Usage–based approaches to SLA[1]

Nick C. Ellis

University of Michigan

ncellis@umich.edu

Stefanie Wulff

University of Florida

swulff@ufl.edu

What is the Theory and What are the Major Constructs?

Various approaches to Second Language Acquisition (SLA) can be labeled as "usage–based". What unites these approaches is their commitment to two working hypotheses:

(1) Language learning is primarily based on learners' exposure to their second language (L2) in use, that is, the linguistic input they receive.

Ellis, N. C. & Wulff, S. (2014). Usage-based approaches to SLA. In B. VanPatten & J. Williams (Eds.)., *Theories in second language acquisition: An introduction* (2nd Edn), chapter 6. NY: Routledge.

(2) Learners induce the rules of their L2 from the input by employing cognitive mechanisms that are not exclusive to language learning, but that are general cognitive mechanisms at work in any kind of learning, including language learning.

In the following, we will look at the following major constructs of usage–based approaches to SLA in more detail:

- Constructions: language learning is the learning of constructions, pairings of form and meaning or function. Constructions range from simple morphemes like *–ing* to complex and abstract syntactic frames such as Subject–Verb–Object–Object (as in *Nick made Steffi a sandwich*).
- Associative language learning: learning constructions means learning the association between form and meaning or function. The more reliable the association between a form and its meaning or function, the easier it is to learn. For example, the sound sequence /ˈsæn(d)wɪtʃ/ is reliably associated with a particular meaning ("slices of meet and/or cheese between two slices of bread"). The form *–ing*, in contrast, has different meaning/functions in different contexts, making it comparatively harder to learn.
- Rational cognitive processing: language learning is rational such that a learner's knowledge of a given form–meaning pair at any point in their language development is a reflection of how often and in what specific contexts the learner has encountered that form–meaning pair.
- Exemplar–based learning: language learning is in large parts implicit in the sense of taking place without the learner being consciously aware of it. The learner's brain engages simple learning mechanisms in distributional analyses of the exemplars of a given form–meaning

pair that take various characteristics of the exemplar into consideration, including how frequent it is, what kind of words and phrases and larger contexts it occurs with, etc.

- Emergent relations and patterns: language learning is a gradual process in which language emerges as a complex and adaptive (in the sense of continuously fine–tuning) system from the interaction of simple cognitive learning mechanisms with the input (and in interaction with other speakers in various social settings).

Constructions

The basic units of language representation are **constructions**. Constructions are pairings of form and meaning or function. By that definition, we know that simple words like, say, *squirrel*, must be constructions: a form – that is, a particular sequence of letters or sounds – is conventionally associated with a meaning (in the case of *squirrel*, something like 'agile, bushy–tailed, tree–dwelling rodent that feeds on nuts and seeds'). In Construction Grammar, constructions are not restricted to the level of words (Goldberg, 2006). Instead, these form–function pairings are assumed to pervade all layers of language. Simple morphemes such as –*licious* (roughly meaning 'delightful or extremely attractive') are constructions. Idiomatic expressions such as *I can't wrap my head around this* (meaning 'I do not fully comprehend this') are constructions. Even abstract syntactic frames are constructions: sentences like *Nick gave the squirrel a nut*, *Steffi gave Nick a hug*, or *Bill baked Jessica a cake* all have a particular form (Subject–Verb–Object–Object) that, regardless of the specific words that realize its form, share at least one stable aspect of meaning: something is being transferred (nuts, hugs, and cakes). Some constructions do not have a meaning in the traditional sense, but serve more functional purposes; passive constructions, for example, serve to shift what is in attentional focus by defocusing the agent of the action (compare an active sentence such as *Bill baked Jessica a cake* with its passive counterpart *A cake was baked for Jessica*).

Constructions can be simultaneously represented and stored in multiple forms and at various levels of abstraction: *table* + *s* = *tables*; [Noun] + (morpheme *–s*) = 'plural things'). Ultimately, constructions blur the traditional distinction between lexicon and grammar. A sentence is not viewed as the application of grammatical rules to put a number of words obtained from the lexicon in the right order; a sentence is instead seen as a combination of constructions, some of which are simple and concrete while others are quite complex and abstract. For example, *What did Nick give the squirrel?* comprises the following constructions:

- *Nick*, *squirrel*, *give*, *what*, *do* constructions
- VP, NP constructions
- Subject–Verb–Object–Object construction
- Subject–Auxiliary inversion construction

We can therefore see the language knowledge of an adult as a huge warehouse of constructions. Constructions vary in their degree of complexity and abstraction. Some of them can be combined with one another while others cannot; their combinability largely depends on whether their meanings/functions are compatible, or can at least be coerced into compatibility, given the specific context and situation in which a speaker may want to use them together. The more often a speaker encounters a particular construction, or combination of constructions, in the input, the more entrenched that (arrangement of) constructions becomes.

Associative Learning Theory

Constructions that are frequent in the input are processed more readily than rare constructions. This empirical fact is compatible with the idea that we learn language from usage in an associative manner. Let's stick to words for now, though the same is true for letters, morphemes, syntactic patterns, and all other types of constructions. 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*.

When a learner notices a word in the input for the first time, a memory is formed that binds its features into a unitary representation, such as the phonological sequence /wʌn/ or the orthographic sequence *one*. Alongside this representation, a so–called detector unit is added to the learner's perceptual system. The job of the detector unit is to signal the word's presence whenever its features are present in the input. Every detector unit 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 previous one – the detector is primed. This means it will need less activation from the environment in order to reach threshold and fire the next time. 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 /wʌn/ is properly interpreted as *one*, the strength of this connection is incremented. 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 language, is tuned such that the spread of activation upon hearing the formal cue /wʌn/ reflects prior probabilities of its different interpretations.

Many additional factors qualify this simple picture. Firstly, 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 (see DeKeyser, this volume). Secondly, the amount of learning induced from an experience of a construction depends upon the salience of the form (that is, how much it stands out relative to its context) and the importance of understanding it correctly. Thirdly, the learning of a construction is interfered with if the learner already knows another form that cues that interpretation, or conversely, if the learner knows another interpretation for that form. Fourthly, 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, and it is not just first–order probabilities that are important – sequential probabilities matter a great deal as well, because context qualifies interpretation. For example, the interpretation of /wʌn/ in the context *Alice in wun …* is already clear after the learner has heard *Alice in …*; in other words, *Alice in* and /wʌn/ are highly reliably associated with each other. If a sentence starts out with *I /wʌn/ …*, in contrast, several competing interpretations are co–activated

(*I wonder …*, *I won …*, *I once …*, etc.) because the first person pronoun *I* is a much less reliable cue for the interpretation of /wʌn/ than *Alice*.

Rational Language Processing

These associative underpinnings allow language users to be rational in the sense of having a mental model of their language that is custom–fit to their linguistic experience at any given time (Ellis, 2006a). 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 readily available to them by their language processing systems. Their unconscious language representation systems are adaptively 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. As a field of research, the **rational analysis of cognition** 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 cue–interpretation mappings it must solve (Anderson, 1989).

Exemplar–based Learning

Much of our language use is formulaic, that is, we recycle phrasal constructions that we have memorized from prior use (Wulff, 2008). However, we are obviously not limited to these 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 based on **exemplars**. Every time the language learner encounters an exemplar of a construction, the language system compares this exemplar with memories of previous encounters of either the same or a sufficiently similar exemplar in order to retrieve the correct interpretation. According to exemplar theory, constructions such as *Good* + (time of day), [Adj Noun], or [NounStem–PL] all gradually emerge over time as the learner's language system, processing exemplar after exemplar, identifies the regularities that exemplars share, and makes the corresponding abstractions.

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 sparrows as birds (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 that reflects the central tendencies of the distributions of the relevant features of these memorized exemplars. We don't walk around consciously counting these features, but yet we 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 plural forms. In fact, world and language categorization go hand in hand: 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" plurals, the ones that go against the central tendency and that are exemplified by fewer types, such as [plural + 'NounStem*s*' = 'NounStem*s*–*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. This particular form has several potential interpretations: it could be the plural (*squirrels*), it could indicate possession (*Nick's hat*), it could mark third person singular present (*Steffi sleeps*), etc. Therefore, if we evaluate –*s* as a cue for any one of these outcomes, it is clear that the cue will be abundantly frequent in learners' input, yet neither of the cues are reliably associated with their interpretation or outcome. A similar picture emerges when we reverse the directionality of our thinking: plural –*s*, third person singular present –*s*, and possessive –*s* all have variant expression as the allomorphs [*s*], [*z*], and [ɨz].Thus if we evaluate just one of these, say [ɨz] 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 the cue–interpretation associations suggests that they will not be readily learnable. High frequency grammatical functors are often highly ambiguous in their interpretations (Goldschneider & DeKeyser, 2001).

Connectionism is one strand of research in SLA that seeks to investigate how simple associative learning mechanisms such as the kind of contingency analysis mentioned above meets the complex language evidence available to a learner in their input and output. The term "connectionist" reflects the idea that mental and behavioral models are in essence interconnected networks of simple units. Connectionist models are typically run as computer simulations. The simulations are data–rich and process–light: massively parallel systems of artificial neurons use simple learning processes to statistically generalize over masses of input data. 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 large–scale, maximally representative digital collections of authentic language (these are often called databanks or **corpora**). Connectionist simulations show how prototypes emerge 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, for example), also emerge as smaller, less powerful attractors. Connectionism provides the computational framework for testing usage–based theories as simulations, for investigating how patterns appear from the interactions of many language parts.

Emergent relations and patterns

Complex systems are those that involve the interactions of many different parts, such as ecosystems, economies, and societies. All complex systems share the key aspect that many of their systematicities are **emergent**: they develop over time in complex, sometimes surprising, dynamic, and 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 (Beckner et al., 2009; Ellis & Larsen–Freeman, 2009, Larsen–Freeman, this volume). 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, constructions, interactions, and discourses), different human conglomerations (individuals, social groups, networks, and cultures), and different timescales (evolutionary, epigenetic, ontogenetic, interactional, neuro–synchronic, diachronic). 'Emergentists believe that simple learning mechanisms, operating in and across the human systems for perception, motor–action and cognition as they are exposed to language data as part of a communicatively–rich human social environment by an organism eager to exploit the functionality of language, suffice to drive the emergence of complex language representations.' (Ellis 1998: 657).

Two languages and language transfer

Our neural apparatus is highly plastic in its initial state. It is not entirely an empty slate, since there are broad genetic constraints on the usual networks of system–level connections and on the broad timetable of maturation. Nevertheless, the cortex of the brain is broadly equipotent in terms of the types of information it can represent (Elman et al., 1996). From this starting point, the brain quickly responds to the input patterns it receives, and through associative learning, it optimizes its representations to model the particular world of an individual's experience. The term "neural plasticity" summarizes the fact that the brain is tuned by experience. Our neural endowment provides a general purpose cognitive apparatus that, constrained by the make–up of our human bodies, filters and determines our experiences. In the first few years of life, the human learning mechanism optimizes its representations of the first language (L1) being learned. Thousands of hours of L1 processing tunes the system to the cues of the L1 and automatizes its recognition and production. It is impressive how rapidly we start tuning into our ambient language and disregarding cues that are not relevant to them (Kuhl, 2004). 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. Our later experience is shaded by prior associations; it is perceived through the memories of what has gone before. Since the optimal representations for the L2 do not match those of the L1, SLA is impacted by various types of L1 interference. Transfer phenomena pervade SLA (Flege, 2002; Jarvis & Pavlenko, 2008; Lado, 1957; MacWhinney, 1997; Odlin, 1989; Weinreich, 1953).

Associative aspects of transfer: Learned attention and interference

Associative learning provides the rational mechanisms for L1 acquisition from input–analysis and usage, allowing just about every human being to acquire fluency in their native tongue. Yet although L2 learners too are surrounded by language, not all of it 'goes in', and SLA is typically limited in success. 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? If L1 acquisition is rational, is SLA fundamentally irrational? No. Paradoxically perhaps, it is the very achievements of L1 acquisition that limit the input analysis of the L2. Associative learning theory explains these limitations too, because associative learning in animals and humans alike is affected by what is called **learned attention**.

We can consider just one example of learned attention here. Many grammatical form–meaning relationships are both low in salience and redundant in the understanding of the meaning of an utterance. It is often unnecessary, for instance, to interpret inflections that mark 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). This redundancy is much more influential in second rather than L1 acquisition. Children learning their native language only acquire the meanings of temporal adverbs quite late in development. But L2 learners already know about adverbs from their L1 experience, and adverbs are both salient and reliable in their communicative functions, while tense markers are neither (see VanPatten, this volume on the "Lexical Preference Principle"). Thus, the L2 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 volume).

This is an example of the associative learning phenomenon of "blocking", where redundant cues are overshadowed because the learners' L1 experience leads them to look elsewhere for the cues to interpretation (Ellis, 2006b). Under normal L1 circumstances, usage optimally tunes the language system to the input; under these circumstances of low salience of L2 form and blocking, however, all the extra input in the world can sum to nothing, with interlanguage sometimes being described as having "fossilized" (Han & Odlin, 2006). Untutored adult associative L2 learning from naturalistic usage can thus stabilize 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 or not present at all (Klein, 1998).

The usual social–interactional or pedagogical reactions to such non–nativelike utterances involve an interaction partner (Gass & Mackey, this volume; Long, 1983; Mackey, Abbuhl, & Gass, 2011) or instructor (Doughty & Williams, 1998) who intentionally brings additional evidence to the learner’s attention by some means of attentional focus that helps the learner to “notice” the cue (Schmidt, 2001). This way, SLA can be freed from the bounds of L1—induced selective attention: a focus on form is provided in social interaction (Lantolf, Thorne & Poehner, this volume; Tarone, 1997) that recruits the learner’s **explicit** conscious **processing**. We might say that the input to the associative network is “socially gated” (Kuhl, 2007).

What counts as evidence?

Like other enterprises in cognitive science and cognitive neuroscience, usage–based approaches are not restricted to one specific research methodology or evidential source. Indeed different approaches require different methods, and often a combination of different qualitative and quantitative methods. As mentioned earlier, many usage—based analyses employ data from large digitized collections of language, so—called corpora; computational modeling is at the heart of rational cognition analysis, exemplar theory, and emergentist analyses alike. Other relevant research methods include classroom field research, psycholinguistic studies of processing, and dense longitudinal recording.

Corpus–based analysis constitutes a particularly growing trend across usage–based paradigms (McEnery & Hardie, 2012; Sinclair, 1991). If language learning is in the social–cognitive linguistic moment of usage, we need to capture all these moments so that we can objectively study them. We need large, dense, longitudinal corpora of language use, with audio, video, transcriptions and multiple layers of annotation, for data sharing in open archives. We need these in sufficient dense mass so that we can chart learners' usage history and their development (Tomasello & Stahl, 2004). We need them in sufficient detail that we can engage in detailed analyses of the processes of interaction (Kasper & Wagner, 2011). MacWhinney has long been working towards these ends, first with CHILDES (MacWhinney, 1991), a corpus of L1 acquisition data, and later with Talkbank (MacWhinney, 2007), a corpus that also covers language data from L2 learners. Alongside these and other corpora, a growing number of computer tools are becoming available that assist the researcher in analyzing corpus data. These corpus tools can help researchers interested in the most diverse areas of SLA by covering the full range from qualitative data analysis, such as a fine–grained conversation analysis of individual corpus files (say, a transcribed conversation between a student and an ESL teacher), to semi–quantitative analysis of a representative sample of attestations of a particular phenomenon (such as the use of the *–ing* morpheme by English language learners), to large–scale quantitative analysis of distributional patterns (for example, the association strength between verbs and the larger constructions they occur in – see the exemplary study below or Gries & Wulff, 2005, 2009 for examples).

What are some common misunderstandings about the theory?

Broad frameworks, particularly those that revive elements of no–longer–fashionable theories such as behaviorism or structuralist approaches to linguistics, open the potential for misunderstanding. Common misconceptions include that connectionism is the new behaviourism; that connectionist models cannot explain creativity, and have no regard for internal representation; and that cognitive approaches deny influence of social factors, motivational aspects, and other individual differences between learners. At the heart of most of these misunderstandings is the idea that usage–based analyses only do number–crunching, with too much of a focus on the effects that the frequency of constructions and other cues plays in the learning process. While it is true that most usage–based approaches will discuss frequency as one of several factors, no usage–based theorist would claim that frequency is the only factor impacting SLA. In fact, there is a lively debate among usage–based theorists about the exact role that frequency effects play in what is conceived of as a complex network of factors that can mute and amplify each other in complex ways (Ellis & Larsen-Freeman, 2006). At an even more fundamental level, what constitutes a frequency effect in the first place is a question we are far from having answered. Without going into too much detail here, there is ample empirical evidence, for instance, that we cannot always define a frequency effect by the rule 'the more frequent, the more salient/important/relevant' – by that rationale, English articles and prepositions, which are the most frequent words in the English language, should not pose such an obstinate challenge to the average language learner! Instead, it seems that frequency effects come in different kinds (as absolute frequencies, ratios, association strengths, and other distributional patterns), and they will have differently weighted impacts depending on the target structure under examination, and, crucially, depending on the state of the learner's language development. An emergentist/complex systems approach views SLA as a dynamic process in which regularities and systems emerge from many of the processes covered in this volume – from the interaction of people, brains, selves, societies, and cultures using languages in the world (Beckner et al., 2009; Ellis, 2008) – while at the same time investigating component processes in a rigorous fashion.

An Exemplary Study: Ellis, O'Donnell, & Römer (2014a)

Research Questions

While previous studies were able to demonstrate that frequency, prototypicality, and contingency are factors that impact L2 learners' constructional knowledge, most of these studies have considered only one of these factors at a time. This study wanted to determine whether and how these factors jointly affect L2 learners' constructional knowledge. The specific kind of constructions this study focused on are so–called verb–argument constructions (VACs). VACs are semi–abstract patterns that comprise verbs and the arguments they occur with, such as 'V *across* N' or 'V *of* N'; in this study, the authors examined VACs that another team of researchers previously identified using corpus analysis (Francis, Hunston, & Manning, 1996).

Methods

One hundred and thirty one German, 131 Spanish, and 131 Czech advanced L2 learners of English as well as 131 native speaker of English were engaged in a free association task: they were shown 40 VAC frames such as 'V *across* N or '*it* V *of the* N' and asked to fill in the verb slot with the first word that came to mind. The learners' responses were compared with results obtained from two native speaker databases. In order to get an impression of the frequencies with which different verbs occur in the VACs examined, and in order to calculate how strongly each verb is associated with the individual VACs, the authors consulted the British National Corpus (BNC). The BNC is a 100–million word corpus of British English that strives to be representative of language use across different registers and genres. In order to obtain the verb type frequencies, one can simply run a search for the VACs in the BNC and count how often each verb type occurs. In order to calculate the association strength between each verb type and each VAC, the authors used a specific association measure called DeltaP (for more information on how DeltaP works, see Ellis, 2006a). In order to see how prototypical the verbs selected by the participants would be for each VAC, the authors consulted a second data base called WordNet (Miller, 2009). WordNet is a lexical database, so unlike the BNC, it is not a collection of cohesive and complete texts and dialogues, but rather a thesaurus–like database that groups words together based on their meanings. Using sophisticated computational techniques, the authors used this information to generate semantic networks for each of the VACs examined. For the 'V *across* N' VAC, for instance, the verbs in the center of the network are *go*, *move*, *run*, and *travel*, while verbs like *shout*, *splash*, and *echo* constitute less prototypical verbs in that VAC.

Main Findings

A multifactorial analysis (that is, a statistic that can gauge the impact of more than one factor on a specific outcome at a time; in our example, it measured the potential impact of frequency, prototypicality, and contingency on speakers' associations) revealed that for all of the VACs examined, each factor made an independent contribution to learners' and native speakers' associations alike:

1) the more frequently a particular verb occurred in a specific VAC in the native speaker corpus data, the more likely it was elicited as a response for that VAC in the word association experiment;
2) the more strongly a verb and a VAC were associated with each other as expressed in their DeltaP association scores, the more likely that verb was elicited as a response for that VAC in the word association experiment;
3) the more prototypical a verb was for the VAC as indicated by its position in the semantic networks the authors generated for each VAC, the more likely it was elicited as a response for that VAC in the word association experiment.

Theoretical Implications

Based on the statistical analyses, the authors concluded that advanced L2 learners' knowledge of VACs involves rich associations that are very similar in kind and strength to those of native speakers (Ellis, O'Donnell & Römer, 2014b). The word associations generated in the experiment testified to learners having rich associations for VACs that are tuned by verb frequency, verb prototypicality, and verb–VAC contingency alike – factors that, in combination, interface across syntax and semantics.

Why/how this theory provides an adequate explanation of observable phenomena in SLA

Observation # 1. Exposure to input is necessary for SLA.

Usage based approaches are input–driven, emphasizing the associative learning of constructions from input. 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. Usage is necessary, and it is sufficient for successful L1 acquisition though not for SLA. This is because the initial state for SLA is knowledge of an L1, and the learner's representations, processing routines, and attention to language are tuned and committed to the L1.

Observation # 2. A good deal of SLA happens incidentally.

The majority of language learning is implicit. Implicit tallying is the raw basis of human pattern recognition, categorization, and rational cognition. 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 without us being aware of it (Ellis, 2002). We know, or can be shown to be sensitive in our processing, far too many linguistic regularities for us to have explicitly learned them. Usage–based approaches believe that incidental associative learning provides the rational mechanisms and is sufficient for L1 acquisition from input–analysis and usage, allowing just about every human being to acquire fluency in their native tongue. They do not suffice for SLA because of learned attention.

Observation # 3. Learners come to know more than what they have been exposed to in the input.

The study of implicit human cognition shows us to know far more about the world than we have been exposed to or have been explicitly taught. Prototype effects are one clear and ubiquitous example of this: learners who have never been exposed to the prototype of a category nevertheless classify it faster and more accurately than examples further from the central tendency, and name it with the category label with great facility. The same is true for language, where learners go beyond the input in producing **U–shaped learning**, with novel errors (like *goed* instead of *went*) and other systematicities of stages of interlanguage development in L2 acquisition, for example of negation or question formation. These creations demonstrate that the learners' language system is constantly engaged in making generalizations and finding abstractions of systematicities.

Observation # 4. Learners' output (speech) often follows predictable paths with predictable stages in the acquisition of a given structure.

As in L1 acquisition, SLA is characterized not by complete idiosyncrasy or variability but rather by predictable errors and stages during the course of development: interlanguage is systematic. Usage–based approaches hold 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 of like–type, and that are salient, are likely to be acquired earlier than those that do not have these features (Ellis, 2007).

Observation # 6. Second language learning is variable across linguistic subsystems

The learners' mental lexicon is diverse in its contents, spanning lexical, morphological, syntactic, phonological, pragmatic, and sociolinguistic knowledge. 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, and reliability, along with the various other associative learning factors that affect the emergence of attractors in the problem space. Some aspects of these problem spaces are simpler than others. Second language learning is piecemeal development from a database of exemplars with patterns of regularity emerging dynamically.

Observation # 7. There are limits on the effects of frequency on SLA

This is explicitly addressed above under *Associative aspects of transfer: Learned attention and interference.*

Observation # 8. There are limits on the effect of a learner's first language on SLA

The effect of a learner's L1 is no longer considered the exclusive determinant of SLA as proposed in the Contrastive Analysis Hypothesis. However, its significance in the language learning process is uncontroversial. At every level of language, there is evidence of L1 influence, both negative and positive. The various cross–linguistic phenomena of learned selective attention, overshadowing and blocking, latent inhibition, perceptual learning, interference, and other effects of salience, transfer, and inhibition all filter and color the perception of the L2.

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. Attention to language form is sometimes necessary to allow learners to notice some blocked, overshadowed, or otherwise non–salient aspect of the language form. Reviews of the empirical studies of instruction demonstrate that social recruitment of learners' conscious, 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 always 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 for a large variety of reasons. Explicit instruction can be ill–timed and out of synchrony with development (Pienemann, 1998, this volume); it can be confusing; it can be easily forgotten; it can be dissociated from usage, 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.

Observation # 10. There are limits on the effects of output (learner production) on language acquisition.

There is 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. However, you cannot learn an L2 by output alone.

The Explicit/Implicit Debate

Learning a new symbol, for example the French word for this symbol: ★, *étoile*, initially involves explicit learning: you are consciously aware of the fact that you did not know the French word for 'star' before, and that now you do (Ellis, 1994). Some facts about how to use *étoile* properly you may not know yet, such as its proper pronunciation, its grammatical gender, synonymous forms, words, phrases, and idioms that *étoile* is associated with, etc. Some of these facts you will learn by making a conscious effort, that is, via explicit learning; other facts you will not consciously figure out, but rather learn implicitly. Without you being aware of it, your language system is hard at work, upon each subsequent encounter of *étoile*, to fill in these knowledge gaps and fine–tune the mental representation you have for this construction.

In spite of the fact that many of us go to great lengths to engage in explicit language learning, the bulk of language acquisition is implicit learning from usage. Most knowledge is tacit knowledge; most learning is implicit; the vast majority of our cognitive processing is unconscious. Implicit learning supplies a distributional analysis of the problem space: our language system implicitly figures out how likely a given construction is in particular contexts, how often they instantiate one sense or another, how these senses are in turn associated with different features of the context, and so on. To the extent that these distributional analyses are confirmed time and again through continuous exposure to more input, generalizations and abstractions are formed that are also largely implicit.

Implicit learning would not do the job alone. Some aspects of an L2 are unlearnable—or at best are acquired very slowly—from implicit processes alone. In cases where linguistic form lacks perceptual salience and so goes unnoticed by learners, or where the L2 semantic/pragmatic concepts to be mapped onto the L2 forms are unfamiliar, additional attention is necessary in order for the relevant associations to be learned. In order to counteract the L1 attentional biases to allow implicit estimation procedures to optimize induction, all of the L2 input needs to be made to count (as it does in L1 acquisition), not just the restricted sample typical of the biased intake of L2 acquisition.

Ellis (2005) reviews the instructional, psychological, social, and neurological dynamics of the interface by which explicit knowledge of form–meaning associations impacts upon implicit language learning. ‘The interface is dynamic: It happens transiently during conscious processing, but the influence upon implicit cognition endures thereafter. Explicit memories can also guide the conscious building of novel linguistic utterances through processes of analogy. Patterned practice and declarative pedagogical grammar rules both contribute to the conscious creation of utterances whose subsequent usage promotes implicit learning and proceduralization. Flawed output can prompt focused feedback by way of recasts that present learners with psycholinguistic data ready for explicit analysis’ (Ellis, 2005, p. 305). Once a construction has been represented in this way, its use in subsequent implicit processing can update the statistical tallying of its frequency of usage and probabilities of form–function mapping.

So we believe that learners’ language systematicity emerges from their history of interactions of implicit and explicit language learning, from the statistical abstraction of patterns latent within and across form and function in language usage. The complex adaptive system of interactions within and across form and function is far richer than that emergent from implicit or explicit learning alone (Ellis, in press).

Conclusion

In the terms of Chapter 1, the usage–based approaches touched upon here are too broad and far—ranging to qualify as a theory. Instead, they are a framework for understanding many of the complex agents that underlie language 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 things a theory should explain.

Discussion Questions

YOU DECIDE ON THESE, I BELIEVE

Suggested Further Readings

Ellis, N. C. (Ed.). (1994). *Implicit and explicit learning of languages*. San Diego, CA: Academic Press. An edited collection focusing upon the explicit/implicit debate.

Ellis, N. C. & Larsen–Freeman, D. (Eds). (2009). Language as a complex adaptive system. Special issue. *Language Learning*, *59,* Supplement 1. A special issue gathering experts from various language domains who share the CAS perspective.

Ellis, N. C. & Cadierno, T. (2009). Constructing a second language. *Annual Review of Cognitive Linguistics, 7.* A special issue taking SLA to the cognitive linguists.

Collins, L. & Ellis, N. C. (2009). Input and second language construction learning: frequency, form, and function. Special issue. *Modern Language Journal, 93* (2). SLA has always emphasized input. This special issue presents new ways of analyzing its latent properties.

Rebuschat, P. (Ed.). (2014). *Implicit and explicit learning of language*. Amsterdam: John Benjamins. A reprise summarizing research twenty years on.

Tomasello, M. (2003). *Constructing a language*. Boston, MA: Harvard University Press. A thorough usage–based account of child language.

Trousdale, G., & Hoffmann, T. (Eds.). (2013). *Oxford Handbook of Construction Grammar*. Oxford: Oxford University Press. A range of contemporary usage–based linguistic theories.

Tyler, A. (2012). *Cognitive linguistics and second language learning*. New York and London: Routledge. SLA from a cognitive–linguistic perspective, referencing many usage–based studies and research into pedagogical applications of a usage–based SLA.

References

Anderson, J. R. (1989). A rational analysis of human memory. In H. L. I. Roediger & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 195–210). Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Beckner, C., Blythe, R., Bybee, J., Christiansen, M. H., Croft, W., Ellis, N. C., . . . Schoenemann, T. (2009). Language is a complex adaptive system. Position paper. *Language Learning, 59 Supplement 1*, 1–26.

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

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

Ellis, N. C. (1994). Vocabulary acquisition: The implicit ins and outs of explicit cognitive mediation. In Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 211–282). San Diego, CA: Academic Press.

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

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

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

Ellis, N. C. (2006a). Language acquisition as rational contingency learning. *Applied Linguistics*, *27*, 1–24.

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

Ellis, N. C. (2007). Dynamic Systems and SLA: The wood and the trees. *Bilingualism: Language & Cognition, 10*, 23–25.

Ellis, N. C. (2008). The dynamics of second language emergence: Cycles of language use, language change, and language acquisition. *Modern Language Journal, 41*, 232–249.

Ellis, N. C. (in press). Implicit AND explicit learning of language. In P. Rebuschat (Ed.), *Implicit and explicit learning of language* (Vol.) John Benjamins: Amsterdam.

Ellis, N. C., & Larsen–Freeman, D. (2006). Language emergence: Implications for Applied Linguistics (Introduction to the Special Issue). *Applied Linguistics, 27*, 558–589.

Ellis, N. C., & Larsen–Freeman, D. (Eds.). (2009). *Language as a Complex Adaptive System*. Boston, MA: Wiley–Blackwell.

Ellis, N. C., O'Donnell, M. B., & Römer, U. (2014a). Second language verb-argument constructions are sensitive to form, function, frequency, contingency, and prototypicality.

Ellis, N. C., O’Donnell, M. B., & Römer, U. (2014b). The processing of verb-argument constructions is sensitive to form, function, frequency, contingency, and prototypicality. *Cognitive Linguistics*, *25*(1), 55-98.

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.

Flege, J. (2002). Interactions between the native and second–language phonetic systems. In P. Burmeister, T. Piske & A. Rohde (Eds.), *An integrated view of language development: Papers in honor of Henning Wode* (pp. 217-244). Trier, Germany: Wissenschaftlicher Verlag Trier.

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

Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford: Oxford University 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.

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

Gries, St. Th., & Wulff, S. (2009). Psycholinguistic and corpus–linguistic evidence for L2 constructions. *Annual Review of Cognitive Linguistics*, *7*, 164–187.

Han, Z.–H., & Odlin, T. (Eds.). (2006). *Studies of fossilization in second language acquisition*. Clevedon: Multilingual Matters.

Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. New York: Routledge.

Kasper, G., & Wagner, J. (2011). A conversation–analytic approach to second language acquisition. Alternative approaches to second language acquisition. In D. Atkinson (Ed.), *Alternative approaches to second language acquisition* (pp. 117–142.).

Klein, W. (1998). The contribution of second language acquisition research. *Language Learning, 48*, 527–550.

Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nature Reviews Neuroscience,* 5, 831–843.

Kuhl, P. K. (2007). Is speech–learning gated by the 'social brain'? *Developmental Science, 10*, 110–120.

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

Long, M. H. (1983). Linguistic and conversational adjustments to non–native speakers. *Studies in Second Language Acquisition, 5*, 177–193.

Mackey, A., Abbuhl, R., & Gass, S. (2011). Interactionist approach. In S. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition*. New York: Routledge.

MacWhinney, B. (1991). *The CHILDES project: Tools for analyzing talk*: (1991). xi, 360pp.

MacWhinney, B. (1997). Second language acquisition and the Competition Model. In A. M. B. De Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 113–142). Mahwah, NJ: Lawrence Erlbaum Associates.MacWhinney, B. (2007). The TalkBank Project. Retrieved from http://repository.cmu.edu/cgi/viewcontent.cgi?article=1175&context=psychology.

McEnery, T., & Hardie, A. (2012). *Corpus linguistics*. Cambridge: Cambridge University Press.

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

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

Pienemann, M. (1998). *Language processing and second language development: Processability theory*. Amsterdam: Benjamins.

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

Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.

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

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

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

Weinreich, U. (1953). *Languages in contact*. The Hague: Mouton.

Wulff, S. (2008). *Rethinking idiomaticity: A usage–based approach.* London and New York: Continuum.

Glossary

Construction

Central notion of Construction Grammar. Constructions are pairings of form and meaning or function that span from morphemes to words and abstract syntactic frames.

Contingency

When the presence and/or specific realization of a form A depends on the presence and/or specific realization of another form B, then A is contingent on B. Contingency can vary in strength. Some cues (like lightening -> thunder) are highly predictive and so have a high contingency, some cues are less reliable (summer days -> fine weather).

Corpus

A large and structured collection of transcribed spoken and/or written language data in digital format.

Emergentism

A system is emergent if it is some way more than sum of the properties of the system's parts.

Exemplar

Exemplars are specific examples of a category. The word "house", for instance, is an exemplar of the category NOUN in English.

Explicit learning

Explicit learning is the learning of information in a conscious and often effortful manner.

Implicit learning

Implicit learning is the learning of information in an incidental and unconscious manner.

Learned attention

People learn to attend to the cues that are relevant to a problem-space, and this increases speed of acquisition and automaticity of processing. While selective attention benefits acquisition, it can also lead to distortions of knowledge that are evident when the learner transfers to novel problem-spaces. Learners continue to attend to the old cues, even when these are no longer optimal, and can ignore relevant new cues, especially when they are lacking in salience.

Prototype

A prototype is the most typical member of a category by combining the most representative attributes of that category.

Rational analysis of cognition

Rational analysis is an empirical program that attempts to explain the function and purpose of cognitive processes. Unlike in traditional cognitive science, in which the cognitive system is often treated as an arbitrary assortment of mechanisms with likewise arbitrary limitations, rational analysis views cognition as intricately adapted to its environment and the problems it faces.

U-shaped learning

U-shaped learning denotes one frequent developmental path when new cognitive skills are developed. Imagine a curve shaped like the letter "U" in a graph, with the x-axis depicting time, and the y-axis depicting the learner's level of skill. Learners often start out with seemingly high levels of skill, but then go through a phase in which their proficiency plummets before it rises again. U-shaped learning characterizes the learning of new words, high-level mathematic algorithms, and even the building muscle strength, among many other skills. Early high levels of performance often reflect memorized, unanalyzed responding; middle lower-levels of performance reflect the development of analyzed systematic responding (whereby irregular responses are now over-generalized).