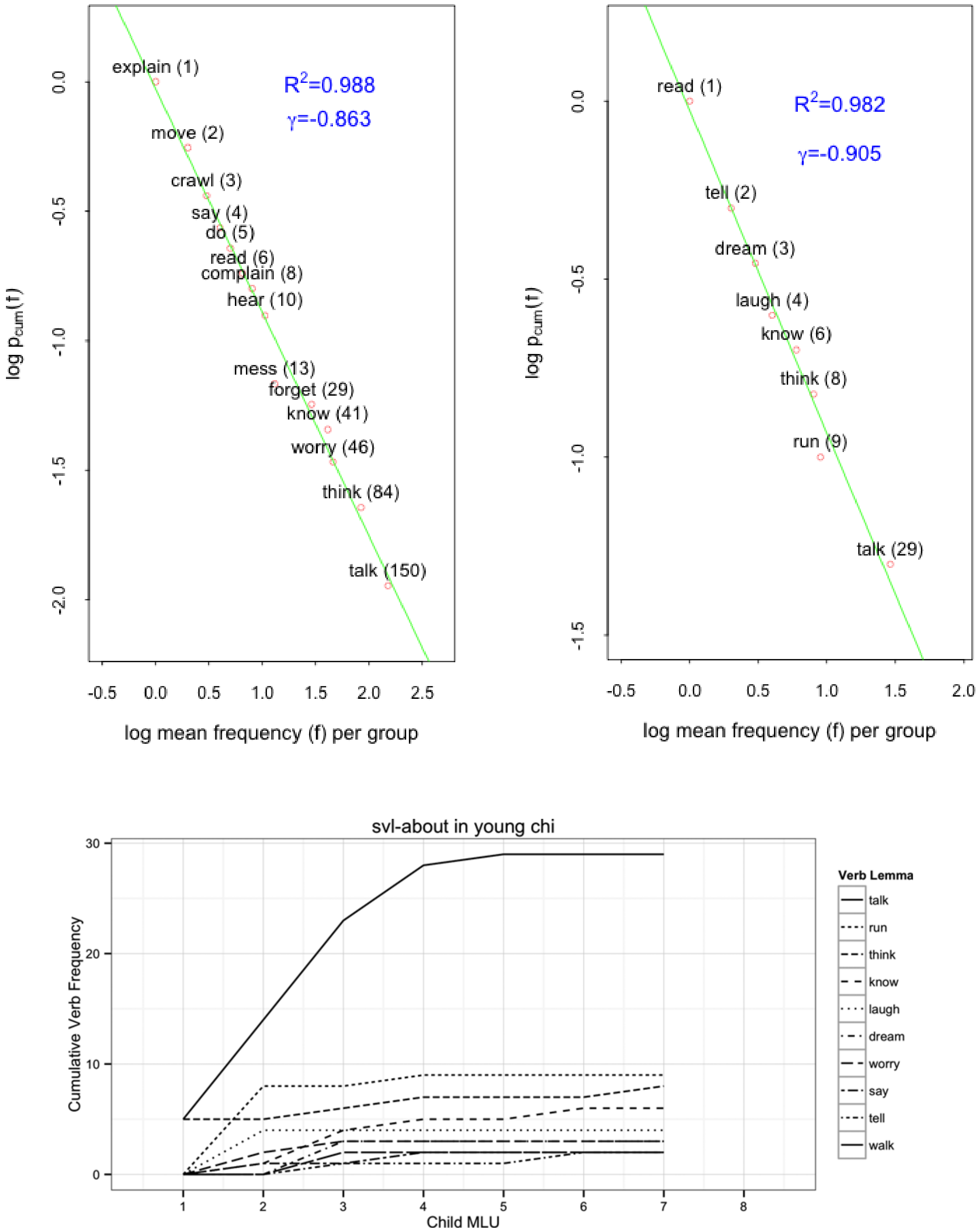

Fig. 3. Child-directed speech verb type distribution for “V *about* n” (left), child verb type distribution for “V *about* n” (right), child cumulative frequency of verb type in “V *about* n” against increasing mean length of utterance (MLU, bottom).

Children (right). The lower plot shows the very clear correspondence between child and parent use, with the prototypical locative verbs *go*, *come*, and *look* leading the way. Similar patterns are found with the other more schematic constructions studied. As *go* is the lead verb in schematic VL, so *put* is in VOL, and *give* in VOO.

In sum, these analyses replicate and extend the findings of Goldberg et al. (2004) to show (a) the distribution of verb types in a range of schematic as well as more specific VACs is near-Zipfian; (b) VACs are selective in their constituency; and (c) VACs are semantically coherent. Children's VAC acquisition follows these patterns, being affected by input frequency, contingency, and semantic prototypicality. Child VACs are seeded by the more frequent and semantically prototypical verbs that occupy the VAC in the input and these verbs continue ever to lead VAC acquisition.

The input from which learners acquire language is far from unstructured, unhelpful, or barren. The evidence of language usage is rich in latent structure. Child language acquisition follows the patterns in the parents' usage.

4. VAC processing

This section addresses studies of VAC processing. The psycholinguistic study of the factors which affect comprehension and production typically focuses upon online, automatic processing in real time. However, language is used to share, shape, and align speaker and listener consciousness, and sometimes that alignment involves quite

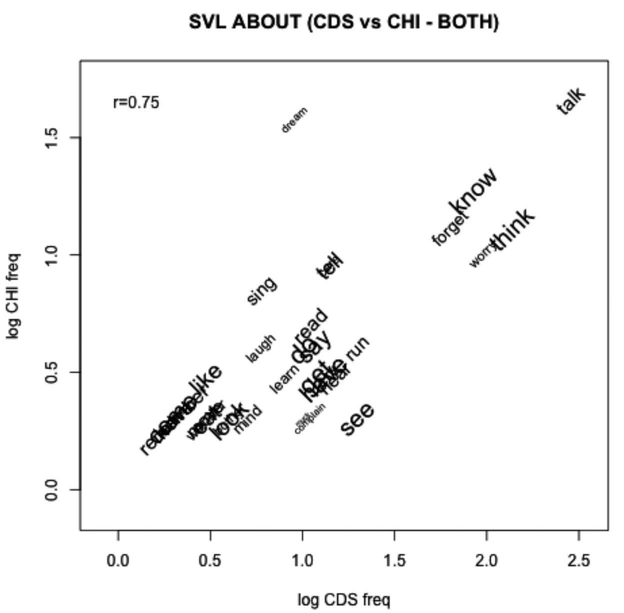

Fig. 4. Child vs. child-directed speech verb type distribution "V *about* n."

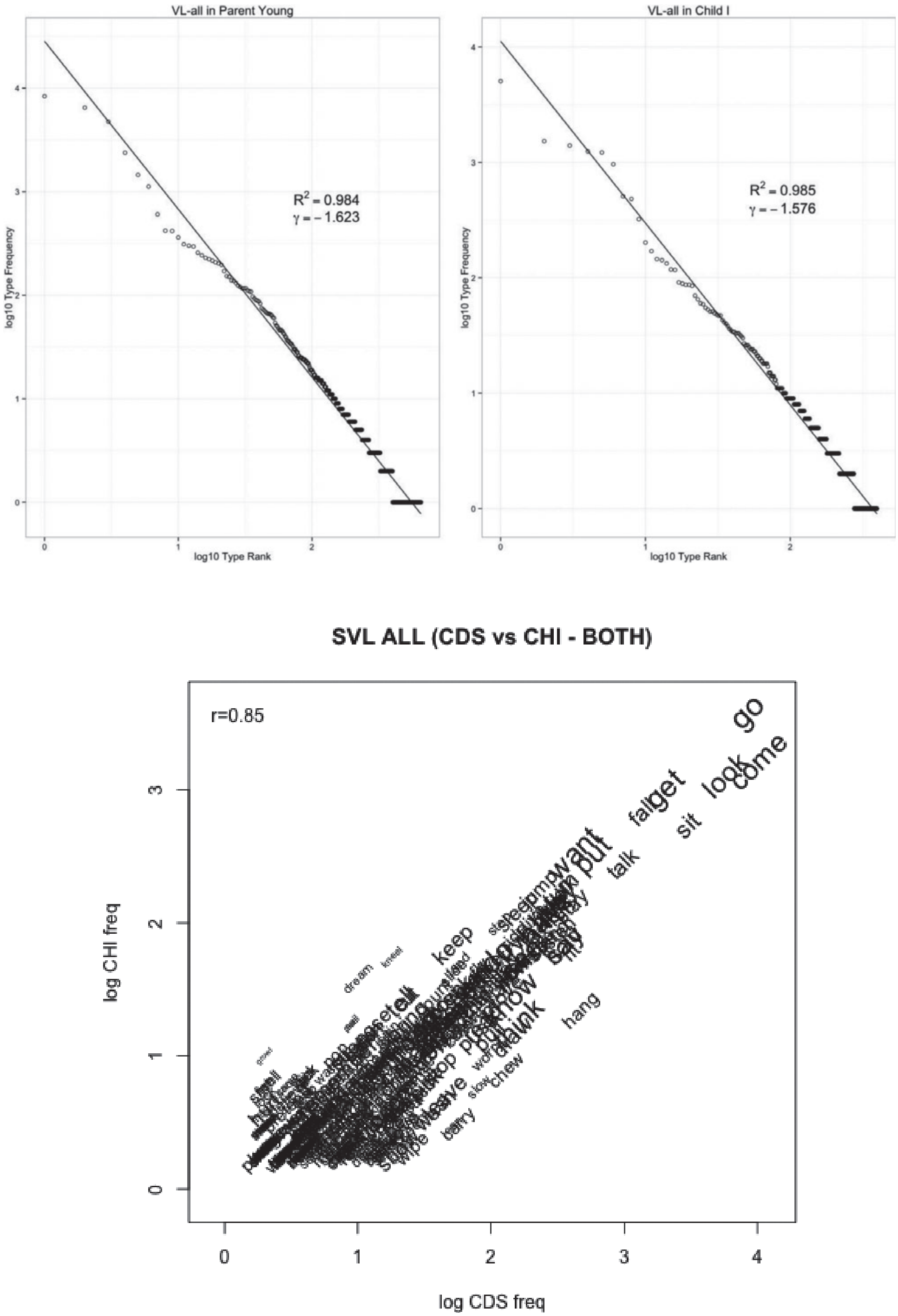

Fig. 5. Verb type distribution for the schematic verb locative construction in child-directed speech (CDS) (left), Young Child (right), and Child vs. CDS verb type distribution (bottom).

conscious processes of understanding and expression. The use of considered free-association tasks to investigate VAC processing therefore has some ecological validity. Section 4.1 reports studies of conscious L1 processes, and Section 4.2 summarizes a range of online studies adapted from lexical processing research to investigate more automatic VAC processing. These experiments were designed to investigate the effects upon VAC processing of (a) verb frequency in the language; (b) verb frequency in the VAC; (c) VAC–verb contingency; and (d) verb prototypicality in terms of centrality within the VAC semantic network.

4.1. Conscious L1 VAC processing

Ellis, O'Donnell, and Römer (2014) used free-association tasks to investigate the ways in which VAC processing is sensitive to the statistical patterns of usage (verb type-token frequency distribution, VAC–verb contingency, verb–VAC semantic prototypicality) measured in the BNC. Forty English speakers generated as many verbs that fit the frame as they could think of in a minute. Each prepositional VAC (*about*, *across*, *after*, *against*, *among*, *around*, *as*, *at*, *between*, *for*, *in*, *into*, *like*, *of*, *off*, *over*, *through*, *towards*, *under*, and *with*) was tested twice, once with an animate (*he/she*) and once with an inanimate subject (*it*).

For each VAC, the verbs generated were compared with their frequencies in the corpus analyses of usage. Fig. 6 shows the results for "V *about* n," where the log frequencies of generation are correlated with the log frequencies of the verb in the VAC in usage ($r = .69$). The same semantic strands are present in the generations as are in the semantic network of Fig. 2—the communicate group (*talk*, *ask*, *speak*, *read*, *tell*), the mentation group (*think*, *know*, *wonder*), as well as more literal motion (*run*, *dance*, *jump*), etc.

The frequency of verbs used in each VAC frame was tested against the BNC statistics for the complete dataset of verb–VAC pairings. Multiple regression analyses showed independent contributions of entrenchment (verb token frequencies in those VACs in usage experience), contingency (how faithful verbs are to particular VACs in usage experience), and semantic prototypicality (the centrality of the verb meaning in the semantic network of the VAC in usage experience). The major predictor was contingency ΔPcw (relative importance 0.41), followed by BNC verb frequency in that VAC (0.30), followed by verb betweenness centrality (0.29). Ellis et al. (2014) concluded that the effects of entrenchment and contingency, which are standard across the associative learning literature, show that VACs are learned implicitly from usage and that language users have knowledge of these usage statistics and use them in VAC processing. They interpreted the effects of semantic prototypicality as effects of spreading activation across the VAC semantic network, with more prototypical hubs being the most connected and receiving the most reminding from other nodes. The combined effects of usage, form, and semantic factors suggest that VAC processing involves rich associations, tuned by verb type and token frequencies and their contingencies of usage, which interface syntax, lexis, and semantics.

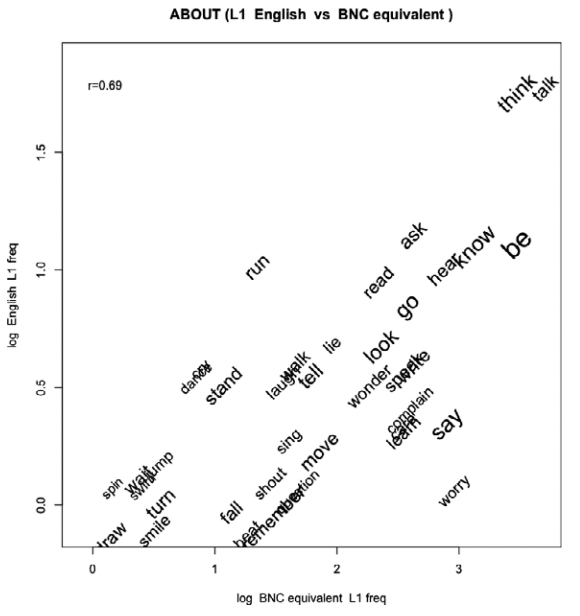

Fig. 6. L1 English log10 frequency of verb generation in the 1 min task against log10 verb frequency in that verb–argument construction in the British National Corpus (BNC) for "V *about* n." Verb font size is proportional to overall verb token frequency in the BNC as a whole.

4.2. Online L1 processing

Free-association tasks can involve conscious rather than automatic processing, especially those achieved over the timespan of a minute. Various deliberate search strategies can come to play in building ad hoc categories (Barsalou, 2010) in order to engage in the association task. Ad hoc categories are novel and do not reside as knowledge structures in long-term memory waiting to be retrieved. The usage effects upon free-association tasks described in Section 4.1 show the symbolic nature of VACs and the echoes of their prior usage in conscious processing. However, psycholinguistic research into mental representation, rather than conscious flow, is more concerned with online, automatic, ballistic processing—the implicit representations and dynamics which lead to the election of the contents of consciousness in the first place.

Ellis (2016a,b) reports the results of online processing experiments involving (a) masked perceptual recognition threshold, (b) naming, (c) successive lexical decision, (d) interposed lexical decision, and (e) meaning evaluation. All of the experiments used the same stimuli involving the prepositions *about*, *across*, *against*, *among*, *around*, *between*, *for*, *into*, *like*, *of*, *off*, *over*, *through*, *towards*, *under*, *with* in VACs that were stripped down from "V(erb) preposition n(oun phrase)" to their bare minimum, that is, the verb preposition collocation. The same generalized linear mixed model design was used across experiments to assess the independent effects of five predictors (Stimulus Length in Letters, log10 Verb Frequency, log10 Verb–VAC frequency, log10 VAC–verb contingency

(ΔPcw), and log10 Verb–VAC semantic prototypicality) determined using the procedures outlined in Section 2. The outcome variables were log 10 recognition threshold in experiment 1 and log 10 response time (RT) in Experiments 2–5.

There were robust frequency and conditional frequency effects in all of the processing tasks. Contingency was additionally influential in recognition and in meaning decision. Semantic prototypicality was influential in both successive and interposed lexical decision. Such results encourage the conception of a unified construction where words and VACs alike are symbolic representations, acquired from usage, with their subsequent processing tuned by verb type and token frequencies, their contingencies, and their histories of interpretations, both specific and prototypical, which interface syntax, lexis, and semantics.

5. Summary discussion

This research demonstrated that VACs as multiword sequences have notable properties.

In broad usage, (a) The frequency distribution for the types occupying the verb slot of each VAC is Zipfian, with the most frequent verb taking the lion's share of the distribution. (b) The most frequent verbs in each VAC are prototypical of that construction's functional interpretation, albeit generic in its action semantics. (c) VACs are selective in their verb form family occupancy: individual verbs select particular constructions; particular constructions select particular verbs; there is high contingency between verb types and constructions. (d) VACs are coherent in their semantics.

In child language acquisition, these patterns apply too. VAC–verb frequency distribution is near-Zipfian. VACs are selective: lead verbs differ in different syntactic patterns. The distribution of verbs is not trivial but corresponds meaningfully with the semantics of the VAC. Child VACs are seeded by the more frequent and semantically prototypical verbs that occupy the VAC in the input. These verbs lead VAC acquisition. Child verb–VAC frequency of usage follows parents' verb–VAC frequency of usage with r values $\approx .8$. With language, as with other cognitive realms, experiences conspire to give competence. Lexis, syntax, and semantics are highly connected in the soft-assembled, dynamic, symbolic exchanges of shared, cooperative usage.

In language processing (both considered and online automatic) there is sensitivity to the statistical forms of the syntagmatic construction, the contingency of their association with meaning, and spreading activation and prototypicality effects in semantic reference. Language users have rich implicit knowledge of the statistics of multiword sequences.

In other words, "language is never, ever, ever random" (Kilgarriff, 2005). Not in its usage, not in its acquisition, and not in its processing. These are initial investigations, briefly summarized here. Further detail is available in Ellis, Römer, and O'Donnell (2016). Much remains to be done. Nevertheless, it seems clear that theories of language acquisition and processing that ignore the regularities of multiword usage are missing important characteristics of the problem space. Structure abounds. Multiword constructions give language users plenty to think about.

References

Adamic, L. A. (2002). Zipf, power-laws, and pareto—A ranking tutorial from http://www.hpl.hp.com/research/idl/papers/ranking/ranking.html. Accessed 1,11,2016.

Ambridge, B., & Lieven, E. (2011). *Child language acquisition: Contrasting theoretical approaches*. Cambridge, UK: Cambridge University Press.

Anderson, J. R. (1992). Automaticity and the ACT-super(*) theory. *American Journal of Psychology*, *105*(2), 165–180. doi:10.2307/1423026

Arnon, I., & Christiansen, M. H. (2017, this issue). The role of multiword building blocks in explaining L1 − L2 differences. *Topics in Cognitive Science*, *9*.

Arnon, I., & Cohen Priva, U. (2013). More than words: The effect of multi-word frequency and constituency on phonetic duration. *Language and speech*, *56*(3), 349–371. doi:10.1177/0023830913484891

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language*, *62*, 67–82. doi:10.1016/j.jml.2009.09.005

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.

BNC (2007). BNC XML edition. Available at: http://www.natcorp.ox.ac.uk/corpus/. Accessed 7, 6, 2015.

Bybee, J. (2003). Mechanisms of change in grammaticalization: The role of frequency. In R. D. Janda & B. D. Joseph (Eds.), *Handbook of historical linguistics* (pp. 602–623). Oxford, UK: Blackwell.

Bybee, J., & Hopper, P. (Eds.) (2001). *Frequency and the emergence of linguistic structure*. Amsterdam: Benjamins.

Clauset, A., Newman, M. E. J., & Moore, C. (2004). Finding community structure in very large networks. *Physical Review E*, *70*, 066111. doi:10.1103/physreve.70.066111

Croft, W. (2012). *Verbs: Aspect and causal structure*. Oxford, UK: Oxford University Press.

Crossley, S. A., Salsbury, T., & Mcnamara, D. S. (2015). Assessing lexical proficiency using analytic ratings: A case for collocation accuracy. *Applied Linguistics*, 570-590, 1–22. doi:10.1093/applin/amt056

Culicover, P. W., Jackendoff, R., & Audring, J. (2017, this issue). Multiword constructions in the grammar. *Topics in Cognitive Science*, *9*.

de Nooy, W., Mrvar, A., & Batagelj, V. (2010). *Exploratory social network analysis with Pajek*. Cambridge, UK: Cambridge University Press.

Elio, R., & Anderson, J. R. (1984). The effects of information order and learning mode on schema abstraction. *Memory & Cognition*, *12*(1), 20–30. doi:10.3758/bf03196994

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. doi:10.1017/s0272263102002024

Ellis, N. C. (2012). Formulaic language and second language acquisition: Zipf and the phrasal teddy bear. *Annual Review of Applied Linguistics*, *32*, 17–44. doi:10.1017/s0267190512000025

Ellis, N. C. (2016a). On-line processing of Verb–Argument Constructions: Lexical decision and meaningfulness. *Language and Cognition*, *8*, 391–420.

Ellis, N. C. (2016b). On-line processing of Verb–Argument Constructions: Visual recognition thresholds and naming. *Review of Cognitive Linguistics*, *14*, 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. doi:10.1075/arcl.7.08ell

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. (2014). The processing of Verb–Argument Constructions is sensitive to form, function, frequency, contingency, and prototypicality. *Cognitive Linguistics*, *25*, 55–98. doi:10.1515/cog-2013-0031

Ellis, N. C., Römer, U., & O'Donnell, M. B. (2016). *Language usage, acquisition, and processing: Cognitive and corpus investigations of construction grammar*. Malden, MA: Wiley-Blackwell.

Francis, G., Hunston, S., & Manning, E. (Eds.) (1996). *Grammar patterns 1: Verbs*. London: Harper Collins.

Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago, IL: University of Chicago Press.

Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford, UK: Oxford University Press.

Goldberg, A. E., Casenhiser, D. M., & Sethuraman, N. (2004). Learning argument structure generalizations. *Cognitive*. *Linguistics*, *15*, 289–316. doi:10.1515/cogl.2004.011

Hebb, D. O. (1949). *The organization of behaviour*. New York: John Wiley & Sons.

Jurafsky, D., Bell, A., Gregory, M., & Raymond, W. D. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 229–254). Amsterdam: Benjamins.

Kilgarriff, A. (2005). Language is never, ever, ever, random. *Corpus Linguistics and Linguistic Theory*, *1*(2), 263–276. doi:10.1515/cllt.2005.1.2.263

Kiss, G. R. (1973). Grammatical word classes: A learning process and its simulation. *The Psychology of Learning and Motivation*, *7*, 1–41. doi:10.1016/s0079-7421(08)60064-x

Levin, B. (1993). *English verb classes and alternations: A preliminary analysis*. Chicago, IL: Chicago University Press.

MacDonald, M. C., & Seidenberg, M. S. (2006). Constraint satisfaction accounts of lexical and sentence comprehension. In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed., pp. 581–611). London: Elsevier Inc.

MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk* (3rd ed). Mahwah, NJ: Lawrence Erlbaum Associates.

McCauley, S. M., & Christiansen, M. H. (2017, this issue). Computational investigations of multiword chunks in language learning. *Topics in Cognitive Science*, *9*.

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. doi:10.1017/s0272263112000113

Miller, G. A. (2009). WordNet—About us, Princeton University. Available at http://wordnet.princeton.edu. Accessed March 1, 2010.

Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 191–225). London: Longman.

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, *77*, 353–363. doi:10.1037/h0025953

Reali, F., & Christiansen, M. H. (2007). Processing of relative clauses is made easier by frequency of occurrence. *Journal of Memory and Language*, *57*, 1–23. doi:10.1016/j.jml.2006.08.014

Redington, M., Chater, N., & Finch, S. (1998). Distributional information: A powerful cue for acquiring syntactic categories. *Cognitive Science*, *22*(4), 425–469. doi:10.1207/s15516709cog2204_2

Römer, U., O'Donnell, M., & Ellis, N. C. (2013). Using COBUILD grammar patterns for a large-scale analysis of Verb–Argument 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.

Sagae, K., Davis, E., Lavie, A., MacWhinney, B., & Wintner, S. (2007). High-accuracy annotation and parsing of CHILDES transcripts. *Proceedings of the ACL-2007 workshop on cognitive aspects of computational language acquisition* (pp. 25–32). Prague: Czech Republic.

Segalowitz, N. (2010). *The cognitive bases of second language fluency*. New York: Routledge.

Shanks, D. R. (1995). *The psychology of associative learning*. New York: Cambridge University Press.

Sinclair, J. M. (1991). *Corpus, concordance, collocation*. Oxford, UK: Oxford University Press.

Theakston, A. L., & Lieven, E. (2017, this issue). Multiunit sequences in first language acquisition. *Topics in Cognitive Science*, *9*.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Boston: Harvard University Press.

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge, UK: Cambridge University Press.

Zipf, G. K. (1949). *Human behaviour and the principle of least effort: An introduction to human ecology*. Cambridge, MA: Addison-Wesley.