5

The Associative–Cognitive CREED

Nick C. Ellis
The University of Michigan

THE THEORY AND ITS CONSTRUCTS

SLA has been actively studied from a *cognitive* perspective for the last two or three decades, and researchers within this tradition share basic goals, methods, and constructs. As in any active field of scientific enquiry, our theories are continuously measured against the observable phenomena, refined, and developed. The position outlined in this chapter is fairly typical of the beliefs shared by psychologists. The Associative–Cognitive CREED holds that SLA is **C**onstruction-based, **R**ational, **E**xemplar-driven, **E**mergent, and **D**ialectic. Some of these terms may be new to you but each will be explained in detail below, and it is worth sticking with them because of their currency in the different subdivisions of contemporary cognitive science.

A fundamental tenet of SLA 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 the same as those involved in 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 introduced in chapter 1) and *cognitive* (the wider range of learning processes studied within cognitive psychology, including more conscious, explicit, deductive, or tutored processes). You will see commonalities between this chapter and others in this volume by cognitively minded SLA researchers such as DeKeyser, VanPatten, Gass and Mackey, and Pienemann, although we clearly have different emphases too.

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. These usage-based theories of language acquisition hold that we learn constructions by engaging in communication. Thus an individual's creative linguistic competence emerges from the combination of two things: the memories of all of the utterances encountered in communicative situations, and the induction of regularities in those utterances based on frequency (Ellis, 2002).

Many of the constructions we know are quite specific and are 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 patterns 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 interpretation of *one* over *won*.

[MAB1] ***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 realization of its whole form, such as the spoken sequence "w∧n" or the orthographic sequence *one*. As a result of this, a detector unit for that word, whose job is to signal the word's presence, or "fire," whenever its features are present in the input, is added to the learner's perception system. Every word detector has a resting level of activation and some threshold level that, 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, thus 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. Features that occur frequently acquire consistently 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 "wʌn" is properly interpreted as *one* this connection is strengthened. Each time "wʌn" 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 "wʌn" 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 diminish thereafter (see DeKeyser, chapter 6, this volume, for discussion of the power law of practice). 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 that cues that interpretation (e.g., "Yesterday I walked"; see VanPatten's *Lexical Primacy Principle*, this volume), or an- [MAB2]
other interpretation for an ambiguous form (e.g., the preposition *a* in Spanish meaning *to* but also used as a direct object case marker). Some cues are much more reliable signals of an interpretation than others (e.g., word order is a more reliable cue to the subject role in English than is animacy). Knowing the most probable interpretation of individual cues is not sufficient; context is important too, with cue interpretation probabilities combining sequentially to qualify interpretation. Thus, for example, the interpretation of 'wʌn' in the context "Alice in wʌn . . ." is already clear.

Rational Language Processing

Indeed, it has been argued that such associative underpinnings allow language users to be *rational* in the sense that their understanding of the way language works is the best mental model possible, given their linguistic experience 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 is 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 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 (Ellis, in press-a). [MAB3]

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. Learners have to figure language out. 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, 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 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 and other rule-like regularities of language are *exemplar-based*, with generalizations coming from frequency-biased abstraction of regularities from similar constructions.

[MAB4] 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 sparrows as birds 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, with the prototype reflecting the central tendencies of the distributions of the relevant features. 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 case that is exemplified by many types, and are slower and less accurate at generating irregular cases, those that go against the central tendency and that have few "friends" (e.g., exemplars similar in form that operate in similarly deviant manners; compare the plurals of moose, goose, and noose). 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*. This is illustrated in Figure 5–1.

Connectionist models of language acquisition are used to investigate the representations that result when simple associative learning mechanisms are exposed to complex language evidence of this type. Connectionist simulations are massively parallel systems of artificial neurons that use simple learning processes. From the large numbers of stored exemplars, these processes abstract information (generalizations) based on frequency. Connectionist simulations show how the prototype case emerges as the prominent underlying structural

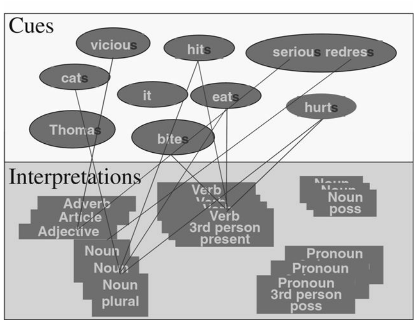

FIGURE 5–1. The variety of contingencies between the *–s* morpheme in English and its functional interpretations make this a relatively low reliability cue.

regularity, and how minority subpatterns of inflection regularity—such as the English plural subpatterns discussed above—also emerge as smaller, less powerful generalizations. These generalizations are less powerful because they are less frequent; yet they are nonetheless powerful enough to attract friends that are structurally just like them (e.g., swim, swam, swum and drink, drank, drunk).

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. Meteorologists have developed rules and principles of the phenomena of the planet and its atmosphere that allow the prediction of weather. Geologists have outlined 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—all 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 emergentist study of language acquisition examines how these regularities emerge as learners' perceptual, cognitive, motor, and social functions induce structure.

Emergent language representation: From blank slate to language transfer. Our neural apparatus is highly plastic in its initial state. It is not entirely a blank slate, because there are broad genetic constraints on the usual networks of system-level connections and on the broad timetable of maturation and myelination. Nevertheless the cortex of the brain is broadly equipotent in terms of the types of information it can represent (Elman et al., 1996). But from this starting point, 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 no longer a plastic system; it is one that is already tuned and committed to the L1. Transfer phenomena thus pervade SLA (James, 1980; Lado, 1957; Odlin, 1989).

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 may be surrounded by language, not all of it "goes in," and SLA is typically much less successful than first language acquisitions (L1A). This is Corder's distinction between input (the available target language) and intake (that subset of input that actually gets in and that 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, does this mean that SLA is 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 (Ellis, in press-b).

We can consider just one example here. Many grammatical meaning-form [MAB5] relationships are both low in salience as well as 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, p. 59; see also VanPatten, this volume). This redundancy is much more influential in second rather than first language acquisition. Children learning their native language only ac- [MAB6] quire 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, 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," whereby 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; however, in the L2 situation, forms of low salience may be blocked by prior L1 experience, and all the extra input in the world may not result in advancement.

Dialectic

Associative L2 learning from naturalistic usage can thus fall far short of a native-like endstate. The usual social–interactional or pedagogical reactions to such non-nativelike utterances involve an interaction partner or instructor intentionally bringing additional evidence to the attention of the learner. In these ways, SLA can be freed from the bounds of L1-induced selective attention by some means of intervention that is socially provided (Lantolf & Thorne, chapter 11, this volume) and that recruits the learner's explicit conscious processing. Thus SLA is also *dialectic*, 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.

Cognitive Contributions to SLA

Reviews of the experimental and quasi-experimental investigations into the effectiveness of explicit learning and L2 instruction (Ellis & Laporte, 1997; Spada, 1997), particularly the comprehensive meta-analysis of Norris & Ortega (2000), 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.

The Associative–Cognitive CREED is particularly concerned with the ways in which explicit and implicit learning processes interact in SLA. Ellis (2005) argues that the primary mechanism of explicit learning is the initial registration of pattern recognizers for constructions that are then tuned and integrated into the system by implicit learning during subsequent input processing.

WHAT COUNTS AS EVIDENCE?

Like other enterprises in cognitive science and cognitive neuroscience, the Associative–Cognitive CREED does not adhere to any specific research methodology or evidence. Data come from diverse approaches to research: educational experimental designs, psycholinguistic studies of sentence processing, analyses of learner production, and brain imaging and neuroscience. It is important that theories of SLA accord with what else we know about human learning. Thus, the CREED may draw upon evidence from studies related to comprehension, production, interaction, and computer simulations in order to build its case for the associative nature of L2 learning.

COMMON MISUNDERSTANDINGS

Broad frameworks, particularly those that revive elements of no-longer-fashionable theories such as behaviorism, open the potential for misunderstandings, of which we will mention a few here. Common misconceptions include that connectionism is the new behaviorism; that connectionist models cannot explain creativity and that connectionist models have no regard for internal representation; and that cognitive approaches deny influence of social factors, or motivational aspects, or—paradoxically—learners' beliefs, desires, or experiences of language learning on SLA. These misunderstandings are the result of focusing on one piece of the larger whole. The pieces of that whole interact, and when we acknowledge these interactions, we see how the separate components can limit each other, can mediate and moderate each other, and in other cases can even amplify each other in positive feedback relationships. A clear example of the interactions of frequency, salience, and reliability of mapping can be seen in the choice of an exemplary study that follows here. A more general moral is that it is important to adopt a complex systems framework that views SLA as a dynamic process in which regularities and system emerge from the interaction of people, their conscious selves, and their brains, using language in their societies, cultures, and world.

AN EXEMPLARY STUDY: GOLDSCHNEIDER AND DEKEYSER (2001)

As introduced by VanPatten and Williams in chapter 2, SLA follows predictable and reliable developmental sequences. Goldschneider and DeKeyser (2001) wanted to determine the degree to which one well-attested order of acquisition, that concerning grammatical morphemes in English, is the result of aspects of the input such as frequency, perceptual salience, morphophonological regularity, and semantic complexity. Frequency, the reliability of the mappings between cues and their interpretations, and salience, as explained above, are essential determinants of associative learning in humans and animals alike. Therefore, can this aspect of SLA be equally understood in these terms?

Their study is a meta-analysis, a "study of studies." Meta-analysis is an important and powerful research tool that allows the systematic and replicable overview of findings relating to a particular question by means of the quantitative pooling of the sizes of the effects found in all of the relevant empirical studies. It is the best way to pull together the research data into one [MAB7] large body, review the findings, and summarize its trends (Norris & Ortega, in press).

Goldschneider and DeKeyser (2001) analyzed 12 morpheme order studies that, in the 25 years following Brown (1973), 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*, and established a common order of acquisition for these structures in L2 that was broadly reliable across learners of different age and first-language backgrounds.

Brown (1973) had investigated the English L1 order of acquisition of grammatical morphemes in a naturalistic corpus of the speech of three children, Adam, Eve, and Sarah. Dulay and Burt (1973) followed this by investigating the order of acquisition in 151 Spanish-speaking children, aged 6–8, learning English as an L2 in the United States. Sentences were elicited via the Bilingual Syntax Measure (BSM), which involved cartoon descriptions to establish the percentage suppliance of different grammatical morphemes in obligatory contexts. Bailey, Madden, and Krashen (1974) subsequently tested the order of acquisition in 73 adults, 33 of whom were Spanish-speakers and the rest represented 11 other languages. These adult L2 learners showed similar orders for the acquisition of English morphemes, regardless of their mother tongue:

1. plural *–s*	"Book*s*"
2. progressive *–ing*	"John go*ing*"
3. copula *be*	"John *is* here"/"John'*s* here"
4. auxiliary *be*	"John *is* going"/"John'*s* going"
5. articles *the/a*	"*The* books"
6. irregular past tense	"John *went*"
7. third person *–s*	"John like*s* books"
8. possessive *'s*	"John'*s* book"

These and ten other studies showed remarkable commonality in the orders of acquisition of these functors across a wide range of learners of English as a second (and first) language. Although a number of possible explanations of these factors have been proposed, including frequency in the input, semantic complexity, grammatical complexity, phonological form, and perceptual salience, with input frequency being the favored major cause (Larsen-Freeman, 1975), nevertheless, Larsen-Freeman concluded that "[a] single explanation seems insufficient to account for the findings" (p. 419).

Goldschneider and DeKeyser (2001) pooled the studies, developed objective operationalizations of each of these aspects of the input, and analyzed their weights as individual independent causes of the observed acquisition order using multiple regression analysis.

Input frequency was assessed from a corpus of spoken English. Perceptual salience was composed from three subfactors: the number of phones in the functor (phonetic substance), the presence or absence of a vowel in the surface form (syllabicity), and the total relative sonority of the functor. As explained above, Goldschneider and DeKeyser's factor of morphophonological regularity relates to reliability of form-function mapping because conditioned phonological variation (for example, the [s, z, əz] allomorphs of plural *–s*, possessive *–s*, and third person singular *–s*) results in multiple forms of the cue and thus a less clear mapping between the interpretation and one particular cue, while homophony with other grammatical factors results in a less clear mapping between the cue and one particular interpretation. Semantic complexity measured the number of meanings expressed by a given form (e.g., third person singular *–s* expresses person, number, and present tense). Finally, syntactic category weighted the morphemes according to whether they were lexical or functional items, and within each of these groups whether they were free or bound morphemes (free morphemes are typically acquired before bound morphemes, again suggesting explanations in terms of salience).

Oral production data from the 12 studies, together involving 924 subjects, were pooled. On their own, each of these factors significantly correlated with acquisition order: perceptual salience $r = 0.63$, frequency $r = 0.44$, morphophonological regularity $r = -0.41$, syntactic category $r = 0.68$, and semantic complexity $r = -0.41$. When these five factors were combined in a multiple regression analysis, they jointly explain 71% of the variance in acquisition order, with perceptual salience having the highest predictive power on its own.

Each of the factors of frequency, reliability of form-function mapping, and perceptual salience is a significant predictor of acquisition order; together they explain a substantial amount of acquisition difficulty. That these stimulus factors are the major determinants of associative learning in animals (Rescorla & Wagner, 1972) strongly supports explanations in terms of associative learning from input analysis and usage-based acquisition.

Goldschneider and DeKeyser's analysis could not investigate the modulating effects of transfer from learners of different L1 backgrounds (Ellis, in press-b) because, as they explain themselves on p. 31, the original studies did [MAB8]
not provide sufficient details of this. Nor, of course, could they look at the de-

tailed sentence contexts of the morphemes as they had appeared in the learners' particular input histories.

EXPLANATION OF OBSERVED FINDINGS IN SLA

Observation # 1. Exposure to input is necessary for SLA. The Associative–Cognitive CREED is input driven. As with other statistical estimations, a large and representative sample of language is required for the learner to abstract a rational model that is a good fit to the language data. Input is necessary, and it is sufficient for successful L1A but not for SLA. This is because the initial state for SLA is a no longer a blank slate: The learner's language representations, processing routines, and attention to language, are tuned and committed to the L1. The optimal solution for L2 is not the same as for L1, and the shortcomings of SLA stem from these various aspects of transfer and learned attention.

Observation # 2. A good deal of SLA happens incidentally. The Associative–Cognitive CREED holds that the large majority of language learning is implicit. All of the counting that underpins the setting of thresholds and the tuning of the system to the probabilities of the input evidence is unconscious. So also is the emergence of structural regularities, prototypes, attractors, and other system regularities. At any one point we are conscious of one particular communicative meaning; yet meanwhile the cognitive operations involved in each of these usages are tuning the system, unconsciously, without us being aware of it (Ellis, 2002). We "know" far too many linguistic regularities for us to have explicitly learned them.

Observation # 3. Learners come to know more than what they have been exposed to in the input. The Logical Problem of Language Acquisition argues that second language learners' grammars, like those of children learning their first language, are said to be underdetermined by the input, and some linguists therefore propose that learners have internal mental representations of how the language works (putative abstract rules that govern sentence structure) that contain information that could not have been learned by exposure. Productivity, abstraction, and generalization are natural features of connectionist systems—we have already seen how, like children, they produce overgeneralization errors such as *goed*. The processes of this generativity are demonstrable by simulation. The study of implicit human cognition shows that we know far more about the world than we are aware of or have been explicitly taught. Prototype effects are one clear and ubiquitous example of this: We have never seen the prototype deer but we will classify it faster and more accurately than examples further from the central tendency, and we will call it a deer, and not a horse, with great facility. Different notions, different constructions. VanPatten and Williams in chapter 1 cite the UG argument that the acquisition of *wanna* contractions shows that learners have abstract rules involving *wh*-traces, whereby contraction is not possible "across a trace," rules that are neither taught in ESL classes nor evidenced from the language that learners are exposed to. However, other linguists explain that transitive *want NP to* and intransitive *want to* are different lexicalized constructions, and *wanna* is only possible (or is very much more probable) as a contraction of the latter, a very natural shortening given that the sequence *want* + *to* happens much more frequently in the intransitive construction than in the transitive one (Bybee, 2002; Sag & Fodor, 1996). Again, different notions, different constructions. We do not have explanations of the emergence of all of the linguistic representations that play their part in UG accounts of language competence, any more than generativists have explained the inheritance and neural instantiation of these. The future will tell whether the logical problem will be solved better by inheritance or emergence. Meanwhile, it is clear that structure-dependence and hierarchical organization, factors which have been traditionally held up as unique and defining features of language, are actually properties of all sequential routines, scripts, and motor behaviors—they, at least, are naturally emergent properties.

Observation # 4. Learners' output (speech) often follows predictable paths with predictable stages in the acquisition of a given structure. As in first language acquisition, SLA is characterized not by complete idiosyncrasy or randomness but rather by predictable errors and stages during the course of development: Interlanguage is systematic. The Associative–Cognitive CREED believes that these systematicities arise from regularities in the input; for example, constructions that are much more frequent, that are consistent in their mappings and exhibit high contingency, that have many friends, and that are salient are likely to be acquired earlier than those that do not have these features. The Goldschneider and DeKeyser (2001) study featured earlier clearly demonstrates how such factors interact in determining the acquisition order of one specific SLA system. Connectionist simulations allow the investigation of how systematicities emerge from the conspiracy of such factors.

Observation # 5. Second language learning is variable in its outcome. Our linguistic systems are not dictionaries or grammars; they are not books or lists or reference tables, frozen in time. 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 activity, dance, and dynamic patterns of a large community of agents in patterns of activity that are context- and perspective-sensitive, reflecting both past and current activity of the language systems in interaction with the rest of the cognitive system (Ellis, 2005). For example, (a) priming effects show how learners are more likely to use constructions that have been recently used; (b) 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; and (c) sociolinguistic and pragmatic factors spill over into the construction process. Different contexts, different dynamics. Each of the components of the system can affect the process and outcome success.

Observation # 6. Second language learning is variable across linguistic subsystems. Language representation is diverse in its contents and spans lexicon, morphology, syntax, phonology, pragmatics, and sociolinguistics, among others. Within any of these areas of language, learners may master some structures before they acquire others. Such variability is a natural consequence of input factors such as exemplar type and token frequency, recency, context, salience, contingency, regularity, reliability, and conspiracy of friends: enemies along with the various other associative learning factors that affect the emer-
[MAB9] gence of regularities.

Thus development, which is for the most part gradual and incremental, also evidences sudden changes in performance, suggesting occasional fundamental restructuring of the underlying grammar. Consider the first language acquisition of past tense marking. Learners initially fail to mark past tense; their first marking involves frequent irregular verbs such as *came* and *went*; next appear regular marking (addition of the default ending) in verbs such as *talked* and *cooked*, and the productivity of this schema is evidenced by the disappearance of irregulars from the interlanguage as they are replaced by overextensions (incorrect forms that have regular endings like *goed, wented*). The stage at which irregulars disappear and are replaced by regularized forms is sudden and suggests that learners' grammars are restructuring themselves to make everything regular even though such forms as *goed* are not part of the input. Eventually, of course, the irregulars reappear.

Observation # 7. There are limits on the effects of frequency on SLA. Although there is no simple and direct correlation between input frequency and acquisition, broadly, high frequency elements are acquired earlier than low frequency elements. Frequency interacts with many other factors in language learning, such as salience, complexity, and individual differences. The study of associative learning concerns the factors that determine how much is acquired from each learning experience and how each increment is integrated into a representational system (Shanks, 1995).

Observation # 8. There are limits on the effect of a learner's first language on SLA. At every level of language, there is evidence of L1 influence, both negative and positive. Although it is no longer considered the clear and direct influence proposed in the Contrastive Analysis Hypothesis, its significance in the language learning process seems incontrovertible. The Associative–Cognitive CREED considers how various cross-linguistic phenomena of learned selective attention—blocking, latent inhibition, perceptual learning, interference and other effects of salience, transfer, and inhibition—all filter and color the perception of the second language (Ellis, in press-a, in press-b). [MAB10]

Observation # 9. There are limits on the effects of instruction on SLA. L1-tuned learned attention limits the amount of intake from L2 input, thus restricting the endstate of SLA. The cognitive and dialectic aspects of the CREED consider how attention to language form is sometimes necessary to allow learners to notice some blocked aspect of the language form. Reviews of the empirical studies of instruction demonstrate that tapping into learners' conscious and explicit learning processes can be effective.

However, instruction is not always effective. Any classroom teacher can provide anecdotal evidence that what is taught is not learned. But this observation can be made for all aspects of the curriculum, not just language. Explicit knowledge about language is of a different stuff from that of the implicit representational systems, and it need not impact upon acquisition from a large variety of reasons. Explicit instruction can be ill-timed and out of synchrony with development (Pienemann, 1998, chapter 8, this volume); it can be con- [MAB11] fusing; it can be easily forgotten; it can be dissociated from usage (Ellis, 1993), lacking in transfer-appropriateness and thus never brought to bear so as to tune attention to the relevant input features during usage; it can be unmotivating; it can fail in so many ways. Ineffective instruction is all too easy; the challenge is the reverse.

Observation # 10. There are limits on the effects of output (learner production) on language acquisition. The CREED holds that there are a variety of mechanisms by which the demands of output can encourage creative construction. The conscious processes involved here allow a dynamic interface whereby explicit knowledge can influence implicit language learning (Ellis, 2005, section 4). Output is also a driving force of chunking, proceduralization, and automatization. There are thus a variety of roles of output in the SLA. However, you cannot learn a second language by output alone.

CONCLUSION

In the terms of chapter 1, the Associative–Cognitive CREED is perhaps too broad to qualify as a theory. But it does aspire to being a useful framework that relates many of the complex agents that underlie the dynamic nature of consciousness and the interface between explicit and implicit learning. No single factor alone is a sufficient cause of SLA. 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, different human configurations, and different timescales. Take out any one of these levels and a different pattern emerges, a different conclusion is reached. But nevertheless, like other complex dynamic systems, there are many systematicities that, like observations 1–10, emerge to form the central phenomena of SLA.

ACKNOWLEDGMENTS

I thank David Ingledew for the credo. Thanks also to Robert DeKeyser, John Williams, Diane Larsen-Freeman, and the editors of this volume for constructive comments on prior drafts. [MAB12]

DISCUSSION QUESTIONS

1. One critique of the type of approach Ellis takes is that it is an updated version of behaviorism. Do you agree with this criticism? How would Ellis respond?
2. Explain the difference between rule-based and rule-like behavior.
3. How does the Associative–Cognitive CREED address explicit and implicit learning and the nature of their interface?
4. Consider the case of the acquisition order of the present perfect and the pluperfect from Bardovi-Harlig (chapter 4). The functionalist approach offers one explanation for the order of emergence. How would the Associative–Cognitive CREED account for the order? What evidence might distinguish the two interpretations?
5. As we saw in chapter 3, the principle foundation of the approach White takes is the *poverty of the stimulus* (POS) situation or the *logical problem of language acquisition*. How does the Associative–Cognitive CREED view the POS? (You might want to review the examples in White's chapter before answering.)

SUGGESTED FURTHER READINGS

Croft, W., & Cruise, A. (2004). *Cognitive linguistics*. Cambridge: Cambridge University Press.

A comprehensive introduction to the fast-growing field of cognitive linguistics.

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.

This book marked a sea change in our understanding of learning and development. Six coauthors representing cognitive psychology, connectionism, neurobiology, and dynamical-systems theory synthesize a new emergentist framework for cognitive development with special focus on language acquisition.

Larsen-Freeman, D., & Ellis, N. C. (Eds.) (2006). *Language emergence: Implications for applied linguistics*. *Applied Linguistics, 27*, (4).

A special issue focusing on emergentist accounts of topics including complexity in interlanguage, the multilingual lexicon, the development of complex syntactic structures, bilingual language acquisition, metaphor in discourse, and language evolution, with commentaries from cognitive and sociocultural perspectives.

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

Proceedings of the founding conference for Emergentist Approaches to Language. Child language researchers, linguists, psycholinguists, and modelers using a range of formalisms present emergentist accounts of a wide range of linguistic phenomena.

Robinson, P. (Ed.). (2001). *Cognition and second language instruction*. Cambridge: Cambridge University Press.

Tutorial reviews of second language instruction and acquisition as seen from cognitive psychology. Topics include attention, memory, processing, learnability, connectionism and the competition model, and individual differences.

Robinson, P., & Ellis, N. C. (Eds.). (in press). *Handbook of cognitive linguistics and second language acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates. [MAB13]

A handbook of cognitive linguistics. The first half brings together key figures in the development of cognitive linguistics as it relates to first language acquisition. The second half has second-language researchers develop these themes for SLA.

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

A thorough account of child language acquisition from the perspectives of psychology and construction grammar.

REFERENCES

Bailey, N., Madden, C., & Krashen, S. (1974). Is there a 'natural sequence' in adult second language learning? *Language Learning, 24*, 235–243.

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

Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Bybee, J. (2002). Phonological evidence for exemplar storage of multiword sequences. *Studies in Second Language Acquisition, 24*, 215–221.

Corder, S. P. (1967). The significance of learners' errors. *International Review of Applied Linguistics, 5*, 161–169.

Dulay, H. C., & Burt, M. K. (1973). Should we teach children syntax? *Language Learning, 23*, 245–258.

Ellis, N. C. (1993). Rules and instances in foreign language learning: Interactions of explicit and implicit knowledge. *The European Journal of Cognitive Psychology, 5*, 289–318.

———. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition, 24*, 143–188.

———. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition, 27*, 305–352.

———. (in press-a). Language acquisition as rational contingency learning. *Applied Linguistics, 27*. [MAB14]

———. (in press-b). Selective attention and transfer phenomena in SLA: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics, 27*. [MAB15]

[MAB16] Ellis, N. C., & Laporte, N. (1997). Contexts of acquisition: Effects of formal instruction and naturalistic exposure on second language acquisition. In A. M. DeGroot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 53–83). Mahwah, NJ: Lawrence Erlbaum Associates.

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.

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.

James, C. (1980). *Contrastive analysis*. London: Longman.

Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. Ann Arbor: University of Michigan Press.

Larsen-Freeman, D. (1975). The acquisition of grammatical morphemes by adult ESL students. *TESOL Quarterly, 9*, 409–419.

Norris, J., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning, 50*, 417–528.

Norris, J., & Ortega, L. (Eds.). (in press). *Synthesizing research on language learning and* [MAB17] *teaching*: Benjamins. [MAB18]

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

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.

Sag, I. A., & Fodor, J. D. (1996). A traceless account of extraction phenomena. *Languages, 30*, 8–31.

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

Smith, L. B., & Samuelson, L. K. (1997). Perceiving and remembering: category stability, variability, and development. In K. Lamberts & D. Shanks (Eds.), *Knowledge, concepts and categories* (pp. 161–195). Hove: Psychology Press. [MAB19]

Spada, N. (1997). Form-focused instruction and second language acquisition: A review of classroom and laboratory research. *Language Teaching Research, 30*, 73–87.

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

[MAB1]Why is this paragraph set in italics and indented? If it is an extracted quote, it needs a citation and it should not be set in italics. If this sentence is to be a level-4 head, it is the only one in this section and should be removed and the paragraph styled in roman.

[MAB2]Is this in chapter 7? Cite chapter number here.

[MAB3]Insert date if available.

[MAB4]This heading was deleted because it was the only level-4 head in this section.

[MAB5]Provide date if it is available.

[MAB6]Provide chapter number.

[MAB7]Provide date if available.

[MAB8]Provide date if available.

[MAB9]Note sure if this phrase after the colon makes sense.

[MAB10]Provide dates for both citations if available.

[MAB11]Pienemann, 1998 not included in references.

[MAB12]Because this is the only chapter acknowledgments, I suggest including these in the acknowledgments in the front matter.

[MAB13]Provide date if available.

[MAB14]Provide date if available.

[MAB15]Provide date if available.

[MAB16]Need page numbers if available.

[MAB17]Provide date if available.

[MAB18]Need place of publication.

[MAB19]City/state of publication?