

Constructing a Second Language: Analyses and Computational Simulations of the Emergence of Linguistic Constructions From Usage

Nick C. Ellis

University of Michigan

with

Diane Larsen-Freeman

University of Michigan

This article presents an analysis of interactions in the usage, structure, cognition, coadaptation of conversational partners, and emergence of linguistic constructions. It focuses on second language development of English verb-argument constructions (VACs: VL, verb locative; VOL, verb object locative; VOO, ditransitive) with particular reference to the following: (a) Construction learning as concept learning following the general cognitive and associative processes of the induction of categories from experience of exemplars in usage obtained through coadapted micro-discursive interaction with conversation partners; (b) the empirical analysis of usage by means of corpus linguistic descriptions of native and nonnative speech and of longitudinal emergence in the interlanguage of second language learners; (c) the effects of the frequency and Zipfian type/token frequency distribution of exemplars within the Verb and other islands of the construction archipelago (e.g., [Subj V Obj Obj_{path/loc}]), by their prototypicality, their generic coverage, and their contingency of form-meaning-use mapping, and (d) computational (emergent connectionist) models of these various factors as they play out in the emergence of constructions as generalized linguistic schema.

We thank our colleagues in this research: Steffi Wulff, Fernando Ferreira-Junior, and Ute Römer.

Correspondence concerning this article should be addressed to Nick C. Ellis or Diane Larsen-Freeman, University of Michigan, 500 E. Washington Street, Ann Arbor, MI 48104. Internet: ncellis@umich.edu or dianelf@umich.edu

The Emergence of Linguistic Constructions From Usage

Language has come to represent the world as we know it; it is grounded in our perceptual experience. Language is used to organize, process, and convey information from one person to another, from one embodied mind to another. Language is also used to establish and maintain social relationships and to enact functions. Language and its use are mutually inextricable; they determine each other.

Learning language involves determining structure from usage and this, like learning about all other aspects of the world, involves the full scope of cognition: the remembering of utterances and episodes, the categorization of experience, the determination of patterns among and between stimuli, the generalization of conceptual schema and prototypes from exemplars, and the use of cognitive models, metaphors, analogies, and images in thinking. At the same time, there is an all-important social dimension to the process. There is nothing that so well characterizes human social action as language. It is in the coadaptation in the micro-discursive encounters between conversation partners that learners experience relevant and accessible exemplars from which they will learn. Cognition, consciousness, experience, embodiment, brain, self, human interaction, society, culture, and history—in other words, phenomena at different levels of scale and time (Larsen-Freeman & Cameron, 2008)—are all inextricably intertwined in rich, complex, and dynamic ways in language, its use, and its learning. So we require perspectives on dynamic interactions at all levels, perspectives provided by general approaches such as Emergentism (Ellis, 1998; Ellis & Larsen Freeman, 2006a, 2006b; Elman, et al., 1996; MacWhinney, 1999), Chaos/Complexity Theory (Holland, 1992, 1998; Larsen-Freeman, 1997, 2002b; Larsen-Freeman & Cameron), and Dynamic Systems Theory (de Bot, Lowie, & Verspoor, 2007; Ellis, 2007, 2008a; Port & Van Gelder, 1995; Spivey, 2006; Thelen & Smith, 1994; van Geert, 1991) as they apply to usage-based theories of language (Barlow & Kemmer, 2000; Bod, Hay, & Jannedy, 2003; Bybee, 2005; Bybee & Hopper, 2001; Croft & Cruise, 2004) and first language acquisition (Goldberg, 1995, 2003, 2006; Tomasello, 1998, 2003) and second language acquisition (Ellis, 2002, 2005; Larsen-Freeman, 2003; Larsen-Freeman & Long, 1991; Robinson & Ellis, 2008).

This article applies these approaches to investigate linguistic constructions, their cognition, and their development. We focus on the second language development of English verb-argument constructions (VACs: VL verb locative; VOL, verb object locative; VOO, ditransitive) with particular reference to the following:

1. construction learning as concept learning following the general cognitive and associative processes of the induction of categories from experience of exemplars in conversational interaction;
2. the empirical analysis of usage by means of corpus linguistic descriptions of English native-speaker and nonnative speaker speech over time;
3. the islands (Tomasello, 1992) comprising each construction and the effects of frequency and type/token frequency distribution of their constituent exemplars, their prototypicality, and their contingency of form-meaning-use mapping;
4. computational (connectionist) models of these various factors as they play out in the emergence of constructions as generalized linguistic schema.

In addition to the general approaches we have just enumerated, our theoretical framework is also informed by cognitive linguistics, particularly constructionist perspectives (e.g., Bates & MacWhinney, 1987; Goldberg, 1995, 2003, 2006; Lakoff, 1987; Langacker, 1987; Ninio, 2006; Robinson & Ellis, 2008; Tomasello, 2003), corpus linguistics (Biber, Conrad, & Reppen, 1998; Sinclair, 1991, 2004), and psychological theories of cognitive and associative learning as they relate to the induction of psycholinguistic categories from social interaction (Ellis, 1998, 2002, 2003, 2006a, 2006b, 2006c). The basic tenets are as follows: Language is intrinsically symbolic. It is constituted by a structured network of constructions as conventionalized form-meaning-use combinations used for communicative purposes. As speakers communicate, they coadapt their language use on a particular occasion. From such repeated encounters, stable language-using patterns (Larsen-Freeman & Cameron, 2008) emerge. The patterns are eventually broken down and their form-meaning-use is extended in novel ways. Usage leads to these becoming entrenched in the speaker's mind and for them to be taken up by members of the speech community.

Constructions are of different levels of complexity and abstraction; they can comprise concrete and particular items (as in words and idioms), more abstract classes of items (as in word classes and abstract grammatical constructions), or complex combinations of concrete and abstract items (as mixed constructions). The acquisition of constructions is input-driven and depends on the learner's experience of these form-meaning-use combinations in interactions with others. They develop following the same cognitive principles as the learning of other categories, schemata, and prototypes (Cohen & Lefebvre, 2005; Murphy, 2003). Creative linguistic usage emerges from the collaboration of the memories of all of the utterances in a learner's entire history of language use and the frequency-biased abstraction of regularities within them (Ellis, 2002).

Cognitive linguistics, corpus linguistics, and psycholinguistics are alike in their realizations that we cannot separate grammar from lexis, form from function, form from meaning, meaning from context, nor structure from usage.

Constructions specify the morphological, syntactic, and lexical form of language and the associated semantic, pragmatic, and discourse functions (Figure 1). Any utterance is comprised of a number of constructions that are nested. Thus, the expression *Today he walks to town* is constituted of lexical constructions such as *today*, *he*, *walks*, and so forth, morphological constructions such as the verb inflection *s* signaling third-person singular present tense, abstract grammatical constructions such as Subj, VP, and Prepositional Phrases, the intransitive motion Verb-Locative (VL: [Subj V Obl_{path/loc}]) verb-argument construction (VAC), and so forth. The function of each of these forms contributes in communicating the speaker's intention.

Psychological analyses of the learning of constructions as form-meaning-use combinations is informed by the literature on the associative learning of cue-outcome contingencies for which the usual determinants include the following: factors relating to the form such as frequency and salience (Ellis, 2002; Larsen-Freeman, 1976); factors relating to the meaning such as significance in the comprehension of the overall utterance, prototypicality, ambiguity, generality, redundancy, and surprise value; factors relating to use such as the social value of particular forms or their value in discourse construction (Celce-Murcia & Larsen-Freeman, 1999; Larsen-Freeman, 2002a, 2003); factors relating to the contingency of form and meaning and use; and factors relating to learner attention, such as automaticity, transfer, and blocking (Ellis, 2002, 2003, 2006b, 2008b). These various factors conspire in the acquisition and use of any linguistic construction.

Whereas some constructions, like *walk*, are concrete, imageable, and specific in their interpretation, others are more abstract and schematic. For example, the caused motion construction, (e.g., X causes Y to move Z_{path/loc} [Subj V Obj Obl_{path/loc}]) exists independently of particular verbs; hence "Tom sneezed the paper napkin across the table": is intelligible despite "sneeze" being usually intransitive (Goldberg, 1995). How might verb-centered constructions develop these abstract properties? Semantic bootstrapping accounts suggest that they inherit their schematic meaning from the conspiracy of the particular types of verb that appear in their verb island (Pinker, 1989). The verb is a better predictor of sentence meaning than any other word in the sentence and plays a central role in determining the syntactic structure of a sentence (Tomasello, 1992). There is a close relationship between the types of verbs that typically appear within constructions (in this case, *put*, *move*, *push*, etc.); hence, their meaning

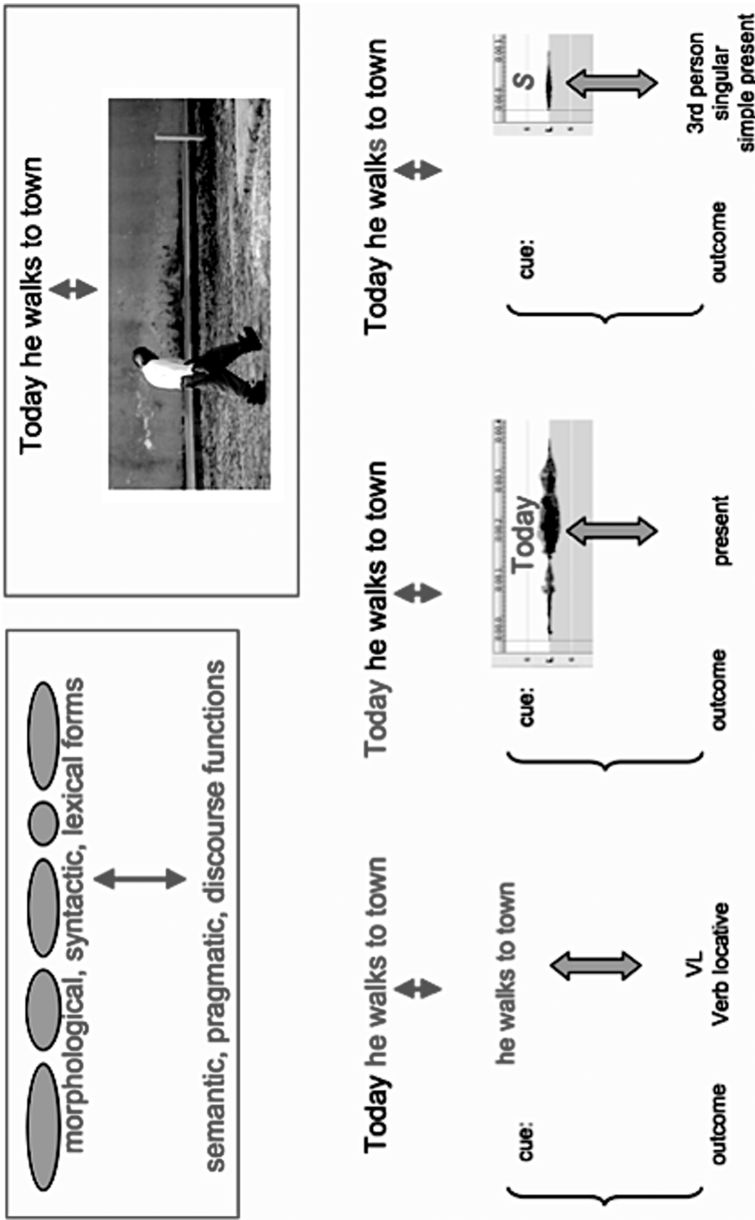


Figure 1 Constructions as form-function mappings. Any utterance comprises multiple nested constructions. Some aspects of form are more salient than others; the amount of energy in *today* far exceeds that in *s*.

as a whole is inducible from the lexical items experienced within them. Ninio (1999) argued that in child language acquisition, individual “pathbreaking” semantically prototypic verbs form the seeds of verb-centered argument-structure patterns, with generalizations of the verb-centered instances emerging gradually as the verb-centered categories themselves are analyzed into more abstract argument structure constructions.

Learning grammatical constructions thus involves the distributional analysis of the language stream and the contingent analysis of perceptual activity following general psychological principles of category learning. Categories have graded structures, with some members being better exemplars than others. The prototype is the best example, the benchmark against which surrounding “poorer,” more borderline instances are categorized. 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.

Frequency promotes learning, and psycholinguistics demonstrates that language learners are exquisitely sensitive to input frequencies of patterns at all levels (Ellis, 2002). In the learning of categories from exemplars, acquisition is optimized by the introduction of an initial, low-variance sample centered on prototypical exemplars (Elio & Anderson, 1981, 1984; Posner & Keele, 1968, 1970), which allows learners to get a “fix” on what will account for most of the category members. Then the bounds of the category can later be defined by experience of the full breadth of exemplars. Goldberg, Casenhiser, and Sethuraman (2004) demonstrated that in samples of child language acquisition, for each VAC there is a strong tendency for one single verb to occur with very high frequency in comparison to other verbs used, a profile that closely mirrors that of the mothers’ speech to these children. Dale and Spivey (2006) also showed how the child and his or her caregiver produce sequences of words or syntactic phrases during a conversation that match those being heard, a process they call “syntactic coordination.” Interesting from our point of view is that the researchers found a Zipf-like distribution in the patterns that were shared with each child and caregiver pair. In other words, there are highly frequent sequences of word classes guiding the recurrent patterns in conversation (Larsen-Freeman & Cameron, 2008). Additionally, in second language acquisition, there is evidence of coadaptation of conversation partners (“foreigner talk discourse”; Larsen-Freeman & Long, 1991) and with teachers and students in classrooms, with the result that learners receive an optimal sample of language from which to learn. In natural language, too, Zipf’s law (Zipf, 1935) describes how the highest frequency words disproportionately account for the most linguistic tokens. Goldberg et al. (2004) showed that Zipf’s

law applied within VACs, too, and 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:

- 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% of the time, *take* 10% of the time, and *do/pick* 6% of the time, 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 Obj₁ 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.

Consider language as it passes, utterance by utterance, as illustrated in Figure 2. Learners with a history of exposure to this profile of natural language

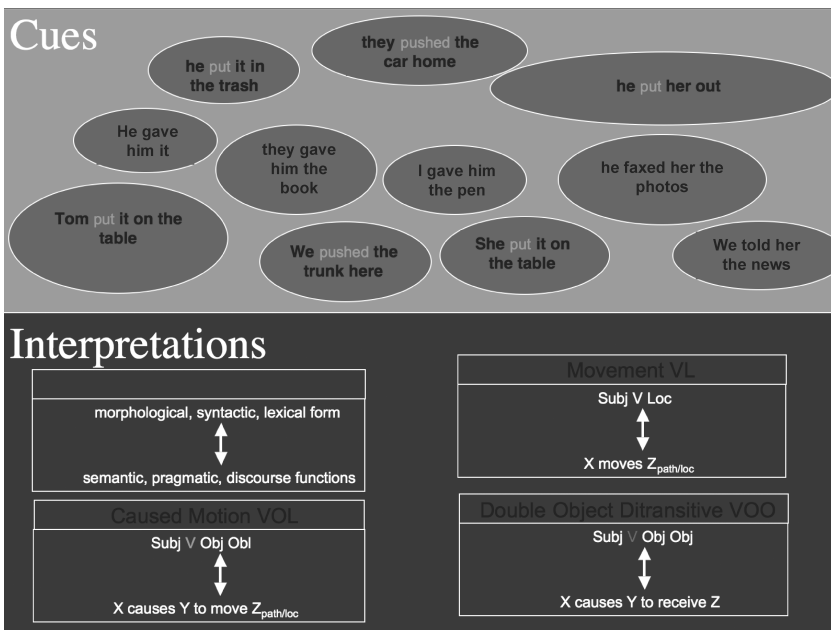


Figure 2 Verb island occupancy as cues to VAC membership.

might thus successfully categorize the different utterances as examples of different VAC categories on the basis of the occupants of the verb islands.

However, if the verbs were the only cues that were available, then VACs could have no abstract meaning above that of the verb itself. For “Tom sneezed the paper napkin across the table” to make sense despite the intransitivity of *sneeze*, the hearer has to make use of additional information from the syntactic frame. In considering how children learn lexical semantics, Gleitman (1990) argued that they made use of clues from syntactic distributional information—nounlike things follow determiners, prepositions most often prepose a noun phrase in English, and so forth. The two alternatives of *semantic* and *syntactic bootstrapping* are by no means mutually exclusive; indeed, they reinforce and complement each other.

In the identification of the caused motion construction (X causes Y to move $Z_{\text{path/loc}}$ [Subj V Obj Obl_{path/loc}]), the whole frame as an archipelago of islands is important. The Subj island helps to identify the beginning bounds of the parse. More frequent, more generic, and more prototypical occupants will be more easily identified. Pronouns, particularly those that refer to animate entities, will more readily activate the schema (Childers & Tomasello, 2001; Wilson, 2003). As illustrated in Figure 3, the Obj island, too, will be more readily identified when occupied by more frequent, more generic, and more prototypical lexical items (pronouns like *it*, required by discourse constraints, rather than nouns such as *napkin*). So, too, the locative will be activated more readily if opened by a prepositional island populated by a high-frequency, prototypical exemplar such as *on* or *in* (Tomasello, 2003, p. 153). Activation of the VAC schema arises from the conspiracy of all of these features, and arguments about Zipfian type/token distributions and prototypicality of membership extend to all of the islands of the construction.

Thus, frequency of usage defines construction categories. However, there is one additional qualification to be borne in mind. Some lexical types are very specific in the VACs that they occupy; the vast majority of their tokens occur in just one VAC and so they are very reliable and distinctive cues to it. Other lexical types are more widely spread over a range of constructions, and this promiscuity means that they are not faithful cues. *Put* occurs almost exclusively in VOL; it is defining in the acquisition of this VAC and a distinctive and reliable cue in its subsequent recognition. *Turn*, however, occurs both in VL and VOL and is less distinctive in distinguishing between these two. Similarly, *send* is attracted to both the VOO and VOL constructions and so is a less discriminating cue for these categories. Consider the other islands too. It is clear that however useful they are at defining the beginning region of interest in the VAC parse,

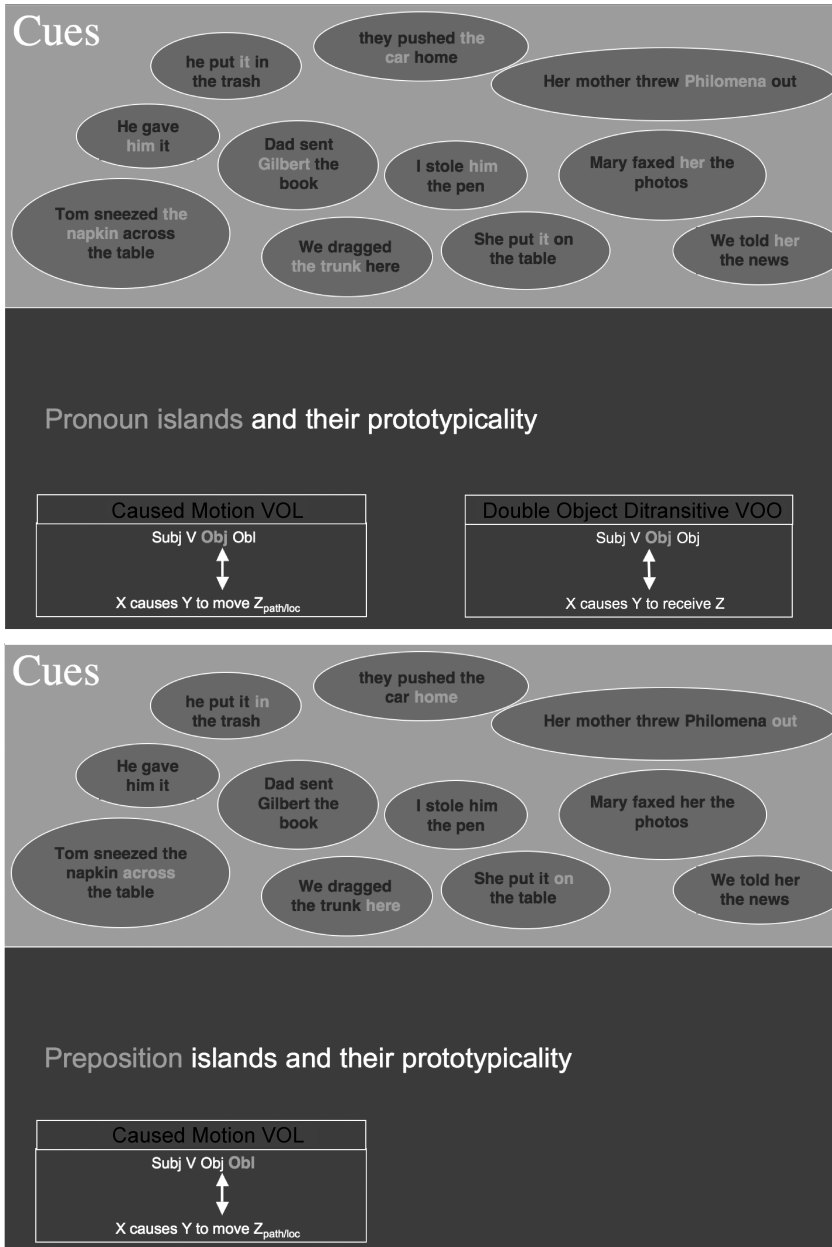


Figure 3 Other syntactic islands and their occupants as cues to VAC identity.

subject pronouns freely occupy any VAC with hardly any discrimination except that concerning animacy of agent. Prepositions are substantially selective for locatives, but as a class, they do not distinguish between the transitive and intransitive VACs, and so on.

The associative learning literature has long recognized that although frequency is important, so, too, is contingency of mapping. Consider how, in the learning of the category of birds, although eyes and wings are equally frequently experienced features in the exemplars, it is wings that are distinctive in differentiating birds from other animals. Wings are important features to learning the category 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. Contingency, or reliability of form-function mapping, is a driving force of all associative learning (Rescorla, 1968). It, and its associated aspects of predictive value, information gain, and statistical association, is therefore central in psycholinguistic theories of language acquisition too (Ellis, 2006b, 2006c, 2008b; Gries & Wulff, 2005; MacWhinney, 1987; Wulff, Ellis, Römer, Bardovi-Harlig, & LeBlanc, 2009).

Taken together, these considerations of language acquisition as the associative learning of schematic constructions from experience of exemplars in usage, adjusted for comprehension/learning, generate a number of hypotheses concerning VAC acquisition:

- H1. The frequency distribution for the types occupying the verb island of each VAC will be Zipfian.
- H2. The first verbs to emerge in each VAC will be those which appear more frequently in that construction in the input.
- H3. The pathbreaking verb for each VAC will be much more frequent than the other members.
- H4. The first verbs to emerge in each VAC will be prototypical of that construction's interpretation.
- H5. The first verbs to emerge in each construction will be those which are more distinctively associated with that construction in the input.

We also assume similar contributions relating to H1–H5 from the other islands in each VAC, although perhaps to a lesser degree. Ellis and Ferreira-Junior (2009a, 2009b) tested these hypotheses using corpus data of English. Their methods and findings are reported in depth in those articles. What we do here is summarize the findings in order to lay the foundations for emergentist simulations designed to understand the interactions of these factors in development.

The Naturalistic Acquisition of English VACs

Methods

Ellis and Ferreira-Junior (2009a) analyzed the speech of second language learners of English VACs in the European Science Foundation (ESF) corpus (<http://www.mpi.nl/world/tg/lapp/esf/esf.html>; Dietrich, Klein, & Noyau, 1995; Feldweg, 1991; Perdue, 1993), which collected the spontaneous and elicited second language of adult immigrants recorded longitudinally in interviews every 4–6 weeks for approximately 30 months. They focused on seven English as a second language (ESL) learners living in Britain whose native languages were Italian ($n = 4$) or Punjabi ($n = 3$). Data from 234 sessions were gathered and transcribed for these ESL learners and their native-speaker (NS) conversation partners from a range of activities.

They performed semiautomated searches through the transcriptions to identify the VACs of interest and to tag them as VL, VOL, or VOO following the operationalizations described in the work of Goldberg, Casenhiser & Sethuraman (2004); for example:

- a) you come out of my house. [come] [VL]
- b) Charlie say # shopkeeper give me one cigar [give] [VOO]
- c) no put it in front # thats it # yeah [put] [VOL]

For the NS conversation partners, they identified 14,574 verb tokens (232 types), of which 900 tokens were identified to occur in VL (33 types), 303 in VOL (33 types), and 139 in VOO constructions (12 types). For the ESL learners, they identified 10,448 verb tokens (234 types), of which 436 tokens were found in VL (39 types), 224 in VOL (24 types), and 36 in VOO constructions (9 types).

Hypotheses and Findings

H1. The frequency distributions for the types occupying the verb island of each VAC are Zipfian

The frequency distributions of the verb types in the VL, VOL, and VOO constructions produced by the interviewees and the learners are shown in Figure 4. For the NS interviewees, *go* constituted 42% of the total tokens of VL, *put* constituted 35% of VOL use, and *give* constituted 53% of VOO. After this leading exemplar, subsequent verb types decline rapidly in frequency. For the ESL learners, again, for each construction there was one exemplar that accounted for the lion's share of total productions of that construction: *go* constituted 53% of VL, *put* constituted 68% of VOL, and *give*

constituted 64% of VOO. Plots of these frequency distributions as log verb frequency against log verb rank produced straight-line functions explaining in excess of 95% of the variance, thus confirming that Zipf's law is a good description of the frequency distributions with the frequency of any verb being inversely proportional to its rank in the frequency table for that construction, the relationship following a power function.

H2. The verbs to emerge first in each VAC are those which appear more frequently in that construction in the input

The order of emergence of verb types in the learner constructions followed the frequencies in the interviewer data. Correlational analyses across all 80 verb types that featured in any of the NS and/or NNS constructions confirmed this. For the VL construction, frequency of lemma use by learner correlated with that by NS interviewer, $r(78) = 0.97$, $p < .001$. The same analysis for VOL resulted in $r(78) = 0.89$, $p < .001$, and for VOO it resulted in $r(78) = 0.93$, $p < .001$.

H3. The pathbreaking verb for each VAC is much more frequent than the other members

Go was the first-learned verb for VL, *put* for VOL, and *give* for VOO. The Zipfian frequency profiles (Figure 4) for the types/tokens confirm H3. The emergent curves (Figures 5–7, left-hand panels; right-hand shows results from simulations to be described in the last part of the article) showed in each case that the verb to first emerge seeded the construction and predominated in its cumulative usage but thereafter the construction grew in membership as verbs similar in meaning to the pathbreaker joined one at a time.

H4. The first-learned verbs in each VAC are prototypical of that construction's interpretation

In order to determine the degree to which different verbs matched the prototypical semantics of the three VACs, Ellis and Ferreira-Junior (2009b) had native English speakers rate the verbs for the degree to which they matched a VL schema (the movement of someone or something to a new place or in a new direction), a VOL schema (someone causes the movement of something to a new place or in a new direction), or a VOO schema (someone causes someone to receive something). They then assessed the association between verb-acquisition order and prototypicality so measured.

For the VL construction the most used verb, *go*, was rated as 7.4 out of 9 in terms of the degree to which it matched the prototypical schematic meaning.

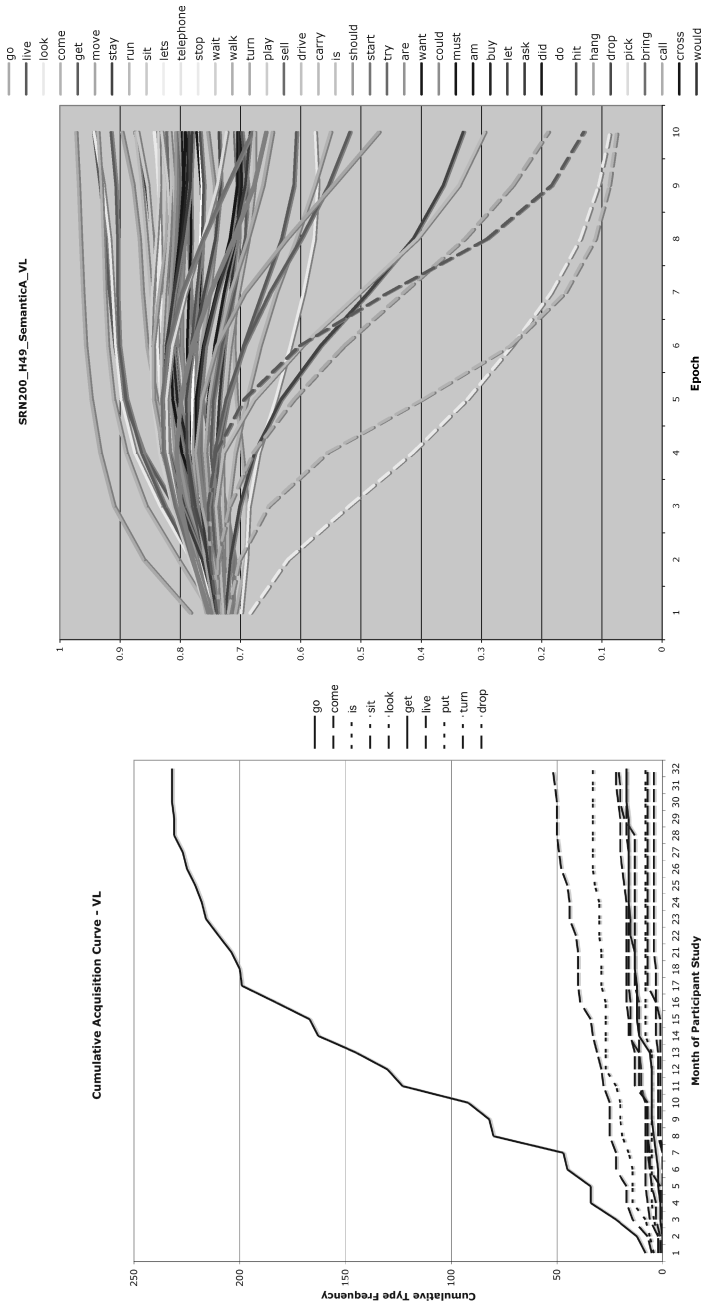


Figure 5 Learner use of verb types in the VL construction as a function of study month (left panel) alongside activation of the VL pattern by different verb types as a function of epoch of training in Simulation 1 (right panel).

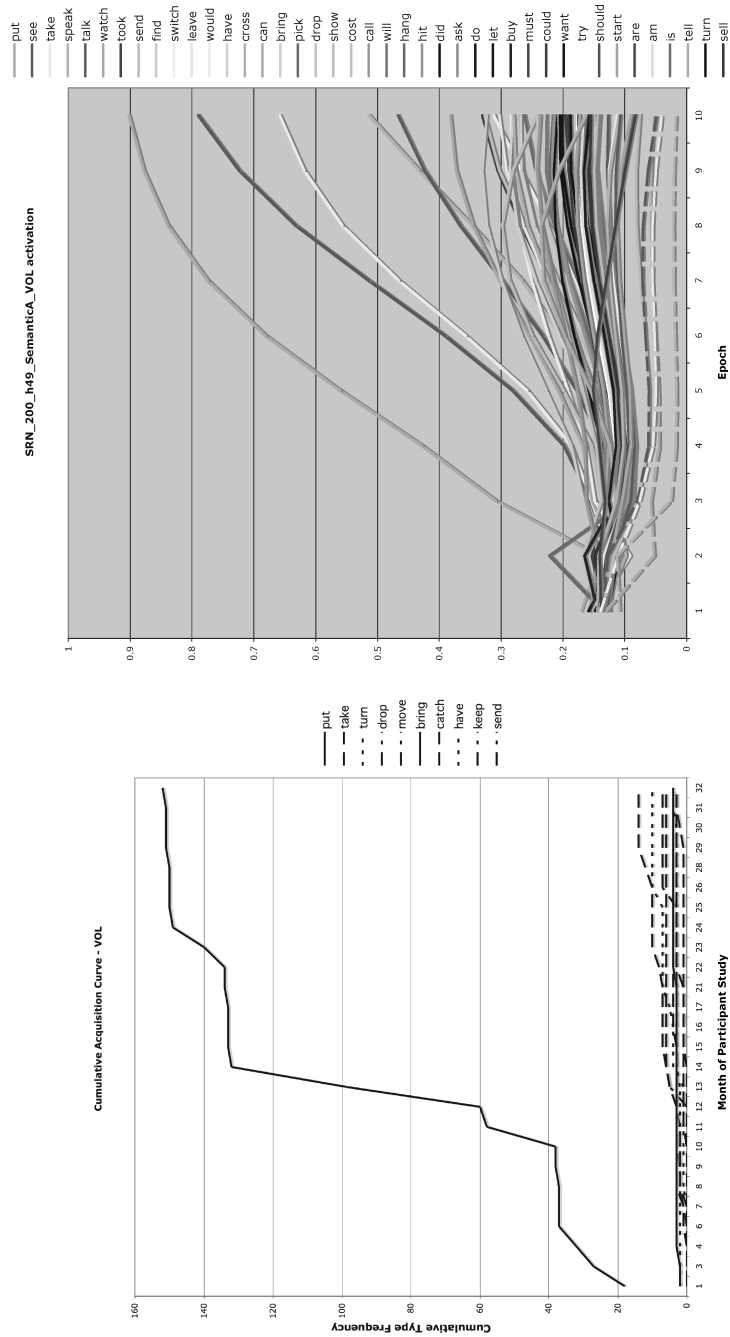


Figure 6 Learner use of verb types in the VOL construction as a function of study month (left panel) alongside activation of the VOL pattern by different verb types as a function of epoch of training in Simulation 1 (right panel).

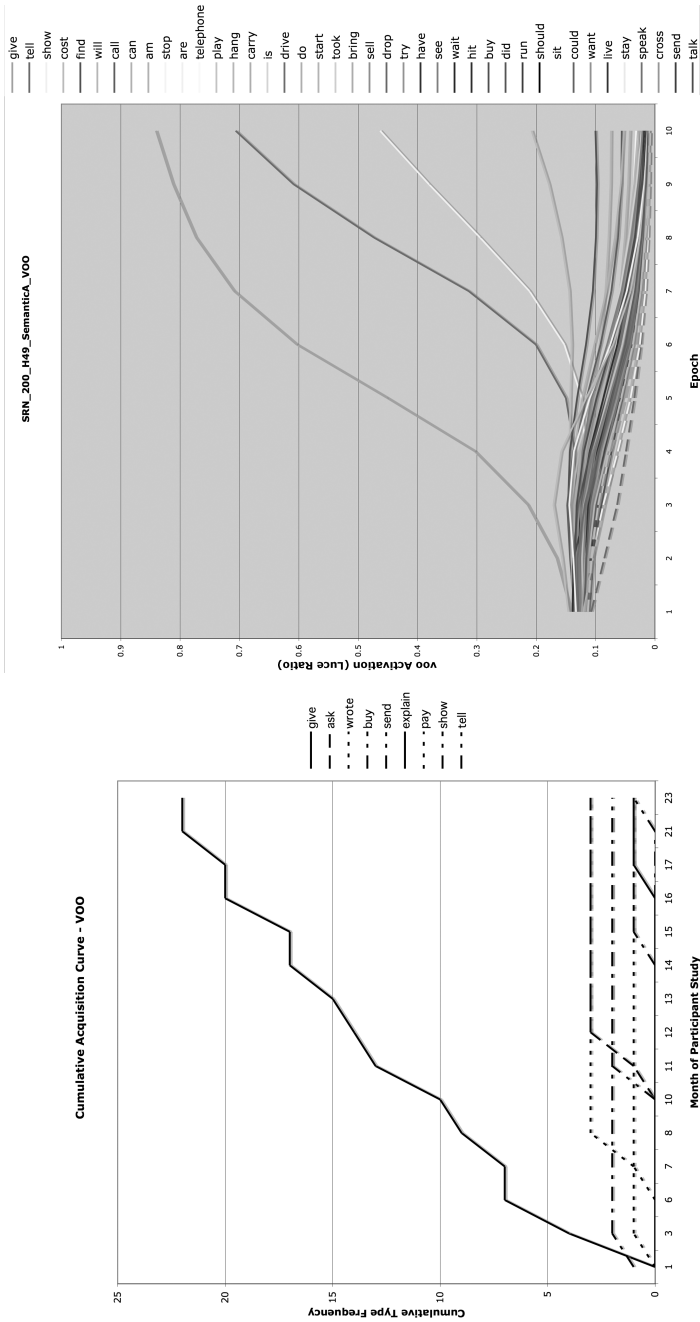


Figure 7 Learner use of verb types in the VOO construction as a function of study month (left panel) alongside activation of the VOO pattern by different verb types as a function of epoch of training in Simulation 1 (right panel).

The correlation between prototypicality of verb meaning and log frequency of learner use was VL $\rho(78) = 0.44$, $p < .001$. They had expected a higher correlation than this but realized that 10 other verbs surpassed *go* in this rating (*walk* [9.0], *move* [8.8], *run* [8.8], *travel* [8.8], *come* [8.4], *drive* [8.2], *arrive* [8.0], *jump* [8.0], *return* [8.0], and *fall* [7.8]). These match the schemata very well, but their additional specific action semantics limit the generality of their use. What is special about *go* is that it is prototypical and generic—thus widely applicable. The same pattern held for the other constructions. For VOL, the most used verb *put* was rated 8.0 in terms of how well it described the construction schema. For the VOO construction, the most used verb *give* was rated 9.0 in terms of how well it described the VOO schema.

In sum, these data demonstrate that learner VAC development is seeded by the highest frequency, prototypical, and generic exemplar across learners and VACs. These are the exemplars that are provided in NS-nonnative speaker interaction. The use of such exemplars presumably facilitates comprehension in the micro-discursive moment and perhaps their subsequent emergence and ultimate acquisition.

Ellis and Ferreira-Junior (2006a) extended these analyses, first to include the dimension of contingency/distinctiveness of form-meaning association and, second, to investigate the contribution of the other islands in the VAC archipelago.

H5. The first verbs to emerge in each construction are those which are more distinctively associated with that construction in the input

To assess the association strength between the verbs and the VACs in which they occur, they used collexeme strength (the log to the base 10 of the p -value of the Fisher-Yates exact test), a measure of contingency from collostructional analysis (Gries & Stefanowitsch, 2004; Stefanowitsch & Gries, 2003).

As already described under H2, learner usage was strongly associated with frequency in the NS speech (over the 80 verbs, VL, $r = .97$; VOL, $r = .89$; VOO, $r = .93$). Their analyses under H5 showed that, if anything, learner uptake was predicted even more so by collexeme strength in the NS speech (over the 80 verbs, VL, $r = .96$; VOL, $r = .97$; VOO, $r = .97$).

H6. The frequency distribution for the types occupying each of the islands of each VAC is Zipfian

Ellis and Ferreira-Junior (2009b) determined the frequency distributions of the types occupying each (nonverb) island in the VL (Subj, Prep, Locative), VOL (Subj, Obj, Prep, Locative), and VOO (Subj, Obj₁, Obj₂) constructions

produced by the interviewers and the learners. For each construction, the frequency distribution for each island was Zipfian. In each case, for NS and NNS both, the lead exemplars took the lion's share of instances in that island, and the distribution was a power function as indexed by the log frequency versus log type rank regression being linear.

H7. The first types to emerge in each VAC island are those which appear more frequently in that construction island in the input

There was a clear correspondence between the types used in each island by the NNSs and the types that occupy them in the NS speech. The interviewers filled the Subj island of VL with the following top eight types, in decreasing order: *you, to* [verb in infinitive phrase], *implied you* [imperative], *I, he, they, we, us*. The corresponding list for the learners was as follows: *implied you* [imperative], *I, you, he, they, to* [verb in infinitive phrase], *she, we*. A similar profile was found for the Subj island for VOL: NS (*you, implied you* [imperative], *to, I, they, he, we, she*); NNS (*implied you* [imperative], *I, you, to* [verb in infinitive phrase], *he, the, bag, they*); and for VOO, top four NSs (*I, you, implied you* [imperative], *to* [verb in infinitive phrase]), NNS (*they, I, she, implied you* [imperative]). Although a potentially infinite range of nouns could occupy the Subj islands in these different constructions, in NS and learner alike, they were populated by far by a few high-frequency generic forms, the pronouns, both to honor discourse constraints and perhaps as a consequence of NSers making adjustments to facilitate comprehension.

The top eight occupants of the Prep island were in NS VLs (*to, in, at, there, from, into, out, back*) and in NNS VLs (*to, in, out, on, down, there, inside, up*). Similar profiles occurred for the Prep island of VOL: NS (*in, on, there, off, out, up, from, to*); NNS (*in, on, there, the table, up, from, the bag, down*). Although a wide range of directions or places could occupy the postverbal island in these two constructions, in NS and learners alike, it was occupied by far by a few high-frequency generic prepositions.

Finally, we look at the Obj islands of VOO. For Obj₁, the interviewers' top five occupants were (*you, me, him, her, it*) and the NNS learners' top three were (*me, you, him*). For Obj₂, the NS top eight were (*AMOUNTMONEY* [like 20 pounds, 3 pounds, etc.], *the names, a bit, money, a book, a picture, something, the test*) and the NNS top eight were (*money, a letter, hand, something, the money, a bill, a cheque, a lot*).

The general pattern, then, for each island of each VAC, is that there was high correspondence between the top types used in each island by the learners and the types that occupy them in NS input typical of their experience. Although we

do not claim that NSs make conscious choices to facilitate acquisition by NSSs, we do believe that there is coadaptation between interlocutors that facilitates comprehension and, therefore, potentially scaffolds acquisition.

H8. The first pathbreaking type for each VAC island is more frequent than the other members

The qualitative patterns summarized under H7 demonstrates that, unlike for the verbs that center the semantics of each VAC, there was no single pathbreaker that initially takes over each of the other islands of the VAC exclusively. Nevertheless, for each construction, there was a high overlap between NS and NNS use of the top 5–10 occupant types, which together make up the predominance of its inhabitation.

H9. The first-learned types in each VAC island are prototypical of that island's contribution to the construction's interpretation

The 5–10 major occupant types for each island described under H7 do indeed seem to be prototypical in role. Although a very wide range of nouns could occupy the Subj islands in the VL, VOL, and VOO constructions, in NS and NNS learner alike, these were occupied by the most frequent, prototypical and generic forms for this slot: pronouns such as *I, you, it, we*, and so forth. The Prep islands in VL and VOL were clearly identified with high frequency prototypical generic prepositions such as *in, on, there, to, and off*. Likewise, the Obs in VOO are stereotypic in their interpretations and there is a broad overlap between NS and NNS use: Because of their informational status, people, deictically present (as pronouns), routinely give people (as pronouns) money, letters, bills, or books.

H10. The first types to emerge in each VAC island are those more distinctively associated with that construction island in the input

Their analyses showed that certain subjects were more significantly associated with certain VACs (i.e., *it* and *I* for VOO and implied *you* in the imperative for VOL). Nevertheless, comparison of the data under H5 showed that verbs are generally much more distinctively associated with these VACs than Subjs in terms of collexeme strength. Thus, although the occupants of Subj do follow a Zipfian distribution led by pronouns and thus could indeed signal the beginning of a VAC parse, they tend not to be associated with any particular VAC. Prepositions were much more like the verbs in their selectivity; *to, back, in, and out* were distinctively associated with VL, *on, off*, and *up* were strongly selective of VOL, and all of these prepositions were repulsed by VOO. For

the Obj₁ islands, any Obj₁ repulsed VL, *it*, *money*, *them* and *that* were very significantly distinctive of VOL, and the object pronouns *you*, *me*, *him* and *her* were distinctive recipients in VOO.

Together, these analyses demonstrated that although the verb island is most distinctive, the constituency of the other islands is by no means negligible in determining VAC identity. In particular, VL and VOL are highly selective in terms of their Prep occupancy, and Obj₁ types clearly select among VOO, VOL, and VL.

Interim Conclusions

These findings demonstrate a range of influences in the emergence of linguistic constructions. For each VAC island there is the following:

1. the frequency and frequency distribution of the form types;
2. the frequency, the frequency distribution, the prototypicality and generality of the semantic types, their importance in interpreting the overall construction;
3. the reliabilities of the mapping between items 1 and 2;
4. the degree to which the different elements in the VAC sequence (such as Subj V Obj Obl) are mutually informative and form predictable chunks.

There are many factors involved, and so far, all we have done is to look at each, hypothesis by hypothesis, variable by variable, one at a time. However, they interact. What we really want is a model of usage and its effects on acquisition. We can measure these factors individually. However, such counts are vague indicators of how the demands of human interaction affect the content and ongoing coadaptation of discourse, how this is perceived and interpreted, how usage episodes are assimilated into the learner's system, and how the system reacts accordingly. We need a model of learning, development, and emergence. Learning is dynamic; it takes place during processing, as Hebb (1949), Craik and Lockhart (1972), Elman et al. (1996), and Bybee and Hopper (2001) have variously emphasized from their neural, cognitive, connectionist, and linguistic perspectives, and the units of learning are thus the units of language processing episodes. Before learners can use constructions productively, they have to encounter useful exemplars and analyze them, to identify their linguistic form and to map it to meaning and use. Each construction has its form, its meaning, its use, and its contingency of mapping among them. Our analyses here have shown that the input that learners get is biased so that they frequently experience forms that are distinctively associated with prototypical functions or construals. People's actions in the world, their categorization of

the world, and their talk about these actions and classifications occur in broadly parallel relative frequencies. We believe that these parallels make constructions learnable, but we need a method for pursuing these ideas.

Connectionist (Emergent) Simulations of Acquisition

Although decontextualized, computer simulation allows the investigation of the dynamic interactions of these factors in language learning, processing, and use. In the remainder of this article we present two different connectionist architectures for the simulation of the emergence of the VACs described here.

Architecture 1

We use serial connectionist models. Simple recurrent networks (SRNs) have a proven utility in simulating language learning: allowing the identification of word boundaries from sequences of phonemes, word classes from sequences of words in small language samples, and phrase structure and lexical semantics from large usage corpora (Borovsky & Elman, 2006; Christiansen & Chater, 2001; Elman, 1990, 1998, 2004; Redington & Chater, 1998).

In SRNs the input to the network is the current item (letter, phoneme, word, phrase, or whatever) in a language stream, and the output represents the network's best guess as to the next item. The difference between the predicted state and the correct subsequent state (the target output) is used by the learning algorithm to adjust the weights in the network at every time step. In this way, the network improves its accuracy with experience. A common architecture involves an input layer (a layer of processing units that receive inputs from sources external to the network itself, whose units code the set of items in the language and whose activity identifies which item is currently being experienced), an output layer (which codes the language in the same way and which sends signals outside the network itself), and a hidden layer (whose units communicate between the inputs and outputs and whose activity represents the internal state of the model). A context layer is a special subset of inputs that receives no external input but which feeds the result of the previous processing back into the internal representations. Thus, at time 2, the hidden layer processes both the input of time 2 and, from the context layer, its own prior state of processing at time 1, and so on, recursively. It is by this means that SRNs capture the sequential nature of temporal inputs.

Elman (1990) trained a network of 31 input nodes, 31 output nodes, and hidden and context vectors of 150 units, each with sequences of words following a simple grammar. A 27,534-word sequence formed the training set

and the network had to learn to predict the next word in the sequence. At the end of training, Elman cluster-analyzed the representations that the model had formed across its hidden unit activations for each word + context vector. The resultant dendrogram demonstrated that the network had discovered several major categories of words: large categories of verbs and nouns, smaller categories of inanimate or animate nouns, smaller still categories of human and nonhuman animals, and so forth. This graded, soft, and implicit category structure had emerged from the language input without any semantics or real-world grounding.

Our network architecture is shown in Figure 8. There is a $15 \times 14 = 210$ -unit input layer, which is used to code the most frequent words in the NS

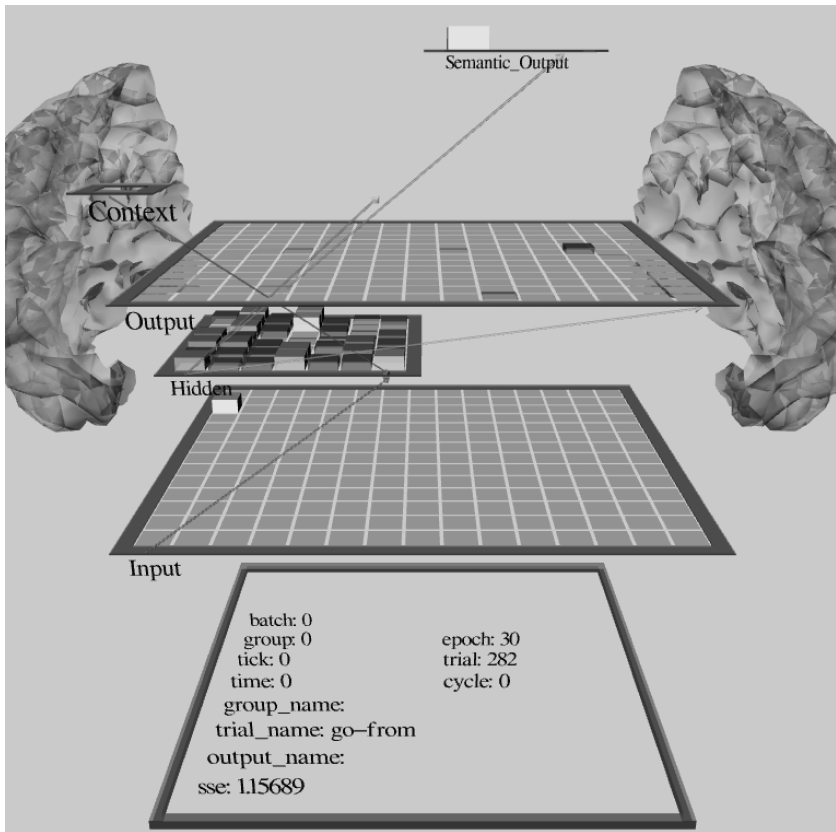


Figure 8 Emergent architecture 1.

constructions. These words and their codings are shown in the Appendix. Input units have activity 0 or 1. In Figure 8, input unit 166 is active; in our coding, this represents the word “go” There are 49 hidden units (HUs) and a context layer (iconified in the diagram for simplicity) of the same dimensions. There is a $15 \times 14 = 210$ -unit output layer that codes the most frequent words in the NS constructions in the same way as the input. There is additionally a three-unit Semantic output layer whose units code the semantics associated with these utterances (1 for VL, 2 for VOL, 3 for VOO). Figure 8 shows the network in a trained state, in which, on experiencing “go,” the model is predicting that corresponding VL semantics are likely and that several words could probably follow, including “to,” “from,” “left,” and so forth (shown by raised columns on the corresponding output units).

Simulations were run using the “Emergent” software environment (Aisa, Mingus, & O’Reilly, 2008a, 2008b). The models were feedforward and learned by backpropagation of error. The learning rate was 0.1, no momentum was used, and training was carried out in epochs of 5,980 sweeps (one sweep corresponding to presentation of one word). The epoch sequence was constructed as follows: 1,341 NS Interviewer constructions (900 VL, 302 VOL, and 139 VOO) in the analysis of the ESF data were randomly sorted and then coded into input-output word pairs; for example, the construction “I gave you the money” became ###-I, I-give, give-you, you-the, the-money (### = start of construction marker, verbs were lemmatized). The codes followed the system shown in the Appendix; for example, *go* as a string of 210 zeros (aligned to the input units) with the exception of unit 166 with value 1, ### as a 1 followed by 209 zeros, and so forth. Low-frequency words, which do not appear in the Appendix, are coded with a 1 on unit 201, indicating “word not known.”

The model is initialized with random weights. Imagine an epoch where the first construction, by random sort, is “I gave you the money,” a VOO construction. Trial 1 shows the input “###” and the model makes some outputs according to its random weights. It is then shown on the output units that the next word is *I* with corresponding VL semantic activity. Its internal state is copied to the context units. It adjusts its weights by backprop to better approximate this outcome in the future, exploiting the connections through its hidden units. On the next trial, the model experiences “I” in the updated context and it makes some prediction according to its current weights. It is then shown on the output units that the next word is “go” with corresponding VL semantic activity, and so on.

The model as a whole was trained over 10 such epochs. At the end of each, it was tested by giving it, without feedback, the 200 words shown in

the Appendix, one by one, and for each, its predictions of the corresponding semantics were recorded along with the internal state of activation across the hidden units. We also tested it for its generalized responses to the test patterns “you go to the shop,” “put it in there,” and “you give me money” in all variants where each word is substituted in turn by a wug pattern (that becomes a “not known” code).

Results

The amount of activation for the different verbs after each of the 10 epochs of training is shown for the VL, VOL, and VOO units in the right-hand panels of Figures 5–7.

In the simulation data in Figure 5 it is clear that the model learns early on that VL is the most probable (default) construction with a baseline activation of around 0.75. The verb that seeds this construction is *go*, followed somewhat later by *come*. Comparison of the learner data and simulation data shows a very similar development profile. The rank order correlation between the VL verb emergence order in the learners and VL activation for those verbs in the simulations at epoch 10 was $\rho = 0.77$.

In Figure 6, for learner and model alike, *put* is the verb which leads the development of VOL. Again, there are similar profiles across the learners and simulations. The rank order correlation between VOL verb emergence in the learners, and verb VOL activation in the simulations at epoch 10 was $\rho = 0.78$.

In Figure 7, there is a clear pacemaker for the VOO construction in simulation and learner alike, *give*. The rank order correlation between learner VOO verb emergence order, and verb VOO activation in the simulations at epoch 10 was $\rho = 0.81$.

Thus, an SRN exposed to NS usage acquires these VACs using the verbs at their center as the primary cues to construction category. The development of the different constructions and the different verbs in each demonstrate that model and learner alike are sensitive to the frequency and frequency distribution of the verbs, of the semantic types, and of the reliabilities of mapping between them.

What of the other cues? Figure 9 shows that the simulation, like the learners, comes to learn that some prepositions are more reliable cues to particular VACs than others (*to*, *at*, and *into* for VL, *on*, *up*, *under* and *over* for VOL, all such prepositions inhibit VOO activation; subject pronouns *they*, *you* and *he* for VL, inanimate object pronouns *it*, *them*, *these* for VOL, animate object pronouns *me*, *him* for VOO).

Finally, all of these test activations as reported here for the simulations were for individual words, out of context. Yet the driving force of these investigations

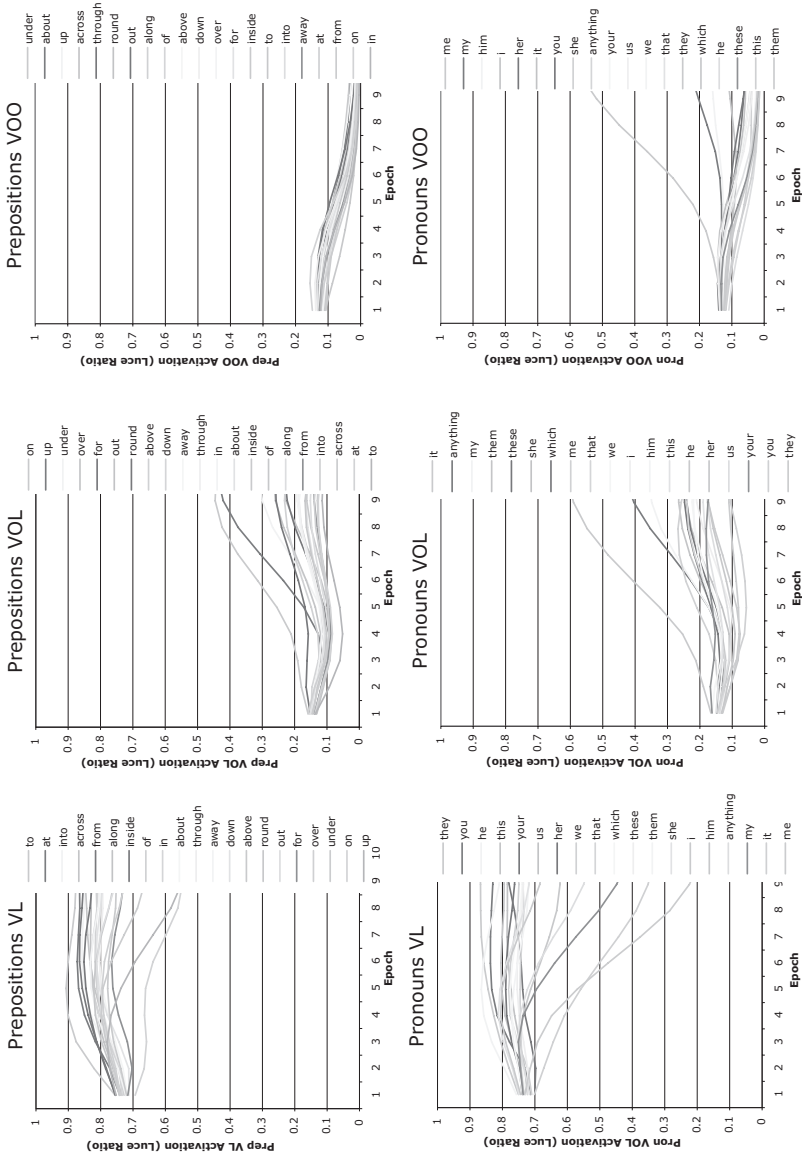


Figure 9 Activation of the VL, VOL, and VOO patterns by different Prepositions and Pronouns in Simulation 1.

was to understand their conspiracy as cue follows cue, and these different features interact dynamically in the activation of abstract VAC schemata. Can the model identify these in the absence of specific word information, in the same way that we know that “Tom wugged the paper napkin across the table” is a VOL? Figure 10 shows VL activation as the prototypical sentence “you go to the shop” is successively experienced, and as possible alternative permutations where one word is replaced by a wug pattern (i.e., an unknown word) and the same for VOL activation with “put it in there” and VOO activation with “you give me money.” In each case, it can be seen that activation is cumulative over words, that each construction is successfully activated even when individual component words are wugged, and that the greatest decrement is when the cues from the verb are absent; these are abstract schema, but they have been built from collaborative experience of individual verb islands.

These simulations show how simple general learning mechanisms, exposed to the coadapted language usage typical of NSs as they speak with NNSs, learn abstract verb argument constructions in the same order of emergence as NNSs and using the same cues. The factors that we measured in the first part of this article conspire in the emergence of these constructions from usage.

Architecture 2 (No Semantics)

One response to initial presentations of the results of these simulations is that we give too much to these models by including a semantic layer. Perhaps this information serves as TRICS (The Representations it Crucially Supposes) that cryptoembody rules within the connectionist network so that no real learning is necessary (Lachter & Bever, 1988). We find this a difficult criticism to credit, as any alternative would deny any processes of Semantic Bootstrapping. Language without meaningful reference is no language at all, and anything of any complexity would not be learnable. Nevertheless, it is an interesting exercise to see just what structure is learnable by processes of Syntactic Bootstrapping alone. For that reason we ran the same simulations with the same architecture except for the elimination of the semantic layer. The model was trained with the same input patterns but simply had to predict the next word.

In the absence of semantic output units, there is no explicit way of directly testing its accuracy of categorization of the patterns. However, we can investigate the patterning of its internal states on the hidden units as it experiences different words. We ran this model for 100 epochs, testing the HU activations on experiencing the 200 different test words out of context without feedback after every 10 epochs. Figure 11 shows the dendrograms from a cluster analysis of these patterns for the key verbs of interest. It can be seen, as in the

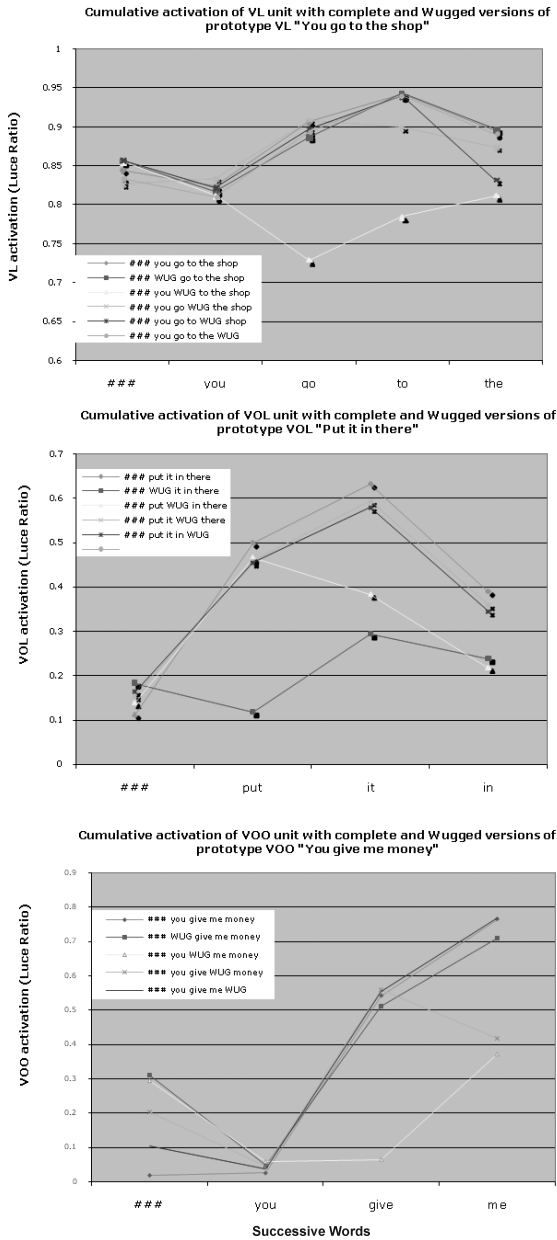


Figure 10 Dynamic activation of the VL, VOL, and VOO patterns by different prototypical sequences with each word wugged in turn after 10 epochs of training of Simulation 1.

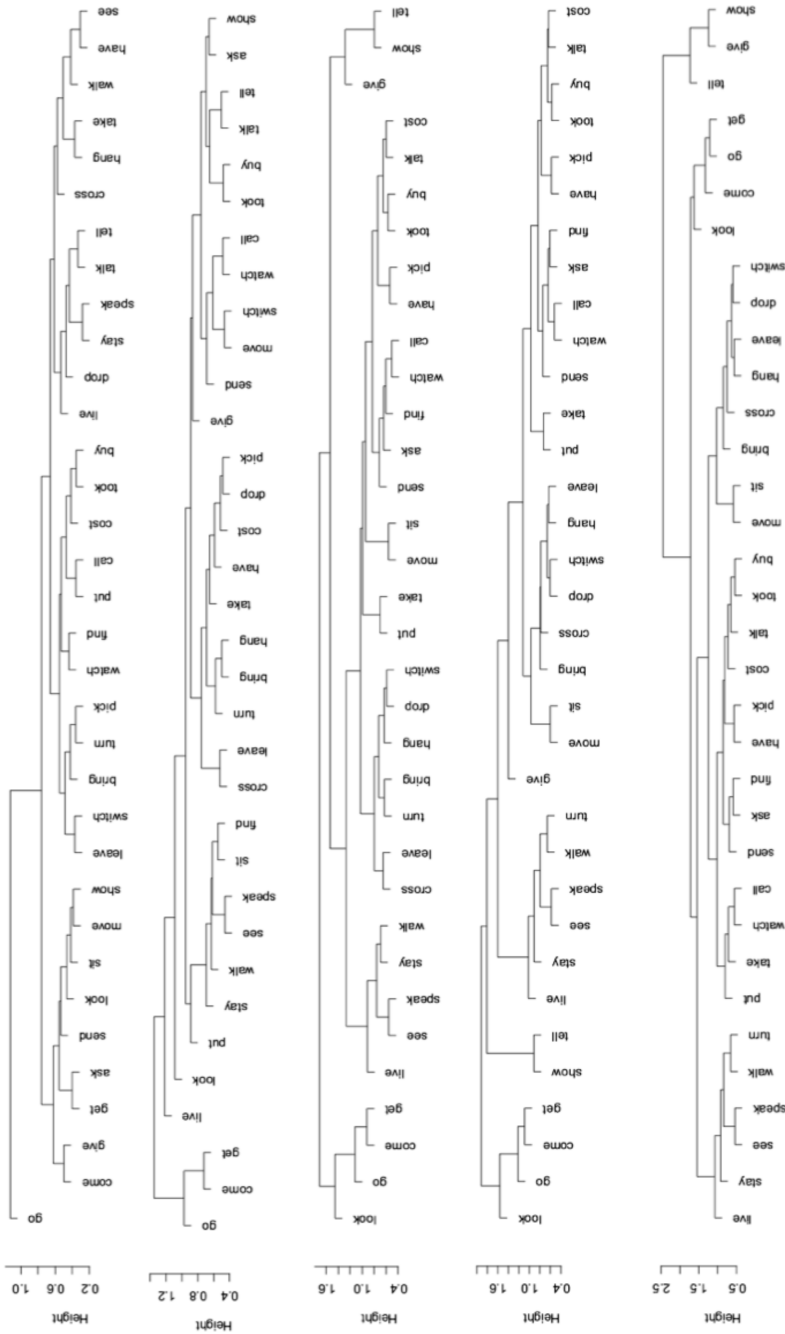


Figure 11 Dendrogram of hidden unit representations from Simulation 2 after 10, 30, 50, 70, and 90 epochs of training when tested.

demonstrations of Elman (1990) and Borovsky & Elman (2006), that structure is emergent. At Epoch 10, *go* is categorized separately from the other verbs; by epoch 30, it is joined by *come* and *get*, and by epoch 50, there is a clear cluster of verbs that we could label VL: *go, come, look, get*. At epoch 30, a nascent VOO category is forming over on the right, including *give, buy, tell, ask, and show*. After epoch 50, *give, show, and tell*, the prototypical VOO verbs, are in a clear category of their own. *Put* and *take* separate out by epoch 50 and move with other VOL patterns thereafter. Syntagmatic patterning alone is sufficient to allow the model to learn these different categories, albeit more slowly than when semantic information is also available.

Conclusions and Future

Our findings provide empirical support for the hypothesis that the emergence of linguistic constructions can be understood according to psychological principles of category learning and the social principle of coadaptation. Learning is sensitive to input frequency, reliabilities of form-meaning-use mapping, and prototypicality and generality of function. However, there is more to it than that. The structure of language reflects these principles too. It is doubtful that these parallels are accidental—more likely they emerge through usage. A consequence, we believe, is that in natural language, the type-token frequency distributions of construction islands, their prototypicality and generality of function in these roles, and their reliability of mappings between these, together, conspire to optimize learning.

We intend in future simulations to try to tease out the different roles of these factors. In the same way that Simulation 2 denies the role of semantics, so in future models we will investigate the emergence of these constructions in the absence of each factor in turn: What is the effect of providing the model with a flat type-token frequency profile rather than a Zipfian one? What is the effect of making all exemplars equal members of a semantic category rather than populating a radial structure with exemplars varying in prototypicality, etc.? These simulations parallel cognitive neuroscientific connectionist investigations, where the effects of lesioning different parts of the model architecture are investigated, but here we are lesioning simulated language itself rather than brain structure.

Like the other authors of articles in this issue, we believe that the functions of language in human communication have resulted in the evolution through usage of a system that optimally maps human sociocognition onto communicatively effective language form. The result is a system that is readily acquired. We are

only at the beginnings of understanding the dynamic emergence of this complex system, but we are sure, at least, that this is the appropriate approach.

Revised version accepted 1 June 2009

References

- Aisa, B., Mingus, B., & O'Reilly, R. C. (May 2008a). [Computer Software]. Emergent. Retrieved from http://grey.colorado.edu/emergent/index.php/Main_Page/
- Aisa, B., Mingus, B., & O'Reilly, R. C. (2008b). The emergent neural modeling system. *Neural Networks*, *21*, 1146–1152.
- Barlow, M., & Kemmer, S. (Eds.). (2000). *Usage based models of language*. Stanford, CA: CSLI Publications.
- Bates, E., & MacWhinney, B. (1987). Competition, variation, and language learning. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 157–193). Hillsdale, NJ: Lawrence Erlbaum.
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. New York: Cambridge University Press.
- Bod, R., Hay, J., & Jannedy, S. (Eds.). (2003). *Probabilistic linguistics*. Cambridge, MA: MIT Press.
- Borovsky, A., & Elman, J. L. (2006). Language input and semantic categories: A relation between cognition and early word learning. *Journal of Child Language*, *33*, 759–790.
- Bybee, J. (2005, January). *From usage to grammar: The mind's response to repetition*. Paper presented at the Linguistic Society of America, Oakland, CA.
- Bybee, J., & Hopper, P. (Eds.). (2001). *Frequency and the emergence of linguistic structure*. Amsterdam: Benjamins.
- Celce-Murcia, M., & Larsen-Freeman, D. (1999). *The grammar book: An ESL/EFL teacher's course* (2nd ed.). Boston: Heinle/Cengage.
- Childers, J. B., & Tomasello, M. (2001). The role of pronouns in young children's acquisition of the English transitive construction. *Developmental Psychology*, *37*, 739–748.
- Christiansen, M. H., & Chater, N. (Eds.). (2001). *Connectionist psycholinguistics*. Westport, CT: Ablex.
- Cohen, H., & Lefebvre, C. (Eds.). (2005). *Handbook of categorization in cognitive science*. Mahwah, NJ: Elsevier.
- Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*, 671–684.
- Croft, W., & Cruise, A. (2004). *Cognitive linguistics*. Cambridge: Cambridge University Press.
- Dale, R., & Spivey, M. (2006). Unraveling the dyad: Using recurrence analysis to explore patterns of syntactic coordination between children and caregivers in conversation. *Language Learning*, *56*, 391–430.

- de Bot, K., Lowie, W., & Verspoor, M. (2007). A dynamic systems theory to second language acquisition. *Bilingualism: Language and Cognition*, *10*, 7–21.
- Dietrich, R., Klein, W., & Noyau, C. (Eds.). (1995). *The acquisition of temporality in a second language*. Amsterdam: Benjamins.
- Elio, R., & Anderson, J. R. (1981). The effects of category generalizations and instance similarity on schema abstraction. *Journal of Experimental Psychology: Human Learning & Memory*, *7*(6), 397–417.
- Elio, R., & Anderson, J. R. (1984). The effects of information order and learning mode on schema abstraction. *Memory & Cognition*, *12*(1), 20–30.
- 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. (2003). Constructions, chunking, and connectionism: The emergence of second language structure. In C. Doughty & M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 33–68). Oxford: Blackwell.
- Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, *27*, 305–352.
- Ellis, N. C. (2006a). Cognitive perspectives on SLA: The Associative Cognitive CREED. *AILA Review*, *19*, 100–121.
- Ellis, N. C. (2006b). Language acquisition as rational contingency learning. *Applied Linguistics*, *27*(1), 1–24.
- Ellis, N. C. (2006c). 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. (2007). Dynamic Systems and SLA: The wood and the trees. *Bilingualism: Language & Cognition*, *10*, 23–25.
- Ellis, N. C. (2008a). The dynamics of language use, language change, and first and second language acquisition. *Modern Language Journal*, *41*(3), 232–249.
- 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., & Ferreira-Junior, F. (2009a). Constructions and their acquisition: Islands and the distinctiveness of their occupancy. *Annual Review of Cognitive Linguistics*, *7*, 188–221.
- Ellis, N. C., & Ferreira-Junior, F. (2009b). Construction Learning as a function of Frequency, Frequency Distribution, and Function *Modern Language Journal*, *93*, 370–386.
- Ellis, N. C., & Larsen Freeman, D. (Eds.). (2006a). Language emergence: Implications for Applied Linguistics [Special issue]. *Applied Linguistics*, *27*(4).

- Ellis, N. C., & Larsen Freeman, D. (2006b). Introduction. *Applied Linguistics*, 27(4), 558–589.
- Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14, 179–211.
- Elman, J. L. (1998). Generalization, simple recurrent networks, and the emergence of structure. In M. A. Gernsbacher & S. J. Derry (Eds.), *Proceedings of the twentieth annual conference of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum.
- Elman, J. L. (2004). An alternative view of the mental lexicon. *Trends in Cognitive Science*, 8, 301–306.
- Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.
- Feldweg, H. (1991). *The European Science Foundation Second Language Database*. Nijmegen: Max-Planck-Institute for Psycholinguistics.
- Gleitman, L. (1990). The structural source of verb meaning. *Language Acquisition*, 1, 3–55.
- Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.
- Goldberg, A. E. (2003). Constructions: A new theoretical approach to language. *Trends in Cognitive Science*, 7, 219–224.
- Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.
- Goldberg, A. E., Casenhiser, D. M., & Sethuraman, N. (2004). Learning argument structure generalizations. *Cognitive Linguistics*, 15, 289–316.
- Gries, S. T., & Stefanowitsch, A. (2004). Extending collocation analysis: A corpus-based perspective on “alternations.” *International Journal of Corpus Linguistics*, 9, 97–129.
- 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.
- Hebb, D. O. (1949). *The organization of behaviour*. New York: Wiley.
- Holland, J. H. (1992). *Hidden order: How adaptation builds complexity*. Reading, MA: Addison-Wesley.
- Holland, J. H. (1998). *Emergence: From chaos to order*. Oxford: Oxford University Press.
- Lachter, J., & Bever, T. (1988). The relation between linguistic structure and associative theories of language learning: A constructive critique of some connectionist learning models. *Cognition*, 28, 195–247.
- 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. (1976). An explanation for the morpheme acquisition order of second language learners. *Language Learning*, 26, 125–134.
- Larsen-Freeman, D. (1997). Chaos/complexity science and second language acquisition. *Applied Linguistics*, 18, 141–165.
- Larsen-Freeman, D. (2002a). The grammar of choice. In E. Hinkel & S. Fotos (Eds.), *New perspectives on grammar teaching* (pp. 105–120). Mahwah, NJ: Lawrence Erlbaum.
- Larsen-Freeman, D. (2002b). Language acquisition and language use from a chaos/complexity theory perspective. In C. Kramsch (Ed.), *Language acquisition and language socialization* (pp. 33–46). London: Continuum.
- Larsen-Freeman, D. (2003). *Teaching language: From grammar to grammaring*. Boston: Heinle & Heinle.
- Larsen-Freeman, D., & Cameron, L. (2008). *Complex systems and applied linguistics*. Oxford: Oxford University Press.
- Larsen-Freeman, D., & Long, M. (1991). *An introduction to second language acquisition research*. New York: Longman.
- MacWhinney, B. (1987). The Competition model. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 249–308). Hillsdale, NJ: Lawrence Erlbaum.
- MacWhinney, B. (Ed.). (1999). *The emergence of language*. Hillsdale, NJ: Lawrence Erlbaum.
- Murphy, G. L. (2003). *The big book of concepts*. Boston: MIT Press.
- 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: Oxford University Press.
- Perdue, C. (Ed.). (1993). *Adult language acquisition: Crosslinguistic perspectives*. Cambridge: Cambridge University Press.
- Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: Bradford Books.
- Port, R. F., & Van Gelder, T. (1995). *Mind as motion: Explorations in the dynamics of cognition*. Boston: MIT Press.
- Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, 77, 353–363.
- Posner, M. I., & Keele, S. W. (1970). Retention of abstract ideas. *Journal of Experimental Psychology*, 83, 304–308.
- Redington, M., & Chater, N. (1998). Connectionist and statistical approaches to language acquisition: A distributional perspective. *Language and Cognitive Processes*, 13, 129–192.
- 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.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Sinclair, J. (2004). *Trust the text: Language, corpus and discourse*. London: Routledge.
- Spivey, M. (2006). *The continuity of mind*. Oxford: Oxford University Press.
- Stefanowitsch, A., & Gries, S. T. (2003). Collocations: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics*, 8, 209–243.
- Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge MA: MIT Press.
- Tomasello, M. (1992). *First verbs: A case study of early grammatical development*. New York: Cambridge University Press.
- Tomasello, M. (2003). *Constructing a language*. Boston: Harvard University Press.
- Tomasello, M. (Ed.). (1998). *The new psychology of language: Cognitive and functional approaches to language structure*. Mahwah, NJ: Lawrence Erlbaum.
- van Geert, P. (1991). A dynamic systems model of cognitive and language growth. *Psychological Review*, 98, 3–53.
- Wilson, S. (2003). Lexically specific constructions in the acquisition of inflection in English. *Journal of Child Language*, 30, 75–115.
- Wulff, S., Ellis, N. C., Römer, U., Bardovi-Harlig, K., & LeBlanc, C. (2009). The acquisition of tense-aspect: Converging evidence from corpora, cognition, and learner constructions. *Modern Language Journal*, 93, 354–370.
- Zipf, G. K. (1935). *The psycho-biology of language: An introduction to dynamic philology*. Cambridge, MA: MIT Press.