Second language learning of morphology

NICK C. ELLIS (iD)

ABSTRACT

Second language (L2) speakers have especial difficulty learning and processing morphosyntax. I present a usage-based analysis of this phenomenon. Usage-based approaches to language learning hold that we learn constructions (form-function mappings, conventionalized in a speech community) from language usage by means of general cognitive mechanisms (exemplar-based, rational, associative learning). An individual's language system emerges from the conspiracy of these associations.

I take the broad theoretical framework of language as a complex adaptive system and then focus in upon several subcomponents of the ecology including: cognitive linguistics and construction grammar; the psychology of implicit and explicit learning; effects of contingency and salience upon associative learning; the linguistic cycle: Zipf's law, shortening, and grammaticalization; the low learnability of morphology that results from language change; learned attention, blocking, and transfer; form-focused dialogic feedback or instruction in second language acquisition (L2A); the types and tokens of exemplars that lead the learning of particular morphemes from usage and the distributional effects of frequency, reliability, and formulaic contexts.

In terms of fundamental principles of associative learning: Low salience, low contingency, and redundancy all lead to morphological form-function mappings being less well learned. Compounding this, adult L2 acquirers show effects of learned attention and blocking as a result of L1-tuned automatized processing of language. I review a series of experimental studies of learned attention and blocking in L2A. I describe educational interventions targeted upon these phenomena. Form-focused instruction recruits learners' explicit, conscious processing capacities and allows them to notice novel L2 constructions. Once a construction has been represented as a form-function mapping, its use in subsequent implicit processing can update the statistical tallying of its frequency of usage and probabilities of form-function mapping, consolidating it into the system.

CORRESPONDING AUTHOR:

Nick C. Ellis

University of Michigan, US

ncellis@umich.edu

Table 1 Outline structure and hyperlinks.

1 SECOND LANGUAGE LEARNERS HAVE ESPECIAL DIFFICULTY IN LEARNING MORPHOSYNTAX

Second language (L2) speakers have especial difficulty using morphosyntax and grammatical functors. L2 learning of grammatical morphology is slow, piecemeal, and often incomplete. Milestone descriptions of this phenomenon in the SLA literature include.

Schmidt's (1984) case study of Wes, a naturalistic L2 learner, who was described as very fluent, with high levels of strategic competence, but low levels of grammatical accuracy: "using 90% correct in obligatory contexts as the criterion for acquisition, none of the grammatical morphemes counted has changed from unacquired to acquired status over a five year period" (Schmidt, 1984, p. 5). Many adult EFL learners never acquire total control of L2 morphology, even after tens of years of English immersion (e.g., DeKeyser, 2005; Johnson & Newport, 1989).

The European Science Foundation crosslinguistic and longitudinal research project (Perdue, 1993) examined how 40 adult learners picked up the language of their social environment (Dutch, English, French, German or Swedish) by everyday communication. Analysis of the interlanguage of these L2 learners resulted in its being described as the 'Basic Variety'. All learners, independent of source language and target language, developed and used it, with about one-third of them fossilizing at this level in that although they learned more words, they did not further complexify their utterances in respects of morphology or syntax. In this Basic Variety, most lexical items stem from the target language, but they are uninflected. "There is no functional morphology. By far most lexical items correspond to nouns, verbs and adverbs; closed-class items, in particular determiners, subordinating elements, and prepositions, are rare, if present at all... Note that there is no functional inflection whatsoever: no tense, no aspect, no mood, no agreement, no case marking, no gender assignment" (Klein, 1998, pp. 544–545).

Input Processing theory (Van Patten, 2007) considers how learners process linguistic input as they try to comprehend it. The theory includes "The Primacy of Content Words Principle" that learners process content words in the input before they process other linguistic features, along with its corollary "Lexical Preference Principle" that learners will depend on lexical items to extract meaning as opposed to grammatical form.

The longitudinal study of ESL children by Jia and Fuse (2007) demonstrated that the acquisition of a morpheme such as the third-person singular present-tense -s can take five years or more to go from zero to 80% provision in obligatory contexts.

Slabakova's (2009) comparison of L2 learners' abilities in inflectional morphology, syntax, the syntax-semantics interface, the syntax-discourse interface, and the semantics-pragmatics interface led her to propose a "Bottleneck Hypothesis" whereby it is difficulties in acquiring functional morphology that limits the rate of L2 acquisition. This reflects, of the two alternatives considered by White (2003, chapter 4), "morphology-before-syntax" rather than "syntax-before-morphology".

These L2 difficulties with morphsyntax can affect language change: "Languages are 'streamlined' when history leads them to be learned more as second languages than as first ones, which abbreviates some of the more difficult parts of their grammars" (McWhorter, 2004, p. 51). Around the world, English is now spoken more as an L2 than it is an L1: Graddol (2000) estimated there to be 375 million L1 speakers compared to 750 million EFL and 375 million ESL speakers. The preponderance of L2 learning of English changes the nature of its varieties. When Seidlhofer (2004, p. 236) catalogued such changes in English as a Lingua Franca, she observed first and foremost "'dropping' the third person present tense -s (as in "She look very sad")." These large-scale descriptions have now been supported by detailed experimental evidence in support of the hypothesis that iterated imperfect learning leads to language simplification (Atkinson, Smith, & Kirby, 2018; Berdicevskis & Semenuks, 2022).

Grammatical morphemes pervade language usage: each day of experience provides tens to thousands of receptive experiences of functional morphemes, and tens to thousands of contexts requiring their productive use, yet, as these studies show, L2 provision is far from consistent. Especial L2 difficulty in morphosyntax is ubiquitous (albeit to varying degrees) across languages, learners, and usage. It is so robust a finding that it warrants status as a Law of SLA.

2 WHY? HOW SMALL A THOUGHT IT TAKE TO FILL A WHOLE LIFE

Why is it so? This question has filled much of my scholarly life. In what follows, I will describe a variety of studies and their conclusions. As you will see, I have at times found myself focusing upon 'as small a thought' as the English third person present tense *-s*. It serves as 'a canary in coal mine'—a tell-tale indicator that illuminates the flow of the English language and the fragility of morphology in second language acquisition wherever it sings or falls silent.

Some laws follow simple cause-effect rules, where the effect is the outcome of a necessary and sufficient cause, as with a binary switch or parameter-setting. I have not found L2 difficulties in acquiring morphosyntax to be of this type where the regularities are rule-driven: there are no simple mechanisms for top-down governance nor are there plausible innate linguistic universals with well-specified mechanisms that could coordinate such phenomena. Instead, L2 difficulty in morphosyntax emerges from the dynamics of the ecology of language usage (Beckner et al., 2009; Bybee & Hopper, 2001; Ellis, 1998, 2019; Ellis & Larsen Freeman, 2006; MacWhinney & O'Grady, 2015). As complex adaptive systems, languages emerge, evolve, and change over time as they are honed by social discourse. They adapt to their speakers, to their contexts, to natural laws of cognition, learning, social interaction, cooperation, and communication in the ecology of language as a whole (Douglas Fir Group, 2016; Ellis, 2022; Titone & Tiv, 2022).

In what follows here in trying to understand L2 difficulty with morphosyntax, we will engage with several of these domain-specific laws and see how each plays a part. The structure of the argument is shown in Table 1 where the headers serve as hyperlinks. In order to illustrate the big picture, each section is sketched in broad brushstrokes along with references to one or two articles giving further detail. I tend to cite myself because these articles mark milestones in my thinking, I know intimately how the relevant ~10,000 words of the argument go, and each of these papers cites >100 significant scholars who have seriously influenced me. I hope I have properly followed their leads. Ideas also adapt to their speakers, to their contexts, to natural laws of cognition, learning, social interaction, cooperation, and communication in the ecology of scholarship as a whole.

3 COGNITIVE LINGUISTICS AND CONSTRUCTION GRAMMAR

Learning a language involves the learning of its constructions – the symbolic form-meaning mappings that are conventionalized in its speech community. Constructions include morphemes – the smallest pairing of form and meaning in language – as well as words, phrases, and syntactic frames (Goldberg, 2006; Trousdale & Hoffmann, 2013). Simple morphemes such as the English noun plural *–s* morpheme are constructions in the same way as simple words like *word*, idioms like *to break one's word,* and abstract syntactic frames like the Subject-Verb-Object-Object verb-argument construction (which signals that something is being transferred to someone, as realized in sentences as diverse as *she gave him her word, the company sent the complainant a lengthy denial*, etc.). Including abstract syntactic frames admits that not all constructions carry meaning in the traditional sense, but rather serve a functional or meaningful purpose, as with the passive construction which encourages a shift of attentional focus from the agent of the action to the patient undergoing the action (compare the passive *A cake was baked for the birthday-child* with its active counterpart *They baked the birthday-child a cake*). Ellis, Römer, and O'Donnell (2016) present a monograph on the learning, processing, and lexical-syntactic-semantic interactions of verb-argument constructions in L1 and L2 speakers and large corpora of usage.

Constructions are stored simultaneously in multiple forms that differ in their level of complexity and abstraction. For instance, the word *word* and the plural *–s* morpheme are both simple constructions that are also stored as constituent parts of the more complex construction *words* [*word* + plural *-s*] as well as the completely schematic [Noun + plural *-s*]. While the vast majority of base words are free to stand on their own, as free morphemes, the smaller number of bound morphemes like the English third person present tense *-s* and the past-tense *-ed* can only appear attached to base words. Different levels of constructional abstraction/schematization are similarly evident as we go from the fully lexicalized formula

F-word, through the strongly associated collocation *four-letter word*, the partially schematized slot-and-frame greeting pattern [*Good* + (time of day)] which renders lexicalized phrases like *Good afternoon* and *Good evening*, up to the completely schematic [Adjective + Noun Phrase] construction.

This wide definition of constructions blurs the traditional division between lexicon and grammar, or what generative approaches labeled words and rules: from a construction grammar perspective, a sentence is not the product of applying a rule that strings several words into a particular order, but the product of combining a number of constructions – some simple, some complex, some lexically specific, some abstract – in a particular way. A sentence like '*What did the company send the complainant?*', for instance, potentially combines the following constructions:

- *complain (v)* and *-ant (suffix [= one who (v)erbs])* morphological constructions
- *company, complainant, send, what, did, the* lexical constructions
- *what did, the company* collocational constructions
- VP, NP constructions
- Subject–Verb–Object–Object construction
- Subject–Auxiliary inversion question construction

Cognitive linguistics investigates the ways in which constructions overlap in making meaning (= conspire in a speaker's memories of meaningful language usage) and the indivisible interplay of language function, structure, learning and usage. Robinson and Ellis (2008) provided an early book-length edited introduction to Cognitive Linguistics and SLA; for more recent collections see Dabrowska and Divjak (2015) and Trousdale and Hoffmann (2013). An adult's knowledge of their language(s) is a huge collection of constructions (the 'constructicon'), which vary in terms of complexity and abstraction. Constructions have emergent properties that specify if and how they can combine with other constructions; these properties are mostly semantically and/or functionally motivated such that constructions can only be combined if their meanings/ functions are compatible, or at least can temporarily attain compatibility in a specific context or discourse situation (Goldberg, 2006). Constructional compatibility is crucially solidified by the frequency with which they are used (and therefore, heard) together: the more often they co-occur, the more entrenched that particular arrangement becomes. Likewise, L2 learners will acquire constructions first in the contexts of the constructions that they most often co-occur with in the input before they gradually expand the repertoire of combinations to less frequent combinations and even acceptable novel combinations (Ellis, 2002; Ellis et al., 2016). The system emerges from usage and its rich structure rationally reflects language usage. This is why language learning can be profitably understood as statistical learning, aka associative learning, or rational contingency learning (Ellis, 2006b).

4 LANGUAGE LEARNING INVOLVES GENERIC COGNITIVE AND ASSOCIATIVE MECHANISMS

Usage-based theories hold that domain-general cognitive mechanisms drive the learning of linguistic constructions and the emergence of generalizations (Ellis & Wulff, 2014). Human cognition rests upon separable complementary systems for implicit and explicit learning (for reviews see Ellis, 2017b; Ellis, 1994b; Rebuschat, 2015).

4.1 EXPLICIT LEARNING

Explicit learning is essential in consolidating new explicit memories, particularly episodic or declarative memories that allow us to learn a new vocabulary form and its referent or to recognize new constructions and the objects or events playing out in concurrent perceptual/ imaginal world and to bind their features cross-modally (think of the conscious processes involved in your initial fast-mapping of the concept of COVID-19, images of the spikes on the coronavirus, the spoken word-form as you heard the morning news, …). Ellis (2019, p. 48) analyzes the factors that engender Representation Quality as a result of Embodied, Enacted, Embedded, and Extended (4E) Cognition. Explicit learning requires attention, Schmidt's (1990) "noticing", conscious processing of multimodal representations in working memory, neural explicit memory systems involving the hippocampus – "explicit cognitive mediation" (Ellis, 1994a, 2001).

4.2 IMPLICIT LEARNING

Contrast implicit learning which is essential in tuning our knowledge to attain the competence, fluency, prediction, and idiomaticity of language expertise. Implicit learning occurs in various of our perceptual and motor systems for language – "the implicit ins and outs" (Ellis, 1994a). Ellis (2002) reviewed the evidence of frequency effects in the learning and processing of all levels of language representation: phonology and phonotactics, reading, spelling, lexis, morphosyntax, formulaic language, language comprehension, grammaticality, sentence production, and syntax. Given that we never consciously count our use of different linguistic constructions (I've said "a" *n* million times, "aardvark" thirty seven times, "abandon all hope" three hundred and seven times, …), these effects of usage frequency upon language learning must reflect implicit (i.e., unconscious) learning. Implicit learning rationally tunes us to the likelihood of perceiving things in the world and the things that likely co-occur with them. It gives us rational cognition (Ellis, 2006b) – rational in the sense that it allows us to make predictions about what to expect next ("once upon a …", "they all lived happily ever …"). Usage-based linguistics shows that language cognition rests on thousands of hours of implicit learning from usage from which emerge language-relevant representations and their associations.

4.3 EXPLICIT AND IMPLICIT LEARNING

Implicit and explicit knowledge are dissociable but cooperative. Both are necessary. Neither is sufficient. Without implicit learning there is no proficiency nor integration of a construction into the language system (Krashen, 1985). Without explicit noticing there is no chance (or at least significantly reduced chance) of further implicit learning (Schmidt, 1990). Ellis (2005, 2015) reviews various psychological and neurobiological processes 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 consolidation of a construction and its subsequent repeated usage/practice/processing/prediction results in better learning outcomes in terms of accuracy/entrenchment/automatization/fluency/breadth/depth/richness/precision/ idiomaticity and nativelike selection in collocation and phraseology/proficiency/pragmatic competence.[1]

5 THE ASSOCIATIVE LEARNING OF LINGUISTIC CONSTRUCTIONS

Constructions are the symbolic form-meaning mappings that are conventionalized in a speech community. Learning constructions is the learning of these form-meaning associations. A century of research into associative learning has elucidated a number of laws that are as relevant to language learning in humans as they are to classical conditioning and reinforcement learning in animals.

5.1 ASSOCIATIVE LEARNING DEPENDS UPON CONTINGENCY

Learning associations between cues (forms) and outcomes (meanings) depends upon the contingency of the relationship. In classical conditioning it is the reliability of the bell as a predictor of food that determines the ease of acquisition of this association (Rescorla, 1968). In language learning it is the reliability of the form as a predictor of an interpretation that determines its acquisition and processing (Ellis, 2006c; Gries & Ellis, 2015; MacWhinney, 1987). Psychological investigations into human sensitivity to the contingency between cues and outcomes (Shanks, 1995) demonstrates that when given sufficient exposure to a relationship, people's judgments match the contingency specified by ΔP (the one-way dependency statistic, Allan, 1980) which measures the directional association between a cue and an outcome, as illustrated in Table 2.

	OUTCOME	NO OUTCOME
Cue	*a*	*b*
No cue	*c*	*d*

Table 2 A contingency table showing the four possible combinations of events showing the presence or absence of a target Cue and an Outcome.

1   These are some of the various cause and outcome keywords that are common in current theories of language learning.

a, b, c, d represent frequencies, so, for example, *a* is the frequency of conjunctions of the cue and the outcome, and c is the number of times the outcome occurred without the cue.

Δ*P* is the probability of the outcome given the cue *P(O|C)* minus the probability of the outcome in the absence of the cue *P(O|¬C)*, calculated using this formula:

$$\Delta P = P(O|C) - P(O|\neg C) = \frac{a}{a+b} - \frac{c}{c+d}$$

When the outcome is just as likely when the cue is present as when it is not, there is no covariation between the two events and Δ*P* = 0. Δ*P* approaches 1.0 as the presence of the cue increases the likelihood of the outcome. A learnable cue is one where when the cue is there, the outcome is there, and when the cue is not there, neither is the outcome, i.e., where *a* and *d* are large and *b* and *c* are small.

There are rarely 1:1 mappings between constructional forms and their interpretations. The less reliably a form is associated with a function or interpretation, the more difficult learning becomes (Ellis, 2006b). Cues with multiple interpretations are ambiguous and so hard to resolve; cue-outcome associations of high contingency are reliable and readily processed. Consider how, in the learning of the category of birds, while eyes and wings are equally frequently experienced features in the exemplars, it is wings which are distinctive in differentiating birds from other animals. Wings are important features to learning the category of birds because they are reliably associated with class membership while being absent from outsiders. Raw frequency of occurrence (*a, b, c, d* in Table 1 when considered independently) is therefore less informative than the contingency between cue and interpretation (*a, b, c, d* when considered in interaction as Δ*P*). But of course, one cannot assess contingency without first tallying the raw frequencies.

Low contingency of association will prove to be central to the issue of why L2 learners have especial difficulty in learning morphosyntax: as I explain in section 6.4, it comes as a natural consequence of language change.

5.2 ASSOCIATIVE LEARNING DEPENDS UPON SALIENCE

Learnability also depends on salience: less salient cues are less readily learned than highly salient ones (Ellis, 2017d; Rescorla & Wagner, 1972). Salience refers to the property of a stimulus to stand out from the rest. Salient items or features are more likely to be perceived, to be attended to, and more likely to enter into subsequent cognitive processing and learning. Salience can be independently determined by physics and the environment, and by our knowledge of the world.

(1) The physical world, our embodiment, and our sensory systems come together to cause certain sensations to be more intense (louder, brighter, heavier, etc.) than others.

(2) As we experience the world, we learn from it, and our resultant knowledge values some associations higher than others. These associations can make a stimulus cue "dear" (Hi Kim, Gabriel, Aspen, Tanner!). A loved one stands out from the crowd, as does a stimulus with weighty associations ($5.00 vs. $0.05, however similar the amount of pixels, characters, or ink in their sensation), or one which matches a motivational state (a meal when hungry but not when full). The units of perception are influenced by prior association (James, 1890, p. 82). Psychological salience is experience-dependent: hotdog, sushi, and 寿司 mean different things to people of different cultural and linguistic experience. This is why, contra sensation, the units of perception cannot simply be measured in physical terms. They are subjective. Hence Miller's definition of the units of perception and short-term memory as "chunks": "We are dealing here with a process of organizing or grouping the input into familiar units or chunks, and a great deal of learning has gone into the formation of these familiar units" (Miller, 1956, p. 91).

Rescorla and Wagner (1972) presented a formal model of learning which expresses the capacity of any cue (Conditioned Stimulus, CS, for example a bell in Pavlovian conditioning) to become associated with an outcome (Unconditioned Stimulus, US, for example food in Pavlovian conditioning) on any given experience of their pairing. This formula summarized over

eighty years of research in associative learning; it elegantly encapsulates the three factors of physical salience, psychological salience, and surprisal. The role of US surprise and of CS and US salience in the process of conditioning can be summarized as follows:

$$dV = ab(L - V)$$

The associative strength of the US to the CS is referred to by the letter *V* and the change in this strength which occurs on each trial of conditioning is called *dV*. On the right-hand side, *a* is the salience of the CS, *b* is the salience of the US, and L is the amount of processing given to a completely unpredicted, surprising, US. Thus both the salience of the cue (a) and the psychological importance of the outcome (b) are essential factors in any associative learning. As for (*L – V*), the more a CS is associated with a US, the less additional association the US can induce. As Beckett (1954) put it: "habit is a great deadener". Alternatively, with novel associations where *V* is close to zero, there is much surprisal, and consequently much learning: first impressions, first love, first time...

This is arguably the most influential formula in the history of learning theory. Physical salience, psychological salience, and surprisal interactively affect what we learn from our experiences of the world. I recognize that some of this content concerning classical conditioning and associative learning theory might seem foreign to some readers from a pure linguistics background, but to understand the A in SLA, there is value to be had in open-minded exploration of the parallels between cues and outcomes (the CSs and USs of learning theory) and form-meaning relations (constructions in theories of cognitive linguistics and construction grammar; signifiers and signifieds in Saussurian linguistics).

Low salience will prove to be central to the issue of why L2 learners have especial difficulty in learning morphosyntax: as I explain in sections 6.2 and 6.3, it comes as a natural consequence of language change.

5.3 THE ROUTINE DETERMINANTS OF CONSTRUCTION LEARNING

Contingency, salience, and surprisal affect the learning of all types of association in all cognitive species. Linguistic constructions have particular content that modulates these factors, as has been extensively demonstrated, for example, in the learning of words. An early study of the psycholinguistic factors involved in foreign-language vocabulary learning by Ellis and Beaton (1993) showed contributions of phonological regularity, semantic content, word class, imageability of concept, word frequency, meaningfulness, orthographic factors, word length, and familiarity of grapheme-to-phoneme mappings. Recent large-scale studies (Berger, Crossley, & Skalicky, 2019; Crossley, Subtirelu, & Salsbury, 2013; Peters, 2019) support the relevance of these factors and add others such as contextual diversity and contextual distinctiveness, age of acquisition, and neighborhood density. Phonological and orthographic factors and word-length contribute to the salience of the word form; while meaningfulness, concreteness, imageability (Paivio, 1986), grounding in perceptual symbol systems (Barsalou, 1999), and contextual factors contribute to the salience of the function. These play out in each and every experience of their pairing.

Usage provides many cumulative experiences of related constructions, and these experiences conspire in the emergence of the language system. The frequency and variability of these experiences determines how well they are learned and entrenched as constructions and concepts (Ellis, 2002). There are well-established laws relating the effects of frequency upon acquisition ('the power law of practice', Anderson, 1982), as well as the effects of time interval upon retention ('the forgetting function', Ebbinghaus, 1885). Likewise, we understand how the balance of type- and token-frequency and the proportion of 'friends: enemies'[2] affects the balance of item-learning and generalization: high token-frequency leads to item knowledge (which allows the survival of high-frequency 'irregular' items that go against the flow); high-type frequency leads to generalization and productivity (Bod, Hay, & Jannedy, 2003; Bybee & Hopper, 2001).

---

2   Friends share mappings from stem to tense (e.g., walk-walked, talk-talked) and pull together; enemies use different mappings (e.g., throw-threw, swim-swam) and pull apart (Marchman, Wulfeck, & Weismer, 1999).

5.4 CONNECTIONIST LEANING AND THE FREQUENCY BY REGULARITY INTERACTION

Many of these regularities have been modelled using connectionist models in which generalizations emerge from the conspiracy of usage events (Christiansen & Chater, 2001; Ellis, 2003; MacWhinney & O'Grady, 2015; Seidenberg & Plaut, 2014; Shirai, 2018). For example, Ellis and Schmidt (1998) investigated adult learning of morphosyntax in a novel language where frequency and regularity were factorially combined. They were particularly interested in frequency by regularity interactions in the learning and processing of morphosyntax – effects which had become a crux issue in the debate between single-system connectionist accounts and dual-mechanism hybrid accounts which posit that regular inflections are computed by an affixation rule in a neurally based symbol-manipulating syntactic system, while irregular verbs are retrieved from an associative memory (Pinker & Prince, 1994). Ellis and Schmidt demonstrated frequency effects for both regular and irregular forms early in the acquisition process. However, as learning progressed, the frequency effect on regular items diminished whereas it remained for irregular items. Performance of a simple connectionist system, when trained on the same materials, showed a close correspondence to the human acquisition data. Ellis and Schmidt showed that the regularity by frequency interaction is a natural consequence of the power law of practice, and thus is entirely consistent with associative learning processes.

5.5 THE ASSOCIATIVE-COGNITIVE CREED

In sum, there are well-attested laws of associative learning and cognition which have reliable and robust consequences for the learning of lexical and morphosyntactic constructions. Ellis (2006a) gathered these ideas into a framework, the 'Associative-Cognitive **CREED**', which promotes that language acquisition is **C**onstruction-based, **R**ational, **E**xemplar-driven, **E**mergent, and **D**ialectic. Subsequent research has demonstrated the relevance of this approach to the learning and processing of vocabulary, morphosyntax, multi-word sequences and formulaic language, and verb-argument and other syntactic constructions in L1 and L2 (Ellis & Ogden, 2017; Ellis et al., 2016; Robinson & Ellis, 2008; Wulff & Ellis, 2018). Just as these phenomena affect everyday psycholinguistic processing, so they affect the shape of language too (Hundt, Pfenninger, & Mollin, 2017): "Language and cognition are like the shore-line and the sea" (Ellis, 2019, p. 39). We turn next to these matters of language change.

6 THE LINGUISTIC CYCLE: LANGUAGE CHANGE AS A FUNCTION OF USAGE

Languages change over time as their native speakers, sharing similar conventionalized codes, do their well-practised dialogic dance; practice has made perfect: They know the steps and transitions inside out, and they perform their discourse automatically, concentrating on meanings, while the associated forms play out naturally, unhindered by conscious control. The study of language change shows that, whatever the language, whoever its current speakers, languages change in a variety of law-like ways that are as regular as the laws of other ecologies and complex adaptive systems.

6.1 HIGH FREQUENCY OF USE LEADS TO CHUNKING AND FORMULAIC PATTERNS

The basic principles of skill-acquisition that apply to all kinds of motor activities and productions (like playing a musical instrument, a sport, cooking, or learning a new song) are that, through repetition, sequences of units that were previously independent come to be processed as a single unit or chunk (Ellis, 1996, 2017a). "Words used together fuse together" (Bybee, 2003) (after Hebb's (1949) research often summarized by the phrase "Cells that fire together, wire together"). These processes result in the memorized formulaic sequences, collocations, idioms, and phraseology that pervade language (Ellis, 2008b, 2012; Ellis & Ogden, 2017; Simpson-Vlach & Ellis, 2010). Repeated practice also leads to their proceduralization, automatization and fluency of production (Anderson, 1992; DeKeyser, 2001; Segalowitz, 2010).

6.2 HIGH FREQUENCY OF USE LEADS TO SHORTENING — ZIPF'S LAW

The more frequently they use a form or chunk, the more speakers abbreviate it: this is a law-like relationship across languages (Ellis, 2017c; Zipf, 1935). Zipf (1949) summarized this in the principle of least effort – speakers want to minimize articulatory effort and hence encourage brevity and phonological reduction. They tend to choose the most frequent words, and the more they use them, automatization of production causes their shortening. Highly frequently used words become shorter and phonologically eroded.

6.3 ZIPF'S LAW PARTICULARLY IMPACTS GRAMMATICAL FUNCTORS

Grammatical functors are the most frequent words of a language. In informal and rapid speech, this tendency to give short shrift to function words and bound morphemes, exploiting their frequency and predictability, deforms their phonetic structure and blurs the boundaries between these morphemes and the words that surround them. Of the strong syllables in a corpus examined by Cutler and Carter (1987), 86% occurred in open class words and only 14% in closed-class words. The pattern was reversed for weak syllables, with 72% in closed-class words and 28% in open-class words. Thus grammatical function words and bound inflections tend to be short and low in stress, even in speech that is produced slowly and deliberately (Bates & Goodman, 1997) or in speech directed to children (Goodman, Nusbaum, Lee, & Broihier, 1990), with the result that these cues are difficult to perceive.

6.4 HIGH FREQUENCY OF USE ALSO LEADS TO HOMOPHONY AND AMBIGUITY

Ambiguity is a loss of communicative capacity that arises if individual sounds are linked to more than one meaning. The most frequent words of the language tend also to be the most ambiguous ones (Polikarpov, 2006). This is a natural result of shortening, where initially-distinct longer words erode down to crash into the same sound-form in a tightly-packed space of alternatives, resulting in various different meanings being hung onto the same short sound. Many of the most frequently used words of English are ambiguous in their homophony and polysemy (e.g., *to, too, two; there, their, they're; I, eye, aye*). Morphemes too: the 3rd person present *–s* morpheme is a remnant of a much fuller verb inflectional system in Old English where all verb forms (all numbers and persons) took inflectional endings; we see remnants of this up through Early Modern English in expressions such as "what sayest thou?". All these inflectional endings have been lost except the 3rd person *–s* (see, e.g., Williams, 1986, p. 262). As noted in section 1, many varieties of English (including African American English and many world varieties of English) have already lost this *–s* as well (Seidlhofer, 2004).

This pattern generalizes across languages: the greater the number of monosyllabic words in the lexicon of a language, the greater the degree of homophony (Ke, 2006).

6.5 GRAMMATICALIZATION

Words which frequently co-occur together come to be cognitively processed as single chunks and then evolve to be individual words. Frequent usage typically encourages four processes of change: *Desemanticization* – broadening or abstraction of meaning or content; *Extension* – use in new contexts; *Decategorialization* – loss of morphosyntactic properties; and *Erosion* – loss of phonetic substance (Hopper & Traugott, 2003). Hopper and Traugott's famous pattern for the cline of grammaticalization illustrates the various stages of the form:

content word → grammatical word → clitic → inflectional affix.

This is part of the very predictable "Linguistic Cycle" (Givón, 1971; Hodge, 1970) whereby a notion that is first expressed in

discourse → syntax → morphology → morphophonemics → zero.

The extensive scholarship on language change and grammaticalization (e.g., Bybee, 2010; Hilpert, 2021; Hundt et al., 2017; Traugott & Trousdale, 2013) deserves much more than the cursory treatment that is possible here, but I hope that for our present purposes of placing L2 difficulties with morphosyntax in ecological perspective, it suffices to recognize the universality of the Linguistic Cycle: "The mechanisms and principles involved in grammaticalization conform to a complex process of coding and organization of language which is universally

applicable to describe the evolution of grammatical forms" (Wischer & Diewald, 2002). These universal principles emerge from dynamic processes of cognition and diachrony: "For a theory of grammaticalization, it is both unjustified and impractical to maintain a distinction between synchrony and diachrony," (Heine, Claudi, & Hünnemeyer, 1991, chapter 9), and of usage, discourse, and social-interaction: "Grammar is not absolutely formulated and abstractly represented, but always anchored in the specific form of an utterance… Its forms are not fixed templates, but are negotiable in face-to-face interaction in ways that reflect individual speakers' past experience of these forms, and their assessment of the present context," (Hopper, 1998, p. 142). "Grammar is always emergent and never present" (Hopper, 1998, p. 148).

6.6 THE LINGUISTIC CYCLE AS A PANCHRONIC PRINCIPLE

From these dynamic processes over all diachronic timescales and all synchronic states, there emerge what Saussure (1916, p. 135) termed *Panchronic* principles, generalizations of language that exist independently of time, of a given language, or of any concrete linguistic facts. There are indeed laws of language change: Like the laws of the natural sciences, they are emergent too.

7 LANGUAGE MEETS LEARNING IN CONSPIRING TO MAKE MORPHOLOGY ESPECIALLY DIFFICULT TO LEARN

Morphosyntax pervades language and so these frequency-driven principles of language change (section 6) particularly impact morphological constructions, reducing their learnability following the routine determinants (section 5). Let me fill in a little further detail.

7.1 MORPHOLOGY – LOW CONTINGENCY

One factor determining the learning of construction form is contingency of association (section 5.1). Cue-outcome reliability can be reduced in two directions: forms can have multiple interpretations (polysemy and homophony) and interpretations can be realized by more than once form (synonymy). The same usage-phenomenon whereby frequently used words become shorter drives grammatical functors towards homophony since different functions associated with forms that were originally distinct eventually merge into the same shortened form. An example is the [s] suffix in English: in modern English, it has come to encode a plural form (*kids*), it indicates possession (*Tanner's guitar*), it marks third person singular present (*Aspen thinks*); this in addition to the many English words that just happen to end in -s. The [s] form is abundantly frequent in learners' input, but not reliably associated with any/just one of these meanings/functions (increasing *b* in Table 2). Conversely, the plural, possessive, and third person singular constructions are all realized by more than one form: they are all variably expressed by the allomorphs [s], [z], and [ɨz]. 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 (*c* in Table 2). Thus, the low cue-interpretation contingency makes plural -s difficult to learn.

This fact, that many high frequency grammatical constructions are highly ambiguous in their interpretations, poses a challenge to language learners (DeKeyser, 2005; Ellis, 2008a; Goldschneider & DeKeyser, 2001).

7.2 MORPHOLOGY – LOW SALIENCE

Another factor determining the learning of construction form is salience (section 5.2). In his landmark study of first language acquisition, Brown (1973) breaks down the measurement of perceptual salience, or "clarity of acoustical marking" (p. 343), into "such variables as amount of phonetic substance, stress level, usual serial position in a sentence, and so on" (p. 463). Prepositional phrases, temporal adverbs, and lexical linguistic cues are salient and stressed in the speech stream. Verb inflections are usually not. Many grammatical form-function relationships in English, like grammatical particles and inflections such as the third person singular -s, are of low salience in the language stream.

Because grammatical function words and bound inflections are short and unstressed, they are difficult to perceive from the input. When grammatical function words (*by, for, no, you*, etc.) are clipped out of connected speech and presented in isolation at background noise levels where

their open-class equivalents (*buy, four, know, ewe*, etc.) are perceived 90 to 100% correctly, adult native speakers can recognize them only 40% to 50% of the time (Herron & Bates, 1997). Clitics, accent-less words or particles that depend accentually on an adjacent accented word and form a prosodic unit together with it, are the extreme examples of this: the /s/ of 'he's', /l/ of 'I'll' and /v/ of 'I've' can never be pronounced in isolation.

These factors make grammatical functors extremely difficult to perceive from bottom-up auditory evidence alone. Fluent language processors can perceive these elements in continuous speech because their language knowledge provides top-down support. But this is exactly the knowledge that L2 learners lack: they haven't had sufficient experience to develop a sufficiently schematized knowledge system (constructicon) that would offer the same levels of top-down support as in fluent L1 processing. Thus the low psychophysical salience of grammatical functors contributes to L2 learners' difficulty in learning them (Ellis, 2017d).

7.3 REDUNDANCY

These effects of low salience and low contingency are compounded by redundancy. Grammatical morphemes often appear in redundant contexts where their interpretation is not essential for correct interpretation of the sentence (Schmidt, 2001; Terrell, 1991; Van Patten, 1996). Tense markers often appear in contexts where other cues have already established the temporal reference (e.g., "*yesterday* Gabriel walk*ed...*", "*Last winter, there was so much snow,* Kim and the kids sculpt*ed* an ice-horse"), plural markers are accompanied by quantifiers or numerals (e.g, "*3 kids...*"), etc. Hence their neglect does not result in communicative breakdown, they carry little psychological importance of the outcome (term *b* in the Rescorla-Wagner equation), and the Basic Variety "satisfices" (Simon, 1957) for everyday communicative purposes.

In terms of linguistic constructions and the difficulty distance separating a novice L2/FL learner from an L1 speaker, it is in morphological constructions where the gap is largest. The high frequency of L1 morphology entails maximal L1 automaticity; the lack of salience and contingency of morphology makes it maximally difficult for an L2 learner to perceive and analyze in terms of its associated meanings and functions.

7.4 ENOUGH, THOUGH THERE IS MORE

The dynamics described in this section whereby universal laws of learning interact with panchronic universals of linguistic construction[3] are sufficient to understand why morphosyntax is more difficult to acquire than open class words. These factors apply to any learner. But there is yet more that particularly prejudices the learning of morphosyntax in second language learners. The way we perceive the world depends upon our habits of attention: new cognition builds upon cumulative prior cognition. So, L2 learners have L1-biased learned attention and L1-tuned automatized processing of language that contribute to the blocking of morphosyntax being implicitly learned by adult naturalistic L2 learners whose attentional focus is on communication. This is the subject of section 8.

8 LEARNED ATTENTION, BLOCKING, TRANSFER

L2A is subject to attentional biases which result from L2 learners' knowledge of a prior language and the routine ways they have come to attend and process it. Transfer effects are evidenced in all aspects (phonological, morphological, lexical, collocational, syntactic, and pragmatic) of language learning (Odlin, 2022).

8.1 BLOCKING

Ellis (2006b, 2006c) presents a theoretical analysis which attributes L2 difficulties in acquiring inflectional morphology to an effect of learned attention known as "blocking" (Kamin, 1968; Kruschke & Blair, 2000; Mackintosh, 1975). Blocking is an associative learning phenomenon, occurring in animals and humans alike, that shifts learners' attention to input as a result of prior experience (Rescorla & Wagner, 1972; Shanks, 2010).

---

3   I should emphasize that these "laws" are not doing anything in themselves – they are simply summary observations of phenomena that reliably emerge in the complex adaptive system of usage.

Knowing that a particular cue is associated with a particular outcome (such as past temporality) makes it harder to learn that another cue (e.g., L2 English *-ed*), subsequently paired with that same outcome, is also a good predictor of it. The prior association "blocks" further associations. ALL languages have lexical and phrasal means of expressing temporality (the equivalents of lexical adverbs (e.g., *now*, *next*, *yesterday*, *tomorrow*), prepositional phrases (*in the morning*, *in the future*), and calendric reference (*May 12*, *Monday*), etc.). So ANYONE with knowledge of ANY first language is aware that that there are reliable and frequently used lexical cues to temporal reference (words like German *gestern*, French *hier*, Spanish *ayer*, English *yesterday*). Such are cues to look out for in an L2 because of their frequency, their reliability of interpretation, their simplicity, and their salience. In face-to-face "here and now" communication, there are also available other gestural and pragmatic means of expressing temporality (e.g., serialization: presenting events in their order of occurrence along with appropriate pointing or counting off of fingers). Learned attention theory holds that, once known, such cues block the acquisition of less salient and less reliable verb tense morphology from analysis of redundant utterances such as *Yesterday I walked*.

8.2 EXPERIMENTAL DEMONSTRATIONS OF BLOCKING

A series of experimental investigations involving the learning of a small number of Latin expressions and their English translations have explored the basic mechanisms of learned attention in SLA. Ellis and Sagarra (2011) illustrates the core design. There were three groups: Adverb Pretraining, Verb Pretraining, and Control. In Phase 1, Adverb Pretraining participants learned two adverbs and their temporal reference – *hodie* today and *heri* yesterday; Verb Pretraining participants learned verbs (shown in either first, second, or third person) and their temporal reference – e.g., *cogito* present or *cogitavisti* past; the Control group had no such pretraining. In Phase 2, all participants were shown sentences which appropriately combined an adverb and a verb (e.g., *heri cogitavi, hodie cogitas, cras cogitabis*) and learned whether these sentences referred to the past, the present, or the future. In Phase 3, the Reception test, all combinations of adverb and verb tense marking were presented individually and participants were asked to judge whether each sentence referred to the past, present, or future. The logic of the design was that in Phase 2 every utterance contained two temporal references – an adverb and a verb inflection. If participants paid equal attention to these two cues, then in Phase 3 their judgments should be equally affected by them. If, however, they paid more attention to adverb (/verb) cues, then their judgments would be swayed towards them in Phase 3.

The results showed that the three groups reacted to the cues in very different ways – the Adverb pretraining group followed the adverb cue, the Verb pretraining group tended to follow the verb cue, and the Control group lay in between. For example, multiple regression analyses, one for each group, where the dependent variable was the group mean temporal interpretation for each of the Phase 3 strings and the independent variables were the information conveyed by the adverbial and verbal inflection cues showed in standardized ß coefficients, Adverb Group Time = 0.99Adverb – 0.01Verb; Control Group Time = 0.93Adverb + 0.17Verb; Verb Group Time = 0.76Adverb + 0.60Verb.

This experiment demonstrated how short-term instructional manipulations could affect attention to language.

8.3 BLOCKING AND LANGUAGE-SPECIFIC TRANSFER EFFECTS

Ellis and Sagarra (2010) Experiment 2 and Ellis and Sagarra (2011) Experiments 2 and 3 also illustrated long-term language transfer effects whereby the nature of learners' first language (+/– verb tense morphology) biased the acquisition of morphological vs. lexical cues to temporal reference in the same subset of Latin. First language speakers of Chinese (no tense morphology) were less able than first language speakers of Spanish or Russian (rich morphology) to acquire inflectional cues from the same language experience where adverbial and verbal cues were equally available, with learned attention to tense morphology being in standardized ß coefficients: Chinese (–0.02) < English (0.17) < Russian (0.22) < Spanish (0.41) (Ellis & Sagarra, 2011, Table 4). These findings demonstrate long-term attention to language, a processing bias affecting subsequent cue learning that comes from a lifetime of prior L1 usage.

Ellis, Hafeez, et al. (2014) replicated Ellis & Sagarra (2010) in demonstrating short-term learned attention in the acquisition of temporal reference in L2 Latin in EFL learners, extending the investigation using eye-tracking indicators to determine the extent to which these biases are overt or covert. Eye-tracking measures showed that prior experience of particular cue dimensions affected what participants overtly focused upon during subsequent language processing, and how, in turn, this overt study resulted in covert attentional biases in comprehension and in productive knowledge. These learned attention effects have elements of both positive and negative transfer. Prior use of adverbial cues causes participants to pay more attention to adverbs – positive effects of entrenchment of the practiced cue. Additionally, increased sensitivity to adverb cues is accompanied by a reduced sensitivity to morphological cues – blocking. A meta-analysis of the combined results of Ellis and Sagarra (2010, 2011) demonstrated that the average effect size of entrenchment was large (+1.23) and that of blocking was moderate (–0.52).

While these learned attention demonstrations concern the first hour of learning Latin, Sagarra and Ellis (2013) show the results of blocking over years of learning in intermediate and advanced learners of Spanish. 120 English (poor morphology) and Romanian (rich morphology) learners of Spanish (rich morphology) and 98 English, Romanian and Spanish monolinguals read sentences in L2 Spanish (or their L1 for the monolinguals) containing adverb-verb or verb-adverb congruencies/incongruencies. Eye-tracking data revealed significant effects for sensitivity (all participants were sensitive to tense incongruencies), cue location in the sentence (participants spent more time at their preferred cue), and L1 experience (morphologically rich L1 learners and monolinguals looked longer at verbs than morphologically poor L1 learners and monolinguals).

8.4 LEARNED ATTENTION AND TRANSFER IN L2 MORPHOLOGY

Transfer phenomena pervade SLA (Jarvis & Pavlenko, 2008; Lado, 1957; MacWhinney, 1997; Odlin, 2022). As a result, second language learning is rarely entirely native-like, even if the learner is surrounded by ambient input. Since everything is filtered through the attentional lens of the L1, not all of the relevant input is in fact taken advantage of (hence Corder's (1967) classic distinction between input and intake).

One of the most extensive investigations of transfer in L2 morphology by Murakami and Alexopoulou (2016) investigated the L2 acquisition order of six English grammatical morphemes by learners from seven L1 groups across five proficiency levels who provided approximately 10,000 written exam scripts from the Cambridge Learner Corpus. The study established clear L1 influence on the absolute accuracy of morphemes and their acquisition order, and showed that L1 influence is morpheme specific, with morphemes encoding language-specific concepts most vulnerable to L1 influence.

It is important to emphasize that the limitations of L2 learning described here do not license the conclusion that L2 learning is qualitatively different from L1 learning – second language learners employ the same statistical learning mechanisms that they employed when they acquired their first language. First language learners have learned to attend to their language environment in one particular way. L2 learners are tasked with reconfiguring the attentional biases of having acquired their first language (Slobin, 1996).

9 FORM-FOCUSSED INSTRUCTION IN L2A

9.1 FORM-FOCUSED INSTRUCTION (FFI)

The fact that L2 learners must learn to adjust their L1-shaped attention biases has consequences for effective L2 instruction. Schmidt's (2001) *Noticing Hypothesis* holds that conscious attention to linguistic forms in the input is an important precondition to L2 learning: "people learn about the things they attend to and do not learn much about the things they do not attend to" (Schmidt, 2001, p.30). In order to successfully acquire specific aspects of their L2, learners must pay conscious and selective (i.e., focused) attention to the target structures (Ellis, 2017b). This holds in particular for the morphosyntactic forms in the L2 that are redundant and/or lack perceptual salience. Form-focused Instruction (FFI) attempts to encourage noticing, drawing learners' attention to linguistic forms that might otherwise be ignored (Ellis, 1993, 1995, 2007, 2008c; Ellis & Laporte, 1997).

Norris and Ortega's (2000) milestone meta-analysis comparing the outcomes from studies that employed differing levels of explicitness of L2 input demonstrated that FFI instruction results in substantial target-oriented L2 gains, that explicit types of instruction are more effective than implicit types, and that the effectiveness of L2 instruction is durable. More recent meta-analyses of effects of type of instruction by Spada and Tomita (2010) and Goo, Granena, Yilmaz, and Novella (2015) likewise report large advantages of explicit instruction in L2 acquisition. Variants of FFI vary in the degree and the manner in which they recruit learner consciousness and in the role of the learner's metalinguistic awareness of the target forms (Ellis, 2016; Long, 2020).

9.2 HOW FFI OVERCOMES BLOCKING: PROCESS AND PROCESSING ANALYSES AT THE INTERFACE

Cintrón-Valentín and Ellis (2015) used eye-tracking to investigate the attentional processes whereby different types of FFI overcome learned attention and blocking in learners' online processing of L2 input. English and Chinese (no L1 verb-tense morphology) native speakers viewed Latin utterances combining lexical and morphological cues to temporality under control conditions (CC) and three types of explicit Focus on Form (FonF): verb grammar instruction (VG), verb salience with textual enhancement (VS), and verb pretraining (VP). All groups participated in three phases: exposure, comprehension test, and production test. VG participants viewed a short lesson on Latin tense morphology prior to exposure. VS participants saw the verb inflections highlighted in bold and red during exposure. VP participants had an additional introductory phase where they were presented with solitary verb forms and trained on their English translations. When the verb is presented on its own like this, rather than in potentially redundant combination with adverbial cues, there is less scope for blocking. CC participants were significantly more sensitive to the adverbs than verb morphology. Instructed participants showed greater sensitivity to morphological cues in comprehension and production. Eye-tracking revealed how FonF engages learners' attention to cues which might otherwise be ignored during online processing and thus modulated long-term blocking of verb morphology.

Cintrón-Valentín and Ellis (2016) extended these investigations of three aspects of salience: the physical form of language, learner associative experience, and instructional focus on form (FFI). Experiment 1 replicated the 2015 findings in Chinese native speakers and revealed the attentional processes whereby learners' prior linguistic experience can shape their attentional focus toward cues in the input, and by which FFI helps learners overcome the long-term blocking of verbal morphological cues.

Experiment 2 additionally examined the role of modality of input presentation – aural or visual – in L1 English learners' attentional focus on morphological cues and the effectiveness of different FFI manipulations. Again, participants learning under CC showed greater sensitivity toward the adverb than the verb cues. FFI was effective in increasing attention to verbal morphology, and learning morphological cues was considerably more difficult under aural than under visual presentation. The most effective FFI was grammar instruction. The effectiveness of morphological salience-raising varied across modality: VS was effective under visual exposure, but not under aural exposure.

This study showed that the visual modality of instruction was particularly effective in focusing learners' attention upon low-salience forms by means of textual enhancement to facilitate the consciousness-raising that allows their initial apprehension and consolidation. There is now a wide body of research confirming these effects of multimodal input and the effectiveness of captioning (Cintrón-Valentín, García-Amaya, & Ellis, 2019; Montero Perez, 2020; Muñoz, Pujadas, & Pattemore, 2021; Pattemore & Muñoz, 2020; Sharwood-Smith, 1981; Vanderplank, 2016).

9.3 EXPLICIT AND IMPLICIT L2 MORPHOSYNTAX ACQUISITION

Such results demonstrate how salience in physical form, learner attention, and instructional focus all variously affect the success of L2 acquisition. Form-focused instruction recruits learners' explicit, conscious processing capacities and allows them to consolidate unitized form-function bindings of novel L2 constructions. That's a good and necessary start, but it's only a precondition: there's a lot more usage needed yet to achieve "nativelike selection and nativelike fluency" (Pawley & Syder, 1983). Once a construction has been explicitly represented, its use in subsequent implicit processing can update the statistical tallying of its frequency of usage and probabilities of form-function mapping. There needs be lots of this.

10 LEARNING A PARTICULAR MORPHEME

We have considered usage-based theories concerning how domain-general cognitive mechanisms drive the learning and generalization of linguistic constructions and how acquisition is modulated by factors affecting attention and memory, including exemplar type- and token-frequency, contingency of form-function mapping, salience of form and of function, and the proportion of friends: enemies in quasi-regular domains, etc. We have also seen that the acquisition of a morpheme such as the 3rd-person singular -s can take five years or more to go from zero to 80% provision in obligatory contexts for ESL children (Jia & Fuse, 2007). Five years of English usage involves many thousands of receptive experiences of high frequency functional morphemes, and many thousands of contexts requiring their productive use, yet provision is variable. This suggests that the system is learned incrementally, and that regularities/generalization/productivity emerge from the combined experience of usage.

But are all experiences of the morpheme equally potent or, for any given morpheme, is it the case that some exemplars are more easily recognized in the input and produced earlier in acquisition? If so, what are these exemplars that are more likely to be learned early and preferentially processed? And why these bellwethers? Are they special in their distributional statistics, for example, in terms of their frequency, or their form-function contingency, or their formulaicity? Are they special in their meanings or functions?

More specifically, in the five years during which L2 learners are learning to produce 3rd-person singular -s, do experiences of particular -s inflected verbs play a role in the acquisition of the system more than others? Likewise, for the even more extended period during which L2 learners are learning to produce regular past-tense -ed, are particular -ed inflected verbs more potent exemplars than others? Two recent research projects have investigated these questions separately (1) for online processing in listening and speaking, and (2) for controlled conscious processing for written composition.

10.1 AUTOMATIC ONLINE PROCESSING IN LISTENING AND SPEAKING

Guo and Ellis (2021) investigated how statistical distributions at different linguistic levels – morphological and lexical (Experiments 1 and 2), and phrasal (Experiment. 2) – contribute to the ease with which morphosyntax is processed and produced by second language learners. We analyzed Chinese ESL learners' knowledge of four English inflectional morphemes: -ed, -ing, and 3rd-person -s on verbs, and plural -s on nouns. In Elicited Imitation Tasks, participants listened to length- and difficulty-matched sentences each containing one target morpheme and typed the whole sentence as accurately as they could after a short delay. Experiment 1 investigated lexical and morphemic levels, testing the hypotheses that a morpheme is expected to be more easily processed when it is 1) highly *available* (i.e., occurring in frequent word-forms), and 2) highly *contingent* (i.e., occurring in lemma words that are consistently conjugated in the form containing this morpheme). Thirty sentences were made for each morpheme, divided into three Availability-Reliability Distribution (ARD) groups on the basis of corpus analysis in the Corpus of Contemporary American English (COCA; Davies, 2008-): 10 target words high in availability, 10 high in contingency, and 10 low in both contingency and availability. Responses were scored on whether the target morpheme was accurately reproduced given the provision of the correct lemma. Generalized linear mixed-effects logit models (GLMM) revealed significant effects of morpheme type, availability, and contingency on the accuracy of morpheme provision. There were no effects of lemma frequency – this is a common finding in psycholinguistic studies of morphosyntax processing and it carries important theoretical implications that argue against the idea that regular inflections are computed by procedural application of an affixation rule in a neurally based symbol-manipulating syntactic system.[4]

Experiment 2 successfully replicated these results and extended the investigation to explore phrasal formulaicity by manipulating the frequency of the 4-word strings in which the morpheme was embedded. GLMMs replicated the effects of word-form availability and contingency and additionally revealed independent *phrase-superiority effects* where morphemes were better reproduced in phrasal contexts of higher string-frequency. These are the phrasal equivalents

---

4   According to these accounts, first the lemma is accessed (with a reaction time proportional to its frequency), and then the affixation routine is applied (which should take the same time for all stems).

of word-superiority effects (Wheeler, 1970) whereby recognition of a letter is more accurate when it is part of a meaningful word than when it is alone. Taken together, these findings demonstrated that morpheme acquisition reflects the distributional properties of learners' experience and the mappings therein between lexis, morphology, phraseology, and semantics. They support an emergentist view of the statistical symbolic learning of morphology where language acquisition involves the satisfaction of competing constraints across multiple grain-sizes of units.

10.2 CONTROLLED CONSCIOUS PROCESSING FOR WRITTEN COMPOSITION

Murakami and Ellis (2022) investigated whether the accuracy of grammatical morphemes in second language (L2) learners' writing is associated with these same types of usage-based distributional factors in EF-Cambridge Open Language Database (EFCAMDAT; Geertzen, Alexopoulou, & Korhonen, 2014), a partially error-tagged large-scale longitudinal learner corpus which includes approximately 1.2 million writings by 175,000 learners (147 million words) in a wide range of nationality groups.

Specifically, we examined whether the accuracy of L2 English inflectional morphemes is associated with the availability (i.e., token frequency) and contingency (i.e., token frequency relative to other forms with the same lemma) of the inflected word form, as well as the formulaicity of the context it occurs in (i.e., predictability of the form given the surrounding words). Data drawn from the learner corpus indicated that contingency is a robust predictor of accuracy of morpheme provision in written composition and that its relationship with accuracy does not necessarily lessen when learners' proficiency rises. Contrary to online processing, availability and formulaicity were not generally identified as predictors of accuracy of morpheme production in writing.

10.3 MORPHEMES ARE BETTER PROCESSED WITH LEMMAS RELIABLY CONJUGATED IN THIS FORM IN THE LANGUAGE

These two studies demonstrate that the ways various psycholinguistic factors determine morphosyntax accuracy can depend upon the task demands (automatic vs. controlled; reception vs. production), with memorized associative knowledge having more effect in online tasks. However, both studies suggest that morphemes are more easily processed when they occur with lemmas that are reliably conjugated in this form.

10.4 WHY RELIABILITY, PARTICULARLY?

Why is reliability of association a more potent determinant of acquisition than availability? We can make sense of this from the three different perspectives adopted here (1) learning theory (see sections 4 and 5), (2) cognitive linguistics (section 3), and (3) SLA theory (sections 8, 9, 10).

(1) Associative learning theory demonstrates that contingency of association trumps mere frequency (as described in section 5.1). In operationalizing reliability in the Guo and Ellis (2021) and Murakami and Ellis (2022) studies, we focused upon how likely it is that a linguistic cue (a morpheme) reliably co-occurs with another (a lemma). But morphemes and lemmas are more than mere forms, they are linguistic constructions with particular functions and meanings: they are symbolic.

(2) Cognitive linguistic theories of construction grammar (section 3) view lexical, morphological, and syntactic forms as symbolic form-function pairings and hold that we learn language from usage. When learners are processing usage, they are tallying the associations between forms, between interpretations, and between forms and their interpretations. Verbs have interpretations. Morphemes have interpretations. Verbs and morphemes can be more or less reliably associated. The matrix of association goes beyond mere forms; in full it involves:
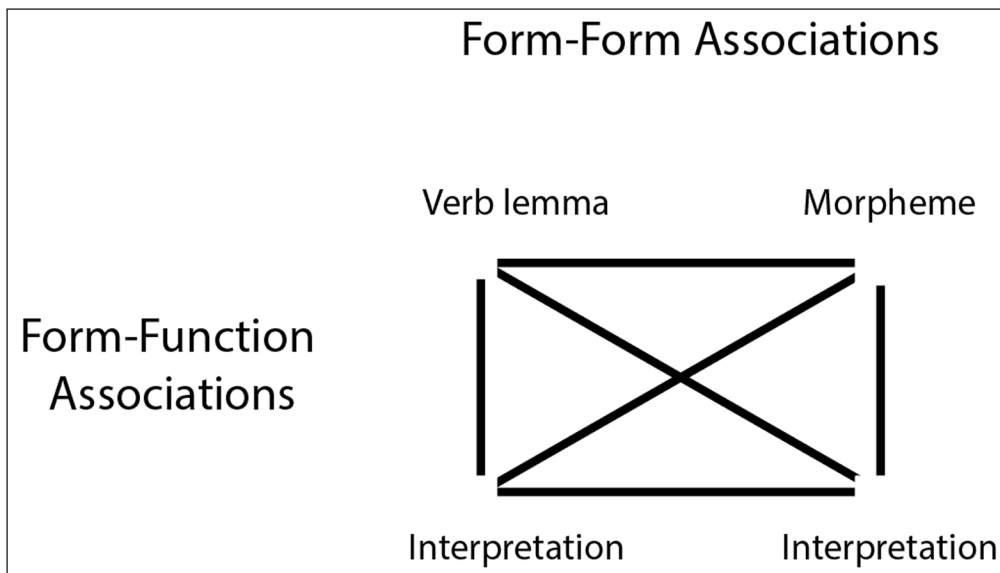
The original "thought-sound" (de Saussure, 1916, p. 111) better illustrates and expresses the dynamics.

(3) Functional theories of SLA emphasize the interplay of form and meaning in acquisition. One much-researched example for morphology is the Aspect Hypothesis (AH) (Andersen & Shirai, 1996; Bardovi-Harlig, 2000; Bardovi-Harlig & Comajoan-Colomé, 2020). The AH builds on three main constructs: tense, grammatical aspect, and lexical aspect. Tense establishes the location of an event in time with respect to the moment of speech or some other reference point. Grammatical aspect allows for "ways of viewing the internal temporal constituency of a situation" (Comrie, 1976, p. 3). For instance, in English, a contrast in grammatical aspect is found between simple past "Phoebe walked" and past progressive "Phoebe was walking". In contrast, lexical aspect refers to semantic differences in verbs and their arguments (Dowty, 1979), such as whether a predicate has inherent duration like "walk", "sleep", and "kid", is punctual like "recognize", "broke", or "sigh", or has elements of both duration and culmination like "walk a mile" and "paint a picture". The AH predicts that "second language learners will initially be influenced by the inherent semantic aspect of verbs or predicates" (Andersen & Shirai, 1996, p. 533). "In its simplest form, the AH for SLA predicts that in the initial stages of the acquisition of tense-aspect morphology by adults, the acquisition of past morphology will be influenced by lexical aspectual categories. Namely, verbal morphology will be attracted to and will occur with predicates with similar semantics. Perfective past will occur with telic predicates (predicates with inherent endpoints); in contrast, imperfective will occur with unbounded predicates, and progressive will occur with ongoing activities" (Bardovi-Harlig & Comajoan-Colomé, 2020, p. 1139). Bardovi-Harlig & Comajoan-Colomé conclude from their review of perhaps thirty different studies of the AH over the last twenty years that the AH accurately predicts the adult L2 acquisition of past morphology in a number of languages.

Guo and Ellis (2021) and Murakami and Ellis (2022) demonstrated effects of distributional learning – particularly the privileged processing of reliably associated lemma-morpheme pairings – on morphology acquisition. The question that naturally follows is to wonder why language is distributed this way. Why do particular lemmas appear in the language more reliably associated with particular morphemes? Cognitive linguistics more generally, and the AH in particular, suggest that for the case of tense-aspect morphology, there are semantic and functional motivations. Likewise, for noun number, we suspect that inherent number, pluralia tantum, and prototypically plural count nouns might lead the way. There is good reason and plenty of scope to study a broad range of morphology in a range of languages in this way.

10.5 MEANING: A MORPHEME SHALL BE KNOWN BY THE COMPANY IT KEEPS

A lesson from vocabulary learning (section 5.3) is that words with more meaningful, imageable, concrete, and salient referents are those that are more easily learned. A lesson from the cognitive grammar of verb-argument constructions (section 3) is that those exemplars that are closest to the construction meaning prototype, high in the Zipfian frequency profile thus to take the lion's share of experiences, and well-connected in the semantic network, are more easily learned (Ellis, O'Donnell, & Römer, 2014; Ellis et al., 2016; Wulff, Ellis, Römer, Bardovi-Harlig, & LeBlanc, 2009). The results reviewed here in section 10 suggest a similar lesson concerning semantic motivations in the learning of morphosyntax.

Whereas lexical constructions get much of their meanings directly from being grounded in rich, imageable, multimodal experiences, this is less true for morphosyntax which has been subject to prior grammaticalization processes of *Desemanticization* and *Extension* (section 6.5). However, the unit of meaning in language goes well beyond individual words – as corpus linguistics has shown us, ultimately it lies in "the phrase, the whole phrase, and nothing but the phrase" (Sinclair, 2008); (see further Hoey, 2005; Sinclair, 1996, 2008; Stubbs, 2009). Thus individual words derive further meaning from their collocations and contexts: "You shall know a word by the company it keeps" (Firth, 1957). Elman (2004, 2011) reviews psycholinguistic research on the dynamics of sentence processing, and proposes that, rather than words *having meaning*, as in traditional views of the mental lexicon, they are better viewed as *cues to meaning* in the interactive conspiracy with the other available cues as they unfold in sentence processing: a word's 'meaning' lies in the dynamic causal effects it can have on mental states. In the same way, it is from experience of the lemmas with which they have high contingency of association and their meaning correlates in phrasal contexts that morphemes derive their status as cues to meaning: Morphemes become known by the company they keep.

11 CONCLUSION

Everything discussed here warrants replication, clarification, and further investigation. Much of it is Anglocentric and needs extending cross-linguistically. Some details are possibly wrong, or wrong in their degree of generalization. There is plenty to do.

I began by framing this piece in terms of Language as a Complex Adaptive System, something the Five Graces group worked on together in 2009: "Cognition, consciousness, experience, embodiment, brain, self, human interaction, society, culture, and history are all inextricably intertwined in rich, complex, and dynamic ways in language. Everything is connected. Yet despite this complexity, despite its lack of overt government, instead of anarchy and chaos, there are patterns everywhere. Linguistic patterns are not preordained by God, genes, school curriculum, or other human policy. Instead, they are emergent—synchronic patterns of linguistic organization at numerous levels (phonology, lexis, syntax, semantics, pragmatics, discourse, genre, etc.), dynamic patterns of usage, diachronic patterns of language change (linguistic cycles of grammaticalization, pidginization, creolization, etc.), ontogenetic developmental patterns in child language acquisition, global geopolitical patterns of language growth and decline, dominance and loss, and so forth. We cannot understand these phenomena unless we understand their interplay" (Beckner et al., 2009, p. 18). I have tried here to fill in some relevant detail regarding L2 morphology learning.

I recognize that these studies focus on but a grain of sand on the morphosyntactic shoreline where the seas of cognition and social usage meet the lands of linguistic structure in the languages of the world, now, past, future, synchronic, diachronic, and panchronic. In the header of section 2 of this paper, in its subtitle, I misquoted Wittgenstein (1980, p. 50e) "How small a thought it takes to fill a whole life". It seems that many facets of language usage, learning, structure, and change are coherently reflected in this *-s*. As I have explained, L2 learners of English are less likely to have spotted the missing morpheme than L1 speakers. Still, not much of any consequence could stem from the omission. I'm sure no one was prevented from getting at Wittgenstein's original intent, one sublimely celebrated in Reich's (1995) Proverb.

ACKNOWLEDGEMENTS

AUTHOR AFFILIATION

Nick C. Ellis ⓘ orcid.org/0000-0002-4235-6481
University of Michigan, US

REFERENCES

Allan, L. G. (1980). A note on measurement of contingency between two binary variables in judgment tasks. *Bulletin of the Psychonomic Society, 15*, 147–149. DOI: https://doi.org/10.3758/BF03334492

Andersen, R. W., & **Shirai, Y.** (1996). The primacy of aspect in first and second language acquisition: The pidgin-creole connection. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 527–570). San Diego, CA: Academic Press. DOI: https://doi.org/10.1016/B978-012589042-7/50018-9

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89*(4), 369–406. DOI: https://doi.org/10.1037/0033-295X.89.4.369

Anderson, J. R. (1992). Automaticity and the ACT-super(*) theory. *American Journal of Psychology, 105*(2), 165–180. DOI: https://doi.org/10.2307/1423026

Atkinson, M., Smith, K., & **Kirby, S.** (2018). Adult Learning and Language Simplification. *Cognitive Science, 42*(8), 2818–2854. DOI: https://doi.org/10.1111/cogs.12686

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

Bardovi-Harlig, K., & **Comajoan-Colomé, L.** (2020). The aspect hypothesis and the acquisition of L2 past morphology in the last 20 years: A state-of-the-scholarship review. *Studies in Second Language Acquisition, 42*(5), 1137–1167. DOI: https://doi.org/10.1017/S0272263120000194

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*, 577–660. DOI: https://doi.org/10.1017/S0140525X99002149

Bates, E., & **Goodman, J. C.** (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia and real-time processing. *Language and Cognitive Processes, 12*, 507–586. DOI: https://doi.org/10.1080/016909697386628

Beckett, S. (1954). *Waiting for Godot.* New York: Grove Press.

Beckner, C., Blythe, R. A., Bybee, J., Christiansen, M. H., Croft, W., Ellis, N. C., Holland, J., Ke, J., Larsen-Freeman, D., & **Schoenemann, T.** (2009). Language is a complex adaptive system. Position paper. *Language Learning, 59*, 1–26. DOI: https://doi.org/10.1111/j.1467-9922.2009.00533.x

Berdicevskis, A., & **Semenuks, A.** (2022). Imperfect language learning reduces morphological overspecification: Experimental evidence. *PLoS ONE, 17*(1), e0262876. DOI: https://doi.org/10.1371/journal.pone.0262876

Berger, C., Crossley, S., & **Skalicky, S.** (2019). Using lexical features to investigate second language lexical decision performance. *Studies in Second Language Acquisition, 41*(5), 911–935. DOI: https://doi.org/10.1017/S0272263119000019

Bod, R., Hay, J., & **Jannedy, S.** (Eds.) (2003). *Probabilistic linguistics*. Cambridge, MA: MIT Press. DOI: https://doi.org/10.7551/mitpress/5582.001.0001

Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press. DOI: https://doi.org/10.4159/harvard.9780674732469

Bybee, J. (2003). Mechanisms of change in grammaticalization: The role of frequency. In R. D. Janda & B. D. Joseph (Eds.), *Handbook of historical linguistics* (pp. 602–623). Oxford: Blackwell. DOI: https://doi.org/10.1002/9780470756393.ch19

Bybee, J. (2010). *Language, usage, and cognition*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511750526

Bybee, J., & **Hopper, P.** (Eds.) (2001). *Frequency and the emergence of linguistic structure*. Amsterdam: Benjamins. DOI: https://doi.org/10.1075/tsl.45

Christiansen, M. H., & **Chater, N.** (Eds.) (2001). *Connectionist psycholinguistics*. Westport, CO: Ablex.

Cintrón-Valentín, M., & **Ellis, N. C.** (2015). Exploring the Interface: Explicit Focus-on-Form Instruction and Learned Attentional Biases in L2 Latin. *Studies in Second Language Acquisition, 37*, 197–235. DOI: https://doi.org/10.1017/S0272263115000029

Cintrón-Valentín, M., & **Ellis, N. C.** (2016). Salience in Second Language Acquistion: Physical form, learner attention, and instructional focus. *Frontiers in Psychology, section Language Sciences, 7*, 1284. DOI: https://doi.org/10.3389/fpsyg.2016.01284

Cintrón-Valentín, M., García-Amaya, L., & **Ellis, N. C.** (2019). Captioning and grammar learning in the L2 Spanish classroom. *The Language Learning Journal, 47*(4), 439–459. DOI: https://doi.org/10.1080/09571736.2019.1615978

Comrie, B. (1976). *Aspect: An Introduction to the Study of Verbal Aspect and Related Problems*. Cambridge, NY: Cambridge University Press.

Corder, S. P. (1967). The significance of learners' errors. *International Review of Applied Linguistics, 5*, 161–169. DOI: https://doi.org/10.1515/iral.1967.5.1-4.161

Crossley, S. A., Subtirelu, N., & **Salsbury, T.** (2013). Frequency effects or context effects in second language word learning: What predicts early lexical production? *Studies in Second Language Acquisition, 35*(4), 727–755. DOI: https://doi.org/10.1017/S0272263113000375

Cutler, A., & **Carter, D. M.** (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech & Language, 2*(3), 133–142. DOI: https://doi.org/10.1016/0885-2308(87)90004-0

Dabrowska, E., & **Divjak, D.** (Eds.) (2015). *Handbook of cognitive linguistics*. Berlin: Mouton DeGruyter. DOI: https://doi.org/10.1515/9783110292022

Davies, M. (2008-). The Corpus of Contemporary American English: 450 million words, 1990-present. http://www.americancorpus.org

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

DeKeyser, R. (2001). Automaticity and automatization. In P. Robinson (Ed.), *Cognition and second language acquisition*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9781139524780.007

DeKeyser, R. (2005). What Makes Learning Second-Language Grammar Difficult? A Review of Issues. *Language Learning, 55*(S1), 1–25. DOI: https://doi.org/10.1111/j.0023-8333.2005.00294.x

Douglas Fir Group. (2016). A Transdisciplinary Framework for SLA in a Multilingual World. *The Modern Language Journal, 100*(S1), 19–47. DOI: https://doi.org/10.1111/modl.12301

Dowty, D. (1979). *Word meaning and Montague grammar*. Reidel. DOI: https://doi.org/10.1007/978-94-009-9473-7

Ebbinghaus, H. (1885). *Memory: A contribution to experimental psychology* (H.A. Ruger & C.E. Bussenius, Trans.). New York: Teachers College, Columbia University.

Ellis, N. C. (1993). Rules and instances in foreign language learning: Interactions of explicit and implicit knowledge. *The European Journal of Cognitive Psychology, 5*(3), 289–318. DOI: https://doi.org/10.1080/09541449308520120

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

Ellis, N. C. (Ed.). (1994b). *Implicit and explicit learning of languages*. San Diego, CA: Academic Press.

Ellis, N. C. (1995). Consciousness in second language acquisition: A review of field studies and laboratory experiments. *Language Awareness, 4*(3), 123–146. DOI: https://doi.org/10.1080/09658416.1995.9959876

Ellis, N. C. (1996). Sequencing in SLA: Phonological memory, chunking, and points of order. *Studies in Second Language Acquisition, 18*(1), 91–126. DOI: https://doi.org/10.1017/S0272263100014698

Ellis, N. C. (1998). Emergentism, connectionism and language learning. *Language Learning, 48*(4), 631–664. DOI: https://doi.org/10.1111/0023-8333.00063

Ellis, N. C. (2001). Memory for language. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 33–68). New York: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9781139524780.004

Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition, 24*(2), 143–188. DOI: https://doi.org/10.1017/S0272263102002024

Ellis, N. C. (2003). Constructions, chunking, and connectionism: The emergence of second language structure. In C. Doughty & M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 33–68). Oxford: Blackwell.

Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition, 27*, 305–352. DOI: https://doi.org/10.1017/S027226310505014X

Ellis, N. C. (2006a). Cognitive perspectives on SLA: The Associative Cognitive CREED. *AILA Review, 19*, 100–121. DOI: https://doi.org/10.1075/aila.19.08ell

Ellis, N. C. (2006b). Language acquisition as rational contingency learning. *Applied Linguistics, 27*(1), 1–24. DOI: https://doi.org/10.1093/applin/ami038

Ellis, N. C. (2006c). Selective attention and transfer phenomena in SLA: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics, 27*(2), 1–31. DOI: https://doi.org/10.1093/applin/aml015

Ellis, N. C. (2007). The weak interface, consciousness, and form-focussed instruction: Mind the doors. In S. Fotos & H. Nassaji (Eds.), *Form focused instruction and teacher education: Studies in honour of Rod Ellis* (pp. 17–34). Oxford: Oxford University Press.

Ellis, N. C. (2008a). The dynamics of second language emergence: Cycles of language use, language change, and language acquisition. *Modern Language Journal, 41*(3), 232–249. DOI: https://doi.org/10.1111/j.1540-4781.2008.00716.x

Ellis, N. C. (2008b). Phraseology: The periphery and the heart of language. In F. Meunier & S. Granger (Eds.), *Phraseology in language learning and teaching* (pp. 1–13). Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/z.138.02ell

Ellis, N. C. (2008c). The psycholinguistics of the interaction hypothesis. In A. Mackey & C. Polio (Eds.), *Multiple perspectives on interaction in SLA: Second language research in honor of Susan M. Gass* (pp. 11–40). New York: Routledge.

Ellis, N. C. (2012). Formulaic language and second language acquisition: Zipf and the phrasal teddy bear. *Annual Review of Applied Linguistics, 32*, 17–44. DOI: https://doi.org/10.1017/S0267190512000025

Ellis, N. C. (2015). Implicit AND explicit language learning: Their dynamic interface and complexity In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 3–23). Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/sibil.48.01ell

Ellis, N. C. (2017a). Chunking. In M. Hundt, S. Mollin & S. Pfenninger (Eds.), *The Changing English Language* (pp. 113–147). Cambridge: Cambridge University Press.

Ellis, N. C. (2017b). Implicit and explicit knowledge about language. In J. Cenoz & D. Gorter (Eds.), *Language Awareness and Multilingualism. Encyclopedia of Language and Education* (pp. 113–124). Switzerland: Springer International Publishing. DOI: https://doi.org/10.1007/978-3-319-02240-6_7

Ellis, N. C. (2017c). Salience. In M. Hundt, S. Pfenninger & S. Mollin (Eds.), *The changing English language* (pp. 71–92). Cambridge: Cambridge University Press.

Ellis, N. C. (2017d). Salience in Usage-based SLA. In S. Gass, P. Spinner & J. Behney (Eds.), *Saliency in Second Language Acquisition*. NY: Routledge. DOI: https://doi.org/10.4324/9781315399027-2

Ellis, N. C. (2019). Essentials of a theory of language cognition. *Modern Language Journal, 103* (Supplement 2019), 39–60. DOI: https://doi.org/10.1111/modl.12532

Ellis, N. C. (2022). Bilingual language cognition as a complex adaptive system. *Bilingualism: Language and Cognition*, 1–2. DOI: https://doi.org/10.1017/S1366728922000372

Ellis, N. C., & **Beaton, A.** (1993). Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning, 43*(4), 559–617. DOI: https://doi.org/10.1111/j.1467-1770.1993.tb00627.x

Ellis, N. C., Hafeez, K., Martin, K. I., Chen, L., Boland, J., & **Sagarra, N.** (2014). An eye-tracking study of learned attention in Second Language Acquisition. *Applied Psycholinguistics, 35*, 547–579. DOI: https://doi.org/10.1017/S0142716412000501

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: Erlbaum.

Ellis, N. C., & **Larsen Freeman, D.** (2006). Language emergence: Implications for Applied Linguistics. *Applied Linguistics, 27*(4), whole issue. DOI: https://doi.org/10.1093/applin/aml028

Ellis, N. C., O'Donnell, M. B., & **Römer, U.** (2014). The Processing of Verb-Argument Constructions is Sensitive to Form, Function, Frequency, Contingency, and Prototypicality. *Cognitive Linguistics, 25*, 55–98. DOI: https://doi.org/10.1515/cog-2013-0031

Ellis, N. C., & **Ogden, D. C.** (2017). Thinking about multiword constructions: Usage-based approaches to acqusition and processing. *Topics in Cognitive Science, 9*, 604–620. DOI: https://doi.org/10.1111/tops.12256

Ellis, N. C., Römer, U., & **O'Donnell, M. B.** (2016). *Language usage, acquisition, and processing: Cognitive and corpus investigations of construction grammar*. Malden, MA: Wiley-Blackwell.

Ellis, N. C., & **Sagarra, N.** (2010). The bounds of adult language acquisition: Blocking and learned attention. *Studies in Second Language Acquisition, 32*(4), 553–580. DOI: https://doi.org/10.1017/S0272263110000264

Ellis, N. C., & **Sagarra, N.** (2011). Learned attention in adult language acquisition: A replication and generalization study and meta-analysis. *Studies in Second Language Acquisition, 33*(4), 589–624. DOI: https://doi.org/10.1017/S0272263111000325

Ellis, N. C., & **Schmidt, R.** (1998). Rules or associations in the acquisition of morphology? The frequency by regularity interaction in human and PDP learning of morphosyntax. *Language & Cognitive Processes, 13*(2&3), 307–336. DOI: https://doi.org/10.1080/016909698386546

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 ed., pp. 75–93). NY: Routledge.

Ellis, R. (2016). Focus on form: A critical review. *Language Teaching Research, 20*(3), 405–428. DOI: https://doi.org/10.1177/1362168816628627

Elman, J. L. (2004). An alternative view of the mental lexicon. *Trends in Cognitive Science, 8*, 301–306. DOI: https://doi.org/10.1016/j.tics.2004.05.003

Elman, J. L. (2011). Lexical knowledge without a lexicon? *The Mental Lexicon, 6*, 1–33. DOI: https://doi.org/10.1075/ml.6.1.01elm

Firth, J. R. (1957). *Papers in linguistics: 1934–1951*. London: Oxford University Press.

Geertzen, J., Alexopoulou, T., & Korhonen, A. (2014). *Automatic linguistic annotation of large scale L2 databases: The EF-Cambridge Open Language Database (EFCAMDAT)*. Paper presented at the Selected proceedings of the 2012 Second Language Research Forum. Building bridges between disciplines.

Givón, T. (1971). Historical syntax and synchronic morphology: an archeologist's field trip. *Chicago Linguistic Society, 7*, 394–415.

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), 1–50. DOI: https://doi.org/10.1111/1467-9922.00147

Goo, J., Granena, G., Yilmaz, Y., & Novella, M. (2015). Implicit and explicit instruction in L2 learning. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 443–482). Amsterdam: Benjamins. DOI: https://doi.org/10.1075/sibil.48.18goo

Goodman, J. C., Nusbaum, H. C., Lee, L., & Broihier, K. (1990). The effects of syntactic and discourse variables on the segmental intelligibility of speech. *Paper presented at the Proceedings of the 1990 International Conference on Spoken Language Processing*, Kobe, Japan.

Graddol, D. (2000). *The future of English*. London: The British Council.

Gries, S. T., & Ellis, N. C. (2015). Statistical Measures for Usage-Based Linguistics. *Language Learning, 65*, 228–255. DOI: https://doi.org/10.1111/lang.12119

Guo, R., & Ellis, N. C. (2021). Language Usage and Second Language Morphosyntax: Effects of Availability, Reliability, and Formulaicity. *Frontiers in Psychology, 12*(1288). DOI: https://doi.org/10.3389/fpsyg.2021.582259

Hebb, D. O. (1949). *The organization of behaviour*. New York: John Wiley & Sons.

Heine, B., Claudi, U., & Hünnemeyer, F. (1991). *Grammaticalization: A conceptual framework*. Chicago: University of Chicago Press.

Herron, D., & Bates, E. (1997). Sentential and acoustic factors in the recognition of open- and closed-class words. *Journal of Memory and Language, 37*, 217–239. DOI: https://doi.org/10.1006/jmla.1997.2514

Hilpert, M. (2021). *Ten lectures on diachronic construction grammar*. DOI: https://doi.org/10.1163/9789004446793

Hodge, C. (1970). The Linguistic Cycle. *Linguistic Sciences, 13*, 1–7.

Hoey, M. (2005). *Lexical priming: A new theory of words and language*. London: Routledge.

Hopper, P. J. (1998). Emergent grammar. In M. Tomasello (Ed.), *The new psychology of language: Cognitive and functional approaches to language structure* (pp. 155–176). Mahwah, NJ: Erlbaum. DOI: https://doi.org/10.4324/9781315085678-6

Hopper, P. J., & Traugott, E. C. (2003). *Grammaticalization* (Second ed.). Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9781139165525

Hundt, M., Pfenninger, S., & Mollin, S. (Eds.) (2017). *The changing English language*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/9781316091746

James, W. (1890). *The principles of psychology* (Vol. 1). New York: Dover. DOI: https://doi.org/10.1037/10538-000

Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. New York: Routledge. DOI: https://doi.org/10.4324/9780203935927

Jia, G., & Fuse, A. (2007). Acquisition of English grammatical morphology by native Mandarin-speaking children and adolescents: Age-related differences. *J Speech Lang Hear Res, 50*(5), 1280–1299. DOI: https://doi.org/10.1044/1092-4388(2007/090)

Johnson, J., & Newport, E. L. (1989). Critical period effects in second language learning and the influence of the maturational state on the acquisition of ESL. *Cognitive Psychology, 21*, 215–258. DOI: https://doi.org/10.1016/0010-0285(89)90003-0

Kamin, L. J. (1968). "Attention-like" processes in classical conditioning. In M. R. Jones (Ed.), *Miami symposium on the prediction of behavior: Aversive stimulation* (pp. 9–31). Miami, FL: University of Miami Press.

Ke, J.-Y. (2006). A cross-linguistic quantitative study of homophony. *Journal of Quantitative Linguistics, 13*, 129–159. DOI: https://doi.org/10.1080/09296170500500850

Klein, W. (1998). The contribution of second language acquisition research. *Language Learning, 48*, 527–550. DOI: https://doi.org/10.1111/0023-8333.00057

Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. London: Longman.

Kruschke, J. K., & Blair, N. J. (2000). Blocking and backward blocking involve learned inattention. *Psychonomic Bulletin & Review, 7*, 636–645. DOI: https://doi.org/10.3758/BF03213001

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

Long, M. H. (2020). Optimal input for language learning: Genuine, simplified, elaborated, or modified elaborated? *Language Teaching, 53*(2), 169–182. DOI: https://doi.org/10.1017/S0261444819000466

Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review, 82*, 276–298. DOI: https://doi.org/10.1037/h0076778

MacWhinney, B. (1987). The competition model. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 249–308). Hillsdale, NJ: Erlbaum.

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., & **O'Grady, W.** (Eds.) (2015). *The handbook of language emergence*. Oxford: Wiley-Blackwell. DOI: https://doi.org/10.1002/9781118346136

Marchman, V. A., Wulfeck, B., & **Weismer, S. E.** (1999). Morphological productivity in children with normal language and SLI: a study of the English past tense. *J. Speech Lang. Hear Res., 42*, 209–219. DOI: https://doi.org/10.1044/jslhr.4201.206

McWhorter, J. (2004). The story of human language Lecture 24: Language Interrupted. *The Great Courses*. The Teaching Company Limited Partnership.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81–97. DOI: https://doi.org/10.1037/h0043158

Montero Perez, M. (2020). Multimodal input in SLA research. *Studies in Second Language Acquisition, 42*(3), 653–663. DOI: https://doi.org/10.1017/S0272263120000145

Muñoz, C., Pujadas, G., & **Pattemore, A.** (2021). Audio-visual input for learning L2 vocabulary and grammatical constructions. *Second Language Research*. DOI: https://doi.org/10.1177/02676583211015797

Murakami, A., & **Alexopoulou, T.** (2016). L1 influence on the acquisition order of English grammatical morphemes: A Learner Corpus Study. *Studies in Second Language Acquisition, 38*(3), 365–401. DOI: https://doi.org/10.1017/S0272263115000352

Murakami, A., & **Ellis, N. C.** (2022). Effects of availability, contingency, and formulaicity in the accuracy of L2 English grammatical morphemes in writing. *Language Learning*. DOI: https://doi.org/10.1111/lang.12500

Norris, J., & **Ortega, L.** (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning, 50*, 417–528. DOI: https://doi.org/10.1111/0023-8333.00136

Odlin, T. (2022). *Explorations of Language Transfer*. Bristol: Multilingual Matters. DOI: https://doi.org/10.21832/ODLIN9547

Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford: Oxford University Press.

Pattemore, A., & **Muñoz, C.** (2020). Learning L2 constructions from captioned audio-visual exposure: The effect of learner-related factors. *System, 93*, 102303. DOI: https://doi.org/10.1016/j.system.2020.102303

Pawley, A., & **Syder, F. H.** (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 191–225). London: Longman.

Perdue, C. (Ed.). (1993). *Adult language acquisition: Crosslinguistic perspectives*. Cambridge: Cambridge University Press.

Peters, E. (2019). Factors affecting the learning of single-word items. In S. Webb (Ed.), *The Routledge Handbook of Vocabulary Studies* (pp. 125–142). London: Routledge. DOI: https://doi.org/10.4324/9780429291586-9

Pinker, S., & **Prince, A.** (1994). Regular and irregular morphology and the psychological status of rules of grammar. In S. D. Lima, R. L. Corrigan & G. K. Iverson (Eds.), *The reality of linguistic rules* (pp. 321–351). Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/slcs.26.21pin

Polikarpov, A. A. (2006). Cognitive model of lexical system evolution and its verification. 2006, from http://www.philol.msu.ru/~lex/articles/cogn_ev.htm

Rebuschat, P. (Ed.). (2015). *Implicit and explicit learning of language*. Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/sibil.48

Reich, S. (1995). Proverb. Nonesuch Records Inc.

Rescorla, R. A. (1968). Probability of shock in the presence and absence of CS in fear conditioning. *Journal of Comparative and Physiological Psychology, 66*, 1–5. DOI: https://doi.org/10.1037/h0025984

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

Robinson, P., & **Ellis, N. C.** (Eds.) (2008). *Handbook of cognitive linguistics and second language acquisition*. London: Routledge. DOI: https://doi.org/10.4324/9780203938560

Sagarra, N., & **Ellis, N. C.** (2013). From seeing adverbs to seeing morphology. Language experience and adult acquisition of L2 tense. *Studies in Second Language Acquisition, 35*, 261–290. DOI: https://doi.org/10.1017/S0272263112000885

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

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

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*, 129–158. DOI: https://doi.org/10.1093/applin/11.2.129

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9781139524780.003

Segalowitz, N. (2010). *The cognitive bases of second language fluency*. New York: Routledge. DOI: https://doi.org/10.4324/9780203851357

Seidenberg, M. S., & **Plaut, D. C.** (2014). Quasiregularity and its discontents: the legacy of the past tense debate. *Cognitive Science, 38*(6), 1190–1228. DOI: https://doi.org/10.1111/cogs.12147

Seidlhofer, B. (2004). Research perspectives on teaching English as a Lingua Franca. *Annual Review of Applied Linguistics, 24*, 209–239. DOI: https://doi.org/10.1017/S0267190504000145

Shanks, D. R. (1995). *The psychology of associative learning*. New York: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511623288

Shanks, D. R. (2010). Learning: From association to cognition. *Annual Review of Psychology, 61*, 273–301. DOI: https://doi.org/10.1146/annurev.psych.093008.100519

Sharwood-Smith, M. (1981). Consciousness raising and the second-language learner. *Applied Linguistics, 2*, 159–168. DOI: https://doi.org/10.1093/applin/2.2.159

Shirai, Y. (2018). *Connectionism and Second Language Acquisition*. London: Routledge. DOI: https://doi.org/10.4324/9780203118085

Simon, H. A. (1957). *Models of man: Social and rational*. New York: Wiley and Sons.

Simpson-Vlach, R., & **Ellis, N. C.** (2010). An Academic Formulas List (AFL). *Applied Linguistics, 31*, 487–512. DOI: https://doi.org/10.1093/applin/amp058

Sinclair, J. M. (1996). The search for units of meaning. *Textus, IX*, 75–106. DOI: https://doi.org/10.1093/applin/amn052

Sinclair, J. M. (2008). The phrase, the whole phrase, and nothing but the phrase. In S. Granger & F. Meunier (Eds.), *Phraseology: An interdisciplinary perspective* (pp. 407–410). DOI: https://doi.org/10.1075/z.139.33sin

Slabakova, R. (2009). What is easy and what is hard to acquire in a second language? *Paper presented at the 10th Generative Approaches to Second Language Acquisition Conference (GASLA 2009)*.

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

Spada, N., & **Tomita, Y.** (2010). Interactions Between Type of Instruction and Type of Language Feature: A Meta-Analysis. *Language Learning, 60*(2), 263–308. DOI: https://doi.org/10.1111/j.1467-9922.2010.00562.x

Stubbs, M. (2009). The search for units of meaning: Sinclair on empirical semantics. *Applied Linguistics, 30*, 115–137. DOI: https://doi.org/10.1093/applin/amn052

Terrell, T. (1991). The role of grammar instruction in a communicative approach. *The Modern Language Journal, 75*, 52–63. DOI: https://doi.org/10.1111/j.1540-4781.1991.tb01083.x

Titone, D., & **Tiv, M.** (2022). Rethinking multilingual experience through a Systems Framework of Bilingualism. *Bilingualism: Language and Cognition*, 1–16. DOI: https://doi.org/10.1017/S1366728921001127

Traugott, E. C., & **Trousdale, G.** (2013). *Constructionalization and constructional changes* (Vol. 6). Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199679898.001.0001

Trousdale, G., & **Hoffmann, T.** (Eds.) (2013). *Oxford handbook of construction grammar*. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/oxfordhb/9780195396683.001.0001

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

Van Patten, B. (2007). Input processing in adult second language acquisition. In B. VanPatten & Williams, J. (Ed.), *Theories in Second Language Acquisition* (pp. 115–135). Mahwah, NJ: Lawrence Erlbaum Associates.

Vanderplank, R. (2016). *Captioned media in foreign language learning and teaching: Subtitles for the deaf and hard-of-hearing as tools for language learning*. Springer. DOI: https://doi.org/10.1057/978-1-137-50045-8

Wheeler, D. D. (1970). Processes in word recognition. *Cognitive Psychology, 1*, 59–85. DOI: https://doi.org/10.1016/0010-0285(70)90005-8

White, L. (2003). *Second language acquisition and Universal Grammar*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511815065

Williams, J. M. (1986). *Origins of the English language*. Simon and Schuster.

Wischer, I., & **Diewald, G.** (Eds.) (2002). *New reflections on grammaticalization*. Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/tsl.49

Wittgenstein, L. (1980). *Culture and value/Ludwig Wittgenstein; edited by G.H. von Wright, in collaboration with Heikki Nyman; translated by Peter Winch (Original work published 1977)*. Oxford: B. Blackwell.

Wulff, S., & **Ellis, N. C.** (2018). Usage-based approaches to SLA. In D. Miller, F. Bayram, J. Rothman & L. Serratrice (Eds.), *Bilingual cognition and language: The state of the science across its subfields.* (pp. 37–56). Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/sibil.54.03wul

Wulff, S., Ellis, N. C., Römer, U., Bardovi-Harlig, K., & **LeBlanc, C.** (2009). The acquisition of tense-aspect: Converging evidence from corpora, cognition, and learner constructions. *Modern Language Journal, 93,* 354–369. DOI: https://doi.org/10.1111/j.1540-4781.2009.00895.x

Zipf, G. K. (1935). *The psycho-biology of language: An introduction to dynamic philology.* Cambridge, MA: The M.I.T. Press.

Zipf, G. K. (1949). *Human behaviour and the principle of least effort: An introduction to human ecology.* Cambridge, MA: Addison-Wesley.

WHITE ROSE
UNIVERSITY PRESS
Universities of Leeds, Sheffield & York