

# Cognitive perspectives on SLA

## The Associative-Cognitive CREED\*

Nick C. Ellis  
University of Michigan

This paper outlines current cognitive perspectives on second language acquisition (SLA). The Associative-Cognitive CREED holds that SLA is governed by the same principles of associative and cognitive learning that underpin the rest of human knowledge. The major principles of the framework are that SLA is Construction-based, Rational, Exemplar-driven, Emergent, and Dialectic. Language learning involves the acquisition of constructions that map linguistic form and function. Competence and performance both emerge from the dynamic system that is the frequency-tuned conspiracy of memorized exemplars of use of these constructions, with competence being the integrated sum of prior usage and performance being its dynamic contextualized activation. The system is rational in that it optimally reflects prior first language (L1) usage. The L1 tunes the ways in which learners attend to language. Learned-attention transfers to L2 and it is this L1 entrenchment that limits the endstate of usage-based SLA. But these limitations can be overcome by recruiting learner consciousness, putting them into a dialectic tension between the conflicting forces of their current stable states of interlanguage and the evidence of explicit form-focused feedback, either linguistic, pragmatic, or metalinguistic, that allows socially scaffolded development. The paper directs the reader to recent review articles in these key areas and weighs the implications of this framework.

SLA has been actively studied from a *Cognitive Psychological* perspective for the last two or three decades, and researchers within this tradition share basic goals, methods, and constructs. My aim in this article is to provide an overview of L2 acquisition in these terms. The position outlined here is fairly typical of the beliefs shared by psychologists: I have been influenced by so many in its development that it must reflect something close to the modal model. The Associative-Cognitive CREED holds that SLA is Construction-based, Rational, Exemplar-driven, Emergent, and Dialectic. Each of these key terms will be explained in detail below.

A fundamental tenet is that we learn language in much the same way as we learn everything else. The cognitive content of language systems is special because the problem of representing and sharing meanings across a serial speech stream is unique to

language, but the processes of learning are cut of the same cloth as the rest of human cognition. Thus SLA is governed by general laws of human learning, both **Associative** (the types of learning first analyzed within the *Behaviorist Tradition*) and **Cognitive** (the wider range of learning processes studied within *Cognitive Psychology*, including more conscious, explicit, deductive, or tutored processes).

### Construction Grammar

The basic units of language representation are **Constructions**. These are form-meaning mappings, conventionalized in the speech community, and entrenched as language knowledge in the learner's mind. Constructions are symbolic in that their defining properties of morphological, syntactic, and lexical form are associated with particular semantic, pragmatic, and discourse functions. Constructions are key components of *Cognitive Linguistic* and *Functional* theories of language. We learn constructions through using language, engaging in communication. *Usage-based theories* of language acquisition hold that an individual's creative linguistic competence emerges from the collaboration of the memories of all of the utterances in their entire history of language use and from the frequency-biased abstraction of regularities within them.

Many of the constructions we know are quite specific, being based on particular lexical items, ranging from a simple 'Wonderful!' to increasingly complex formulas like 'One, two, three', 'Once upon a time', or 'Won the battle, lost the war'. We have come to learn these sequential patterns of sound simply as a result of repeated usage. A major characteristic of the environments that are relevant to human cognition is that they are fundamentally probabilistic: every stimulus is ambiguous, as is any utterance or piece of language. Each of these examples of formulaic constructions begins with the sound 'w^n'. At the point of hearing this initial sound, what should the appropriate interpretation be? A general property of human perception is that when a sensation is associated with more than one reality, unconscious processes weigh the odds, and we perceive the most probable thing. *Psycholinguistic* analyses demonstrate that fluent language users are sensitive to the relative probabilities of occurrence of different constructions in the speech stream. Since learners have experienced many more tokens (particular examples) of 'one' than they have 'won', in the absence of any further information, they favor the unitary interpretation over that involving gain or advantage.

The following reviews provide overviews of the foundation fields of *Cognitive Linguistics* and *Usage-based* models of acquisition (Barlow & Kemmer 2000; Croft & Cruise 2004; Langacker 1987; Tomasello 2003, 1998), formulaic language processing (Ellis 1996; Pawley & Syder 1983; Wray 2002), and *Psycholinguistic* analyses of frequency effects in language processing and SLA (Bod, Hay, & Jannedy 2003; Bybee & Hopper 2001; Ellis 2002a, 2002b; Jurafsky 2002; Jurafsky & Martin 2000).

*The Associative and Cognitive Learning of Constructions*

The fact that high-frequency constructions are more readily processed than low-frequency ones is testament to associative learning from usage. Let's think about words, though the same is true for letters, morphemes, syntactic patterns, and all other types of construction. Through experience, a learner's perceptual system becomes tuned to expect constructions according to their probability of occurrence in the input, with words like one or won occurring more frequently than words like seventeen or synecdoche.

The learner's initial noticing of a new word can result in an explicit memory that binds its features into a unitary representation, such as phonological onset-rime sequence 'w^n' or the orthographic sequence "one". As a result of this, a detector unit for that word is added to the learner's perception system whose job is to signal the word's presence, or 'fire', whenever its features play out in time in the input. Every detector has a set resting level of activation, and some threshold level which, when exceeded, will cause the detector to fire. When the component features are present in the environment, they send activation to the detector that adds to its resting level, increasing it; if this increase is sufficient to bring the level above threshold, the detector fires. With each firing of the detector, the new resting level is slightly higher than the old one — the detector is said to be *primed*. This means it will need less activation from the environment in order to reach threshold and fire the next time that feature occurs. Priming events sum to lifespan-practice effects: features that occur frequently acquire chronically high resting levels. Their resting level of activity is heightened by the memory of repeated prior activations. Thus our pattern-recognition units for higher-frequency words require less evidence from the sensory data before they reach the threshold necessary for firing.

The same is true for the strength of the mappings from form to interpretation. Each time 'w^n' is properly interpreted as 'one', the strength of this connection is incremented. Each time 'w^n' signals 'won', this is tallied too, as are the less frequent occasions when it forewarns of 'wonderland'. Thus the strengths of form-meaning associations are summed over experience. The resultant network of associations, a semantic network comprising the structured inventory of a speaker's knowledge of their language, is so tuned that the spread of activation upon hearing the formal cue 'w^n' reflects prior probabilities.

There are many additional factors that qualify this simple picture: The relationship between frequency of usage and activation threshold is not linear but follows a curvilinear 'power law of practice' whereby the effects of practice are greatest at early stages of learning but eventually reach asymptote. The amount of learning induced from an experience of a form-function association depends upon the salience of the form and the functional importance of the interpretation. The learning of a form-function association is interfered with if the learner already knows another form which cues that interpretation (e.g., Yesterday I walked), or another interpretation for an ambiguous form (e.g. the definite article in English being used for both specific and generic

reference). A construction may provide a partial specification of the structure of an utterance and hence an utterance's structure is specified by a number of distinct constructions which must be collectively interpreted. Some cues are much more reliable signals of an interpretation than others. It is not just first-order probabilities that are important, it's sequential ones too, because context qualifies interpretation, with cues combining according to Bayesian probability theory: thus, for example, the interpretation of 'w^n' in the context 'Alice in w^n ...' is already clear. And so on. These factors, too complex to more than merely acknowledge here, together make the study of associative learning a fascinating business. *Associative Learning Theory* (Pearce 1997; Shanks 1995) has come a long way since the behaviorism of the 1950s, as too have accounts of first and second language acquisition in these terms (Christiansen & Chater 2001; Ellis 2002a, 2002b, in press-b, in press-c; Elman et al. 1996; MacWhinney 1987b, 1999, 2004).

### Rational Language Processing

Indeed, it has been argued that such associative underpinnings allow language users to be **Rational** in the sense that their mental models of the way language works are the most optimal given their linguistic experience and usage to date. The words that they are likely to hear next, the most likely senses of these words, the linguistic constructions they are most likely to utter next, the syllables they are likely to hear next, the graphemes they are likely to read next, the interpretations that are most relevant, and the rest of what's coming next across all levels of language representation, are made more readily available to them by their language processing systems. Their unconscious language representation systems are adaptively probability-tuned to predict the linguistic constructions that are most likely to be relevant in the ongoing discourse context, optimally preparing them for comprehension and production. The *Rational Analysis of Cognition* (Anderson 1989, 1990, 1991; Schooler & Anderson 1997) is guided by the principle that human psychology can be understood in terms of the operation of a mechanism that is "optimally adapted" to its environment in the sense that the behavior of the mechanism is as efficient as it conceivably could be given the structure of the problem space and the input-outputs mapping it must solve.

#### *The Associative Foundations of Rationality*

Language learning is thus an intuitive statistical learning problem, one that involves the associative learning of representations that reflect the probabilities of occurrence of form-function mappings whether these be of the first language (Elman 2004; Jurafsky 2002; Jurafsky & Martin 2000) or the second (Ellis in press-b; MacWhinney 1997). Learners have to figure language out: their task is, in essence, to learn the probability distribution  $P(\text{interpretation}|\text{cue, context})$ , the probability of an interpretation given a formal cue in a particular context, a mapping from form to meaning conditioned by

context. Rational analysis shows that this figuring is achieved, and communication optimized, by considering the *frequency*, *recency*, and *context* of constructions. These are the factors that determine the likelihood of a piece of information being needed in the world. Frequency, recency, and context are likewise the three most fundamental influences on human cognition, linguistic and non-linguistic alike.

### Exemplar-based abstraction and attraction

Although much of language use is formulaic, economically recycling constructions that have been memorized from prior use (Pawley & Syder 1983; Sinclair 1991), we are not limited to these specific constructions in our language processing. Some constructions are a little more open in scope, like the slot-and-frame greeting pattern ['Good' + (time-of-day)] which generates examples like 'Good morning,' and 'Good afternoon.' Others still are abstract, broad-ranging, and generative, such as the schemata that represent more complex morphological (e.g. [NounStem-PL]), syntactic (e.g. [Adj Noun]), and rhetorical (e.g. the iterative listing structure, [the (), the (), the (),..., together they...]) patterns. Usage-based theories investigate how the acquisition of these productive patterns, generative schema, and other rule-like regularities of language is **Exemplar-based**. The necessary generalization comes from frequency-biased abstraction of regularities from constructions of like-type. Constructions form a structured inventory of a speaker's knowledge of language (the constructicon) in which schematic constructions are abstracted over less schematic ones that are inferred inductively by the learner in acquisition: exemplars of similar type (e.g. [plural + 'cat' = 'cat-s'], [plural + 'dog' = 'dog-s'], [plural + 'elephant' = 'elephant-s'], ...) resonate, and from their shared properties emerge schematic constructions [plural + NounStem = NounStem-s]. Thus the systematicities and rule-like processes of language emerge as prototypes or schema, as frequency-tuned conspiracies of instances, as attractors which drive the default case, in the same ways as for the other categories by which we come to know the world.

The following reviews outline *Construction Grammar* (Goldberg 1995, 2003; Tomasello 2003) and *Cognitive Linguistic* analyses of first (Croft & Cruise 2004; J. R. Taylor 2002) and second language (Robinson & Ellis in press 2006).

#### *The Associative bases of Abstraction*

Prototypes, the exemplars that are most typical of their categories, are those that are similar to many members of their category but not similar to members of other categories. People more quickly classify as birds sparrows (or other average sized, average colored, average beaked, average featured specimens) than they do birds with less common features or feature combinations like geese or albatrosses; they do so on the basis of an unconscious frequency analysis of the birds they have known (their usage-history), with the prototype reflecting the central tendencies of the distributions of the

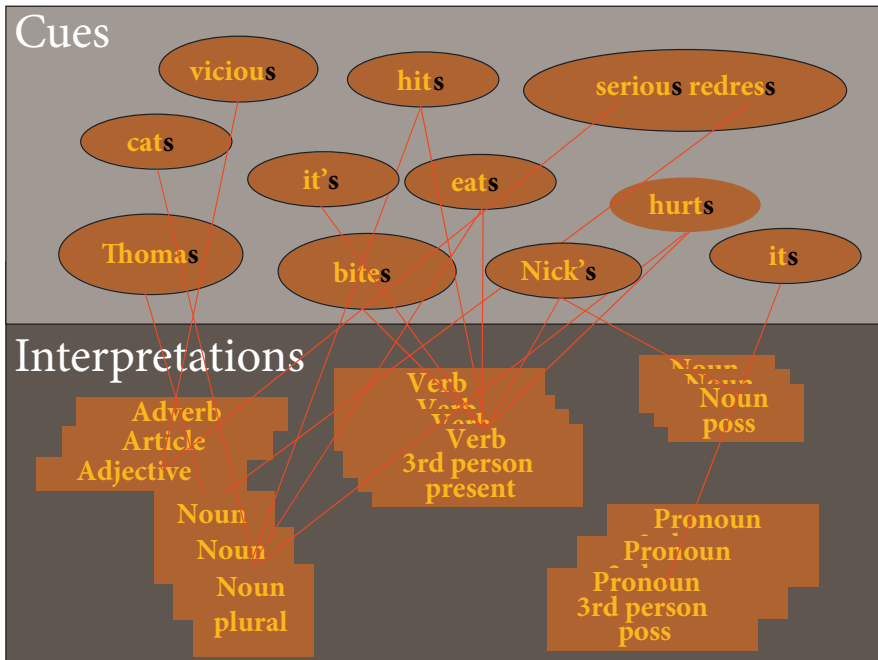


Figure 1. The variety of contingencies between the ‘-s’ morpheme in English and its functional interpretations make this a relatively low reliability cue

relevant features in the conspiracy of these memorized exemplars. Although we don’t go around consciously counting features, we nevertheless have very accurate knowledge of the underlying distributions and their most usual settings.

We are really good at this. Research in *Cognitive Psychology* demonstrates that such implicit tallying is the raw basis of human pattern recognition, categorization, and rational cognition. As the world is classified, so language is classified. As for the birds, so for their plurals. The *sparrows*, *geese*, and *albatrosses* examples illustrate similar processes in the acquisition of patterns of language: *Psycholinguistic* research demonstrates that people are faster at generating plurals for the prototype or default case that is exemplified by many types, and are slower and less accurate at generating ‘irregular’ cases, the ones that go against the central tendency and that have few friends operating in similarly deviant manner, like [plural + ‘NounStems’ = ‘NounStems-*es*’] or, worse still, [plural + ‘moose’ = ?], [plural + ‘noose’ = ?], [plural + ‘goose’ = ?].

These examples make it clear that there are no 1:1 mappings between cues and their outcome interpretations. Associative learning theory demonstrates that the more reliable the mapping between a cue and its outcome, the more readily it is learned. Consider an ESL learner trying to learn from naturalistic input what *-s* at the ends of words might signify. Plural *-s*, third person singular present *-s*, and possessive *-s*, are all homophonous with each other as well as with the contracted allomorphs of copula and auxiliary ‘be’.

Thus, as illustrated in Figure 1, if we evaluate  $-\underline{s}$  as a cue for one particular of these outcomes, it is clear that there are many instances of the cue being there but that outcome not pertaining. Consider the mappings from the other direction too: plural  $-\underline{s}$ , third person singular present  $-\underline{s}$ , and possessive  $-\underline{s}$  all have variant expression as the allomorphs [s, z, əz]. Thus if we evaluate just one of these, say əz as a cue for one particular outcome, say plurality, then it is clear that there are many instances of that outcome in the absence of the cue. Such contingency analysis of the reliabilities of these cue-interpretation associations suggests that they will not be readily learnable. Indeed, the low reliability of possessive  $-\underline{s}$  compounded by interference from contracted 'it is' ensures, as experience of undergraduate essays attests, that even native language learners can fail to sort out some aspects of this system after more than 10 years of experience. The apostrophe is opaque in its [sic] function. High frequency grammatical functors are often highly ambiguous in their interpretations. Consider the range of meanings of the English preposition *in* or the complex semantics and functions of definite and indefinite reference (Diesing 1992; Hawkins 1978; Lyons 1999).

So the simple story of constructions as form-function mappings is given added complexity by frequency and probability of association. Type frequency and the proportion of friends to enemies affect the productivity of patterns. The contingency or reliability with which a form signals an interpretation affects the learnability of constructions and their recruitment in processing. The following reviews outline these effects from the perspectives of linguistics (Bod, Hay, & Jannedy 2003; Bybee & Hopper 2001), natural language processing (Jurafsky & Martin 2000; Manning & Schuetze 1999), and first and second language acquisition (Bates & MacWhinney 1987; Ellis in press-c; Goldschneider & DeKeyser 2001; MacWhinney 1987a, 1997).

*Connectionist* models of language acquisition investigate the representations that result when simple associative learning mechanisms are exposed to complex language evidence. Connectionist simulations are data-rich and process-light: massively parallel systems of artificial neurons use simple learning processes to statistically abstract information from masses of input data as generalizations from the stored exemplars. It is important that the input data is representative of learners' usage history, which is why connectionist and other input-influenced research rests heavily upon the proper empirical descriptions of *Corpus Linguistics*. Connectionist simulations show how the default or prototype case emerges as the prominent underlying structural regularity in the whole problem space, and how minority subpatterns of inflection regularity, such as the English plural subpatterns discussed above or the much richer varieties of the German plural system, also emerge as smaller, less powerful attractors; less powerful because they have fewer friends and many more enemies, yet powerful enough nevertheless to attract friends that are structurally just like them (as in the [plural + 'NounStems' =?] case or [past tense + 'swim' / past tense + 'ring' / past tense + 'bring' /.../ past tense + 'spring' = ?]). Connectionism provides the computational framework for testing usage-based theories as simulations, for investigating how patterns appear from the interactions of many language parts.

The following reviews outline *Connectionist* approaches to first and second language (Christiansen & Chater 2001; Ellis 1998; Elman et al. 1996; Rumelhart & McClelland 1986), the *Competition Model* of language learning and processing (Bates & MacWhinney 1987; MacWhinney 1987a, 1997), and *Corpus Linguistics* (Biber, Conrad, & Reppen 1998; Sampson 2001; Sinclair 1991).

### Emergent relations and patterns

Complex systems, such as the weather, ecosystems, economies, and societies, are those that involve the interactions of many different parts. These share the key aspect that many of their systematicities are **Emergent**: they develop over time in complex, sometimes surprising, dynamic, adaptive ways. Complexity arises from the interactions of learners and problems too. Consider the path of an ant making its homeward journey on a pebbled beach. The path seems complicated as the ant probes, doubles back, circumnavigates and zigzags. But these actions are not deep and mysterious manifestations of intellectual power. Instead the control decisions are simple and few in number. An environment-driven problem solver often produces behavior that is complex because it relates to a complex environment.

Language is a complex adaptive system. It comprises the interactions of many players: people who want to communicate and a world to be talked about. It operates across many different levels (neurons, brains, and bodies; phonemes, morphemes, lexemes, constructions, interactions, and discourses), different human conglomerations (individuals, social groups, networks, and cultures), and different timescales (evolutionary, epigenetic, ontogenetic, interactional, neuro-synchronic, diachronic). As a classically complex system, its systematicities are emergent too. *Chaos/Complexity Theory* serves as the foundations for recent characterizations of theories of the *Emergence of Language*. Conscious reflective reasoning with (and about) language involves our knowledge of the world and our embodiment that constrains this knowledge. It has as a natural basis a lower plane of cognition that is associative and schematic; this apperceptive reasoning involves as a natural basis a lower plane of consciousness that is unreflective and perceptual; these perceptual activities rest upon sensory neural bases; these in turn involve a physico-chemical basis; and so on, as with the fleas,<sup>1</sup> ad infinitum. Each emergent level cannot come into being except by involving the levels that lie below it, and at each higher level there are new and emergent kinds of relatedness that are not found below: language cannot be understood in neurological or physical terms alone, nevertheless, neurobiology and physics play essential roles in the complex interrelations. Liz Bates, the sorely missed founding mother of this field, characterized these interrelationships, as with the turtles,<sup>1</sup> as being 'Emergence all the way down'. Fractal geometry provides a description of much of the world around us such as coastlines, rivers, plant distributions, architecture, wind gusts,<sup>3</sup> music, and the cardiovascular system, that have structure on many sizes. (Mandelbrot 1983).

Meteorologists have developed rules and principles of the phenomena of the



planet and its atmosphere that allow the prediction of weather. Geology has its rules and principles to describe and summarize the successive changes in the earth's crust. But these 'rules' are the descriptions and heuristics of science. They describe emergent patterns. The 'rules' themselves play no causal role in shifting even a grain of sand or a molecule of water. It is the interaction of water and rocks that smoothes the irregularities and grinds the pebbles and sand. Emergentists believe that many of the systematicities of language that are captured in linguistic analyses, for example, the parts of speech used to categorize different words, the syntactic roles used to describe different sentence parts, or the principles and parameters of UG, play a similar role to the meteorologists' or geologists' descriptions of their field and have a similar causal status. They are phenomena to be explained. The brain does not process information by some imaginary computational rule-following any more than the gut does (Searle 1992, see also Scott Kelso, 1997). Our linguistic systems are not dictionaries or grammars, they are not books, or lists or reference tables, frozen in time (Elman 2004). The notion of fixed cognitive categories, linguistic and non-linguistic both, is a myth: conceptual categories are dynamically construed (Smith & Samuelson 1997). Replace these static notions with those of dance, activity, the dynamic patterns (Scott Kelso 1997) of a large community of players in patterns of activity that are context- and perspective-sensitive, reflecting both past and current activity of the constructicon acting in consort with the rest of the cognitive system (Ellis 2005). For example: (1) Priming effects show how learners are more likely to use constructions that have been recently used (Bock & Griffin 2000). (2) The more the working memory demands of a task, the more they will use memorized patterns and formulaic speech; the less the working memory load, the more creative their constructions might be (Kuiper 1996). (3) Sociolinguistic and pragmatic factors spill over into the construction process (Hudson 1990, 1996). Different contexts, different dynamics of language use. Learning too is a dynamic process; it takes place during processing, as Hebb (1949), Craik and Lockhart (1972), Pienemann (1998), and O'Grady (2003) have all reminded us from their neural, cognitive, and linguistic perspectives. The Emergentist study of language acquisition looks to chart the course by which these regularities emerge as learners' perceptual, motoric, cognitive, and social functions induce structure, from undifferentiated novice performance to that remarkably differentiated nativelike competence. It focuses on process, learning and interaction as much as it does content, following a rational analysis of language acquisition in terms of functional models of optimal inference in the presence of uncertainty. Although the study of second language acquisition in these terms is relatively new, the notion of interlanguage has, from its very beginnings (Corder 1967; Selinker 1972), been characterized as reflecting the interactions of many sources of different types of knowledge of the L1 and the L2.

The following reviews provide good introductions to *Emergentist* theories of language (Bates & MacWhinney 1987; Bybee & Hopper 2001; MacWhinney 1998, 1999), *Radical Construction Grammar* and the emergence of syntactic constructions (Croft 2001), and *Chaos/Complexity theory* and first (Cooper 1999) and second (Larsen-Freeman 1997, in press; Larsen-Freeman & Ellis in press, December 2006) language acquisition.

*Emergent language representation: From tabula rasa to tabula repleta*

Our neural apparatus is highly plastic in its initial state. It is not entirely an empty slate, since there are broad genetic constraints upon the usual networks of system-level connections and upon the broad timetable of maturation and myelination, but nevertheless the cortex of the brain is broadly equipotent in terms of the types of information it can represent (Elman et al. 1996; Kandel, Schwartz, & Jessell 2000). But from this starting point, from the very get-go, it quickly responds to the input patterns it receives, and through associative learning, it optimizes its representations to rationally model the particular world of experience of each particular individual. The term 'neural plasticity' summarizes the fact that the brain is tuned by experience and that theories that rely heavily upon the inheritance of detailed knowledge representations are difficult to conceive of in neurological terms. Our neural endowment provides a general purpose cognitive apparatus, embodied within the general human form that filters, constrains, and determines our experience, for each of us to learn about our particular world. In the first few years of life, the human learning mechanism optimizes its representations of first language from the cumulative sample of first language input. One result of this process is that the initial state for SLA is a *tabula repleta*; it is no longer a plastic system, it is one that is already tuned and committed to the L1. Our later experience is shaded by prior associations; it is perceived through the memories of what has gone before. Since the optimal solution for L2 is not that for L1, SLA suffers from various types of L1 interference. Transfer phenomena pervade SLA (James 1980; Lado 1957; MacWhinney 1997; Odlin 1989; Weinreich 1953).

*Associative aspects of transfer: Learned attention and interference*

Associative learning provides the rational mechanisms for first language acquisition from input-analysis and usage, allowing just about every human being to acquire fluency in their native tongue. Yet although second language learners too are surrounded by language, not all of it 'goes in', and SLA is typically much less successful than L1A. This is Corder's distinction between input, the available target language, and intake, that subset of input that actually gets in and which the learner utilizes in some way (Corder 1967). Does this mean that SLA cannot be understood according to the general principles of associative learning that underpin other aspects of human cognition? If L1A is rational, is L2A fundamentally *irrational*? No, paradoxically perhaps, it is the very achievements of associative learning in first language acquisition that limit the input analysis of L2 and that result in the shortcomings of SLA. Associative learning theory explains these limitations too, because associative learning in animals and humans alike is affected by learned attention.

We can consider just one example here. Many grammatical meaning-form relationships are both low in salience and redundant in the understanding of the meaning of an utterance. It is often unnecessary to interpret inflections marking grammatical meanings such as tense because they are usually accompanied by adverbs that indicate

the temporal reference: “if the learner knows the French word for ‘yesterday’, then in the utterance *Hier nous sommes allés au cinéma* (Yesterday we went to the movies) both the auxiliary and past participle are redundant past markers.” (Terrell 1991: 59). This redundancy is much more influential in second rather than first language acquisition. Children learning their native language only acquire the meanings of temporal adverbs quite late in development. But second language learners already know about adverbs from their first language experience, and adverbs are both salient and reliable in their communicative functions, while tense markers are neither. Thus, the second language expression of temporal reference begins with a phase where reference is established by adverbials alone (Bardovi-Harlig 1992; Meisel 1987), and the grammatical expression of tense and aspect thereafter emerges only slowly, if at all (Bardovi-Harlig 2000).

This is an example of the associative learning phenomenon of ‘blocking’, where redundant cues are overshadowed for the historical reasons that learners’ first language experience leads them to look elsewhere for the cues to interpretation. Under normal L1 circumstances, usage optimally tunes the language system to the input; under these circumstances of low salience of L2 form and blocking, all the extra input in the world can sum to naught, and it is tempting to describe the interlanguage as having ‘fossilized’.

Ellis (in press-b; in press-c) reviews the ways in which associative learning induces such phenomena of learned selective attention as overshadowing and blocking, latent inhibition, perceptual learning, interference, and other effects of salience, transfer and inhibition. All of these mechanisms filter and color the perception of the second language. Thus the shortcomings of the L2 endstate are rational when seen through the lenses of the L1, although again, they are the result of many interacting parts and processes.

## Dialectic

Associative L2 learning from naturalistic usage can thus fall far short of a native-like endstate, often stabilizing at a ‘Basic Variety’ of interlanguage which, although sufficient for everyday communicative purposes, predominantly comprises just nouns, verbs and adverbs, with little or no functional inflection and with closed-class items, in particular determiners, subordinating elements, and prepositions, being rare, if present at all (Klein 1998).

The usual social-interactional or pedagogical reactions to such non-nativelike utterances involve an interaction-partner (Gass & Varonis 1994) or instructor (Doughty & Williams 1998) intentionally bringing additional evidence to the attention of the learner by some means of *Focus on Form*, form-focused instruction or consciousness raising that helps the learner to ‘notice’ the cue (Schmidt 2001). Terrell (1991) characterized explicit grammar instruction as the use of instructional strategies to draw the learner’s attention to, or focus on, form and/or structure, with instruction targeted at increasing the salience of inflections and other commonly ignored features by firstly pointing them out and explaining their structure, and secondly by providing

meaningful input that contains many instances of the same grammatical meaning-form relationship. 'Processing Instruction' (VanPatten 1996) similarly aims to alter learners' default processing strategies, to change the ways in which they attend to input data, thus to maximize the amount of intake of data to occur in L2 acquisition. In these ways, SLA can be freed from the bounds of L1-induced selective attention by some means of Focus on Form that is socially provided (Tarone 1997) and that recruits the learner's explicit conscious processing. Thus SLA is also *Dialectic* (Kramsch 2002; Lantolf & Pavlenko 1995; Lantolf & Thorne in press; Swain 2000), involving the learner in a conscious tension between the conflicting forces of their current interlanguage productions and the evidence of feedback, either linguistic, pragmatic, or metalinguistic, that allows socially scaffolded development. There is no better summary of the dialectic and dynamic characters of learning (Scott Kelso 2002) than that of Heraclitus: "All things come into being through opposition, and all are in flux like a river".

### *Cognitive contributions to SLA*

The Associative-Cognitive CREED is especially concerned in this particular interface between explicit and implicit learning in SLA (Ellis in press-a, 1994; Hulstijn & Ellis 2005; Krashen 1982; Reber 1993) and the general involvement of consciousness in cognition (Baars 1988, 1997; Baars, Banks, & Newman 2003; Koch 2004). The various roles of consciousness in second language acquisition (SLA) include: the learner noticing negative evidence; their attending to language form, their perception focussed by social scaffolding or explicit instruction; their voluntary use of pedagogical grammatical descriptions and analogical reasoning; their reflective induction of meta-linguistic insights *about* language; and their consciously guided practice which results, eventually, in unconscious, automatized skill.

Reviews of the experimental and quasi-experimental investigations into the effectiveness of explicit learning and L2 instruction (Doughty & Williams 1998; Ellis & Laporte 1997; Hulstijn & DeKeyser 1997; Lightbown, Spada, & White 1993; Long 1983; Norris & Ortega in press; Spada 1997), particularly the comprehensive meta-analysis of Norris & Ortega (2000) that summarized the findings from 49 unique sample studies experimental and quasi-experimental investigations into the effectiveness of L2 instruction, demonstrate that focused L2 instruction results in large target-oriented gains, that explicit types of instruction are more effective than implicit types, and that the effectiveness of L2 instruction is durable.

Ellis (2005) reviews this range of psychological, educational, and neurological processes by which explicit knowledge of form-meaning associations interfaces upon implicit learning in the emergence of interlanguage. The primary mechanism of explicit learning is the initial registration of pattern recognizers for constructions which are then tuned and integrated into the system by implicit learning during subsequent input processing. Neural systems in the prefrontal cortex involved in working memory provide the neuronal synchrony required for perceptual integration, buildup of coherent representations, attentional selection, awareness, and the unification of consciousness.

These are the mechanisms by which Schmidt's 'noticing' solves Quine's (1960) problem of 'referential indeterminacy'. Explicit learning results in explicit memories. Neural systems in the hippocampus bind disparate cortical representations into unitary episodic representations. By forming unitized memory representations, the hippocampal region performs the information-processing function of forming pattern-recognition units for new stimulus configurations and of consolidating new bindings; these are then adopted by other brain regions in the neocortex where they subsequently partake in implicit tuning (Gluck, Meeter, & Myers 2003). *Neuroscience* too is an increasingly important part of our understanding of implicit and explicit language learning and memory (Kandel, Schwartz, & Jessell 2000, Chapter 62) and in the ways these processes underpin SLA (Gullberg & Indefrey 2006).

There are various additional psychological and neurobiological processes by which explicit knowledge of form-meaning associations can interact with implicit language learning. The dynamics of the interface are complex, and here too there is little scope to go beyond acknowledging them: Explicit memories of formulas can serve analogical reasoning in the conscious construction of novel linguistic utterances and they can provide negative evidence, constraining the hypothesis space from overly-general grammars. Slot-and-frame patterns, drills, mnemonics, and declarative statements of pedagogical grammar likewise all contribute to the conscious creation of utterances which themselves then partake in subsequent implicit learning (Ellis 2005), proceduralization and automatization (DeKeyser 2001). Flawed output can also prompt focused feedback by way of recasts that present learners with psycholinguistic data ready for explicit analysis (Keck, Iberri-Shea, Tracy, & Wa-Mbaleka in press 2005; Swain 2005, 2006). Other processes of acquisition from output include differentiation, analysis, and preemption: Whole formulae, originally explicitly learned as phonological wholes, can later be dissected into their component structural parts (Ellis & Sinclair 1996). Conscious rehearsal in the phonological loop can provide data that evidences non-contiguous associations, discontinuous dependencies which, although out of the scope of implicit learning, can nevertheless thus be scrutinized and conjoined (Ellis, Lee, & Reber 1999). The interface too, like consciousness, is dynamic, situated, and contextualized: It happens transiently during conscious processing, but the influence upon implicit cognition endures thereafter (Ellis 2005).

## Conclusions

This is the *Zeitgeist* from within cognitive science as it strikes me. The Associative-Cognitive CREED is too broad to constitute a theory of SLA. Rather it reflects the resonance of cognate research in cognitive psychology, linguistics, computer science, cognitive neuroscience, education, and sociocultural theory, ideas which engage and interlock in a mutually-supportive framework, and whose interactions, when considered from a dynamic systems viewpoint, throw light on the emergence of many of the essential phenomena of SLA (Ellis in preparation).

But however harmonious, broad frameworks suffer a number of difficulties. They are hard to pin down, to operationalize, and to test. Like other enterprises in cognitive science and cognitive neuroscience, the evidence-base of the Associative-Cognitive CREED is diverse and wide-ranging, variously spanning psychological, linguistic, neurological, social and educational phenomena, and these different areas adopt widely different research techniques and standards of research quality.

Multi-componential frameworks open the potential for misunderstandings too if the contribution of each component to the whole is weighed in isolation. Alone, each is limited in its explanatory power: The CREED revives elements of behaviorism which were roundly attacked in the late 50s (Chomsky 1959). While the effects of usage frequency on learning and processing are coming back into general acceptance, they also have their bounds, as can be seen in the range of responses to the frequency-driven aspects of the CREED in the special issue of *Studies in Second Language Acquisition* volume 24:2. There are limits on the effects of first language transfer as well. Despite the long-standing interest in the Contrastive Analysis Hypothesis (CAH) (James 1980; Lado 1957; Weinreich 1953), there have been many influential and telling criticisms of the predictions of the original, relatively rigid, behaviorist CAH by many of the most influential figures in the field of SLA (Andersen 1983; Corder 1967; Coulon 1982; Gass & Selinker 1983; James 1980; Odlin 1989; Schachter 1974; Schachter & Celce-Murcia 1977; Selinker 1971, 1972, 1990; Spolsky 1979), criticisms which mark the development of the field of SLA as an independent discipline. It is also common to find adherents of different approaches within the framework criticizing each other for essential gaps in their coverage or for imperialist attacks on the field. For example, it has been claimed that cognitive approaches deny influence upon SLA of social factors, or motivational aspects, or paradoxically, learners' beliefs, desires or experiences of language learning, and, vice versa, that sociolinguistic theories ignore the brain and cognitive processes and lack scientific rigor in their methodologies.

Such criticisms are misguided in attacking just one piece of a larger whole. These components interact, they mediate and moderate each other, and, in other cases, can even amplify each other in positive feedback relationships. Consider for example the "morpheme order studies" that, in the 25 years following Brown's (1973) descriptions of first language acquisition, investigated the order of L2 acquisition of the grammatical functors, progressive -ing, plural -s, possessive -s, articles a, an, the, third person singular present -s, and regular past -ed. These studies show remarkable commonality in the orders of acquisition of these functors across a wide range of learners of English as a second language, yet, although each of the factors of input frequency, semantic complexity, grammatical complexity, phonological form, and perceptual salience had been historically considered for their sufficiency of cause, with input frequency being the favored account (Larsen-Freeman 1976), nevertheless, as Larsen-Freeman concluded, "[a] single explanation seems insufficient to account for the findings" (Larsen-Freeman 1975: 419). More recently, Goldschneider and DeKeyser (2001) performed a meta-analysis of these studies, investigating whether instead a *combination* of the five determinants (perceptual salience, semantic complexity, morphophonological regu-

larity, syntactic category, and frequency) could account for the acquisition order. Oral production data from 12 studies, together involving 924 subjects, were pooled. Each of the factors of frequency, salience and contingency alone was a significant predictor of acquisition order, but each independently only explained a small part (16–36%) of the variance: perceptual salience  $r = 0.63$ , frequency  $r = 0.44$ , morphophonological regularity  $r = 0.41$ . Yet when these three factors were combined with semantic complexity and syntactic category in a multiple regression analysis, this combination of five predictors jointly explained a substantial 71% of the variance in acquisition order. Add to these analyses the contribution of language transfer, and more still of the data is accounted for (Ellis *in press-c*; Shin & Milroy 1999; B. P. Taylor 1975)

Such multivariate and meta-analyses are an important step in the investigation of SLA, but there is room for improvement here too: However much they bring together various potential independent variables, they do so assuming linear combinations of their effects. Yet the shape of change in dynamical systems is typically non-linear, and we need to adopt and develop more appropriate ways of analysis to allow for this (Elman et al. 1996, Chapter 4; Scott Kelso 1997).

In sum, the CREED encourages the adoption of an emergentist framework which views SLA as a dynamic process in which regularities and system arise from the interaction of people, brains, selves, societies and cultures using languages in the world. Realizing this goal will present many theoretical and methodological challenges. The field of SLA is still evolving: which flowers thrive, how kempt the borders should be, who is entitled to tend them, and whether there should be a management strategy for this ecology, well, time will tell (Doughty & Long 2003; Firth & Wagner 1997, 1998; Gass 1998; Gregg 1993, 2005; Gregg, Long, Jordan, & Beretta 1997; Jordan 2003, 2004; Lantolf 1996, 2002; Long 1997). Fractally, L2 interlanguage is no more a static representational system than is current SLA research. They are states of being in dynamic systems and any analysis, like this article, is merely a description of the status quo. Your reactions affect the progress of the field, just like your language usage changes language. Quo vadis?

### Acknowledgements

I credit and thank David Ingledew for the credo. Thanks also to Diane Larsen-Freeman, Robert DeKeyser, and John Williams, for constructive comments on prior drafts.

### Notes

\* This article is based in part upon Ellis, N. C. (in press) "SLA: The Associative-Cognitive CREED". In B. VanPatten & J. Williams (Eds.), *Theories in Second Language Acquisition: An Introduction*. Cambridge: Cambridge University Press.

1. Big fleas have little fleas, Upon their back to bite 'em, And little fleas have lesser fleas, And so ad infinitum (De Morgan 1974: 377).
2. According to story, a famous scientist was giving a lecture on astronomy. After the lecture, an elderly lady came up and told him that he had it all wrong. "The world is really a flat plate supported on the back of a giant tortoise." The scientist asked "And what is the turtle standing on?" To which the lady triumphantly replied: "You're very clever, young man, but it's no use — it's turtles all the way down."
3. "Big whorls have little whorls, Which feed on their velocity; And little whorls have lesser whorls, and so on to viscosity" (Richardson 1925)
4. The articles of faith.

## References<sup>4</sup>

- Andersen, R.W. 1983. Transfer to somewhere. In *Language Transfer in Language Learning*, S. Gass and L. Selinker (eds), 250–268. Rowley MA: Newbury House.
- Anderson, J.R. 1989. A rational analysis of human memory. In *Varieties of Memory and Consciousness: Essays in honour of Endel Tulving*, H.L.I. Roediger and F.I.M. Craik (eds), 195–210. Hillsdale NJ: Lawrence Erlbaum Associates.
- Anderson, J.R. 1990. *The Adaptive Character of Thought*. Hillsdale NJ: Lawrence Erlbaum Associates.
- Anderson, J.R. 1991. Is human cognition adaptive? *Behavioral & Brain Sciences* 14(3): 471–517.
- Baars, B.J. 1988. *A Cognitive Theory of Consciousness*. Cambridge: CUP.
- Baars, B.J. 1997. In the theatre of consciousness: Global workspace theory, a rigorous scientific theory of consciousness. *Journal of Consciousness Studies* 4: 292–309.
- Baars, B.J., Banks, W.P. and Newman, J.B. (eds). 2003. *Essential Sources in the Scientific Study of Consciousness*. Cambridge MA: The MIT Press/Bradford Books.
- 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.
- Barlow, M. and Kemmer, S. (eds). 2000. *Usage Based Models of Language*. Stanford CA: CSLI Publications.
- Bates, E. and MacWhinney, B. 1987. Competition, variation, and language learning. In *Mechanisms of Language Acquisition*, B. MacWhinney (ed.), 157–193.
- Biber, D., Conrad, S. and Reppen, R. 1998. *Corpus Linguistics: Investigating language structure and use*. New York NY: Cambridge University Press.
- Bock, J.K. and Griffin, Z.M. 2000. The Persistence of Structural Priming: Transient activation or implicit learning? *Journal of Experimental Psychology: General* 129: 177–192.
- Bod, R., Hay, J. and Jannedy, S. (eds). 2003. *Probabilistic Linguistics*. Cambridge MA: The MIT Press.
- Brown, R. 1973. *A First Language: The early stages*. Cambridge MA: Harvard University Press.
- Bybee, J. and Hopper, P. (eds). 2001. *Frequency and the Emergence of Linguistic Structure*. Amsterdam: John Benjamins.



- Chomsky, N. 1959. Review of the book *Verbal Behavior*. *Language & Cognitive Processes* 35: 26–58.
- Christiansen, M.H. and Chater, N. (eds). 2001. *Connectionist Psycholinguistics*. Westport CO: Ablex.
- Cooper, D. 1999. *Linguistic Attractors: The cognitive dynamics of language acquisition and change*. Amsterdam: John Benjamins.
- Corder, S.P. 1967. The significance of learners' errors. *International Review of Applied Linguistics* 5: 161–169.
- Coulon, R. 1982. Is contrastive analysis better served by behaviourism than by cognitivism? *IRAL* 20: 241–250.
- Craik, F.I.M. and Lockhart, R.S. 1972. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior* 11: 671–684.
- Croft, W. 2001. *Radical Construction Grammar: Syntactic theory in typological perspective*. Oxford: OUP.
- Croft, W. and Cruise, A. 2004. *Cognitive Linguistics*. Cambridge: CUP.
- De Morgan, A. 1974. *A Budget of Paradoxes*, 2nd ed. London: Ayer.
- DeKeyser, R. 2001. Automaticity and automatization. In *Cognition and Second Language Acquisition*, P. Robinson (ed.). Cambridge: CUP.
- Diesing, M. 1992. *Indefinites*. Cambridge MA: The MIT Press.
- Doughty, C. and Long, M. (eds). 2003. *The Handbook of Second Language Acquisition*. Oxford: Blackwell.
- Doughty, C. and Williams, J. (eds). 1998. *Focus on Form in Classroom Second Language Acquisition*. New York NY: CUP.
- Ellis, N.C. 1996. Sequencing in SLA: Phonological memory, chunking, and points of order. *Studies in Second Language Acquisition* 18(1): 91–126.
- Ellis, N.C. 1998. Emergentism, connectionism and language learning. *Language Learning* 48(4): 631–664.
- Ellis, N.C. 2002a. 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. 2002b. Reflections on frequency effects in language processing. *Studies in Second Language Acquisition* 24(2): 297–339.
- 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. In preparation. The associative-cognitive CREED. In *Theories in Second Language Acquisition: An introduction*, B. VanPatten and J. Williams (eds). Cambridge: CUP.
- Ellis, N.C. In press-a. Implicit and explicit knowledge about language. In *Knowledge About Language*, Vol. 6. *Encyclopedia of Language and Education*, J. Cenoz (ed.), Heidelberg: Springer.
- Ellis, N.C. In press-b. Language acquisition as rational contingency learning. *Applied Linguistics* 27(1).
- Ellis, N.C. In press-c. Selective attention and transfer phenomena in SLA: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics* 27(2).
- Ellis, N.C. (ed.). 1994. *Implicit and Explicit Learning of Languages*. San Diego CA: Academic Press.

- Ellis, N.C. and Laporte, N. 1997. Contexts of acquisition: Effects of formal instruction and naturalistic exposure on second language acquisition. In *Tutorials in Bilingualism: Psycholinguistic perspectives*, A.M. DeGroot and J.F. Kroll (eds), 53–83. Mahwah NJ: Lawrence Erlbaum Associates.
- Ellis, N.C., Lee, M.W. and Reber, A.R. 1999. Phonological Working Memory in Artificial Language Acquisition. Ms, University of Wales, Bangor.
- Ellis, N.C. and Sinclair, S.G. 1996. Working memory in the acquisition of vocabulary and syntax: Putting language in good order. *Quarterly Journal of Experimental Psychology A*, 49A(1): 234–250.
- 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. and Plunkett, K. 1996. *Rethinking Innateness: A connectionist perspective on development*. Cambridge MA: The MIT Press.
- Firth, A. and Wagner, J. 1997. On Discourse, communication, and (some) fundamental concepts in SLA research. *The Modern Language Journal* 81: 285–300.
- Firth, A. and Wagner, J. 1998. SLA property: No trespassing! *Modern Language Journal* 82: 91–94.
- Gass, S. 1998. Apples and oranges: Or why apples are not orange and don't need to be: A response to Firth and Wagner. *The Modern Language Journal*, 82: 83–90.
- Gass, S. and Selinker, L. (eds). 1983. *Language Transfer in Language Learning*. Rowley MA: Newbury House.
- Gass, S. and Varonis, E. 1994. Input, interaction and second language production. *Studies in Second Language Acquisition* 16: 283–302.
- Gluck, M.A., Meeter, M. and Myers, C.E. 2003. Computational models of the hippocampal region: Linking incremental learning and episodic memory. *Trends in Cognitive Science* 7: 269–276.
- Goldberg, A.E. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago IL: University of Chicago Press.
- Goldberg, A.E. 2003. Constructions: A new theoretical approach to language. *Trends in Cognitive Science* 7: 219–224.
- Goldschneider, J.M. and DeKeyser, R. 2001. Explaining the natural order of L2 morpheme acquisition in English: A meta-analysis of multiple determinants. *Language Learning* 51: 1–50.
- Gregg, K.R. 1993. Taking explanation seriously; or, let a couple of flowers bloom. *Applied Linguistics* 14: 276–294.
- Gregg, K.R. 2005. A response to Jordan's (2004) 'Explanatory Adequacy and Theories of Second Language Acquisition'. *Applied Linguistics* 26: 121–124.
- Gregg, K.R., Long, M.H., Jordan, G. and Beretta, A. 1997. Rationality and its discontents in SLA. *Applied Linguistics* 18: 539–559.
- Gullberg, M. and Indefrey, P. (eds). 2006. *The Cognitive Neuroscience of Second Language Acquisition*. Oxford: Blackwell.
- Hawkins, J. 1978. *Definiteness and Indefiniteness: A study of reference and grammaticality prediction*. London: Croom Helm.
- Hebb, D.O. 1949. *The Organization of Behaviour*. New York NY: John Wiley & Sons.
- Hudson, R.A. 1990. *English Word Grammar*. Oxford: Blackwell.
- Hudson, R.A. 1996. *Sociolinguistics*, 2nd ed.. New York NY: CUP.

- Hulstijn, J. and DeKeyser, R., (eds). 1997. Testing SLA theory in the research laboratory. *Studies in Second Language Acquisition* 19(2) (Special Issue).
- Hulstijn, J. and Ellis, R., (eds). 2005. Theoretical and empirical issues in the study of implicit and explicit second-language learning. *Studies in Second Language Acquisition* 27(2) (Special Issue).
- James, C. 1980. *Contrastive Analysis*. London: Longman.
- Jordan, G. 2003. *Theory Construction in Second Language Acquisition*. Amsterdam: John Benjamins.
- Jordan, G. 2004. Explanatory adequacy and theories of SLA. *Applied Linguistics* 25: 539–543.
- Jurafsky, D. 2002. Probabilistic modeling in psycholinguistics: Linguistic comprehension and production. In *Probabilistic Linguistics*, R. Bod, J. Hay and S. Jannedy (eds), 39–96. Harvard MA: The MIT Press.
- Jurafsky, D. and Martin, J.H. 2000. *Speech and Language Processing: An introduction to natural language processing, speech recognition, and computational linguistics*. Englewood Cliffs NJ: Prentice-Hall.
- Kandel, E.R., Schwartz, J.H. and Jessell, T.M. 2000. *Principles of Neural Science*, 4th ed.. New York NY: McGraw-Hill.
- Keck, C.M., Iberri-Shea, G., Tracy, N. and Wa-Mbaleka, S. In press 2005. Investigating the empirical link between interaction and acquisition: A quantitative meta-analysis. In *Synthesizing Research on Language Learning and Teaching*, L. Ortega and J. Norris (eds). Amsterdam: John Benjamins.
- Klein, W. 1998. The contribution of second language acquisition research. *Language Learning* 48: 527–550.
- Koch, C. 2004. *The Quest for Consciousness: A neurobiological approach*. Englewood CO: Roberts and Company.
- Kramsch, C. (ed.). 2002. *Language Acquisition and Language Socialization: Ecological perspectives*. London: Continuum.
- Krashen, S.D. 1982. *Principles and Practice in Second Language Acquisition*. Oxford: Pergamon.
- Kuiper, K. 1996. *Smooth Talkers: The linguistic performance of auctioneers and sportscasters*. Mahwah NJ: Lawrence Erlbaum Associates
- Lado, R. 1957. *Linguistics across Cultures: Applied linguistics for language teachers*. Ann Arbor MI: University of Michigan Press.
- Langacker, R.W. 1987. *Foundations of Cognitive Grammar: Vol. 1. Theoretical prerequisites*. Stanford CA: Stanford University Press.
- Lantolf, J. 1996. Second language theory building: Letting all the flowers bloom! *Language Learning* 46: 713–749.
- Lantolf, J. 2002. Commentary from the flower garden : Responding to Gregg 2000. *Second Language Research* 18: 113–119.
- Lantolf, J. and Pavlenko, A. 1995. Sociocultural theory and second language acquisition. *Annual Review of Applied Linguistics* 15: 38–53.
- Lantolf, J. and Thorne, S. In press. *The Sociogenesis of Second Language Development*. Oxford: OUP.
- Larsen-Freeman, D. 1975. The acquisition of grammatical morphemes by adult ESL students. *TESOL Quarterly* 9: 409–419.
- 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. In press. On the need for a new metaphor for language and its development.
- Larsen-Freeman, D. and Ellis, N.C. (eds). In press, December 2006. *Language Emergence: Implications for applied linguistics*.
- Lightbown, P.M., Spada, N. and White, L. 1993. The role of instruction in second language acquisition. *Studies in Second Language Acquisition* 15 (Special issue).
- Long, M.H. 1983. Does second language instruction make a difference? A review of research. *TESOL Quarterly* 17: 359–382.
- Long, M.H. 1997. Construct validity in SLA: A response to Firth and Wagner. *Modern Language Journal* 81: 318–323.
- Lyons, J. 1999. *Definiteness*. Cambridge: CUP.
- MacWhinney, B. 1987a. The competition model. In *Mechanisms of Language Acquisition*, B. MacWhinney (ed.), 249–308. Hillsdale NJ: Lawrence Erlbaum Associates.
- MacWhinney, B. 1997. Second language acquisition and the competition model. In *Tutorials in Bilingualism: Psycholinguistic perspectives*, A.M.B. De Groot and J.F. Kroll (eds), 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. 2004. A multiple process solution to the logical problem of language acquisition. *Journal of Child Language* 31: 883–914.
- MacWhinney, B. (ed.). 1987b. *Mechanisms of Language Acquisition*. Hillsdale NJ: Lawrence Erlbaum Associates.
- MacWhinney, B. (ed.). 1999. *The Emergence of Language*. Hillsdale NJ: Lawrence Erlbaum Associates.
- Mandelbrot, B.B. 1983. *The Fractal Geometry of Nature*. London: W. H. Freeman.
- Manning, C.D. and Schuetze, H. 1999. *Foundations of Statistical Natural Language Processing*. Cambridge MA: The MIT Press.
- Meisel, J. 1987. Reference to past events and actions in the development of natural second language acquisition. In *First and Second Language Acquisition*, C. Pfaff (ed.), New York NY: Newbury House.
- Norris, J. and Ortega, L. 2000. Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning* 50: 417–528.
- Norris, J. and Ortega, L. (eds). 2006. *Synthesizing Research on Language Learning and Teaching*: John Benjamins.
- O'Grady, W. 2003. The radical middle: Nativism without universal grammar. In *Handbook of Second Language Acquisition*, C. Doughty and M.H. Long (eds) Oxford: Blackwell.
- Odlin, T. 1989. *Language Transfer*. New York NY: CUP.
- Pawley, A. and Syder, F.H. 1983. Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In *Language and Communication*, J.C. Richards and R.W. Schmidt (eds), 191–225. London: Longman.
- Pearce, J.M. 1997. *Animal Learning and Cognition*, 2nd ed.. Hove: Psychology Press/ LEA, UK.
- Pienemann, M. 1998. *Language Processing and Second Language Development: Processability theory*. Amsterdam: John Benjamins.
- Quine, W.V.O. 1960. *Word and Object*. Cambridge MA: The MIT Press.

- Reber, A.S. 1993. *Implicit Learning and Tacit Knowledge: An essay on the cognitive unconscious*. Oxford: OUP.
- Richardson, L.F. 1925. Turbulence and vertical temperature difference near trees. *Philosophical Magazine* 49: 81–90.
- Robinson, P. and Ellis, N.C. (eds). In press 2006. *Handbook of Cognitive Linguistics and Second Language Acquisition*. Mahwah NJ: Lawrence Erlbaum Associates.
- Rumelhart, D.E. and McClelland, J.L. (eds). 1986. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol. 2: *Psychological and biological models*. Cambridge MA: The MIT Press.
- Sampson, G. 2001. *Empirical Linguistics*. London: Continuum.
- Schachter, J. 1974. An error in error analysis. *Language Learning* 24: 205–214.
- Schachter, J. and Celce-Murcia, M. 1977. Some reservations concerning error analysis. *TESOL Quarterly* 11: 441–451.
- Schmidt, R. 2001. Attention. In *Cognition and Second Language Instruction*, P. Robinson (ed.), 3–32. Cambridge: CUP.
- Schooler, L.J. and Anderson, J.R. 1997. The role of process in the rational analysis of memory. *Cognitive Psychology* 32(3): 219–250.
- Scott Kelso, J.A. 1997. *Dynamic Patterns: The self-organization of brain and behavior*. Cambridge MA: A Bradford Book/The MIT Press.
- Scott Kelso, J.A. 2002. The complementary nature of coordination dynamics: Self-organization and agency. *Nonlinear Phenomena in Complex Systems* 5: 364–371.
- Searle, J.R. 1992. *The Rediscovery of the Mind*. Cambridge MA: Bradford Books/The MIT Press.
- Selinker, L. 1971. A brief reappraisal of contrastive linguistics. *University of Hawaii Working Papers in Linguistics* 3: 1–10.
- Selinker, L. 1972. Interlanguage. *IRAL* 10: 209–231.
- Selinker, L. 1990. Reframing contrastive analysis. *Georgetown University Round Table on Languages and Linguistics*: 137–143.
- Shanks, D.R. 1995. *The Psychology of Associative Learning*. New York NY: CUP.
- Shin, S.J. and Milroy, L. 1999. Bilingual language acquisition by Korean school children in New York City. *Bilingualism: Language and Cognition* 2: 147–167.
- Sinclair, J. 1991. *Corpus, Concordance, Collocation*. Oxford: OUP.
- Smith, L.B. and Samuelson, L.K. 1997. Perceiving and remembering: Category stability, variability, and development. In *Knowledge, Concepts and Categories*, K. Lamberts and D. Shanks (eds), 161–195. Hove: Psychology Press.
- Spada, N. 1997. Form-focused instruction and second language acquisition: A review of classroom and laboratory research. *Language Teaching Research* 30: 73–87.
- Spolsky, B. 1979. Contrastive analysis, error analysis, interlanguage, and other useful fads. *The Modern Language Journal* 63: 250–257.
- Swain, M. 2000. The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In *Sociocultural Theory and Second Language Learning*, J. Lantolf (ed.), 97–114. Oxford: OUP.
- Swain, M. 2005. The output hypothesis: Theory and research. In *Handbook of Research in Second Language Teaching and Learning*, E. Hinkel (ed.), 471–484. Mahwah NJ: Lawrence Erlbaum Associates.
- Swain, M. 2006. Languaging, agency and collaboration in advanced second language proficiency. In *Proceedings of Gurt 2005*.

- Tarone, E. 1997. Analyzing IL in natural settings: A sociolinguistic perspective of second language acquisition. *Communication and Cognition* 30: 137–150.
- Taylor, B.P. 1975. Adult language learning strategies and their pedagogical implications. *TESOL Quarterly* 9: 391–399.
- Taylor, J.R. 2002. *Cognitive Grammar*. Oxford: OUP.
- Terrell, T. 1991. The role of grammar instruction in a communicative approach. *The Modern Language Journal* 75: 52–63.
- Tomasello, M. 2003. *Constructing a Language*. Boston MA: Harvard University Press.
- Tomasello, M. (ed.). 1998. *The New Psychology of Language: Cognitive and functional approaches to language structure*. Mahwah NJ: Lawrence Erlbaum Associates.
- VanPatten, B. 1996. *Input Processing and Grammar Instruction in Second Language Acquisition*. New York NY: Ablex.
- Weinreich, U. 1953. *Languages in Contact*. The Hague: Mouton.
- Wray, A. 2002. *Formulaic Language and the Lexicon*. Cambridge: CUP.