Nick C. Ellis

Cognitive and Social Aspects of Learning from Usage

Language is essentially human. It is the crowning accomplishment of our social and cognitive competences. Language bridges society and cognition. It is a distributed emergent phenomenon. People and language create each other, grow from each other, and act and change under the influence of the other. Language and cognition are mutually inextricable; they determine each other. 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. 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. Language is used to focus the listener's attention to the world; it can foreground different elements in the theatre of consciousness to potentially relate many different stories and perspectives about the same scene. What is attended is the focus of learning, and so attention controls the acquisition of language itself. The functions of language in discourse determine its usage and learning. Language structure, language acquisition, language processing and usage, and language change are similarly inseparable: they are facets of the same complex adaptive system (Beckner et al., 2009).

These are some of the multiple perspectives importantly represented within usage-based approaches to language (Behrens, 2009; Bybee, 2010; Ellis, O'Donnell, & Römer, 2012; Robinson & Ellis, 2008b; Tomasello, 2003; Trousdale & Hoffmann, 2013) which hold that we learn language while engaging in the "interpersonal communicative and cognitive processes that everywhere and always shape language" (Slobin, 1997). Coming as I do from a background in cognitive processes, I will start from Cognition.

1 Cognition

Some of the basic Cognitive Linguistic tenets of Usage-based approaches and Construction Grammar, many of them explicitly addressed in the beginnings by de Saussure (1916), are:

- Language is intrinsically symbolic, constituted by a structured inventory of *constructions* as conventionalized form-meaning pairings used for communicative purposes.
- Language is intrinsically linked to human cognition and processes of perception, attention, learning, categorization, schematization, and memory.
- Adult language knowledge consists of a continuum of linguistic constructions of different levels of complexity and abstraction. Constructions can comprise concrete and particular items (as in words and idioms), more abstract classes of items (as in word classes and abstract constructions), or complex combinations of concrete and abstract pieces of language (as mixed constructions). No rigid separation exists between lexis and grammar.
- Constructions may be simultaneously represented and stored in multiple forms, at various levels of abstraction (e.g., concrete item: *table*+*s* = *tables* and [Noun] + (morpheme +*s*) = plural things).
- Constructions can thus be meaningful linguistic symbols in their own right, existing independently of particular lexical items. Nevertheless, constructions and the particular lexical tokens that occupy them attract each other, and grammar and lexis are inseparable.
- Language structure emerges ontogenetically from usage in particular contexts. Development is slow and gradual, moving from an initial reliance on concrete items to more abstract linguistic schemata. This process is dependent on the type and token frequencies with which particular constructions appear in the input. Storage of wholes depends on token frequency, the development of abstract linguistic schema depends on type frequency.

These issues are all current in cognitive scientific investigations of language, learning and categorization. Since language learning is about learning the categories of language, the categories of the world, and the mappings between these, constructionist accounts of language acquisition thus involve the distributional analysis of the language stream and the parallel analysis of contingent perceptual activity, with abstract constructions being learned from the conspiracy of concrete exemplars of usage following statistical learning mechanisms (Rebuschat & Williams, 2012) relating input and learner cognition. Psychological analyses of this learning of constructions as form-meaning pairs is informed by the literature on the associative learning of cue-outcome contingencies where the usual deter-

minants include: factors relating to the form such as frequency and salience; factors relating to the interpretation such as significance in the comprehension of the overall utterance, prototypicality, generality, redundancy, and surprise value; factors relating to the contingency of form and function; and factors relating to learner attention, such as automaticity, transfer, overshadowing, and blocking (Ellis, 2002, 2003, 2006c, 2008c). These various psycholinguistic factors conspire in the acquisition and use of any linguistic construction:

1.1 Determinants of Construction Learning

1.1.1 Input Frequency

1.1.1.1 Construction Frequency

Frequency of exposure promotes learning and entrenchment (e.g., Anderson, 2000; Bartlett, [1932] 1967; 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 recently we have experienced something, the stronger our memory for it, and the more fluently it is accessed [hence your reading this sentence more fluently than the preceding one]. 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; so a stimulus becomes associated to a context and we become more likely to perceive it in that context.

Frequency of exposure also underpins statistical learning of categories (Harnad, 1987; Hunt & Aslin, 2010; Lakoff, 1987; Mintz, 2002; Taylor, 1998). Human categorization ability provides the most persuasive testament to our incessant unconscious figuring or 'tallying' (Ellis, 2002). We know that natural categories are fuzzy rather than monothetic. Wittgenstein's (1953) consideration of the concept *game* showed that no set of features that we can list covers all the things that we call games, ranging as the exemplars variously do from soccer, through chess, bridge, and poker, to solitaire. Instead, what organizes these exemplars into the *game* category is a set of family resemblances among these members – son may be like mother, and mother like sister, but in a very different way. And we learn about these families, like our own, from experience. Exemplars are similar if they have many features in common and few distinctive attributes (features belonging to one but not the other); the more similar are two objects on these quantitative grounds, the faster are people at judging them to be similar (Tversky, 1977). 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. The operationalization of this criterion predicts the speed of human categorization performance – people more quickly classify as *dogs* Labradors (or other typically sized, typically colored, typically tailed, typically featured specimens) than they do dogs with less common features or feature combinations like Shar Peis or Neapolitan Mastiffs. Prototypes are judged faster and more accurately, even if they themselves have never been seen before – someone who has never seen a Labrador, yet who has experienced the rest of the run of the canine mill, will still be fast and accurate in judging it to be a dog (Posner & Keele, 1970). Such effects make it very clear that although people don't go around consciously counting features, they nevertheless have very accurate knowledge of the underlying frequency distributions and their central tendencies.

1.1.1.2 Type and Token Frequency

Token frequency counts how often a particular form appears in the input. 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 patterns is a function of type rather than token frequency (Bybee & Hopper, 2001). This is because: (a) the more lexical items that are heard in a certain position in a construction, the less likely it is that the construction is associated with a particular lexical item and the more likely it is that a general category is formed over the items that occur in that position; (b) the more items the category must cover, the more general are its criterial features and the more likely it is to extend to new items; and (c) high type frequency ensures that a construction is used frequently, thus strengthening its representational schema and making it more accessible for further use with new items (Bybee, 2010). In contrast, high token frequency promotes the entrenchment or conservation of irregular forms and idioms; the irregular forms only survive because they are high frequency. There is related evidence for type-token matters in statistical learning research (Aslin & Newport, 2012; Gómez, 2002; Onnis, Monaghan, Christiansen, & Chater, 2004).

1.1.1.3 Zipfian Distribution

In natural language, Zipf's law (Zipf, 1935) describes how the highest frequency words account for the most linguistic tokens. Zipf's law states that the frequency

of words decreases as a power function of their rank in the frequency table. 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: phoneme and letter strings (Kello & Beltz, 2009), words (Evert, 2005), grammatical constructs (Ninio, 2006; O'Donnell & Ellis, 2010), formulaic phrases (O'Donnell & Ellis, 2009) etc. Scale-free laws also pervade language structures, such as scale-free networks in collocation (Bannard & Lieven, 2009; Solé, Murtra, Valverde, & Steels, 2005), in morphosyntactic productivity (Baayen, 2008), in grammatical dependencies (Ferrer i Cancho & Solé, 2001, 2003; Ferrer i Cancho, Solé, & Köhler, 2004), and in networks of speakers, and language dynamics such as in speech perception and production, in language processing, in language acquisition, and in language change (Ellis, 2008a; Ninio, 2006). 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 both language structure and usage. More broadly still, power law behavior like this has 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). It has become a hallmark of Complex Systems theory. Zipfian scale-free laws are universal. 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, 2009a; Ferrer i Cancho & Solé, 2001, 2003; Ferrer i Cancho et al., 2004; Solé et al., 2005). Various usage-based / functionalist / cognitive linguists (Bybee, 2008; Ellis, 2008b; Ellis et al., 2012; Goldberg, 2006; Lieven & Tomasello, 2008) argue that it is the coming together of these distributions across linguistic form and linguistic function that makes language robustly learnable despite learners' idiosyncratic experience.

In first language acquisition, Goldberg, Casenhiser & Sethuraman (2004) demonstrated that there is a strong tendency for verb-argument constructions (VACs) (e.g. VL verb locative, VOL verb object locative, and VOO ditransitive) to be occupied by one single verb (e.g. *go* in VL, *put* in VOL, *give* in VOO, *etc*.) with very high frequency in comparison to other verbs used, a profile which closely mirrors that of the mothers' speech to these children. They argue that this pro-

motes language acquisition: In the early stages of learning categories from exemplars, acquisition is optimized by the introduction of an initial, low-variance sample centered upon prototypical exemplars. This 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. Ogden and Ellis (2014) confirm this in an analysis of the English language data in the CHILDES database (MacWhinney, 2000). In naturalistic second language (L2) acquisition, Ellis and Ferreira-Junior (2009b) investigated type/token distributions in the items comprising the linguistic form of English VACs and likewise showed that VAC verb type/token distribution in the input is Zipfian and that learners first acquire the most frequent, prototypical and generic exemplar.

1.1.2 Linguistic Form (salience and perception)

The general perceived strength of stimuli is commonly referred to as their salience. Low salience cues tend to be less readily learned. Ellis (2006c, 2006d) summarized the associative learning research demonstrating that selective attention, salience, expectation, and surprise are key elements in the analysis of all learning, animal and human alike. As the Rescorla-Wagner (1972) model encapsulates, the amount of learning induced from an experience of a cue-outcome association depends crucially upon the salience of the cue and the importance of the outcome.

Many grammatical meaning-form relationships, particularly those that are notoriously difficult for second language learners like grammatical particles and inflections such as the third person singular *-s* of English, are of low salience in the language stream. For example, some forms are more salient: '*today*' is a stronger psychophysical form in the input than is the morpheme '*-s*' marking 3rd person singular present tense, thus while both provide cues to present time, *today* is much more likely to be perceived, and *-s* can thus become overshadowed and blocked, making it difficult for second language learners of English to acquire (Ellis, 2006d, 2008a; Ellis & Sagarra, 2011; Goldschneider & DeKeyser, 2001)

1.1.3 Function (Embodiment and Perceptual Memories)

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 categories are also those which are the most codable (naming is faster), most coded, and are most frequently utilized. Children acquire basic-category terms like *dog, hammer, apple* earlier than they do their superordinates *animal, tool, fruit*, or subordinates *collie, ball-peen hammer, Granny Smith*. Arguably, 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.

Cognitive linguistics, particularly construction grammar, has since extended these ideas to language as a whole. Meanings are perceptually grounded in our sensory and motor imagery systems (Barsalou, 1999, 2008; Bergen & Chang, 2012). 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 the referents 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 something, someone possessing something, something causing a change of state or location, someone causing a change of possession, something undergoing a change of state or location, something having an effect on someone, etc. (Croft, 2001, 2012; Goldberg, 1995; Levin, 1993).

1.1.4 Function (Attention)

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. How exactly a given meaning is construed depends in large parts on where the language user's attention is being directed. Talmy (2000) describes the building blocks of the attentional system of language; each of around 50 building blocks, or factors, involves a particular linguistic mechanism that increases or decreases attention of a certain type of linguistic entity. Learning a language, then, means learning these various attention-directing mechanisms, which requires L1 learners to develop an attentional system in the first place, and L2 learners to reconfigure the attentional biases of having acquired their first language. 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; MacWhinney, 1998, 1999a; Talmy, 2000; Taylor, 2002).

1.1.5 Function (Prototypicality of Meaning)

Categories have graded structure, with some members being better exemplars than others. As we described in relation to frequency effects, 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, it serves as the benchmark against which surrounding, less representative instances are classified.

Ellis & Ferreira-Junior (2009a) show that the verbs that L2 learners first used in particular VACs are prototypical and generic in function (*go* for VL, *put* for VOL, and *give* for VOO). The same has been shown for child language acquisition, where a small group of semantically general verbs, often referred to as *light verbs* (e.g., *go, do, make, come*) are learned early (E. V. Clark, 1978; Ninio, 1999; Pinker, 1989). Ninio (1999) argues that, because most of their semantics consist of some schematic notion of transitivity with the addition of a minimum specific element, they are semantically suitable, salient, and frequent; hence, learners start transitive word combinations with these generic verbs. Thereafter, as Clark describes, "many uses of these verbs are replaced, as children get older, by more specific terms.... General purpose verbs, of course, continue to be used but become proportionately less frequent as children acquire more words for specific categories of actions" (p. 53).

1.1.6 Interactions between these (Contingency of Form-Function Mapping)

Psychological research into associative learning has long recognized that while frequency of form is important, so too 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. Distinctiveness or reliability of form-function mapping is a driving force of all associative learning, to the degree that the field of its study has been known as 'contingency learning'

since Rescorla (1968) showed that for classical conditioning, if one removed the contingency between the conditioned stimulus (CS) and the unconditioned (US), preserving the temporal pairing between CS and US but adding additional trials where the US appeared on its own, then animals did not develop a conditioned response to the CS. This result was a milestone in the development of learning theory because it implied that it was contingency, not temporal pairing, that generated conditioned responding. Contingency, and its associated aspects of predictive value, information gain, and statistical association, have been at the core of learning theory ever since. It is central in psycholinguistic theories of language acquisition too (Ellis, 2006c, 2006d; Gries, 2012; Gries & Stefanowitsch, 2004; MacWhinney, 1987b), with the most developed account for L2 acquisition being that of the Competition model (MacWhinney, 1987a, 1997, 2001a).

Ellis and Ferreira-Junior (2009b) use a variety of metrics to show that VAC acquisition is determined by their contingency of form-function mapping. They show that the one-way dependency statistic ΔP (Allan, 1980) that is commonly used in the associative learning literature (Shanks, 1995), as well as collostructional analysis measures current in corpus linguistics (Gries & Stefanowitsch, 2004; Stefanowitsch & Gries, 2003) predict effects of form-function contingency upon L2 VAC acquisition. Other researchers use conditional probabilities to investigate contingency effects in VAC acquisition. This is still an active area of inquiry, and more research is required before we know which statistical measures of form-function contingency are more predictive of acquisition and processing (Wiechmann, 2008).

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, and (3) coherent in their semantics.

1.2 Evidence of these factors in L1 Construction Knowledge

Ellis, O'Donnell, and Römer (2014) used free association and verbal fluency tasks to investigate people's knowledge of VACs and the ways in which their processing is sensitive to statistical patterns of usage (verb type-token frequency distribution, VAC-verb contingency, VAC-verb semantic prototypicality). In Experiment 1, 285 native speakers of English (mostly students enrolled at a large mid-western

research university) generated the first word that came to mind to fill the V slot in 40 sparse VAC frames such as '*he __ across the....*', '*it __ of the....*', etc. In Experiment 2, 40 English speakers generated as many verbs that fit each frame as they could think of in a minute. For each VAC, they compared the results from the experiments with the corpus analyses of verb selection preferences in 100 million words of usage in the British National Corpus (BNC) and with the semantic network structure of the verbs in these VACs.

For both experiments, the frequencies of verb types generated for each VAC were affected by three factors:

1. Frequency – verb token frequencies in those VACs in usage experience;
2. Contingency – how faithful verbs are to particular VACs in usage experience;
3. Semantic prototypicality – the centrality of the verb meaning in the semantic network of the VAC in usage experience.

Multiple regression analyses showed that these factors make significant independent contributions. They argue that these factors affect processing in the generation fluency task as follows:

1. Effects of frequency of usage upon language learning and subsequent fluency of linguistic processing reflect entrenchment and the power law of practice (Bybee, 2010; Bybee & Hopper, 2001; Ellis, 2002; MacWhinney, 2001b; Ninio, 2006).
2. Effects of contingency reflect associative learning (Ellis, 2006c, 2006d; MacWhinney, 1987b; MacWhinney, Bates, & Kliegl, 1984; Rescorla & Wagner, 1972; Shanks, 1995)
3. Effects of semantic prototypicality reflect spreading activation (Anderson, 1983). The prototype has two advantages: The first is a frequency factor: the greater the token frequency of an exemplar, the more it contributes to defining the category, and the greater the likelihood it will be considered the prototype (Rosch & Mervis, 1975; Rosch et al., 1976). Thus it is the response that is most associated with the VAC in its own right. But beyond that, it gets the network centrality advantage. When any response is made, it spreads activation and reminds other members in the set. The prototype is most connected at the center of the network and, like Rome, all roads lead to it. Thus it receives the most spreading activation. Likewise in social networks, individuals with high betweenness centrality are key agents in navigating the network – they mediate communication between most other individuals.

These findings promote a usage-based view of L1 VAC processing involving rich associations, tuned by verb type and token frequencies and their contingencies of usage, which interface syntax, lexis, and semantics.

1.3 Second language learning

1.3.1 Reconstructing a language, Transfer, Blocking, and Learned Attention

Usage-based second language acquisition (L2A) is typically less successful than first language acquisition. There is a lot of relearning to do, and a lot of interference from learned attention to both L1 function and L1 form: As Slobin notes, "[f]or the child, the construction of the grammar and the construction of semantic/pragmatic concepts go hand-in-hand. For the adult, construction of the grammar often requires a revision of semantic/pragmatic concepts, along with what may well be a more difficult task of perceptual identification of the relevant morphological elements" (1993, p. 242).

1.3.2 Rethinking L2 Semantic/Pragmatic Concepts

Languages lead their speakers to experience different 'thinking for speaking' and thus to construe experience in different ways (Slobin, 1996). Cross-linguistic research shows how different languages lead speakers to prioritize different aspects of events in narrative discourse (Berman & Slobin, 1994). Because languages achieve these attention-directing outcomes in different ways, learning another language involves learning how to construe the world like natives of the L2, i.e., learning alternative ways of thinking for speaking (Brown & Gullberg, 2008; Brown & Gullberg, 2010; Cadierno, 2008) or learning to 'rethink for speaking' (Robinson & Ellis, 2008a). Transfer theories such as the Contrastive Analysis Hypothesis (Gass & Selinker, 1983; James, 1980; Lado, 1957, 1964) hold that L2 learning can be easier where languages use these attention-directing devices in the same way, and more difficult when they use them differently. To the extent that the constructions in L2 are similar to those of L1, L1 constructions can serve as the basis for the L2 constructions, but, because even similar constructions across languages differ in detail, the acquisition of the L2 pattern in all its detail is hindered by the L1 pattern (Cadierno, 2008; Odlin, 1989, 2008; Robinson & Ellis, 2008b).

1.3.3 Rethinking L2 Form

Naturalistic foreign language acquisition (FLA) tends to stabilize at levels short of nativelike ability. At its most extreme this can present itself as a 'Basic Variety' of interlanguage (Klein, 1998) which, although sufficient for everyday communi-

cative purposes, predominantly comprises just nouns, verbs and adverbs, with closed-class items, in particular grammatical morphemes and prepositions, failing to be put to full nativelike use (Bardovi-Harlig, 1992; Schmidt, 1984; Van Patten, 1996, 2006). FL learners initially make temporal references mostly by use of temporal adverbials, prepositional phrases, serialization, and calendric reference, with the grammatical expression of tense and aspect emerging only slowly thereafter, if at all (Bardovi-Harlig, 1992, 2000).

As discussed above in relation to Linguistic Form, one factor determining cue selection is salience: Prepositional phrases, temporal adverbs, and other lexical cues to time are quite pronounced in the speech stream. Verbal inflections are not (consider "*yesterday* I walk*ed*"). The low salience and low reliability of grammatical cues tends to make them less learnable (Ellis, 2006d; Goldschneider & DeKeyser, 2001), and could underlie late learners' difficulty in processing and producing FL verbal morphology (Jiang, 2004). But salience and reliability affect L1A and FLA alike. There has to be something else which accounts for the limitations in FLA.

Associative learning theory documents a range of effects of transfer and inhibition that shift learners' attention to input as a result of prior experience. Kamin (1969) and Kruschke (2006) describe the phenomenon of blocking. Learning that a particular stimulus is associated with a particular outcome makes it harder to learn that another cue, subsequently paired with that same outcome, is also a good predictor of it. For example, if an animal learns that a conditioned stimulus (CS) is a reliable predictor of an unconditioned stimulus (UCS) (e.g., that a light reliably predicts the onset of some painful stimulus such as a shock), then it will not become conditioned to another CS or learn that any other CS predicts that UCS (e.g., that a bell predicts the onset of the shock the same way the light did). The prior association essentially "blocks" further associations. Blocking is an effect of learned attention (Kamin, 1969; Kruschke & Blair, 2000; Mackintosh, 1975). It is a highly robust and widespread phenomenon, occurring across animal and human learning (Rescorla & Wagner, 1972; Shanks, 1995; Wills, 2005).

Ellis (2006d) reviews the phenomenon as it might apply in second language acquisition. There are many situations in natural language where cues are redundant (Schmidt, 2001; Terrell, 1991; Van Patten, 1996) and thus, as a consequence of blocking, might be less readily learned. Where a learners' L1 experience has led them to look elsewhere for cues to interpretation, they might use these cues where available in FL, and if they do, the principles of associative learning predict that this will be to the detriment of learning other cues that might also be relevant. For example, L1-derived knowledge that there are reliable lexical cues to temporal reference (words like *gestern, hier, ayer, yesterday*) might block the acquisition of verb tense morphology from analysis of utterances such as *Yesterday I walked*.

Ellis and Sagarra (2010, 2011) describe a series of experimental investigations of this effect.

Various theories of SLA incorporate related notions of transfer and learned attention. The Competition Model (MacWhinney, 2001a; MacWhinney & Bates, 1989; MacWhinney et al., 1984) was explicitly formulated to deal with competition between multiple linguistic cues to interpretation. Input Processing theory (Van Patten, 1996) includes the Lexical Preference Principle: "Learners will process lexical items for meaning before grammatical forms when both encode the same semantic information" (Van Patten, 2006, p. 118), which encapsulates the mounting evidence that FL learners prefer lexical to grammatical cues as well as the Preference for Nonredundancy Principle: "Learners are more likely to process nonredundant meaningful grammatical markers before they process redundant meaningful markers" (Van Patten, 2006, p. 119). The Associative-Cognitive CREED (Ellis, 2006a, 2006b, 2006d, 2008c) describes the limited endstate typical of FLA directly in terms of learned attention, salience, overshadowing and blocking.

The human mind is built to integrate new information in a way that is maximally compatible with established knowledge – consequently, L1-attuned expectations and selective attention bias L2 acquisition.

2 Social

The nature of language follows from its role in social interaction. Social interactions are typically characterized by what philosophers of action call shared cooperative activity (Bratman, 1992) or joint actions (H. H. Clark, 1996). Joint actions are dependent on shared cognition, a human being's recognition that she can share beliefs and intentions with other humans. Thus usage-based approaches emphasize how language is learned from participatory experience of processing language during embodied interaction in social and cultural contexts where individually desired outcomes are goals to be achieved by communicating intentions, concepts and meaning with others. Conversation partners scaffold and co-construct meanings. Socially scaffolded 'noticing' (Schmidt, 1990) solves Quine's problem of 'referential indeterminacy' and builds so much more. The dynamics of language learning are inextricably linked to the dynamics of consciousness, in neural activity and in the social world as well (U. Frith & Frith, 2010). Consciousness is co-constructed in social interaction (Ellis, 2005; C. Frith, 2010). In these ways the input to associative learning is *socially gated* (Kuhl, 2007).

Quine (1960) argued that the robustness of language lies in the commonalities of language usage:

> "Each of us learns his [or her] language from other people, through the observable mouthings of words under conspicuously inter-subjective circumstances. Linguistically, and hence conceptually, the things in sharpest focus are the things that are public enough to be talked of publicly, common and conspicuous enough to be talked of often, and near enough to sense to be quickly identified and learned by name; it is to these that words apply first and foremost". (Quine, 1960, p. 1).

> "The uniformity that unites us in communication and belief is a uniformity of resultant patterns overlying a collective subjective diversity of connections between word and experience. Uniformity comes where it matters socially". (Quine, 1960, p. 8).

Thus shared attention, shared cooperative activity, and shared cognition (C. Frith & Frith, 2012) are key to meaningful language usage. In their first two years, infants develop their capabilities of attention detection (gaze following), attention manipulation (directive pointing), intention understanding (the realization that others are goal-directed), and social coordination with shared intentionality (engaging in joint activities with shared interest, negotiating meanings), and these processes are central in child language acquisition (Tomasello, 1999, 2008).

E. V. Clark (2014) analyzed the ways in which adults and children establish the 'Common Ground' that is necessary for the intersubjective circumstances of communication. She argues that children's social experience of their surroundings and highly repetitive routines underpins knowledge of perceptual and conceptual categories, and their participation in reciprocal games, object-exchanges, and proto-turn-taking, establish conceptual and social knowledge combined with a general attentiveness to the other in interaction. Establishing common ground requires joint attention, physical co-presence, and conversational co-presence. This, plus adult feedback when children express something in a non-conventional fashion, shapes the language:

> "Adult feedback consistently provides conventional forms, whether phonological or syntactic, morphological or lexical. These are the forms that children need in order to understand the intentions of others, and to convey their own intentions and be understood. Mastery of these conventions plays a central role for common ground: knowledge of a language and its use offers extensive communal common ground with other users of that language and so allows for more extensive and detailed communication of both needs and interests. Finally, adult reformulations of child errors also attest to the importance of interaction for the acquisition of language. It is in conversation that children master the conventions and so also learn how to use common ground." (E. V. Clark, 2014, p. 21)

Language and language learning is ever thus. Activity theory emphasizes how individual learning is an emergent, holistic property of a dynamic system comprising many influences, both social, individual, and contextual. Action provides a context within which the individual and society, mental functioning and socio-cultural context can be understood as interrelated moments (Wertsch, 1998). Cognitive mechanisms are *culturally gated* too. Tomasello's constructionist approach to language unites with his research in comparative primate cognition, the unique place of social cooperation in humans, and the Vygotskian intelligence hypothesis whereby regular participation in cooperative, cultural interactions during ontogeny leads children to construct uniquely powerful forms of perspectival cognitive representation including language itself (Moll & Tomasello, 2007).

The same holds for second language acquisition. Speech, speakers, identity, and social relationships are inseparable (Lee, Mikesell, Joaquin, Mates, & Schumann, 2009; Norton, 1997; Tarone, 2007). Socio-cultural approaches emphasize how learning takes place in social usage, involving action, reaction, collaborative interaction, intersubjectivity, and mutually assisted performance (Lantolf & Thorne, 2006). Social-interactional approaches analyze how interaction provides comprehensible, negotiated input and reactive feedback (Gass, 1997, 2002, 2003; Gass & Mackey, 2007; Long, 1982; Mackey, 2012). The provision of negative feedback such as a clarification request or a recast, promotes the development of language. Such focus-on-form (Long, 1991) presents learners with psycholinguistic data that is fertile and ready for acquisition because the contrast between the learners' own erroneous utterance and the recast highlights the relevant element of the form at the same time that the desired meaning-to-be-expressed is still active (Doughty, 2001; Doughty & Williams, 1998; Tomasello & Herron, 1988, 1989). Interaction in which participants' attention is focused on resolving a communication problem and the consequent negotiation of form and meaning "connects input, internal learner capacities, particularly selective attention, and output in productive ways" (Long, 1996, p. 452). "Notice this" say conversation partners in their words and their actions, and thus a new wave of explicit analysis is initiated, with consequent benefits for the learner from native speakers and non-native speakers interactions alike (Gass, 1997, p. 107; Gass & Varonis, 1994; Long, 1996).

In these ways, SLA can be freed from the bounds of L1-induced selective attention by some means of form-focus that is socially provided (Tarone, 1997) and that recruits the learner's explicit processing (Ellis, 2005). The learner's conscious tension between the conflicting forces of their current interlanguage productions and the evidence of feedback (either linguistic, pragmatic, or metalinguistic) provides the means of socially scaffolded development. Focus-on-form is socially given and socially determined.

The associative learning and cognitive processes which compute symbolic constructions are embodied, attentionally- and socially-gated, conscious, dialogic, interactive, situated, and encultured. All these factors conspire dynamically in the acquisition and use of linguistic constructions.

3 Future research directions

Usage-based theories hold that an individual's creative linguistic competence emerges from the collaboration of the memories of all the meaningful interactions in their entire history of language usage. What then are some research priorities?

From analyses of large usage corpora, we can analyze the latent structures of language and their roles in the associative and cognitive learning of language (Ellis et al., 2012). An important goal is an understanding of the interaction of memorized/formulaic language and productive schematic patterns (Ellis, 2012; Eskildsen, 2012). This is the stuff of cognitive psychology, associative learning theory, and corpus linguistics. But in addition to construction forms and their frequencies, there are their meanings, embodied, attended, consciously formed in dialogue and dialectic, situated and encultured in social and educational interaction. Here we have the embodied, dynamic mind of modern cognitive science (A. Clark, 1998; Rosch, Varela, & Thompson, 1991).

If language learning is in the social cognitive linguistic moment, we need to capture all these moments, so that we can objectively study them. We need large dense longitudinal corpora of language use, with audio, video, transcriptions and multiple layers of annotation, for data sharing in open archives. We need these in sufficient dense mass that we can chart learners' usage history and their development (Tomasello & Stahl, 2004). We need them in sufficient detail that we can get down to the fine detail of CA analyses of the moment (Kasper & Wagner, 2011; Markee, 2008; Markee & Kunitz, 2013). MacWhinney has long been working towards these ends, first with CHILDES (MacWhinney, 1991), then with Talkbank (MacWhinney, 2007). These projects have developed variously CLAN tools for computer analyses of large bodies of data, right down to, in collaborations with Wagner, tools for a fine grained CABank (MacWhinney & Wagner, 2010). With these types of data, we can study the cognitive alongside the social.

We need more collaboration and more methodological pluralism within SLA (Hulstijn et al., 2014). We need integrative theories too. Cognition, consciousness, experience, embodiment, brain, self, and human interaction, society, culture, and history are all inextricably intertwined in rich, complex, and dynamic ways in language. We require additional perspectives on dynamic interactions between

levels, perspectives provided by approaches such as Complex Adaptive Systems (Ellis & Larsen-Freeman, 2009b), Dynamic Systems Theory (de Bot, Lowie, & Verspoor, 2007; Ellis, 2008a; Spivey, 2006), and Emergentism (Ellis, 1998, 2014; Ellis & Larsen-Freeman, 2006; MacWhinney, 1999b; MacWhinney & O'Grady, in press).

References

Allan, L. G. 1980. A note on measurement of contingency between two binary variables in judgment tasks. *Bulletin of the Psychonomic Society, 15*, 147–149.

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

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

Aslin, R. N., & Newport, E. L. 2012. Statistical learning: From acquiring specific items to forming general rules. *Current Directions in Psychological Science,, 21*, 170–176.

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.

Bardovi-Harlig, K. 1992. The use of adverbials and natural order in the development of temporal expression. *IRAL, XXX*, 299–320.

Bardovi-Harlig, K. 2000. *Tense and aspect in second language acquisition: Form, meaning, and use*. Oxford: Blackwell.

Barsalou, L. W. 1999. Perceptual symbol systems. *Behavioral and Brain Sciences, 22*, 577–660.

Barsalou, L. W. 2008. Grounded cognition. *Annual Review of Psychology, 59*, 617–645.

Bartlett, F. C. ([1932] 1967). *Remembering: A Study in Experimental and Social Psychology*. Cambridge: Cambridge University Press

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.

Behrens, H. 2009. Usage-based and emergentist approaches to language acquisition. *Linguistics, 47*, 383–411.

Bergen, B., & Chang, N. 2012. Embodied Construction Grammar. In G. Trousdale & T. Hoffmann (Eds.), *Oxford Handbook of Construction Grammar* (pp. 168–190). Oxford: Oxford University Press.

Berman, R. A., & Slobin, D. I. (Eds.). 1994. *Relating Events in Narrative: A Crosslinguistic Developmental Study*. Hillsdale, N.J.: Lawrence Erlbaum.

Bratman, M. 1992. Shared cooperative activity. *The Philosophical Review, 101*, 327–341.

Brown, A., & Gullberg, M. 2008. Bidirectional crosslinguistic influence in L1-L2 encoding of manner in speech and gesture. *Studies in Second Language Acquisition, 30*, 225–251.

Brown, A., & Gullberg, M. 2010. Changes in encoding of path of motion after acquisition of a second language. *Cognitive Linguistics, 21*, 263–286.

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.

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

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

Cadierno, T. 2008. Learning to talk about motion in a foreign language. In P. Robinson & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition*. London: Routledge.

Clark, A. 1998. *Being there: Putting brain, body, and world together again*. Cambridge, MA: The MIT Press.

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.

Clark, E. V. 2014. Common Ground. In B. MacWhinney & W. O'Grady (Eds.), *The Handbook of Language Emergence* (Vol. this volume). Boston, MA: Wiley-Blackwell.

Clark, H. H. 1996. *Using language*. Cambridge: Cambridge University Press.

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

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

de Bot, K., Lowie, W., & Verspoor, M. 2007. A dynamic systems theory to second language acquisition. *Bilingualism: Language and Cognition, 10*, 7–21.

de Saussure, F. 1916. *Cours de linguistique générale* (Roy Harris, Trans.). London: Duckworth.

Doughty, C. 2001. Cognitive underpinnings of focus on form. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 206–257). Cambridge: Cambridge University Press.

Doughty, C., & Williams, J. (Eds.). 1998. *Focus on form in classroom second language acquisition*. New York: Cambridge University Press.

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

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. The Associative-Cognitive CREED. In B. VanPatten & J. Williams (Eds.), *Theories in Second Language Acquisition: An Introduction* (pp. 77–96). Mahwah, NJ: Lawrence Erlbaum Associates.

Ellis, N. C. 2006b. Cognitive perspectives on SLA: The Associative Cognitive CREED. *AILA Review, 19*, 100–121.

Ellis, N. C. 2006c. Language acquisition as rational contingency learning. *Applied Linguistics, 27*(1), 1–24.
Ellis, N. C. 2006d. 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. The dynamics of second language emergence: Cycles of language use, language change, and language acquisition. *Modern Language Journal, 41*(3), 232–249.
Ellis, N. C. 2008b. 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*. Oxford: Blackwell.
Ellis, N. C. 2008c. 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. 2012. Formulaic language and second language acquisition: Zipf and the phrasal teddy bear. *Annual Review of Applied Linguistics, 32*, 17–44.
Ellis, N. C. 2014. Cognitive AND social usage. *Studies in Second Language Acquisition, 36*(3), 397–402.
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, 7*, 111–139.
Ellis, N. C., & Larsen-Freeman, D. 2006. Language emergence: Implications for Applied Linguistics (Introduction to the Special Issue). *Applied Linguistics, 27*(4), 558–589.
Ellis, N. C., & Larsen-Freeman, D. (Eds.). 2009a. *Language as a Complex Adaptive System* (Vol. 59 Supplement 1).
Ellis, N. C., & Larsen-Freeman, D. (Eds.). 2009b. *Language as a Complex Adaptive System*. Boston, MA: Wiley-Blackwell.
Ellis, N. C., O'Donnell, M. B., & Römer, U. 2012. Usage-Based Language: Investigating the Latent Structures that Underpin Acquisition. *Currents in Language Learning, 1*, 25–51.
Ellis, N. C., O'Donnell, M. B., & Römer, U. 2014. The Processing of Verb-Argument Constructions is Sensitive to Form, Function, Frequency, Contingency, and Prototypicality. *Cognitive Linguistics, 25*, 55–98.
Ellis, N. C., & Sagarra, N. 2010. The bounds of adult language acquisition: Blocking and learned attention *Studies in Second Language Acquisition, 32*(4), 553–580.
Ellis, N. C., & Sagarra, N. 2011. Learned Attention in Adult Language Acquisition: A Replication and Generalization Study and Meta-Analysis. *Studies in Second Language Acquisition, 33*(4), 589–624.
Eskildsen, S. W. 2012. L2 negation constructions at work. *Language Learning, 62*(2), 335–372.
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. *PNAS, 100*, 788–791.
Ferrer i Cancho, R., Solé, R. V., & Köhler, R. 2004. Patterns in syntactic dependency networks. *Physical Review, E69*, 0519151–0519158.

Frith, C. 2010. What is consciousness for? *Pragmatics and cognition, 18*(3), 497–551.
Frith, C., & Frith, U. 2012. Mechanisms of social cognition. *Annual Review of Psychology, 63*, 287–313.
Frith, U., & Frith, C. 2010. The social brain: allowing humans to boldly go where no other species has been. *Phil. Trans. R. Soc., 365*, 165–175.
Gass, S. 1997. *Input, interaction, and the development of second languages*. Mahwah, NJ: Erlbaum.
Gass, S. 2002. Interactionist perspectives in SLA. In R. Kaplan (Ed.), *Handbook of Applied Linguistics* (pp. 170–181). Oxford: Oxford University Press.
Gass, S. 2003. Input and interaction. In C. Doughty & M. Long (Eds.), *Handbook of second language acquisition* (pp. 224–255). Oxford: Blackwell Publishers
Gass, S., & Mackey, A. 2007. Input, interaction and output: An overview. *AILA Review, 19*, 3–17.
Gass, S., & Selinker, L. (Eds.). 1983. *Language transfer in language learning*. Rowley, MA: Newbury House.
Gass, S., & Varonis, E. 1994. Input, interaction and second language production. *Studies in Second Language Acquisition, 16*, 283–302.
Goldberg, A. E. 1995. *Constructions: A construction grammar approach to argument structure.* Chicago: University of Chicago Press.
Goldberg, A. E. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.
Goldberg, A. E., Casenhiser, D. M., & Sethuraman, N. 2004. Learning argument structure generalizations. *Cognitive Linguistics, 15*, 289–316.
Goldschneider, J. M., & DeKeyser, R. 2001. Explaining the "natural order of L2 morpheme acquisition" in English: A meta-analysis of multiple determinants. *Language Learning, 51*, 1–50.
Gómez, R. 2002. Variability and detection of invariant structure. *Psychological Science, 13*, 431–436.
Gries, S. Th. 2012. Frequencies, probabilities, association measures in usage-/exemplar-based linguistics: some necessary clarifications. *Studies in Language, 11*, 477–510.
Gries, S. Th., & Stefanowitsch, A. 2004. Extending collostructional analysis: a corpus-based perspective on 'alternations'. *International Journal of Corpus Linguistics, 9*, 97–129.
Harnad, S. (Ed.). 1987. *Categorical perception: The groundwork of cognition*. New York: Cambridge University Press.
Hulstijn, J. H., Young, R. F., Ortega, L., Bigelow, M., DeKeyser, R., Ellis, N. C., … Talmy, S. 2014. Bridging the gap: Cognitive and Social Approaches to Research in Second Language Learning and Teaching. *Studies in Second Language Acquisition, 33*(3), 361–421.
Hunt, R. H., & Aslin, R. N. 2010. Category induction via distributional analysis: Evidence from a serial reaction time task. *Journal of Memory and Language, 62*, 98–112.
James, C. 1980. *Contrastive analysis*. London: Longman.
Jiang, N. 2004. Morphological insensitivity in second language processing. *Applied Psycholinguistics, 25*, 603–634.
Kamin, L. J. 1969. Predictability, surprise, attention, and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior* (pp. 276–296). New York: Appleton-Century-Crofts.
Kasper, G., & Wagner, J. 2011. A conversation-analytic approach to second language acquisition. Alternative approaches to second language acquisition. In D. Atkinson

(Ed.), *Alternative approaches to second language acquisition* (pp. 117–142.). Abingdon: Routledge.
Kello, C. T., & Beltz, B. C. 2009. Scale-free networks in phonological and orthographic wordform lexicons. In I. Chitoran, C. Coupé, E. Marsico & F. Pellegrino (Eds.), *Approaches to Phonological Complexity*. Berlin: Mouton de Gruyter.
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.
Klein, W. 1998. The contribution of second language acquisition research. *Language Learning, 48*, 527–550.
Kruschke, J. K. (2006, May 31-June 3, 2006). *Learned Attention.* Paper presented at the Fifth International Conference on Development and Learning, Indiana University.
Kruschke, J. K., & Blair, N. J. 2000. Blocking and backward blocking involve learned inattention. *Psychonomic Bulletin & Review, 7*, 636–645.
Kuhl, P. K. 2007. Is speech-learning gated by the 'social brain'? *Developmental Science, 10*, 110–120.
Lado, R. 1957. *Linguistics across cultures: Applied linguistics for language teachers*. Ann Arbor: University of Michigan Press.
Lado, R. 1964. *Language teaching: A scientific approach*. New York: McGraw-Hill.
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.
Lantolf, J., & Thorne, S. 2006. *Sociocultural theory and the genesis of second language development*. Oxford: Oxford University Press.
Lee, N., Mikesell, L., Joaquin, A. D. L., Mates, A. W., & Schumann, J. 2009. *The interactional instinct: The evolution and acquisition of language*. New York: Oxford University Press.
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*. New York and London: Routledge.
Long, M. H. 1982. Native speaker/non-native speaker conversation in the second language classroom. In M. Long & J. Richards (Eds.), *Methodology in TESOL: A book of readings* (pp. 339–354). New York: Newbury House.
Long, M. H. 1991. Focus on form: A design feature in language teaching methodology. In K. d. Bot, R. Ginsberg & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39–52). Amsterdam: Benjamins.
Long, M. H. 1996. The role of linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisitiom*. San Diego: Academic.
Mackey, A. 2012. *Input, Interaction & Corrective Feedback in L2 Learning* Oxford: Oxford University Press.
Mackintosh, N. J. 1975. A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review, 82*, 276–298.
MacWhinney, B. 1987a. Applying the Competition Model to bilingualism. *Applied Psycholinguistics, 8*(4), 315–327.

MacWhinney, B. 1987b. The Competition Model. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 249–308). Hillsdale, NJ: Erlbaum.
MacWhinney, B. 1991. *The CHILDES project: Tools for analyzing talk*:. 1991. xi, 360pp.
MacWhinney, B. 1997. Second language acquisition and the Competition Model. In A. M. B. De Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 113–142). Mahwah, NJ: Lawrence Erlbaum Associates.
MacWhinney, B. 1998. Models of the emergence of language. *Annual Review of Psychology, 49*, 199–227.
MacWhinney, B. 1999a. The emergence of language from embodiment MacWhinney, B. (Ed). *1999 The emergence of language* (pp 213–256) xvii, 500pp.
MacWhinney, B. 2000. *The CHILDES project: Tools for analyzing talk, Vol 1: Transcription format and programs (3rd ed.)*:. 2000. xi, 366pp.
MacWhinney, B. 2001a. 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. 2001b. Emergentist approaches to language. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 449–470). Amsterdam, Netherlands: Benjamins.
MacWhinney, B. 2007. *The TalkBank Project* Retrieved from http://repository.cmu.edu/cgi/viewcontent.cgi?article=1175&context=psychology
MacWhinney, B. (Ed.). 1999b. *The emergence of language*. Hillsdale, NJ: Erlbaum.
MacWhinney, B., & Bates, E. 1989. *The crosslinguistic study of sentence processing*. Cambridge: Cambridge University Press.
MacWhinney, B., Bates, E., & Kliegl, R. 1984. Cue validity and sentence interpretation in English, German, and Italian. *Journal of Verbal Learning & Verbal Behavior, 23*(2), 127–150.
MacWhinney, B., & O'Grady, W. (Eds.). (in press). *The Handbook of Language Emergence*. Oxford: Wiley-Blackwell.
MacWhinney, B., & Wagner, J. 2010. Transcribing, searching and data sharing: The CLAN software and the TalkBank data repository. *Gespraechsforschung, 11*, 154–173.
Manin, D. Y. 2008. Zipf's Law and Avoidance of Excessive Synonymy *Cognitive Science, 32*, 1075–1098.
Markee, N. 2008. Toward a learning behavior tracking methodology for CA-for-SLA. *Applied Linguistics, 29*, 404–427.
Markee, N., & Kunitz, S. 2013. Doing planning and task performance in Second Language Acquisition: An ethnomethodological respecification. *Language Learning, 63*, 629–664.
Mintz, T. 2002. Category induction from distributional cues in an artificial language. *Memory & Cognition, 30*, 678–686.
Moll, H., & Tomasello, M. 2007. Cooperation and human cognition: The Vygotskian intelligence hypothesis. *Philosophical Transactions of the Royal Society B, 362*, 639–648.
Newman, M. E. J. 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: Oxford University Press.
Norton, B. 1997. Language, identity, and the ownership of English. *Tesol Quarterly, 31*, 409–430.

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, Los Angeles*.

Odlin, T. 1989. *Language transfer*. New York: Cambridge University Press.

Odlin, T. 2008. Conceptual transfer and meaning extensions. In P. Robinson & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition*. London: Routledge.

Ogden, D., & Ellis, N. C. 2014. *Assessing Uptake of Verb-Argument Constructions in a Corpus of Child and Child-Directed Speech*. Paper presented at the Georgetown University Round Table on Languages and Linguistics on Usage-based Approaches to Language, Language Learning, and Multilingualism, Georgetown University, Washington DC.

Onnis, L., Monaghan, P., Christiansen, M. H., & Chater, N. 2004. Variability is the spice of learning, and a crucial ingredient for detecting and generalising nonadjacent dependencies. *Proceedings of the 26th Annual Conference of the Cognitive Science Society*.

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

Posner, M. I., & Keele, S. W. 1970. Retention of abstract ideas. *Journal of Experimental Psychology, 83*, 304–308.

Quine, W. V. O. 1960. *Word and object*. Cambridge, MA: MIT Press.

Rebuschat, P., & Williams, J. N. (Eds.). 2012. *Statistical learning and language acquisition*. Berlin: Mouton de Gruyter.

Rescorla, R. A. 1968. Probability of shock in the presence and absence of CS in fear conditioning. *Journal of Comparative and Physiological Psychology, 66*, 1–5.

Rescorla, R. A., & Wagner, A. R. 1972. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current theory and research* (pp. 64–99). New York: Appleton-Century-Crofts.

Robinson, P., & Ellis, N. C. 2008a. Conclusion: Cognitive Linguistics, Second Language Acquisition and L2 Instruction – Issues for Research. In P. Robinson & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition*. London: Routledge.

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

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.

Rosch, E., Varela, F., & Thompson, E. 1991. *The embodied mind*. Boston, MA: MIT Press.

Schmidt, R. 1984. The strengths and limitations of acquisition: A case study of an untutored language learner. *Language, Learning, and Communication, 3*, 1–16.

Schmidt, R. 1990. The role of consciousness in second language learning. *Applied Linguistics, 11*, 129–158.

Schmidt, R. 2001. Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). Cambridge: Cambridge University Press.

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

Slobin, D. I. 1993. Adult language acquisition: A view from child language study. In C. Perdue (Ed.), *Adult language acquisition: cross-linguistic perspectives* (pp. 239–252). Cambridge: Cambridge University Press.

Slobin, D. I. 1996. From "thought and language" to "thinking for speaking.". In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity.* (pp. 70–96). Cambridge: Cambridge University Press.

Slobin, D. I. 1997. The origins of grammaticizable notions: Beyond the individual mind. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition* (Vol. 5, pp. 265–323). Mahwah, NJ: Erlbaum.

Solé, R. V., Murtra, B., Valverde, S., & Steels, L. 2005. Language Networks: their structure, function and evolution. *Complexity, 15*(6), 20–26.

Spivey, M. 2006. *The continuity of mind*. Oxford: Oxford University Press.

Stefanowitsch, A., & Gries, S. Th. 2003. Collostructions: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics, 8*, 209–243.

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.

Tarone, E. 1997. Analyzing IL in natural settings: A sociolinguistic perspective of second-language acquisition. *Communication and Cognition, 30*, 137–150.

Tarone, E. 2007. Sociolinguistic approaches to second language acquisition research – 1997–2007. *The Modern Language Journal, 91(s1)*, 837–848.

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: Oxford University Press.

Terrell, T. 1991. The role of grammar instruction in a communicative approach. *The Modern Language Journal, 75*, 52–63.

Tomasello, M. 1999. *The cultural origins of human cognition*. Boston, MA: Harvard University Press.

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

Tomasello, M. 2008. *The origins of human communication*. Cambridge, MA: MIT Press.

Tomasello, M., & Herron, C. 1988. Down the garden path: Inducing and correcting overgeneralisation errors in the foreign language classroom. *Applied Psycholinguistics, 9*, 237–246.

Tomasello, M., & Herron, C. 1989. Feedback for language transfer errors: the garden path technique. *Studies in Second Language Acquisition, 11*, 385–395.

Tomasello, M., & Stahl, D. 2004. Sampling children's spontaneous speech: How much is enough? *Journal of Child Language, 31*, 101–121.

Trousdale, G., & Hoffmann, T. (Eds.). 2013. *Oxford Handbook of Construction Grammar*. Oxford: Oxford University Press.

Tversky, A. 1977. Features of similarity. *Psychological Review, 84*, 327–352.

Van Patten, B. 1996. *Input processing and grammar instruction in second language acquisition*. New York: Ablex.

Van Patten, B. 2006. Input processing. In B. Van Patten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 115–136). Mahwah, NJ: Lawrence Erlbaum.
Wertsch, J. V. 1998. *Mind as action*. Oxford: Oxford University Press.
Wiechmann, D. 2008. On the Computation of Collostruction Strength. *Corpus Linguistics and Linguistic Theory, 4*, 253–290.
Wills, A. J. 2005. *New directions in human associative learning*. Mahwah, NJ: Erlbaum.
Wittgenstein, L. 1953. *Philosophical investigations* (G. E. M. Anscombe, Trans.). Oxford: Blackwell.
Zipf, G. K. 1935. *The psycho-biology of language: An introduction to dynamic philology*. Cambridge, MA: The M.I.T. Press.