Usage-Based Language:
Investigating the Latent Structures that Underpin Acquisition

Nick C. Ellis

University of Michigan

Matthew Brook O'Donnell

University of Michigan

Ute Römer

Georgia State University

We thank (1) Katie Erbach, Mary Smith, Lucy Zhao, Gin Corden, Danny Tsu-yu Wu, Liam Considine, Jerry Orlowski, and Sarah Garvey for help in the design, data collection, and analysis of these data; (2) the University of Michigan 'LSA Scholarship/Research' funding opportunity for supporting the project 'Piloting the development of an inventory of usage of English verb grammar'.

Correspondence concerning this article should be addressed to Nick Ellis, Department of Psychology, University of Michigan, 3215 East Hall, 530 Church Street, Ann Arbor, MI 48109-1043, USA. E-mail: ncellis@umich.edu

Abstract

Each of us as language learners had different language experiences, yet somehow we have converged upon broadly the same language system. From diverse, often noisy samples, we have attained similar linguistic competence. How so? What mechanisms channel language acquisition? Could our linguistic commonalities possibly have converged from our shared psychology of learning as applied to the evidence of similar-enough language experience? This article outlines a research program to investigate whether there are sufficient constraints in the dynamics of language to promote robust induction by means of statistical learning over limited samples. It illustrates the approach with regard to English verbs, their grammatical form, semantics, and Zipfian patterns of usage. It explores the emergence of structure from experience using methods from cognitive linguistics, corpus linguistics, learning theory, complex systems, and network science.

Keywords: complex systems; corpus linguistics; emergentism; learning theory; usage- based acquisition; Zipfian distribution

> "Plurality should not be assumed without necessity."
>
> William of Ockham (1495)

> "You generally find that the basic components and the basic laws are quite simple; the complexity arises because you have a great many of these simple components interacting simultaneously. The complexity is actually in the organization—the myriad possible ways that the components of the system can interact."
>
> S. Wolfram, quoted in Waldrop (1992)

> "I think the next century will be the century of complexity".
>
> Stephen Hawking (2000), San Jose Mercury News, January 23

Introduction

This article follows two currents in the language sciences, one moving towards simplicity, the other towards the origins of complexity. Yet we believe they flow together.

We seek an understanding of robust language acquisition: As a child, you engaged your parents and friends talking about things of shared interest using words and phrases that came to mind, and all the while you learned language. We authors of this paper were privy to none of your language circumstances. Yet somehow we have all converged upon a similar-enough 'English' to be able to communicate here. How came this to be?

The first current relates to a 'Simple View of Language Acquisition,' as pursued within cognitive linguistics, construction grammar, child language acquisition, and psycholinguistics. Here a common theme is that language acquisition is simple in Ockham's sense in that it is cut of the same cloth as the rest of cognition[1]: we learn language like we learn anything else. Of course, human cognition is not simple: There is much to the psychology of learning. And the problem-space of language – mapping thoughts to serial sequences of sound - is rather special. But it is parsimonious to assume that language is subject to broadly the same learning mechanisms and cognitive constraints as the rest of our experience. If we follow this course, we might determine exactly what might be learnable of language

by these means. And so far, this stream flows strongly (Bybee, 2010; Goldberg, 2006; Robinson & Ellis, 2008; Tomasello, 2003).

The second current relates to a "Complex Systems" approach as pursued within emergentism, dynamic systems theory, and complex adaptive systems research. Complex adaptive systems (CAS) are characterized by their robustness to different kinds of perturbations and local random variations, by their scale-free properties, and by their structures emerging from the interactions of agents and components at many levels (Page, 2009). These properties make this course attractive in the quest for robust theories of language acquisition, thus the recent explorations of Language as a Complex Adaptive System (Beckner et al., 2009; Ellis, 1998; Ellis & Larsen-Freeman, 2006, 2009b; Larsen-Freeman, 1997; MacWhinney, 1999; Solé, Murtra, Valverde, & Steels, 2005).

We trust that the confluence of these two currents will be abundant. For these approaches promise something rich: a more complete understanding of language cognition, and one that is satisfying in its interdisciplinarity. But whether this will amount to a sea-change in the Language Sciences rests upon a considerable amount of future research. In this article we chart the course we believe necessary, and we illustrate it with some explorations we have undertaken to date.

A significant discovery in the early cognitive analysis of language was that of basic objects in natural categories and how these underpinned the robust acquisition of nouns. Rosch, Mervis, Gray, Johnson and Boyes-Braem (1976) showed how basic categories, those that carry the most information in clustering the things of the world, are those whose members possess significant numbers of attributes in common, are visually imageable with similar shapes, and have associated motor programs which are similar to one another. Basic natural categories are organized around prototypes – the exemplars which are most typical of the category and which are similar to many other category members and not similar to members of other categories. People categorize prototype exemplars (like *robin* as *bird*) faster than

those with less common features or feature combinations like *geese* or *penguins* (Rosch & Mervis, 1975; Rosch, Mervis, Gray, Johnson, & Boyes-Braem 1976). Basic categories are also those which are the most codable (naming is faster), most coded, and most necessary in language (being highly frequent in usage). Children acquire basic-category terms like *dog, bird, hammer,* or *apple* earlier than they do their superordinates *animal, tool, fruit*, or subordinates *poodle, wren, ball-peen hammer,* or *Granny Smith*. It is the reliable coming-together of visual and motor perceptual experience along with frequent and highly-contingent labels which makes these nouns reliably and robustly learnable, despite individual children experiencing their different particular dogs and birds while hearing their own idiosyncratic soundtrack of linguistic experience.

Cognitive linguistics, particularly construction grammar, has since extended these ideas to language as a whole. It is not just that nouns typically relate to the things of the world, but, because language has emerged to describe our experiences of the world, so whole sentences are used to describe the doings of nouns in our world of experiences. Linguistic constructions which correspond to basic sentence types encode as their prototypical senses event types that are basic to human experience – those of something moving, something being in a state, someone causing someone causing a change of possession, something undergoing a change of state or location, something having an effect on someone, etc. (Croft, 2001, 2012; Goldberg, 1995; Levin, 1993).

Corpus and cognitive linguistics have shown that grammar and semantics are reliably associated, and, in turn, that grammatical patterns along with their corresponding events in the world jointly select particular lexis. Syntax, lexis and semantics are inextricably intertwined (Sinclair, 2004). The meaning of the words of a language, and how they can be used in combination, depends on our perception and categorization of the world around us. Since we constantly observe and play an active role in this world, we know a great deal about the entities of which it consists, and this experience and familiarity is reflected in language. The differing degrees of salience or prominence of elements involved in situations

that we wish to describe affect the selection of subject, object, adverbials and other clause arrangement. Figure/ground segregation and perspective-taking, processes of vision and attention, are mirrored in language and have systematic relations with syntactic structure. In language production, what we express reflects which parts of an event attract our attention; depending on how we direct our attention, we can select and highlight different aspects of the frame, thus arriving at different linguistic expressions. The prominence of particular aspects of the scene and the perspective of the internal observer (i.e., the attentional focus of the speaker and the intended attentional focus of the listener) are key elements in determining regularities of association between elements of visuo-spatial experience and elements of phonological form. In language comprehension, abstract linguistic constructions (like locatives, datives, and passives) guide the listener's attention to a particular perspective on a scene while backgrounding other aspects (Langacker, 1987; Talmy, 2000; Taylor, 2002).

Could these associations of form and function allow, by processes of syntactic and semantic bootstrapping, linguistic constructions to be learnable, exemplar-by-exemplar, with abstract schematic patterns being induced from the conspiracy of particular usage patterns and their interpretations? This is an exploration worth pursuit, though it requires a diverse crew. The investigation of form requires structuralist, corpus linguistic analyses, and natural language proceessing. The investigation of function requires functionalist, cognitive linguistic, and psycholinguistic analyses, the study of embodied force dynamics, and a lot else besides. Their association requires quantitative linguistics for the statistical tallying of form and function as well as an understanding of the psychology of learning. The result of this expedition will not be a dictionary, nor a grammar, nor a frequency list. Rather it should be an integrative analysis of the syntactic constructions of a language, the lexis they select, their meanings, and the distributions and mappings of these forms and functions.

In what follows we chart such a journey, illustrating it with some preliminary findings of ongoing surveys of our own where we focus upon verb argument constructions (VACs) in English. As

we do so, it should become clear that learners do not acquire language from unstructured, unhelpful, barren experience. Instead of desert, the evidence of language usage is as rich in latent structure as is a garden. Learners' explorations of this problem-space are grounded and contextualized. There is much latent structure to scaffold development, never mind the fact that learners' observations of this space are often directed, cultivated, and co-constructed in discourse interaction by their interlocutor as a helpful guide.

To VACs, then: Is there a basic variety of verb argument constructions, each with their basic level verb prototype? Despite the fact that you and we have not heard the same input, our experience allows us similar interpretations of novel utterances like "it mandools across the ground" or "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 transfer, 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 previously heard these verbs? Each word of the construction contributes individual meaning, and the verb meanings in these VACs is usually at the core. But the larger configuration of words as a whole carries meaning too. The VAC as a category has inherited its schematic meaning from the conspiracy of all of the examples you have heard. *Mandool* inherits its interpretation from the echoes of the verbs that occupy this VAC – words like *come, walk, move, ..., scud, skitter* and *flit*.

As you read these utterances, you parse them and identify their syntagmatic form: "it mandools across the ground" as a Verb Locative (VL) construction, "the teacher spugged the boy the book" as a double-object (VOO) construction. Then the paradigmatic associations of the types of verb that fill these slots are awakened: for the VL *V across N* pattern *come, walk, move, ..., scud, skitter and flit*, for VOO *give, send, pass*, ..., *read, loan, and fax*. Knowledge of language is based on these types of inference of syntactic and semantic bootstrapping[2], and verbs are the cornerstone of the syntax-semantics interface. The enterprise that we chart is therefore focused upon VACs in English.

First we will sketch the "Simple" – what is it that we learn and how do we learn it? Then we will sketch the Complex – what is the dynamic structure of language usage and how might it support robust language learning?

The Simple View of Language Acquisition

Construction Grammar

We take the Saussurian (1916) view that the units of language are constructions -- form-meaning mappings, conventionalized in the speech community, and entrenched as language knowledge in the learner's mind. They are the symbolic units of language relating the defining properties of their morphological, lexical, and syntactic form with particular semantic, pragmatic, and discourse functions (Goldberg, 1995, 2006). Construction Grammar argues that all grammatical phenomena can be understood as learned pairings of form (from morphemes, words, idioms, to partially lexically filled and fully general phrasal patterns) and their associated semantic or discourse functions: "the network of constructions captures our grammatical knowledge *in toto*, i.e. it's constructions all the way down" (Goldberg, 2006, p. 18). Such beliefs, increasingly influential in the study of child language acquisition, emphasize data-driven, emergent accounts of linguistic systematicities (e.g., Ambridge & Lieven, 2011; Tomasello, 2003).

The Psychology of Learning

Learning and language come together in usage-based approaches which hold that we learn linguistic constructions while engaging in communication (Bybee, 2010). Psycholinguistic research provides the evidence of usage-based acquisition in its demonstrations that language processing is exquisitely sensitive to usage frequency at all levels of language representation from phonology, through lexis and syntax, to sentence processing (Ellis, 2002). That language users are sensitive to the input

frequencies of these patterns entails that they must have registered their occurrence in processing. These frequency effects are thus compelling evidence for usage-based models of language acquisition which emphasize the role of input. Language knowledge involves statistical knowledge, so humans learn more easily and process more fluently high frequency forms and 'regular' patterns which are exemplified by many types and which have few competitors (e.g., MacWhinney, 2001).

Constructionist accounts of language learning involve the distributional analysis of the language stream and the parallel analysis of contingent perceptuo-motor activity, with abstract constructions being learned as categories from the conspiracy of concrete exemplars of usage following statistical learning mechanisms (Bybee & Hopper, 2001; Christiansen & Chater, 2001; Ellis, 2002; Jurafsky & Martin, 2009) relating input and learner cognition.

Determinants of Construction Learning

Psychological analyses of the learning of constructions as form-meaning pairs is informed by the literature on the associative learning of cue-outcome contingencies where the usual determinants include: (1) form frequency in the input (type-token frequency, Zipfian distribution), (2) function (prototypicality of meaning), and (3) interactions between these (contingency of form-function mapping) (Ellis & Cadierno, 2009):

Construction frequency. Frequency of exposure promotes learning and entrenchment (e.g., Anderson, 2000; Ebbinghaus, 1885). 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. The more times we experience conjunctions of features, the more they become associated in our minds and the more these subsequently affect perception and categorization (Harnad, 1987; Lakoff, 1987; Taylor, 1998).

Type and token frequency. Token frequency counts how often a particular form appears in the input. 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. Type frequency, on the other hand, refers to the number of distinct lexical items that can be substituted in a given slot in a construction, whether it is a word-level construction for inflection or a syntactic construction specifying the relation among words. For example, the "regular" English past tense *-ed* has a very high type frequency because it applies to thousands of different types of verbs, whereas the vowel change exemplified in *swam* and *rang* has much lower type frequency. The productivity of phonological, morphological, and syntactic constructions is a function of type rather than token frequency (Bybee & Hopper, 2001).

Zipfian distribution. In natural language, Zipf's law (Zipf, 1935) describes how the highest frequency words account for a disproportionately high number of linguistic tokens. Zipf's law states that the frequency of words decreases as a power function of their rank in the frequency table. Thus in English, the most frequent word (*the* with a token frequency of ~60,000/million words) occurs approximately twice as often as the second most frequent word, three times as often as the third most frequent word, etc. If p_f is the proportion of words whose frequency in a given language sample is f, then $p_f \sim f^{-\gamma}$, with $\gamma \approx 1$. Zipf showed this scaling law holds across a wide variety of language samples. Subsequent research provides support for this law as a linguistic universal. Many language events across scales of analysis follow his power law: words (Evert, 2005), grammatical constructs (Ninio, 2006; O'Donnell & Ellis, 2010), formulaic phrases (O'Donnell & Ellis, 2009), and so forth. Scale-free laws also apply to language structures, such as scale-free networks in collocation (Bannard & Lieven, 2009; Solé et al., 2005), in morphosyntactic productivity (Baayen, 2008), and in grammatical dependencies (Ferrer i Cancho & Solé, 2001, 2003; Ferrer i Cancho, Solé, & Köhler, 2004). Zipfian covering, where, as concepts need to be refined for clear communication, they are split, then split again hierarchically [e.g., animal, canine, dog, retriever, labrador..], determines basic categorization, the structure of

semantic classes, and the language form-semantic structure interface (Manin, 2008; Steyvers & Tennenbaum, 2005). Scale-free laws pervade language structure and usage.

But not just language structure and use: power law behavior like this has since been shown to apply to a wide variety of structures, networks, and dynamic processes in physical, biological, technological, social, cognitive, and psychological systems of various kinds (e.g., magnitudes of earthquakes, sizes of meteor craters, populations of cities, citations of scientific papers, number of hits received by web sites, perceptual psychophysics, memory, categorization, etc.) (Kello et al., 2010; Newman, 2005). This principle has become a hallmark of Complex Systems theory. It is tempting to think of Zipfian scale-free laws as universals. Complexity theorists suspect them to be fundamental, and are beginning to investigate how they might underlie language processing, learnability, acquisition, usage and change (Beckner et al., 2009; Ellis & Larsen-Freeman, 2009b; Ferrer i Cancho & Solé, 2001, 2003; Ferrer i Cancho et al., 2004; Solé et al., 2005).

Various usage-based / functionalist / cognitive linguists (e.g., Boyd & Goldberg, 2009; Bybee, 2008, 2010; Ellis, 2008a; Goldberg, 2006; Goldberg, Casenhiser, & Sethuraman, 2004; Lieven & Tomasello, 2008; Ninio, 1999, 2006) suspect that it is the Zipfian coming together of linguistic form and function that makes language robustly learnable despite learners' idiosyncratic experiences. In first language acquisition, Goldberg, Casenhiser and Sethuraman (2004) demonstrated that there is a strong tendency for VL, Verb Object Locative (VOL), and VOO VACs to be occupied by one single verb with very high frequency in comparison to other verbs used, a profile which closely mirrors that of the mothers' speech to these children. They argued that this might promote language acquisition: The low variance sample allows learners to get a fix on what will account for most of the category members, with the bounds of the category being defined later by experience of the full breadth of exemplar types.

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 et

al., 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 child language acquisition, a small group of semantically general verbs (e.g., *go, do, make, come*) are learned early (Clark, 1978; Goldberg, 2006; Ninio, 1999; Pinker, 1989). Ellis and Ferreira-Junior (2009a) showed the same is true of the second language acquisition of VL, VOL and VOO constructions. The verbs that learners first use in these VACs are prototypical and generic in function (*go* for VL, *put* for VOL, and *give* for VOO). In the early stages of learning categories from exemplars, acquisition might thus be optimized by the introduction of an initial, low-variance sample centered upon prototypical exemplars.

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 neither. Raw frequency of occurrence is less important than the contingency between cue and interpretation (Rescorla, 1968). Contingency/ reliability of form-function mapping and associated aspects of predictive value, information gain, and statistical association, are driving forces of learning. They are central in psycholinguistic theories of language acquisition (Ellis, 2006a, 2006b, 2008b; Gries & Wulff, 2005; MacWhinney, 1987).

Usage-based Acquisition

The primary motivation of construction grammar is that we must bring together linguistic form, learner cognition, and usage. An important consequence is that constructions cannot be defined purely on the basis of linguistic form, *or* semantics, *or* frequency of usage *alone*. All three factors are necessary in their operationalization and measurement. Psychology theory relating to the statistical learning of categories suggests that constructions are robustly learnable when they are (1) Zipfian in their type-token distributions in usage, (2) selective in their verb form occupancy, (3) coherent in their semantics, (4) with a high contingency between form and function.

A necessary research programme, then, is to assess whether language usage provides experience of this type. If so, then VACs as linguistic constructions should be robustly learnable. Is language, shaped by the human brain (Christiansen & Chater, 2008; Deacon, 1997), consequently shaped for the human brain in that the structures latent in language usage make language robustly learnable?

Language Usage as a Complex Adaptive System

This fundamental claim that Zipfian distributional properties of language usage helps to make language learnable has just begun to be explored for a small number of VACs in first (Goldberg, 2006; Goldberg et al., 2004) and second language acquisition (Ellis & Ferreira-Junior, 2009a, 2009b; Ellis & Larsen-Freeman, 2009a). It remains a priority to explore its generality across the wide range of English verbal grammar. (1) We need an integrative analysis of the VACs of a language, the lexis they select, their meanings, and the distributions and mappings of these forms and functions. (2) We need to determine if the latent structures therein are of the type that would promote robust learning. (3) We need to analyze corpora of learner language (L1 and L2) to see if learning is shaped by the input. This is no small task. Here we chart our approaches and the directions we believe worthy of further pursuit.

A Usage-based Grammar of English Verbs and Verb Argument Constructions

A catalogue of VACs. The determination of the semantic associations of particular linguistic forms should start from structuralist definitions of VACs defined by bottom-up means that are semantics-free. There is no one in corpus linguistics who 'trusts the text' more than Sinclair (2004) in his operationalizations of linguistic constructions on the basis of repeated patterns of words in collocation, colligation, and phrases. Therefore we used the VACs presented in the *Verb Grammar Patterns* (Hunston & Francis, 1996) that arose out of the COBUILD project (Sinclair, 1987) for our initial analyses. There are over 700 patterns of varying complexity in this volume. The form-based patterns described in the COBUILD Verb Patterns volume (Francis, Hunston, & Manning, 1996) take the form of word class and lexis combinations, such as the *V across N* pattern:

> The verb is followed by a prepositional phrase which consists of *across* and a noun group.
>
> This pattern has one structure:
> * Verb with Adjunct.
> *I cut across the field.*

Our initial research (for further details see Ellis & O'Donnell, in press; Römer, O'Donnell, & Ellis, in press, 2013) described the methods and findings for an initial convenience sample of 23 VACs, most of which follow the verb – preposition – noun phrase structure, such as *V into N , V after N, V as N* (Goldberg, 2006), but we also included other classic examples such as the *V N N* ditransitive, and the *way* construction.

A large corpus of English. To get a representative sample of usage, one needs a large corpus. We investigated the verb type-token distribution of these VACs in the 100 million word British National Corpus BNC (2007) parsed using the XML version of the BNC using the RASP parser (Briscoe, Carroll, & Watson, 2006). For each VAC, we translated the formal specifications from the COBUILD patterns into queries to retrieve instances of the pattern from the parsed corpus. Using a combination of part-of-

speech, lemma and dependency constraints we constructed queries for each of the construction patterns. For example, the *V across N* pattern was identified by looking for sentences that have a verb form within 3 words of an instance of *across* as a preposition, where there is an indirect object relation holding between across and the verb and the verb does not have any other object or complement relations to following words in the sentence.

The lexical constituency of verbs in VACs. The sentences extracted using this procedure produced verb type frequency distributions like the following one for the V *across N* VAC pattern:

come	483				
walk	203				
cut	199	...			
run	175	veer	4		
...		slice	4	...	
		...		navigate	1
				scythe	1
				scroll	1

These distributions appear to be Zipfian, exhibiting the characteristic long-tail in a plot of rank against frequency. We generated logarithmic plots and linear regressions to examine the extent of this trend using logarithmic binning of frequency against log cumulative frequency. Figure 1 shows such a plot for verb type frequency of the *V across N* construction, Figure 2 shows such the same type of plot for verb type frequency of the ditransitive *V N N* construction. In these Figures we randomly select one verb from each frequency bin for illustration. Both distributions produce a good fit of Zipfian type-token frequency with $R^2 > 0.97$ and slope (γ) around 1. Inspection of the construction verb types, from most frequent down, also demonstrates that the lead member is prototypical of the construction and generic in its action semantics.

Figure 1 about here

Figure 2 about here

Since Zipf's law applies across language, the Zipfian nature of these distributions is potentially trivial. But they are more interesting if the company of verb forms occupying a construction is selective, that is, if the frequencies of the particular VAC verb members cannot be predicted from their frequencies in language as a whole. We measured the degree to which VACs are selective like this using measures of association and the statistic '1-tau' where Kendall's tau measures the correlation between the rank verb frequencies in the construction and in language as a whole. For the VACs studied so far, 1-tau is typically about 0.7, showing that the rankings of verbs in particular VACs differs markedly from the rankings of verbs in the language as a whole. VACs are selective in their verb constituency.

Verb-VAC contingency. Another way of looking at this is to assess verb-VAC contingency. Some verbs are closely tied to a particular construction (for example, *give* is highly indicative of the ditransitive construction, whereas *leave*, although it can form a ditransitive, is more often associated with other constructions such as the simple transitive or intransitive). The more reliable the contingency between a cue and an outcome, the more readily an association between them can be learned (Shanks, 1995), so constructions with more faithful verb members should be more readily acquired. The measures of contingency we adopted are (1) faithfulness – the proportion of tokens of total verb usage that appear this particular construction (e.g., the faithfulness of *give* to the ditransitive is approximately 0.40; that of *leave* is 0.01, and (2) directional mutual information (MI Word → Construction [i.e. the reduction in uncertainty about which construction is involved once the word is known]: *give* 16.26, *leave* 11.73 and MI Construction → Word [i.e. the reduction in uncertainty about which word is involved once the construction is known]:: *give* 12.61 *leave* 9.11). Our analyses for the 23 VACs studied so far show a general pattern whereby that individual verbs tend to select particular constructions (MIw-c) and particular constructions select particular verbs (MIc-w) (for details see Ellis & O'Donnell, in press).

VAC meanings and coherence. Our semantic analyses use WordNet, a distribution-free semantic database based upon psycholinguistic theory which has been in development since 1985

(Miller, 2009). WordNet places words into a hierarchical network. At the top level, the hierarchy of verbs is organized into 559 distinct root synonym sets ('synsets' such as *move1* expressing translational movement, *move2* movement without displacement, etc.) which then split into over 13,700 verb synsets. Verbs are linked in the hierarchy according to relations such as hypernym [verb Y is a hypernym of the verb X if the activity X is a (kind of) Y (to *perceive* is an hypernym of to *listen*], and hyponym [verb Y is a hyponym of the verb X if the activity Y is doing X in some manner (to *lisp* is a hyponym of to *talk*)]. Various algorithms to determine the semantic similarity between WordNet synsets have been developed which consider the distance between the conceptual categories of words, as well as considering the hierarchical structure of the WordNet (Pedersen, Patwardhan, & Michelizzi, 2004). We take the lists of verbs occupying each VAC using the methods described in 3.1.3 and compare the verbs pairwise on these metrics. We then apply networks science, graph-based 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, and so forth are used to assess the cohesion of these semantic networks, and we also apply algorithms for the detection of communities within the networks representing different semantic sets (Clauset, Newman, & Moore, 2004; Danon, Díaz-Guilera, Duch, & Arenas, 2005). The network for V *across N* is shown as an example in Figure 3. 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*, and *travel* – the prototypical *V across N* senses. However, there are also subcommunities, shown in different colors, for example one relating to vision including *look, stare, gaze, face*, another speeded movement with unspecified action semantics: *shoot, scud, race, rush*, etc., and another emphasizing flat contact, *lay, lie, sprawl*, and so on. Note that degree in the network is unrelated to token frequency in the corpus: It simply reflects verb type connectivity within the network.

Figure 3 about here

Across the VACs we have investigated to date (O'Donnell, Ellis, & Corden, 2012), the semantic networks are coherent, with short average path-lengths between the nodes, and they have degree distributions which approximate a Zipfian power function. Satisfaction of these two criteria classifies them as 'small-world networks'. Steyvers and Tennenbaum (2005) showed more generally across semantic networks such as Roget's thesaurus and WordNet that the distributions of the number of connections as a whole follow power laws that indicate a scale-free pattern of connectivity, with most nodes having relatively few connections joined together through a small number of hubs with many connections. Small-world networks are interesting because they are more robust to perturbations and local random variations than are other network architectures (Barabási, 2002). They achieve this through the high connectivity of their hubs.

Is the Latent Structure of VACs of the Type that Would Promote Robust Acquisition Following a "Simple" Theory of Language Cognition?

Our core research questions concern the degree to which VAC form, function, and usage promote robust learning. As we explained earlier, the psychology of learning as it relates to these psycholinguistic matters suggests, in essence, that learnability will be optimized for constructions that are (1) Zipfian in their type-token distributions in usage, (2) selective in their verb form occupancy, and (3) coherent in their semantics. The findings we summarized above confirm that these factors apply.

Remember that the Zipfian distributions described above differ. The first relates to type-token frequency distribution in the language. The second relates to node connectivity (degree distribution) in the semantic network and it has no regard of corpus frequencies. Nevertheless, the high degree items in the semantic distribution tend also to be the high token frequency items in the corpus. It is this coming-together of the two Zipfian distributions, we believe, that makes language robustly learnable. The VAC

pattern is seeded by a high token frequency exemplar that is also prototypical in the action-dynamic construal of events to which it relates. Thereafter the forms and functions of verbs added to the VACs resonate with the category itself. "(T)his process is actually the organization of exemplars of utterances and of verb-specific constructions into clusters of greater or lesser size, with greater or lesser semantic coherence" (Croft, 2012, p. 393).

Our results are still preliminary. They are based on an analysis of 23 constructions to date. It remains to explore a more complete verbal grammar of English. There are over 700 patterns of varying complexity in the *Verb Grammar Patterns* (Hunston & Francis, 1996) – a lot to be done.

There is a considerable amount of statistical analysis and modeling to be done as well. Learning theory makes relative predictions too, and these should inform our understanding of language acquisition and processing. VACs that are higher frequency in the language, with more verb types, and greater semantic cohesion, should be acquired earlier and accessed more readily in speakers' minds. Similarly, the verbs which are more frequent constituents of a VAC, which are more faithful to that VAC, and which are closer to the VAC semantic prototype, should be acquired earlier in that VAC and accessed more readily in speakers' minds when they consider that VAC schema. A considerable amount of statistics and modeling remains necessary to test these hypotheses.

Analyzing Corpora of Learner Language: Is Learning Shaped by the Input?

Our final sketch map concerns proposed methods for investigating the influence of these usage factors upon acquisition. In order to sample child language acquisition of VACs, we searched the longitudinal child acquisition corpora of American and British English available in the Child Language Data Exchange system (CHILDES, MacWhinney, 2000b). We pooled what corpora were available to analyze several hundred hours of naturalistic observations from over 400 families. These included 421 English-speaking children (~190,000 utterances, ~1,100,000 words) and the accompanying Child Directed Speech (CDS) from 506 English-speaking adults. The corpus concerned children between the

ages of 1;0 and 3;6. Like Goldberg (2006), we concentrated upon VL constructions (e.g. *V in N, V across N, V around N*), VOL caused-motion constructions (e.g. *V Obj on N, V Obj in N*), and the ditransitive (*V N N)*. We searched for these separately in the child language and the CDS using CLAN (MacWhinney, 2000a) searches targeted on the %mor part-of-speech tagged tier. We counted the verb types in each VAC and determined their type-token frequency distributions using the same methods as described above. Across over twenty VL VACs (*VL in, on, down, up, out, to, there*, etc.), twenty VOL VACs (*VOL in, on, down, up, out, to, there*, etc.), and the VOO schema, the verb type-token frequency distributions were Zipfian. An example for VL *on* is shown in Figure 4.

Figure 4 about here

Next we analyzed the child language use of verbs in each VAC as a function of usage frequencies in the BNC, in child language as a whole, and in CDS. Figure 5 shows this on log-log axes for child *V on N* against (i) BNC verb frequencies in *V on* (preposition) *N*, (ii) BNC verb frequencies in *V on* (adverb), (iii) verb frequencies in child language as a whole, and (iv) CDS verb frequencies in *V on N*. The top two panes show reasonable correlations ($r = 0.46$ and 0.50) between the log verb frequencies in this VAC in the BNC and those used by children – usage frequencies in a large, mixed sample of English usage predict child language acquisition of verbs in this VAC. The bottom left pane shows a 0.53 correlation between child verb uses in these VACs and their frequencies of verb uses in their language as a whole – some verbs they use pretty frequently across these patterns (e.g., *go on, go in, go up*, etc. in the VLs; *put obj on, put obj up, put obj down,* in the VOLs, etc.). All three of these panels show a fair amount of 'pepper-pot' spread. The bottom right pane, however, looks like this noise has been removed. It shows the relationship between verb uses in child language and the CDS, where there is a substantial correlation of 0.85 – the Zipfian type-token verb distribution in VACs in child-directed speech drives the emergence of these verbs and their VACs in child language acquisition, and these similarly follow Zipfian scale-free profiles (see also Ninio, 2011).

This trend is true across all of the VACs we have studied in child language so far. In each case there is a substantial correlation (> 0.75) between VAC verb constituency in the child and that in the CDS. The spoken input that the child hears structures their experience and drives acquisition. It should come as no surprise that these associations are greater than those between child language and the BNC. The BNC is a catch-all representative corpus that included novels, newspapers, academic writing and all. As a source for general patterns in English as a whole it is difficult to rival. But children hear spoken language from their parents and significant others. Spoken language is different from written language (Leech, 2000) and CDS is a special simplified form of speech which generally selects higher frequency words and simpler syntactic frames (Buttery & Korhonen, 2005). As we more closely estimate the latent structures of usage to which infants are typically exposed, so we see closer parallels relations between the input and the acquisition.

Further Directions and Conclusions

These initial sorties make it clear that usage is intricately structured in ways typical of complex adaptive systems – there are scale-free distributions in verb usage frequency within constructions and scale-free connectivity patterns within semantic networks. There is much latent structure to robustly scaffold development. We believe therefore that the research approaches outlined here are worthy of pursuit across the breadth of English verbal grammar in usage and in child language acquisition.

We also identify the further priorities:

- This work needs to be done for second language acquisition too. We have made a start (Ellis & Ferreira-Junior, 2009a, 2009b), but there is an imperative for larger L2 corpora.
- The psychological reality of VACs needs further work. In various studies we use free association tasks to have people think of the first word that comes to mind to fill the V slot in a particular VAC frame. The responses of adult native and fluent L2 learners are highly predicted by corpus

frequencies, showing that users have implicit knowledge of the occupancy and frequencies of verbs in VACs (Ellis & O'Donnell, 2011; O'Donnell, Römer, & Ellis, in preparation; Römer et al., in press, 2013). Priming methods will also be relevant (Gries & Wulff, 2005).

- The semantic analyses here are very crude; other distributional measures could well be applied alongside techniques for investigating network growth (O'Donnell et al., 2012).
- The acquisition data here are basically correlational. There need to be experimental studies comparing the relative learnability of Zipfian skewed input compared to languages with flatter frequency profiles. We are pursuing this in artificial language learning experiments where we manipulate the type-token frequency profiles of verbs, nouns, and pronouns. Casenhiser and Goldberg (2005) and Goldberg et al. (2004) have made important steps in doing this in children and in adults, but they investigated the learning of just one construction from a small number of trials, and there is need for larger causal studies of the effects of combined Zipfian frequency distributions and Zipfian semantic connectivity upon more complete approximations to natural language.
- To better understand the processes of how these latent structures of usage affect robust acquisition and stable usage, there is need for modeling, both connectionist (Ellis & Larsen-Freeman, 2009a) and agent-based (Beckner et al., 2009).

Meanwhile, it appears that the input from which learners acquire language is far from unstructured, unhelpful, or barren. The evidence of language usage is rich in latent structure. With language, as with other cognitive realms, our experiences conspire to give us competence:

> "And these tend inward to me, and I tend outward to them,
> And such as it is to be one of these, more or less, I am,
> And of these one and all I weave the song of myself."
>
> Walt Whitman (1900 [1855] *Song of Myself*)

Revised version accepted 1 September 2012

Figure captions

Figure 1. Verb type distribution for *V across N* in the BNC

Figure 2. Verb type distribution for *V N N* in the BNC

Figure 3. A semantic network for *V across N* from the BNC using WordNet as a base

Figure 4. Verb type distribution for *V on N* in child language

Figure 5. Child language use of verbs in *V on N* as a function of (i) BNC verb frequencies in *V on* (preposition) *N*, (ii) BNC verb frequencies in V on (adverb), verb frequencies in child language as a whole, (iv) CDS verb frequencies in *V on N*

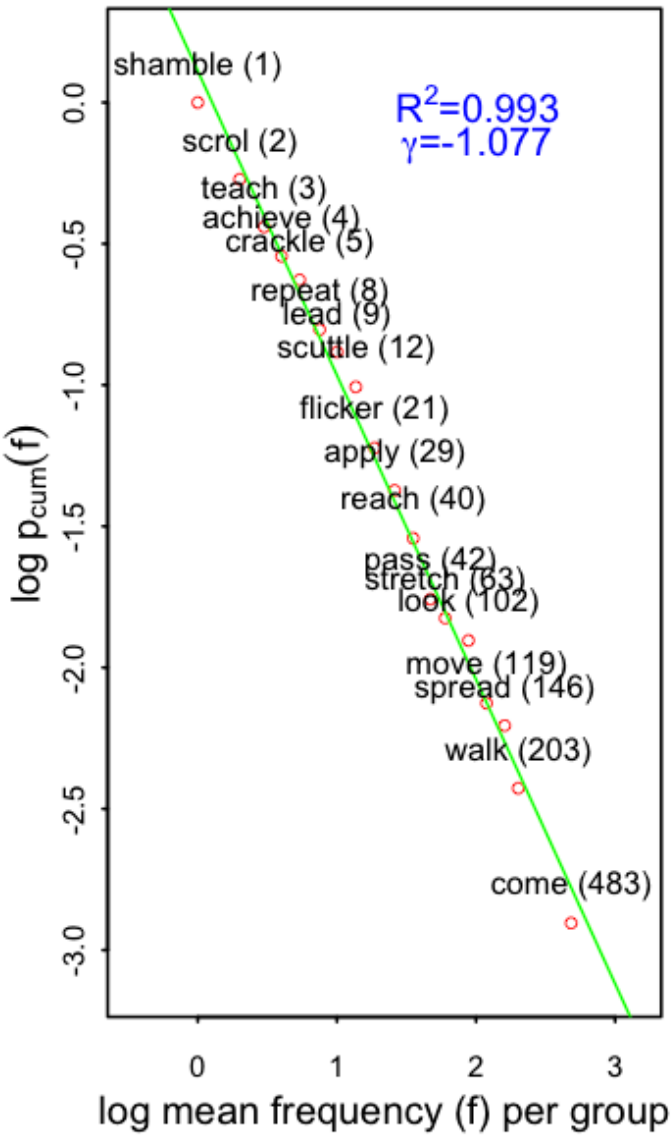

shamble (1)
scrol (2)
teach (3)
achieve (4)
crackle (5)
repeat (8)
lead (9)
scuttle (12)
flicker (21)
apply (29)
reach (40)
pass (42)
stretch (63)
look (102)
move (119)
spread (146)
walk (203)
come (483)
R^2=0.993
γ=-1.077
log p_{cum}(f)
0.0
-0.5
-1.0
-1.5
-2.0
-2.5
-3.0
0
1
2
3
log mean frequency (f) per group

Figure 1

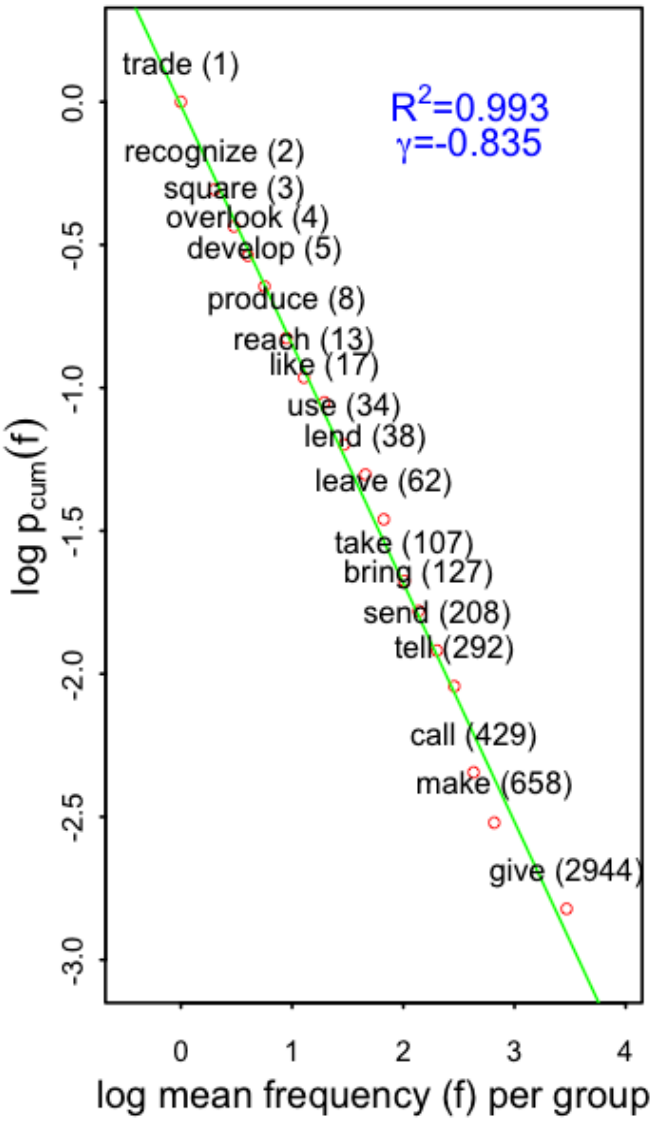

trade (1)
recognize (2)
square (3)
overlook (4)
develop (5)
produce (8)
reach (13)
like (17)
use (34)
lend (38)
leave (62)
take (107)
bring (127)
send (208)
tell (292)
call (429)
make (658)
give (2944)
R^2=0.993
γ=-0.835
log p_{cum}(f)
0.0
-0.5
-1.0
-1.5
-2.0
-2.5
-3.0
0
1
2
3
4
log mean frequency (f) per group

Figure 2

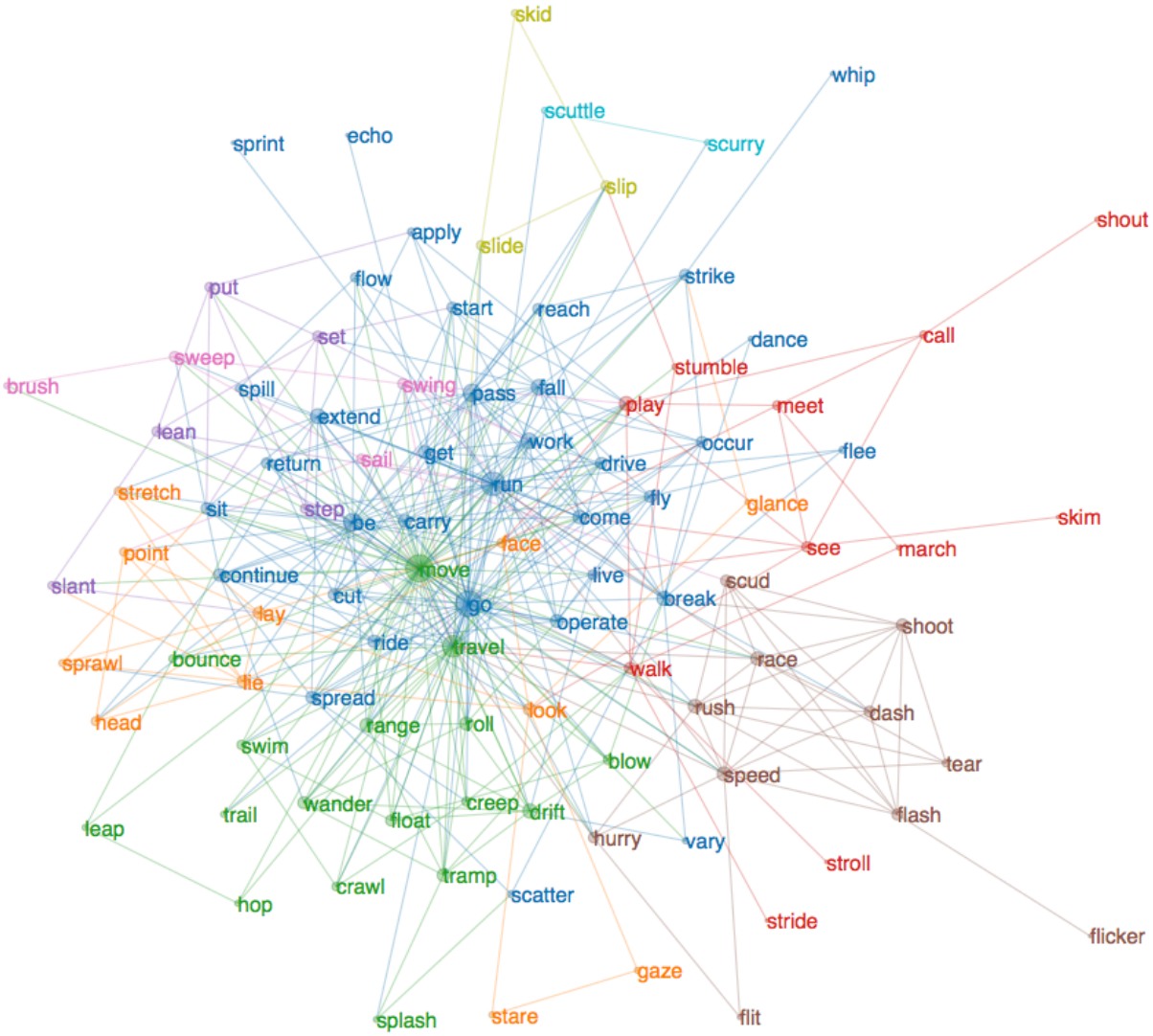

skid
whip
scuttle
sprint
echo
scurry
slip
shout
apply
slide
strike
put
flow
start
reach
dance
call
set
sweep
stumble
brush
spill
swing
pass
fall
play
meet
extend
lean
work
occur
flee
return
sail
get
drive
run
stretch
sit
step
be
carry
come
fly
glance
skim
face
see
march
point
continue
move
live
scud
slant
cut
break
lay
go
operate
shoot
ride
travel
sprawl
bounce
walk
race
lie
spread
look
rush
dash
head
range
roll
swim
blow
tear
speed
wander
creep
drift
flash
trail
float
leap
hurry
vary
stroll
crawl
tramp
scatter
hop
stride
flicker
gaze
splash
stare
flit

Figure 3

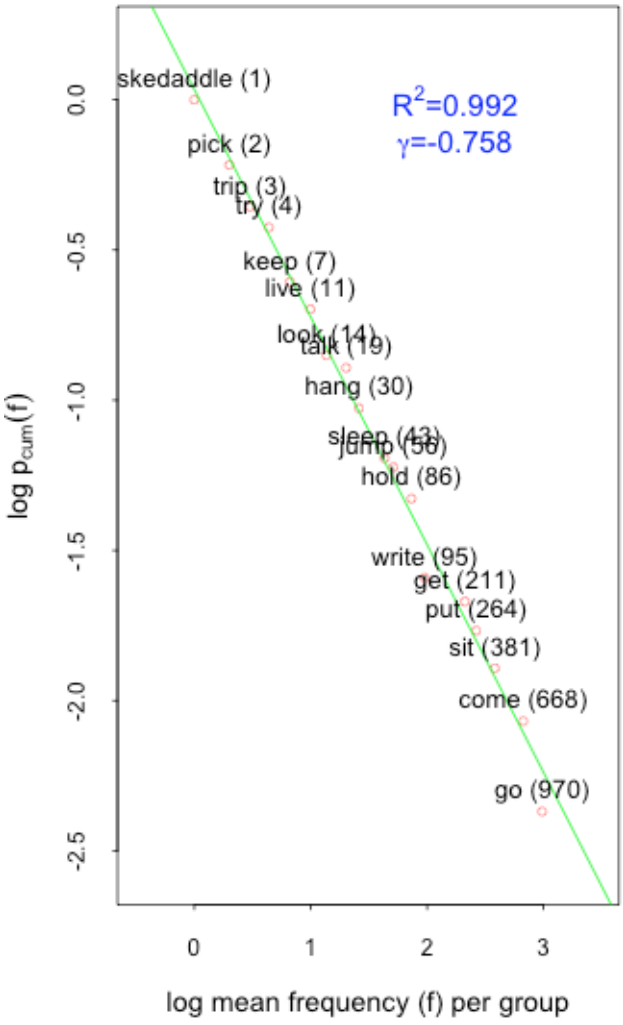
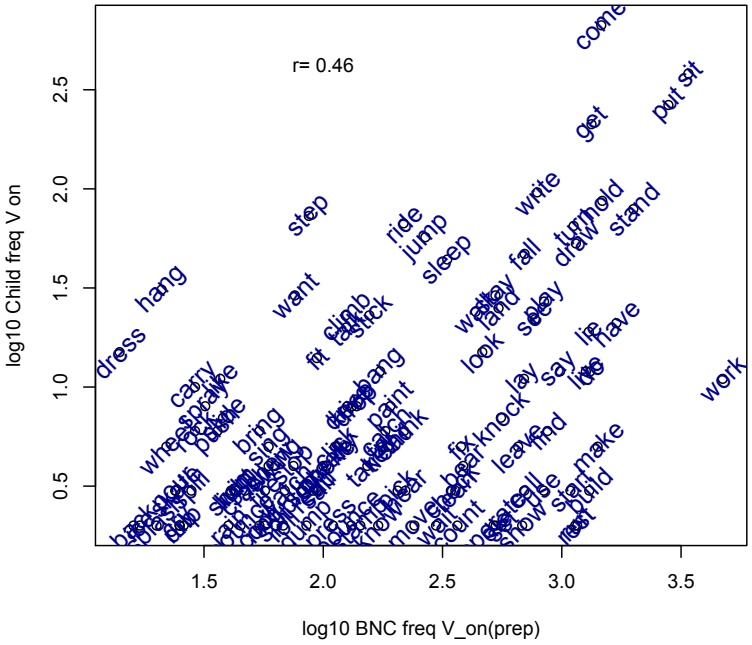
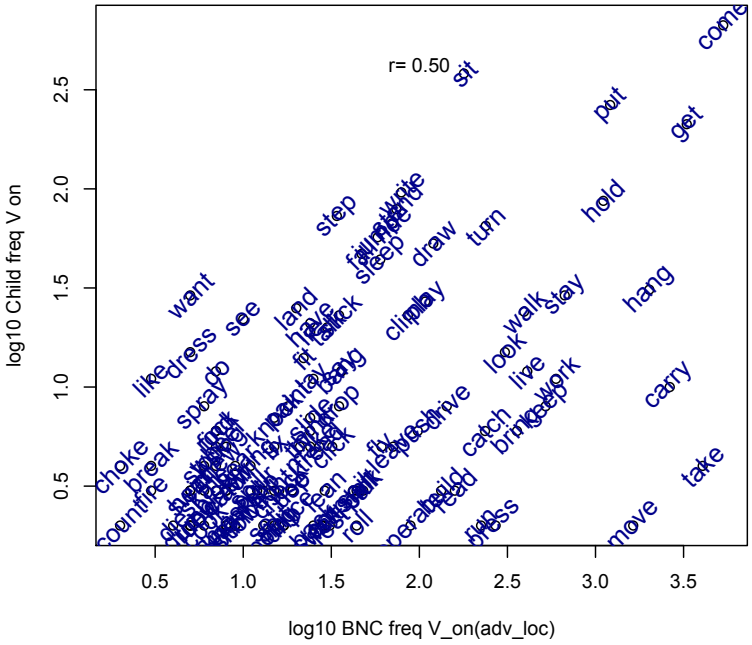
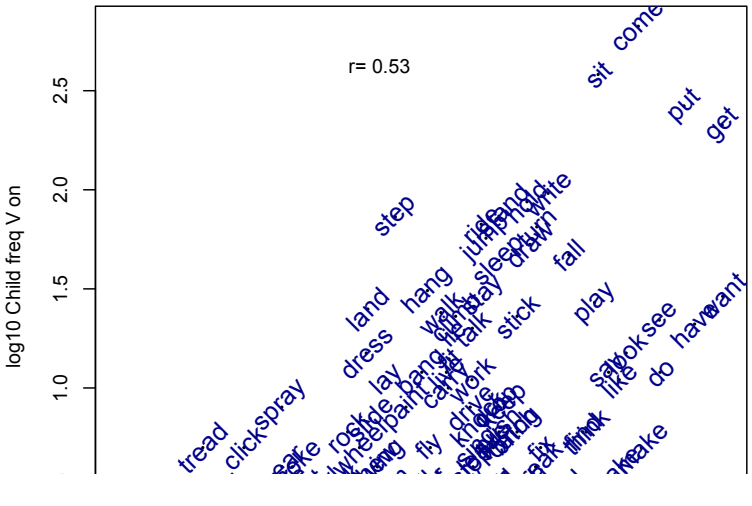
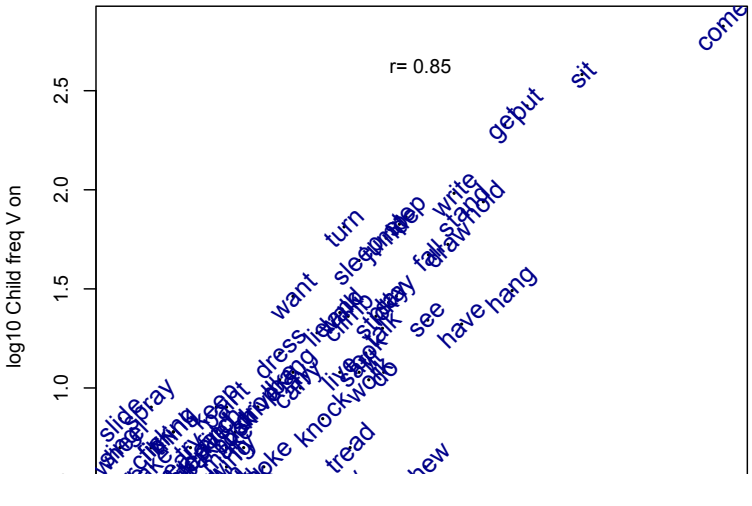

skedaddle (1)
pick (2)
trip (3)
try (4)
keep (7)
live (11)
look (14)
talk (19)
hang (30)
sleep (43)
jump (56)
hold (86)
write (95)
get (211)
put (264)
sit (381)
come (668)
go (970)
R^2=0.992
γ=-0.758
log p_{cum}(f)
0.0
-0.5
-1.0
-1.5
-2.0
-2.5
0
1
2
3
log mean frequency (f) per group

Figure 4


r= 0.46
log10 Child freq V on
log10 BNC freq V_on(prep)
0.5
1.0
1.5
2.0
2.5
3.0
3.5


r= 0.50
log10 Child freq V on
log10 BNC freq V_on(adv_loc)
0.5
1.0
1.5
2.0
2.5
3.0
3.5


r= 0.53
log10 Child freq V on
0.5
1.0
1.5
2.0
2.5


r= 0.85
log10 Child freq V on
0.5
1.0
1.5
2.0
2.5

References

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

Anderson, J. R. (2000). *Cognitive psychology and its implications (5th ed.)*. New York: W.H. Freeman.

Baayen, R. H. (2008). Corpus linguistics in morphology: morphological productivity. In A. Ludeling & M. Kyto (Eds.), *Corpus Linguistics. An international handbook*. Berlin: Mouton De Gruyter.

Bannard, C., & Lieven, E. (2009). Repetition and reuse in child language learning. In R. Corrigan, E. A. Moravcsik, H. Ouali & K. M. Wheatley (Eds.), *Formulaic Language: Volume 2 Acquisition, loss, psychological reality, and functional explanations* (pp. 299-321). Amsterdam: John Benjamins.

Barabási, A.-L. (2002). *Linked: The new science of networks.* Cambridge, MA: Perseus Books.

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.

British National Corpus. (2007). BNC XML Edition http://www.natcorp.ox.ac.uk/corpus/.

Boyd, J. K., & Goldberg, A. E. (2009). Input effects within a constructionist framework. *Modern Language Journal, 93*(2), 418-429.

Briscoe, E., Carroll, J., & Watson, R. (2006). *The Second release of the RASP system.* Paper presented at the Proceedings of the COLING/ACL 2006 Interactive Presentation Sessions, Sydney, Australia.

Buttery, P., & Korhonen, A. (2005). *Large scale analysis of verb subcategorization differences between child directed speech and adult speech.* Paper presented at the Interdisciplinary Workshop on the Identification and Representation of Verb Features and Verb Classes. Saarbrucken, Germany.

Bybee, J. (2008). Usage-based grammar and second language acquisition. In P. Robinson & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition*. London: Routledge. (pp. 216-236).

Bybee, J. (2010). *Language, usage, and cognition*. Cambridge, UK: Cambridge University Press.

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

Casenhiser, D., & Goldberg, A. E. (2005). Fast mapping between a phrasal form and meaning. *Developmental Science, 8*, 500-508.

Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral & Brain Sciences, 31*, 489-509.

Christiansen, M. H., & Chater, N. (Eds.). (2001). *Connectionist psycholinguistics*. Westport, CO: Ablex.

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.

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

Croft, W. (2001). *Radical construction grammar: Syntactic theory in typological perspective.* Oxford, UK: Oxford University Press.

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

Danon, L., Díaz-Guilera, A., Duch, J., & Arenas, A. (2005). Comparing community structure identification methods. *Journal of Statistical Mechanics, 29*, P09008.

de Nooy, W., Mrvar, A., & Batagelj, V. (2010). *Exploratory Social Network Analysis with Pajek*. Cambridge, UK: Cambridge University Press.

Deacon, T. W. (1997). *The symbolic species: The co-evolution of language and the brain*. New York: W.W. Norton.

Ebbinghaus, H. (1885). *Memory: A contribution to experimental psychology* (H. A. R. C. E. B. (1913), Trans.). New York: Teachers College, Columbia University.

Ellis, N. C. (1998). Emergentism, connectionism and language learning. *Language Learning, 48*(4), 631-664.

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. (2006a). Language acquisition as rational contingency learning. *Applied Linguistics, 27*(1), 1-24.

Ellis, N. C. (2006b). Selective attention and transfer phenomena in SLA: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics, 27*(2), 1-31.

Ellis, N. C. (2008a). Optimizing the input: Frequency and sampling in Usage-based and Form-focussed Learning. In M. H. Long & C. Doughty (Eds.), *Handbook of second and foreign language teaching* (pp. 139-158). Oxford: Blackwell

Ellis, N. C. (2008b). Usage-based and form-focused language acquisition: The associative learning of constructions, learned-attention, and the limited L2 endstate. In P. Robinson & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 372-405). London: Routledge.

Ellis, N. C., & Cadierno, T. (2009). Constructing a second language. *Annual Review of Cognitive Linguistics, 7* (Special section), 111-290.

Ellis, N. C., & Ferreira-Junior, F. (2009a). Construction learning as a function of frequency, Frequency distribution, and function. *Modern Language Journal, 93*, 370-386.

Ellis, N. C., & Ferreira-Junior, F. (2009b). Constructions and their acquisition: Islands and the distinctiveness of their occupancy. *Annual Review of Cognitive Linguistics*, 111-139.

Ellis, N. C., & Larsen-Freeman, D. (2006). Language emergence: Implications for Applied Linguistics [Special Issue]. *Applied Linguistics, 27*(4).

Ellis, N. C., & Larsen-Freeman, D. (2009a). Constructing a second language: Analyses and computational simulations of the emergence of linguistic constructions from usage. *Language Learning, 59*(Supplement 1), 93-128.

Ellis, N. C., & Larsen-Freeman, D. (2009b). Language as a Complex Adaptive System [Special Issue]. *Language Learning, 59: Supplement 1*.

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*. Berlin: Mouton de Gruyter.

Evert, S. (2005). *The Statistics of word cooccurrences: Word pairs and collocations.* University of Stuttgart, Stuttgart.

Ferrer i Cancho, R., & Solé, R. V. (2001). The small world of human language. *Proceedings of the Royal Society of London, B., 268*, 2261-2265.

Ferrer i Cancho, R., & Solé, R. V. (2003). Least effort and the origins of scaling in human language. *Proceedings of the National Academy of Sciences, 100*, 788-791.

Ferrer i Cancho, R., Solé, R. V., & Köhler, R. (2004). Patterns in syntactic dependency networks. *Physical Review, E69*, 0519151-0519158.

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

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, UK: Oxford University Press.

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

Gries, S. T., & Wulff, S. (2005). Do foreign language learners also have constructions? Evidence from priming, sorting, and corpora. *Annual Review of Cognitive Linguistics, 3*, 182-200.

Harnad, S. (Ed.). (1987). *Categorical perception: The groundwork of cognition*. New York: Cambridge University Press.

Hunston, S., & Francis, G. (1996). *Pattern grammar: A corpus driven approach to the lexical grammar of English*. Amsterdam: Benjamins.

Jurafsky, D., & Martin, J. H. (2009). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Kello, C. T., Brown, G. D. A., Ferrer-i-Cancho, R., Holden, J. G., Linkenkaer-Hansen, K., Rhodes, T., & Van Orden, G. C. (2010). Scaling laws in cognitive sciences. *Trends in Cognitive Science, 14*, 223-232.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.

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

Larsen-Freeman, D. (1997). Chaos/complexity science and second language acquisition. *Applied Linguistics, 18*, 141-165.

Leech, L. (2000). Grammars of spoken English: New outcomes of corpus-oriented research. *Language Learning, 50*, 675-724.

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

Lieven, E., & Tomasello, M. (2008). Children's first language acquisition from a usage-based perspective. In P. Robinson & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 168-196). New York and London: Routledge.

MacWhinney, B. (1987). The Competition model. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 249-308). Hillsdale, NJ: Erlbaum.

MacWhinney, B. (2000a). *The CHILDES project: Tools for analyzing talk, Vol 1: Transcription format and programs (3rd ed.).* Mahwah, NJ: Lawrence Erlbaum.

MacWhinney, B. (2000b). *The CHILDES Project: Tools for analyzing talk, Vol 2: The database (3rd ed.)*: (2000). . Mahwah, NJ: Lawrence Erlbaum.

MacWhinney, B. (2001). The competition model: The input, the context, and the brain. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 69-90). New York: Cambridge University Press.

MacWhinney, B. (Ed.). (1999). *The emergence of language*. Hillsdale, NJ: Erlbaum.

Manin, D. Y. (2008). Zipf's Law and avoidance of excessive synonymy *Cognitive Science, 32*, 1075-1098.

Miller, G. A. (2009). WordNet - About us. Retrieved March 1, 2010, from Princeton University http://wordnet.princeton.edu

Newman, M. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics, 46*, 323–351

Ninio, A. (1999). Pathbreaking verbs in syntactic development and the question of prototypical transitivity. *Journal of Child Language, 26*, 619-653.

Ninio, A. (2006). *Language and the learning curve: A new theory of syntactic development*. Oxford, UK: Oxford University Press.

Ninio, A. (2011). *Syntactic development, its input and output*. Oxford, UK: Oxford University Press.

O'Donnell, M. B., & Ellis, N. C. (2009, 20-23 July, 2009). *Measuring formulaic language in corpora from the perspective of Language as a Complex System.* Paper presented at the 5th Corpus Linguistics Conference, University of Liverpool.

O'Donnell, M. B., & Ellis, N. C. (2010). Towards an inventory of English verb argument constructions. *Proceedings of the 11th Annual Conference of the North American Chapter of the Association for Computational Linguistics,* (pp. 9-16) Los Angeles, June 1–6, 2010.

O'Donnell, M. B., Ellis, N. C., & Corden, G. (2012). *Exploring semantics in verb argument constructions using community identification algorithms*. Paper presented at the Language & Network Science symposium, The International Conference on Network Science NETSCI 2012, Northwestern University.

O'Donnell, M. B., Römer, U., & Ellis, N. C. (in preparation). *The Psychological reality of verb-argument constructions: Representing form, function, and frequency*.

Page, S. E. (Writer). (2009). *Understanding complexity* [DVD-ROM]. Chantilly, VA: The Teaching Company.

Pedersen, T., Patwardhan, S., & Michelizzi, J. (2004). *WordNet::Similarity – Measuring the* r*elatedness of* c*oncepts.* Paper presented at the Proceedings of Fifth Annual Meeting of the North American Chapter of the Association of Computational Linguistics (NAACL 2004) ,San Jose, CA.

Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: Bradford Books.

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.

Robinson, P., & Ellis, N. C. (Eds.). (2008). *A handbook of cognitive linguistics and second language acquisition*. London: Routledge.

Römer, U., O'Donnell, M., & Ellis, N. C. (in press, 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*. 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.

Saussure, F. d. (1916). *Cours de linguistique générale* [Course in general linguistics] (Roy Harris, Trans.). London: Duckworth.

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

Sinclair, J. (2004). *Trust the text: Language, corpus and discourse*. London: Routledge.

Sinclair, J. (Ed.). (1987). *Looking Up: An account of the COBUILD project in lexical computing*. London: Collins ELT.

Solé, R. V., Murtra, B., Valverde, S., & Steels, L. (2005). Language networks: Their structure, function and evolution. *Trends in Cognitive Sciences, 12*.

Steyvers, M., & Tennenbaum, J. (2005). The large-scale structure of semantic networks: statistical analyses and a model of semantic growth. *Cognitive Science, 29*, 41-78.

Talmy, L. (2000). *Toward a cognitive semantics: Concept structuring systems*. Cambridge, MA: MIT Press.

Taylor, J. R. (1998). Syntactic constructions as prototype categories. In M. Tomasello (Ed.), *The new psychology of language: Cognitive and functional approaches to language structure* (pp. 177-202). Mahwah, NJ: Erlbaum.

Taylor, J. R. (2002). *Cognitive grammar*. Oxford, UK: Oxford University Press.

Tomasello, M. (2003). *Constructing a language*. Boston, MA: Harvard University Press.

Waldrop, M. (1992). *Complexity: The emerging science at the edge of order and chaos*. New York: Simon & Schuster.

Zipf, G. K. (1935). *The psycho-biology of language: An introduction to dynamic philology*. Cambridge, MA: M.I.T. Press.

Notes

[1] ORIGIN late Middle English : from Latin cognitio(n-), from cognoscere ‘get to know.’

[2]“Any [linguistic] creation has to be preceded by an unconscious comparison of materials deposited in the store held by the language, where the sponsoring forms are arranged by syntagmatic and associative relations.” Ferdinand de Saussure (1916)[1972, p. 164]