John Benjamins Publishing Company

Usage-based approaches to second language acquisition

Stefanie Wulff and Nick C. Ellis

University of Florida / University of Michigan

We present an overview of the cognitive underpinnings of usage-based approaches to second language acquisition (L2A). Not all constructions are equally learnable, even after years of (frequent) exposure. We present a usage-based analysis of this phenomenon in terms of fundamental principles of associative learning: Low salience, low contingency, and redundancy all lead to form-function mappings being less well learned. Compounding this, adult acquirers show effects of learned attention and blocking as a result of L1-tuned automatized processing of language. We also describe form-focused instruction studies that aim to recruit learners' explicit, conscious processing capacities for noticing novel L2 constructions before subsequent implicit processing consolidates it into the system. We conclude with further readings which discuss wider coverage of usage-based L2A.

1.   Introduction

"Usage-based" is an umbrella label for a variety of approaches to second language acquisition (L2A) that minimally share these two assumptions:

1. The linguistic input learners receive is the primary source for their second language (L2) learning.
2. The cognitive mechanisms that learners employ in language learning are not exclusive to language learning, but are general cognitive mechanisms associated with learning of any kind.

In this chapter, we describe the constructs and working assumptions that characterize usage-based approaches to language learning, with a particular focus on the cognitive underpinnings of usage-based approaches (for a more detailed discussion of the important contributions of social and interactional factors in usage-based approaches, see, for example, Cadierno & Eskildsen, 2015). We start out by explaining

constructions as the targets of language learning. Then we describe the processes of construction learning in terms of exemplar-based, rational, associative learning. But not all constructions are equally learnable by all learners: Naturalistic second language learners focus more in their language processing upon open-class words than on grammatical cues. We describe a usage-based analysis of this phenomenon in terms of salience, contingency, and redundancy, and explain how adult acquirers show effects of learned attention and blocking which further limit learning. We outline educational interventions targeted upon these phenomena. We conclude with further readings which widen the coverage of usage-based L2A.

2.   The targets of language learning: Constructions

Learning a language involves the learning of constructions. These are the form-function mappings that are conventionalized as ways to express meanings in a speech community. Constructions range from morphemes – the smallest pairing of form and meaning in language – to words, phrases, and syntactic frames (Goldberg, 2006; Trousdale & Hoffmann, 2013). That is, simple morphemes such as –*aholic* (meaning 'being addicted to something') are constructions in the same way as simple words like *nut* (meaning 'a fruit consisting of a hard or tough shell around an edible kernel'), idioms like *It is driving me nuts* (meaning 'It is greatly frustrating me') , and abstract syntactic frames like Subject-Verb-Object-Object (meaning that something is being transferred, as realized in sentences as diverse as *Max gave the squirrel a nut*, *Nick gave Max a hug*, or *Steffi baked Max a cake*, where nuts, hugs, and cakes are being transferred, respectively). As the latter examples illustrate, not all constructions carry meaning in the traditional sense; many constructions rather serve a more functional purpose. The passive construction, for instance, serves the function of shifting the focus of attention in an utterance from the agent of the action to the patient undergoing the action (compare the passive *A cake was baked for Max* with its active counterpart *Steffi baked Max a cake*).

Consequently, constructions have to be simultaneously stored in multiple forms that differ in their level of complexity and abstraction. To give a simple example of different levels of constructional complexity, the words *nut* and the plural -*s* morpheme are simple constructions; both are stored also as constituent parts of the more complex construction *nuts* ('more than one nut'). Different levels of constructional abstraction (also referred to as schematization) are evident for example in the fully lexicalized formula *Thank you* vs. the partially schematized slot-and-frame greeting pattern [*Good* + (time of day)], which renders lexicalized phrases like *Good afternoon* and *Good evening*, and the completely schematic

[Adjective + Noun Phrase] construction, which in turn could be lexically specified as *happy camper*, *sweet tea*, or *grand finale*, to give but three examples.

Given this widely encompassing definition of constructions, the dividing line between the lexicon and the grammar, or words and rules, is blurred: rather than seeing a sentence as the product of applying a rule that strings a number of words into a particular order, a sentence is 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 Max give the squirrel*, for instance, is a combination of the following constructions:

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

An adult speaker's knowledge of their language(s), therefore, can be equated to a huge warehouse of constructions that vary in their degree of complexity and abstraction. Constructions come with properties that define if and how they can be combined with other constructions; for the most part, these properties are semantically and/or functionally motivated such that any two constructions can only be combined if their meanings/functions are compatible, or can at least temporarily attain compatibility in the specific context or discourse situation (Goldberg, 2006). How compatible two constructions are is crucially driven by the frequency with which they are used (and therefore, heard) together: the more often they co-occur, the more entrenched that particular constructional arrangement becomes. Conversely, we can predict that learners will acquire constructions first in the contexts of the constructions that they most often occur with in the input, and then gradually expand the repertoire of combinations to less frequent combinations and even acceptable novel combinations, provided a native-like knowledge of each construction's semantic and functional properties.

3.   The processes of language learning: Exemplar-based rational contingency analysis

In other words, language learning means learning the associations within and between constructions. Constructionist accounts of language acquisition involve the distributional analysis of the language stream and the parallel analysis of contingent perceptual activity, with abstract constructions being learned from the conspiracy of concrete exemplars of usage following statistical learning mechanisms,

relating input and learner cognition (see Rebuschat & Williams, 2012 on statistical learning mechanisms). Psychological analyses of this learning of constructions as form-meaning pairs is informed by the literature on the associative learning of cue-outcome contingencies that hinge on both construction-related and learner-related factors. For constructions, their frequency of experience, salience of form, significance of meaning, prototypicality, redundancy vs. surprise value, and the contingency of form and function seem to matter; for learners, cognitive factors like learned attention, automaticity, transfer, overshadowing, and blocking play a crucial role (Ellis, 2008b). These various psycholinguistic factors conspire in the acquisition and use of any linguistic construction (see Ellis & Wulff, 2015a, 2015b for a detailed discussion of each factor).

Research in psycholinguistics demonstrates that generally, the more frequently a construction (or combination of constructions) is experienced, the earlier it is acquired and the more fluently it is processed (Ellis, 2002). Words such as *one* or *give* occur more frequently than *sixteen* or *syndicate* – and the learner's perceptual system gradually attunes to the probabilities of these constructions in the input.

When a learner notices a construction for the first time, this can result in a unitary representation in memory that binds all its properties (i.e., phonological make-up, spelling, etc.) together. This representation subsequently receives activation whenever the construction's properties are present in the language environment, hence it serves as a form of detector, or pattern-recognition unit. Once the detector unit's activation threshold is met, it will fire. With each firing, the resting level of activation of the detector unit increases (and correspondingly, the threshold for firing decreases) – in other words, it is readied, or primed, for re-activation. This priming effect accrues over a speaker's lifespan such that frequently occurring constructions and the properties associated with them obtain habitually high resting activation levels.

The form-function mappings between a phonological form and its interpretation are strengthened through continued use in the same way: every encounter of /wʌn/ as *one* strengthens the association between the two; every encounter of /wʌn/ signaling *won* is tallied as well; as is the association between /wʌn/ when it is the initial part of *wonderland*.

Once a first memory representation is built, the language system compares each subsequent exemplar that the learner encounters in their language environment against that representation, and gradually alters and adapts it to fit the accumulating experience of that construction, its properties, and its contexts. As encounters with exemplars of a construction manifest similar or identical properties time and again, prototypes emerge that then serve as the basis of comparison for future encounters. Prototypes are knowledge representations that bind together the most typical properties of a construction. They are mental constructs in the sense that

they are abstractions that are the sum of a learner's encounters of sufficiently similar exemplars. Prototypes are the defining center pieces of categories by virtue of being maximally similar to other members of that category and maximally dissimilar to non-members of that category. For example, people are quicker to confirm that sparrows are birds than they are with other kinds of birds like geese or albatrosses. This is because sparrows are more prototypical birds: they unite the most typical features of birds in terms of size, beak shape, wing length, etc.

It is important to note that according to research in cognitive psychology, this adaptive fine-tuning of a learner's language representations is not conscious and explicit in nature, but happens unconsciously and implicitly. As far as properties of categories are concerned (whether it is a conceptual category like *bird* or a linguistic category like *noun phrase*), learners do not consciously take stock of their frequency in the cognitive/linguistic environment; instead, statistical learning happens unconsciously (Ellis, 1994; Rebuschat, 2015).

Another important tenet of usage-based theories in the context of categorization is that no principled distinction is drawn between linguistic and other cognitive categories. In the same way that speakers classify the world around them, so they classify the language that accompanies their experiences. Psycholinguistic research has demonstrated prototypicality, neighborhood, and other categorization effects in learning quasi-regular patterns of construction form. For instance, people are fastest when asked to produce regular forms (like, for example, plural *sparrow + s*), slower and less accurate at generating more marked forms (like *finch + es*), and slowest still to produce irregular forms (such as *geese*) (Chater & Manning, 2006; Seidenberg & Plaut, 2014).

4.   The language system emerges from usage

Over usage experience, form-function mappings are woven into a network of construction forms and their meanings. This language system is sometimes referred to as the "constructicon". Activation spreads through this network as a function of the learned probabilities of the different form-interpretation associations that a speaker has built over his/her lifespan. The resulting mental model that learners build is, at any time in language development, a custom-tailored, adaptively fine-tuned reflection of the learner's summed language experience (Ellis, 2006a). In that sense, language learning is rational as defined in the field of rational cognition: a major impetus for human psychology is to adapt behavior best as possible to its environmental conditions (Anderson, 1989). It is also emergent in the sense that learners employ few and simple learning mechanisms, yet the knowledge networks that arise from employing these mechanisms over time are complex, dynamic, and

adaptive (Ellis, 1998; Beckner et al., 2009; Ellis & Larsen–Freeman, 2009). Language is a complex-adaptive system in the sense that it involves many agents (people who communicate with each other) in many different configurations (individuals, groups, networks, and cultures); and it operates across many different levels of the system architecture (neurons, brains, and bodies; phonemes, constructions, interactions, and discourses) as well as on multiple time scales (evolution, epigenesis, ontogenesis, interactional, neuro-synchronic, and diachronic) (Ellis, Römer, & O'Donnell, 2016; MacWhinney & O'Grady, 2015).

5. Lexical and grammatical constructions in first and second language acquisition

Frequency of usage is a driving force of construction learning. However, not all constructions are equally learnable by all learners. In early stages of acquisition (and for a good number of learners, even after years of naturalistic exposure), learners tend to focus more in their language processing upon open-class words (nouns, verbs, adjectives and adverbs) than on grammatical cues. The limited language attainment of those learners who never move beyond that stage has been described as stabilizing at a "Basic Variety" of interlanguage that is less grammatically sophisticated than that of native-like L1 ability (Klein, 1992; Bardovi-Harlig, 1992). Although naturalistic L2 learners are surrounded by the available target language input, not all of it becomes intake, that subset of input that actually gets in and that the learner utilizes in some way (Corder, 1967). A classic case study illustrating the limitations of intake is that of the naturalistic language learner, Wes, who was described as being 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: 5).

Although the Basic Variety is sufficient for everyday communicative purposes, grammatical morphemes and closed-class words tend not to be put to full use (e.g., Bardovi-Harlig, 1992; Clahsen & Felser, 2006; Schmidt, 1984; Van Patten, 1996, 2006). So, for example, L2 learners initially make temporal references mostly by use of temporal adverbs, prepositional phrases, serialization, and calendric reference, with the grammatical expression of tense and aspect emerging only slowly thereafter, if at all (Bardovi-Harlig, 1992, 2000; Klein, 1998; Lee, 2002; Meisel, 1987; Noyau, Klein, & Dietrich, 1995). L2 learners have been found to prefer adverbial over inflectional cues to tense in naturalistic L2A (e.g., Bardovi-Harlig, 2000; Noyau et al., 1995), training experiments (e.g., Cintrón-Valentín & Ellis, 2015; Ellis et al.,

2014), and studies of L2 language processing alike (e.g., Sagarra & Ellis, 2013; Van Patten, 2007).

A key challenge for L2A research is therefore to explain why grammatical morphemes and closed-class constructions are more difficult to learn than open-class constructions, especially in early stages of language development. Usage-based theories attribute this to three standard phenomena of the psychology of learning: The learnability of a construction is affected by (i) salience, (ii) contingency of form-function association, and (iii) learned attention.

6.   Salience and learning

Learnability depends on salience: less salient cues are less readily learned than highly salient ones (Ellis, 2006c, in press; 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 are 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". A loved one stands out from the crowd, as does a stimulus with weighty associations ($500000.0 vs. $0.000005, 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: 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 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: 91).

Rescorla and Wagner (1972) presented a formal model of conditioning 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, and it elegantly encapsulates the three factors of psychophysical 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 US, *b* is the salience of the CS, 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.

One factor determining the learning of construction form is psychophysical salience. In his landmark study of first language acquisition, Brown breaks down the measurement of perceptual salience, or "clarity of acoustical marking" (1973: 343), into "such variables as amount of phonetic substance, stress level, usual serial position in a sentence, and so on" (1973: 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. This is a result of the well documented effect of frequency and automatization in the evolution of language. The basic principles of automatization that apply to all kinds of motor activities and skills (like playing a sport or a musical instrument) are that through repetition, previously independent sequences of units come to be processed as a single unit or chunk (Ellis, 1996). The more frequently they use a form, the more speakers abbreviate it: this is a law-like relationship across languages. Zipf (1949) summarized this in the principle of least effort – speakers want to minimize articulatory effort, and this leads to brevity and phonological reduction. They tend to choose the most frequent words, and the more they use them, automatization of production causes their shortening. Frequently used words become shorter with use. Grammatical functors are the most frequent

words of a language, thus they lose their emphasis and tend to become abbreviated and phonologically fused with surrounding material (Bybee, 2000; Jurafsky, Bell, Gregory, & Raymond, 2001; Zuraw, 2003). In a corpus study by Cutler and Carter (1987), 86% of strong syllables occurred in open class words and only 14% in closed-class words; for weak syllables, 72% occurred in closed-class words and 28% in open-class words.

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 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.

In sum, grammatical functors are 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 learners lack: they haven't had sufficient experience and corresponding retuning of their L1 system to come up with a sufficiently schematized knowledge system (or 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, 2006c; Goldschneider & DeKeyser, 2001).

These effects are also 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* he walk*ed*"), plural markers are accompanied by quantifiers or numerals ("*10* nut*s*"), 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 for everyday communicative purposes (Simon, 1957).

7. Contingency and learning

The degree to which animals, human and other alike, learn associations between cues and outcomes 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, 2006b; Gries & Ellis, 2015; Gries & Stefanowitsch, 2004; MacWhinney, 1987). The last thirty years of psychological investigation 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 1.

Table 1. A contingency table showing the four possible combinations of events showing the presence or absence of a target cue and an outcome

	Outcome	No outcome
Cue	*a*	*b*
No cue	*c*	*d*

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.

ΔP is the probability of the outcome given the cue $P(O|C)$ minus the probability of the outcome in the absence of the cue $P(O|\neg C)$, 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 $\Delta P = 0$. ΔP 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 forms and their interpretations. The less reliably a form is associated with a function or interpretation, the more difficult learning becomes (Ellis, 2006b; Shanks, 1995). 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 is therefore less important than the contingency between cue and interpretation.

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 (*squirrels*), it indicates possession (*Max's toy*), and it marks third person singular present (*Nick sleeps*). 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 1). 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*]. Thus if we evaluate just one of these, say [*ɨz*], as a cue for one particular outcome, say plurality, then it is clear that there are many instances of that outcome in the absence of the cue (*c* in Table 1). In other words, the low cue-interpretation contingency makes plurals difficult to learn.

This fact that many high frequency grammatical constructions (as well as all other kinds of constructions; see Gries, 2015) are highly ambiguous in their interpretations, poses a challenge to language learners (DeKeyser, 2005; Ellis, 2008a; Goldschneider & DeKeyser, 2001).

8.   Learned attention

L2A is subject to attentional biases which result from L2 learners' history of learning – from their knowledge of a prior language. Ellis (2006a, 2006c) attributes L2 difficulties in acquiring inflectional morphology to an effect of learned attention known as "blocking" (Kamin, 1969; Kruschke, June 2006; 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, 1995; Wills, 2005). Knowing that a particular stimulus is associated with a particular outcome makes it harder to learn that another cue, subsequently paired with that same outcome, is also a good predictor of it. The prior association "blocks" further associations. ALL languages have lexical and phrasal means of expressing temporality. 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, and their salience. 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*.

A number of theories of L2A incorporate related notions of transfer and learned attention. The Competition Model (MacWhinney, 2001; MacWhinney & Bates, 1989) was explicitly formulated to deal with competition between multiple linguistic cues to interpretation. Input Processing (IP) theory (Van Patten, 1996) includes a *Lexical Preference Principle*: "Learners will process lexical items for meaning before grammatical forms when both encode the same semantic information" (Van Patten, 2006: 118), and a *Preference for Nonredundancy Principle*: "Learners are more likely to process nonredundant meaningful grammatical markers before they process redundant meaningful markers" (Van Patten, 2006: 119). Benati (2013) reviews a series of studies showing learners are better able to identify temporal reference when presented with temporal adverbs rather than verbal morphology.

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; Verb Group Time = 0.76Adverb + 0.60Verb; Control Group Time = 0.93Adverb + 0.17Verb.

This experiment demonstrated how short-term instructional manipulations could affect attention to language. 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 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), L1 experience (morphologically rich L1 learners and monolinguals looked longer at verbs than

morphologically poor L1 learners and monolinguals), and L2 experience (inter-mediate learners read more slowly and regressed longer than advanced learners).

Experience with the second language is shaded by attentional biases and other types of interference from the first language. Transfer phenomena pervade SLA (Flege, 2002; Jarvis & Pavlenko, 2008; Lado, 1957; MacWhinney, 1997; Odlin, 1989). As a result of this interference, second language learning is rarely entirely native-like, even if the learner is surrounded by ambient input. Since everything is filtered through the lens of the L1, not all of the relevant input is in fact taken advantage of (hence Corder's distinction between input and intake; Corder, 1967).

It is important to emphasize here that the limitations of L2 learning 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. Rather, first language learning is (nearly always) so marvelously successful that it – paradoxically perhaps – hampers second language learning. First language learners have learned to attend to their language environment in one particular way. L2 learners are tasked with reconfig-uring the attentional biases of having acquired their first language (Slobin, 1996).

9.   Implications for language teaching

The fact that L2 learners have to learn to adjust their attention biases shaped by their L1 has consequences for L2 instruction. Children acquire their first language primarily in an implicit manner. *Implicit learning* is the learning of complex in-formation without selective attention to what is being learned. L2A, in contrast, is characterized in large parts by *explicit learning*. For reviews on implicit and explicit language learning see Ellis (1994); Rebuschat (2015).

Schmidt's (2001) *Noticing Hypothesis* holds that conscious attention to lin-guistic forms in the input is an important precondition to learning: "people learn about the things they attend to and do not learn much about the things they do not attend to" (Schmidt, 2001: 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. Given the bottleneck effect of input vs. intake discussed above even in dense-input, immersive environments, explicit learning and teaching gain even more relevance for the second language learner who receives only limited L2 input (as the typical foreign language learner learning their L2 through instruction does). This holds in particular for aspects of form in the L2 that are redundant and/or lack perceptual salience (like the above-mentioned examples of inflectional morphemes in English). Form-focused Instruction (FFI) attempts to encourage noticing, drawing learners' attention to linguistic forms that might otherwise be

ignored (Ellis, 2012). Variants of FFI vary in the degree and manner in which they recruit learner consciousness and in the role of the learner's metalinguistic awareness of the target forms.

Norris and Ortega's (2000) 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. However, the studies gathered in these meta-analyses used a wide variety of types of instruction, learner, targeted feature, and method of assessment. Future research should control for these factors to see how robust effects of FFI really are.

Cintrón-Valentín and Ellis (2015) and Cintrón-Valentín and Ellis (2016) 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 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 affects learners' attention during online processing and thus modulates long-term blocking of verb morphology.

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 (Ellis, 2005). Once a construction has been represented in this way, its use in subsequent implicit processing can update the statistical tallying of its frequency of usage and probabilities of form-function mapping.

10. Further reading

This chapter has focused upon salience, contingency, and learned-attention in usage-based explanations of L2 morphology acquisition. This is just a part of usage-based approaches to L2A. Ellis and Wulff (2015a, 2015b) and Ortega, Tyler, In Park, and Uno (2016) provide more detail of the broader approach. Ellis et al. (2016) describe a large body of complementary work showing the joint effects of type-token frequency, contingency, and prototypicality in usage-based L1 and L2 acquisition and processing of verb-argument constructions. Cadierno and Eskildsen (2015), Beckner et al. (2009), Douglas Fir Group (Atkinson (2016), and Hulstijn et al. (2014) marry the cognitive of usage-based approaches with the social, since so much of usage and attention-in-usage is socially driven. Robinson and Ellis (2008), Littlemore (2009) and Tyler (2012) give broader overviews of cognitive-linguistic research in L2 learning and teaching.

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:10.3758/BF03334492

Anderson, J. R. (1989). A rational analysis of human memory. In H. L. I Roediger & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 192–210). Hillsdale, NJ: Lawrence Erlbaum.

Bardovi-Harlig, K. (1992). The use of adverbials and natural order in the development of temporal expression. *IRAL*, 30, 299–320. doi:10.1515/iral.1992.30.4.299

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

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

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

Benati, A. (2013). The Input Processing Theory in second language acquisition. In M. del P. García Mayo, M. J. Gutierrez Mangado, M. Martínez Adrián (Eds.), *Contemporary approaches to second language acquisition* (pp. 93–110). Amsterdam: John Benjamins. doi:10.1075/aals.9.08ch5

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

Bybee, J. (2000). *Mechanisms of change in grammaticalization: The role of frequency*. Albuquerque, NM: University of New Mexico.

Cadierno, T., & Eskildsen, S. W. (Eds.). (2015). *Usage-based perspectives on second language learning*. Berlin: De Gruyter Mouton. doi:10.1515/9783110378528

Chater, N., & Manning, C. (2006). Probabilistic models of language processing and acquisition. *Trends in Cognitive Science*, 10, 335–344. doi:10.1016/j.tics.2006.05.006

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:10.1017/S0272263115000029

Cintrón-Valentín, M., & Ellis, N. C. (2016). Salience in second language acquisition: Physical form, learner attention, and instructional focus. *Frontiers in Psychology, section Language Science*, 7, 1284.

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, 27(1), 3–42. doi:10.1017/S0142716406060024

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

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

DeKeyser, R. (2005). What makes learning second-language grammar difficult? A review of issues. *Language Learning*, 55(S1), 1–25. doi:10.1111/j.0023-8333.2005.00294.x

Douglas Fir Group (Atkinson, D., Byrnes, H., Doran, M., Duff, P., Ellis, N., Hall, J. K., Johnson, K., Lantolf, J., Larsen-Freeman, D., Negueruela, E., Norton, B., Ortega, L., Schumann, J., Swain, M., & Tarone, E.). (2016). A transdisciplinary framework for SLA in a multilingual world. *Modern Language Journal*, 100, 19–47.

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

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

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

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

Ellis, N. C. (2006a). The Associative-Cognitive CREED. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 77–96). Mahwah, NJ: Lawrence Erlbaum Associates.

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

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:10.1111/j.1540-4781.2008.00716.x

Ellis, N. C. (2008b). Usage-based and form-focused language acquisition: The associative learning of constructions, learned-attention, and the limited L2 endstate. In P. Robinson & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 372–405). London: Routledge.

Ellis, N. C. (in press). Salience. In M. Hundt, S. Pfenninger & S. Mollin (Eds.), *The changing English language*. Cambridge: Cambridge University Press.

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

Ellis, N. C., & Larsen-Freeman, D. (Eds.). (2009). *Language as a complex adaptive system*. Boston, MA: Wiley-Blackwell.

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:10.1017/S0142716412000501

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: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:10.1017/S0272263111000325

Ellis, N. C., & Wulff, S. (2015a). Second language acquisition. In E. Dabrowska & D. Divjak (Eds.), *Handbook of cognitive linguistics* (pp. 409–431). Berlin: De Gruyter Mouton. doi:10.1515/9783110292022-020

Ellis, N. C., & Wulff, S. (2015b). Usage-based approaches in second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 75–93). London: Routledge.

Ellis, R. (2012). Form-focused instruction and second language learning. In R. Ellis (Ed.), *Language Teaching Research & Language Pedagogy* (pp. 271–306). Malden, MA: Wiley-Blackwell. doi:10.1002/9781118271643.ch9

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

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: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: John Benjamins. doi:10.1075/sibil.48.18goo

Gries, S. T. (2015). Polysemy. In E. Dąbrowska & D. S. Divjak (Eds.), *Handbook of cognitive linguistics* (pp. 472–490). Berlin: De Gruyter Mouton. doi:10.1515/9783110292022-023

Gries, S. T., & Ellis, N. C. (2015). Statistical measures for usage-based linguistics. *Currents in Language Learning*, 2, 228–255. doi:10.1111/lang.12119

Gries, S. T., & Stefanowitsch, A. (2004). Extending collostructional analysis: A corpus-based perspective on alternations. *International Journal of Corpus Linguistics*, 9, 97–129. doi:10.1075/ijcl.9.1.06gri

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

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:10.1006/jmla.1997.2514

Hulstijn, J. H., Young, R. F., Ortega, L., Bigelow, M. H., DeKeyser, R., Ellis, N. C., & Talmy, S. (2014). Bridging the gap: Cognitive and social approaches to research in second language learning and teaching. *Studies in Second Language Acquisition*, 33(3), 473–482.

James, W. (1890). *The principles of psychology* (Vol. 2). New York, NY: Holt. doi:10.1037/11059-000

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

Jurafsky, D., Bell, A., Gregory, M., & Raymond, W. D. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 229–254). Amsterdam: John Benjamins. doi:10.1075/tsl.45.13jur

Kamin, L. J. (1969). Predictability, surprise, attention, and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior* (pp. 276–296). New York, NY: Appleton-Century-Crofts.

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

Kruschke, J. K. (June 2006, May 31–June 3, 2006). Learned attention. Paper presented at the Fifth International Conference on Development and Learning, Indiana University.

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

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

Lee, J. F. (2002). The incidental acquisition of Spanish future morphology through reading in a second language. *Studies in Second Language Acquisition*, 24, 55–80. doi:10.1017/S0272263102001031

Littlemore, J. (2009). *Applying cognitive linguistics to second language learning and teaching*. Houndmills: Palgrave Macmillan. doi:10.1057/9780230245259

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

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

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. (2001). The competition model: The input, the context, and the brain. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 69–90). Cambridge: Cambridge University Press. doi:10.1017/CBO9781139524780.005

MacWhinney, B., & Bates, E. (1989). *The crosslinguistic study of sentence processing*. Cambridge: Cambridge University Press.

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

Meisel, J. (1987). Reference to past events and actions in the development of natural second language acquisition. In C. Pfaff (Ed.), *First and second language acquisition* (pp. 206–224). New York, NY: Newbury House.

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:10.1037/h0043158

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

Noyau, C., Klein, W., & Dietrich, R. (1995). *Acquisition of temporality in a second language*. Amsterdam: John Benjamins.

Odlin, T. (1989). *Language transfer*. Cambridge: Cambridge University Press. doi:10.1017/CBO9781139524537

Ortega, L., Tyler, A. E., In Park, H., & Uno, M. (Eds.). (2016). *The usage-based study of language learning and multilingualism*. Washington, DC: Georgetown University Press.

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

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

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

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

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

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:10.1017/S0272263112000885

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

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:10.1111/cogs.12147

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

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

Slobin, D. (1996). From 'thought and language' to 'thinking for speaking'. In J. J. Gumperz & S. 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, 263–308.   doi:10.1111/j.1467-9922.2010.00562.x

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

Trousdale, G., & Hoffmann, T. (Eds.). (2013). *Oxford handbook of construction grammar*. Oxford: Oxford University Press.

Tyler, A. (2012). *Cognitive linguistics and second language learning: Theoretical basics and experimental evidence*. Abingdon: Routledge.

VanPatten, B. (1996). *Input processing and grammar instruction in second language acquisition*. Norwood, NJ: Ablex.

VanPatten, B. (2006). Input processing. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 115–136). Mahwah, NJ: Lawrence Erlbaum Associates.

VanPatten, 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.

Wills, A. J. (2005). *New directions in human associative learning*. Mahwah, NJ: Lawrence Erlbaum Associates.

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

Zuraw, K. (2003). Probability in language change. In R. Bod, J. Hay, & S. Jannedy (Eds.), *Probabilistic linguistics* (pp. 139–176). Cambridge, MA: The MIT Press.