2

SALIENCE IN USAGE-BASED SLA

Nick Ellis

Salience in Psychology, Learning Theory, and Psycholinguistics

Psychological research uses the term salience to refer to the property of a stimulus to stand out from the rest. Salient items or features are attended, are more likely to be perceived, 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. It is useful to think of three aspects of salience, one relating to psychophysics, the other two to what we have learned:

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. These phenomena are the subject of research in psychophysics (Gescheider, 2013).
2. As we experience the world, we learn from it, and our resultant knowledge values some associations more heavily than others. We know that some stimulus cues are associated with outcomes or possibilities that are important to us, while others are negligible (Gibson, 1977; James, 1890b).
3. We also have expectations about what is going to happen next in known contexts, we are surprised when our expectations are violated, and we pay more attention as a result. These phenomena are the subject of research in associative learning and cognition (Anderson, 2009; Shanks, 1995).

Psychophysical Salience

Loud noises, bright lights, and moving stimuli capture our attention. Salience arises in sensory data from contrasts between items and their context. These stimuli deliver

intense signals in the psychophysics of our data-driven perception. Stimuli with unique features compared to their neighbors (Os in a field of Ts, a red poppy in a field of yellow) "pop out" from the scene, but in a shared feature context will not (Os among Qs) (Treisman & Gelade, 1980). These are aspects of bottom-up processing (Shiffrin & Schneider, 1977).

Salient Associations

Attention can also be driven by top-down, memory-dependent, expectation-driven processing. Emotional, cognitive, and motivational factors affect the salience of stimuli. These associations make a stimulus cue "dear." A loved one stands out from the crowd, as does a stimulus with weighty associations ($500,000.0 versus $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: "The chief cerebral conditions of perception are the paths of association irradiating from the sense-impression, which may have been already formed" (James, 1890a, p. 82). Psychological salience is experience-dependent: *hot dog*, *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 be measured in physical terms. They are subjective. Hence George 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, p. 91).

Context and Surprisal

The evolutionary role of cognition is to predict what is going to happen next. Anticipation affords survival value. The Rational Analysis of Cognition (Anderson, 1990, 1991) is guided by the principle that human psychology can be understood in terms of the operation of a mechanism that is "optimally adapted" to its environment in the sense that the behavior of the mechanism is as efficient as it conceivably could be given the structure of the problem space and the input-outputs mapping it must solve. We find structure in time (Elman, 1990). The brain is a prediction machine (Clark, 2013). One consequence is that it is surprisal, when prediction goes wrong, that maximally drives learning from a single trial. Otherwise, the regularities of the usual course of our experiences sum little by little, trial after trial, to drive our expectations. Cognition is probabilistic, its expectations a conspiracy tuned from statistical learning over our experiences.

Salience and Learning

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 more than 80 years of research in associative learning, and it elegantly encapsulates the three factors of psychophysical salience, psychological importance, 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 side of the equation, 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 US. So 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: "But habit is a great deadener" (Beckett, 1954). Alternatively, with novel associations where V is close to zero, there is much surprisal, and consequently much learning: first impressions, first kiss, first love, first time, etc.

This is arguably the most influential formula in the history of learning theory. Physical salience, psychological salience, and expectation/surprisal all affect what we learn from our experiences of the world.

Cognitive Linguistics and Construction Grammar

Language is intrinsically symbolic. Linguistic forms (cues) are associated with particular meanings or interpretations (outcomes). Cognitive Linguistics calls the units of language 'constructions.' These are form–meaning mappings, conventionalized in the speech community, and entrenched as language knowledge in the learner's mind. Constructions relate the defining properties of their morphological, lexical, and syntactic form with particular semantic, pragmatic, and discourse functions (Goldberg, 1995, 2006). Construction Grammar (Goldberg, 2006; Trousdale & Hoffmann, 2013) argues that all grammatical phenomena can be understood as learned pairings of form (from morphemes, words, and idioms, to partially lexically filled and fully general phrasal patterns) and their associated semantic or discourse functions: "[T]he network of constructions captures our grammatical knowledge in toto, i.e. it's constructions all the way down" (Goldberg, 2006, p. 18). Such beliefs, increasingly influential in the study of child language acquisition, emphasize data-driven, emergent accounts of linguistic systematicities (e.g., Ambridge & Lieven, 2011; Tomasello, 2003).

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 can comprise concrete and particular items (as in words and idioms), more abstract classes of items (as in word classes and abstract constructions), or complex combinations of concrete and abstract pieces of language (as mixed constructions). Constructions may be simultaneously represented and stored in multiple forms,

at various levels of abstraction (e.g., concrete item: *table* + *s* = *tables* and [Noun] + (morpheme + *s*) = plural things). Constructions can thus be meaningful linguistic symbols in their own right, existing independently of particular lexical items. Nevertheless, constructions and the particular lexical tokens that occupy them attract each other, and grammar, morphology, and lexis are inseparable.

Usage-Based Approaches to First and Second Language Acquisition

Usage-based approaches to language learning hold that we learn linguistic constructions throughout our experience of using language to communicate. Psycholinguistic research provides the evidence of usage-based acquisition in its demonstrations that language processing is exquisitely sensitive to usage frequency at all levels of language representation from phonology, through lexis and syntax, to sentence processing (Ellis, 2002). That language users are sensitive to the input frequencies of these patterns entails that they must have registered their occurrence in processing. These frequency effects are thus compelling evidence for usage-based models of language acquisition which emphasize the role of input. Constructionist accounts of language learning involve the distributional analysis of the language stream and the parallel analysis of contingent perceptuo-motor activity, with abstract constructions being learned as categories from the conspiracy of concrete exemplars of usage following statistical learning mechanisms (Bybee & Hopper, 2001; Christiansen & Chater, 2001; Ellis, 2002; Jurafsky & Martin, 2009) relating input and learner cognition. Language knowledge involves statistical knowledge, so humans learn more easily and process more fluently high frequency forms and 'regular' patterns which are exemplified by many types and which have few competitors (e.g., MacWhinney, 2001). The language system emerges from the conspiracy of these associations. Ellis, Römer, and O'Donnell (2016) and Robinson and Ellis (2008) give more detail of usage-based approaches to SLA.

Lexical and Grammatical Constructions in SLA

Not all constructions are equally learnable by all learners. Even after years of naturalistic exposure, adult second language (L2) learners tend to focus more in their language processing upon open-class words (nouns, verbs, adjectives, and adverbs) than on grammatical cues. Their language attainment has been described as stabilizing at a "Basic Variety" of interlanguage that is less grammatically sophisticated than that of nativelike L1 ability (Bardovi-Harlig, 1992; Klein & Perdue, 1992). This phenomenon, if evident over many years, has been termed "fossilization" (Han & Odlin, 2006). Although naturalistic second language learners are surrounded by language input, the available target language, not all of it becomes intake, that subset of input that actually gets in and which the learner utilizes in some way (Corder, 1967). A classic case study 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, p. 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 SLA (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 second language acquisition research is therefore to explain why grammatical morphemes and closed-class constructions are more difficult to learn than open-class constructions. Usage-based theories attribute this to three standard learning phenomena relating to salience: The learnability of a construction is affected by: (1) psychophysical salience, (2) contingency of form-function association, and (3) learned attention.

The Psychophysical Salience of Linguistic Constructions

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, p. 343), into "such variables as amount of phonetic substance, stress level, usual serial position in a sentence, and so on" (1973, 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. 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, sequences of units that were previously independent 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, 2003; 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 90–100% correctly, adult native speakers can recognize them only 40–50% 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. Thus the low psychophysical salience of grammatical functors contributes to L2 learners' difficulty in learning them (Ellis, 2006b; Goldschneider & DeKeyser, 2001).

Contingency and Learning

The degree to which animals 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, 2006a; Gries & Ellis, 2015; Gries & Stefanowitsch, 2004; MacWhinney, 1987). The last 30 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 2.1.

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.

TABLE 2.1 A Contingency Table Showing the Four Possible Combinations of Events Relating the Presence or Absence of a Target Cue and an Outcome

	Outcome	*No Outcome*
Cue	*a*	*B*
No cue	*c*	*D*

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 the formula:

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

When these are the same, 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 ΔP = 0. ΔP approaches 1.0 as the presence of the cue increases the likelihood of the outcome. A learnable cue is one such that when the cue is there, the outcome is there, and when the cue is not there, neither is the outcome; that is, when *a* and *d* are large and *b* and *c* are small.

Construction Contingency

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, 2006a; Shanks, 1995). 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 (*toys*), it indicates possession (*Mary's toy*), and it marks third-person singular present (*Mary 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 2.1). 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 2.1). In other words, the low cue-interpretation contingency makes plurals 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, 2008; Goldschneider & DeKeyser, 2001).

Psychophysical Salience in Second Language Acquisition and Processing

Goldschneider and DeKeyser (2001) performed a detailed meta-analysis of the "morpheme order studies" that, in the 25 years following Brown's (1973) descriptions of first language acquisition, investigated the order of second language (L2)

acquisition of the grammatical functors, progressive *–ing*, plural *–s*, possessive *–s*, articles *a, an, the*, third-person singular present *–s*, and regular past *–ed*. These studies show remarkable commonality in the orders of acquisition of these functors across a wide range of learners of English as a second language (ESL). The meta-analysis investigated whether a combination of five determinants (perceptual salience, semantic complexity, morphophonological regularity, syntactic category, and frequency) could account for the acquisition order. Scores for perceptual salience were composed of three subfactors: the number of phones in the functor (phonetic substance), the presence/absence of a vowel in the surface form (syllabicity), and the total relative sonority of the functor. The major determinants that significantly correlated with acquisition order were: perceptual salience $r = 0.63$, frequency $r = 0.44$, morphophonological regularity $r = 0.41$. When these three factors were combined with semantic complexity and syntactic category in a multiple regression analysis, this combination of five predictors jointly explained 71% of the variance in acquisition order, with salience having the highest predictive power.

To illustrate this, Field (2008) had second language learners of English listen to authentic stretches of spoken English and, when pauses occurred at random intervals, they had to transcribe the last few words. The recognition of grammatical functors fell significantly behind that of lexical words, a finding that was robust across first languages and across levels of proficiency.

It is clear, therefore, that linguistic forms of low psychophysical salience are more difficult to perceive—and, as a consequence, to learn.

Learned Attention

There are other attentional factors which also affect the salience of grammatical functors. The first relates to their redundancy. Grammatical morphemes often appear in redundant contexts in which 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* toy*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 in 2), and the Basic Variety satisfices for everyday communicative purposes (Simon, 1957).

Still again, more importantly so, there are attentional biases that particularly affect L2A. These result from L2 learners' history of learning—from their knowledge of a prior language. Ellis (2006b) attributes L2 difficulties in acquiring inflectional morphology to an effect of learned attention known as "blocking" (Kamin, 1969; Kruschke, 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 SLA 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, p. 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, p. 119).

Learned Attention and Blocking in SLA

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 Control group illustrate the normal state of affairs when learners are exposed to utterance with both cues and learn from their combination. Multiple regression analysis, when the dependent variable was the 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,

Control Group Time = 0.93Adverb + 0.17Verb. The adverb cues far outweighed the verbal inflections in terms of learnability. We believe this is a result of two factors: 1) the greater salience of the adverbial cues, and 2) learned attention to adverbial cues, which blocks the acquisition of verbal morphology.

The two other groups reacted to the cues in quite different ways—the Adverb pretraining group followed the adverb cue, the Verb pretraining group tended to follow the verb cue: Adverb Group Time = 0.99Adverb −0.01Verb; Verb Group Time = 0.76Adverb + 0.60Verb. Pretraining on the verb in non-redundant contexts did allow acquisition of this cue when its processing was task-essential, but still, the adverb predominated.

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 versus 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 under the Control conditions when 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 and 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 early 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.

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. A total of 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 (intermediate learners read more slowly and regressed longer than advanced learners).

Such experiments demonstrate both short-term and long-term effects when sensitivity to lexical cues blocks subsequent acquisition of inflectional morphology. 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).

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 typically limited in success, 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). This is not to say that L2 learning is qualitatively different than 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 reconfiguring the attentional biases of having acquired their first language.

Four Reasons Why L2 Morphology is Difficult to Acquire From Usage-Based Learning

Summing up, grammatical functors abound in the input, but, as a result of: 1) their low salience, 2) their redundancy, 3) the low contingency of their form-function mappings, and 4) adult acquirers' learned attentional biases and L1-tuned automatized processing of language, they are simply not implicitly learned by many naturalistic learners whose attentional focus is on communication.

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 information 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) and Rebuschat (2015).

Schmidt's (2001) Noticing Hypothesis holds that conscious attention to linguistic forms in the input is an important precondition to learning: "[P]eople learn about the things they attend to and do not learn much about the things they do not attend to" (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. With restricted input, too, compared to L1A, explicit learning and teaching gain even more relevance for the second language learner.

This holds in particular for aspects of form in the L2 that are redundant and/or lack perceptual salience (like the previously 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. Long (1991) and Doughty and Long (2003) describe how a focus on meaning can be improved upon by periodic attention to language as object: during otherwise meaning-focused lessons, learners' attention is briefly shifted to linguistic code features, in context, to induce noticing. This is known as *focus on form*. Doughty and Williams (1998) give the following examples of focus-on-form techniques, ranging from less to more explicit: input flood, when texts are saturated with L2 models; input elaboration; input enhancement, when learner attention is drawn to the target through visual highlighting or auditory stress; corrective feedback on error, such as recasting; and input processing, when learners are given practice in using L2 rather than L1 cues.

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. There is need to compare FFI methods upon learning of the same target feature in similar populations of learners.

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 FFI: 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 when they were presented with solitary verb forms and trained on their English translations. 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 FFI 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.

Language Change: The Linguistic Cycle and Grammaticalization

We have been focusing upon language acquisition in particular speakers. Let us now integrate over the community of speakers. From patterns of language usage, processing, and acquisition, dynamic processes over diachronic timescales and synchronic states, there emerge what de 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. One of these is the "Linguistic Cycle" (Givón, 1971; Hodge, 1970; Van Gelderen, 2011) which describes paths of grammaticalization from lexical to functional category followed by renewal. Givón (1979, p. 209) schematized the process as:

"Discourse > syntax > morphology > morphophonemics > zero" and, more memorably, as "Yesterday's syntax is today's morphology."Hopper and Traugott (2003) focus upon morphologicalization as "Lexical item in specific syntactic context > clitic > affix" which leads in turn to "the end of grammaticalization: loss" (p. 140). Sometimes the form alone is lost; more usually, a dying form is replaced by a newer, usually periphrastic form with a similar meaning (p. 172). The periphrastic replacement is salient both psychophysically (it is several lexical items long), and, as an innovation, it is surprising.

Salience in Language Change

Linguistic evolution proceeds by natural selection from among the competing alternatives made available from the idiolects of individual speakers which vary among them (Croft, 2000; Mufwene, 2001, 2008). Since adults are typically less successful than children at language learning, language use by a high proportion of adult language learners typically means simplification, most obviously manifested in a loss of redundancy and irregularity and an increase in transparency (Trudgill, 2002a, 2002b). The 'Basic Variety' of interlanguage (Klein, 1998; Perdue, 1993) shows similarities with pidgins (Schumann, 1978) because pidgins are the languages that result from maximal contact and adult language learning (McWhorter, 2001). Veronique (1999, 2001) and Becker and Veenstra (2003) detail many parallels between the grammatical structures of French-based Creoles and the Basic Variety of interlanguage of learners of French as a second language, particularly in the 1:1 iconicity of their mapping of function and form (Andersen, 1984), their controller-first, focus-last constituent ordering principles, their lack of verbal morphology, and the order of development of their means of temporal reference.

McWhorter argues that the older a language, the more complexity it has—that is, the more it overtly signals distinctions beyond strict communicative necessity.

The most elaborate languages in these respects are those older, more isolated languages that are spoken by groups of people whose interactions are primarily with other speakers of the language and which thus are learned as native languages by children. But their linguistic complexities pose great difficulties to second language learners, prejudiced by L1 transfer, blocking and entrenchment. So some languages are easier for adults to learn, in an absolute sense, than others: "If one were given a month to learn a language of one's choice, I think one would select Norwegian rather than Faroese, Spanish rather than Latin, and Sranan rather than English" (Trudgill, 1983, p. 106). It is no accident that Faroese, as a low-contact language not subject to adult language learning, has maintained a degree of inflectional complexity which Norwegian has lost. Stasis allows a language, left to its own devices, to develop historical baggage—linguistic overgrowths that, however interesting, are strictly incidental to the needs of human exchange and expression (McWhorter, 2001, 2002, 2004).

Consider again English third-person present -s. English is no longer a language spoken primarily as an L1. The 375 million L1 speakers are in a very definite minority compared to the 750 million EFL and 375 million ESL speakers (Graddol, 2000). This preponderance of adult language learning of English is changing its nature. Seidlhofer (2004, p. 236) describes these changes as English is used across the world as a Lingua Franca. First and foremost on her list of observables is "'dropping' the third-person present tense -s (as in 'She look very sad')."

"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). As complex, adaptive systems, languages emerge, evolve, and change over time (Beckner et al., 2009; Croft, 2000; Ellis, 1998; Ellis & Larsen-Freeman, 2006; Ellis et al., 2016; Larsen-Freeman, 1997; MacWhinney & O'Grady, 2015). Just as they are socially constructed, so too they are honed by social usage (Cadierno & Eskildsen, 2015; Douglas Fir Group, 2016; Hulstijn et al., 2014). They adapt to their speakers. Because children are better language learners than adults (at least, as explained in this chapter's section "Learned Attention and Blocking in SLA," as a result of blocking and learned attention), languages that adults can learn are simpler than languages that only children can learn. Second language acquisition by adults changes the very nature of language itself, in ways that are understandable in terms of the psycholinguistics of salience and general principles of associative learning.

So salience pervades SLA, and L, and A, in no particular order of priority.

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.

Ambridge, B., & Lieven, E. (2011). *Child language acquisition: Contrasting theoretical approaches*. Cambridge: Cambridge University Press.

Andersen, R. W. (1984). The one to one principle of interlanguage construction. *Language Learning*, *34*, 77–95.

Anderson, J. R. (1990). *The adaptive character of thought.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Anderson, J. R. (1991). Is human cognition adaptive? *Behavioral & Brain Sciences*, *14*(3), 471–517.

Anderson, J. R. (2009). *Cognitive psychology and its implications* (7th ed.). New York: Worth Publishers.

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

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

Becker, A., & Veenstra, T. (2003). Creole prototypes as basic varieties and inflectional morphology. In C. Dimroth & M. Starren (Eds.), *Information structure and the dynamics of language acquisition* (pp. 235–264). Amsterdam: John Benjamins.

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*, Supplement 1, 1–26.

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

Bybee, J. (2003). Mechanisms of change in grammaticalization: The role of frequency. In B. D. Joseph and J. Janda (Eds.) *The handbook of historical linguistics* (pp. 602–623). Oxford: Blackwell.

Bybee, J., & Hopper, P. (Eds.). (2001). *Frequency and the emergence of linguistic structure.* Amsterdam: Benjamins.

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

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.

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.

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

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*, 181–204.

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

Croft, W. (2000). *Explaining language change: An evolutionary approach.* London: Longman.

Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech & Language*, *2*(3), 133–142.

DeKeyser, R. (2005). What makes learning second-language grammar difficult? A review of issues. *Language Learning*, *55*(S1), 1–25.

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

Doughty, C., & Long, M. (2003). Optimal psycholinguistic environments for distance foreign language learning. *Language Learning and Technology*, 7, 50–80.

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

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

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

Ellis, N. C. (1998). Emergentism, connectionism and language learning. *Language Learning*, *48*(4), 631–664.

Ellis, N. C. (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.

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

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

Ellis, N. C. (2006b). Selective attention and transfer phenomena in SLA: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics*, *27*(2), 1–31.

Ellis, N. C. (2008). The dynamics of second language emergence: Cycles of language use, language change, and language acquisition. *Modern Language Journal*, *41*(3), 232–249.

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.

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

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.

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.

Ellis, R. (2012). *Language teaching research and pedagogy*. Hoboken, NJ: Wiley-Blackwell.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*, 179–211.

Field, J. (2008). Bricks or mortar: Which parts of the input does a second language listener rely on? *TESOL Quarterly*, *42*(3), 411–432.

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

Gescheider, G. A. (2013). *Psychophysics: The fundamentals*. Hove, UK: Psychology Press.

Gibson, J. J. (1977). The theory of affordances. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing* (pp. 67–82). Hillsdale, NJ: Erlbaum.

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

Givón, T. (1979). *On understanding grammar*. New York: Academic Press.

Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure.* Chicago: University of Chicago Press.

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.

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.

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

Gries, S. Th., & Ellis, N. C. (2015). Statistical measures for usage-based linguistics. *Currents in Language Learning, 2*, 228–255.

Gries, S. Th., & Stefanowitsch, A. (2004). Extending collostructional analysis: A corpus-based perspective on alternations. *International Journal of Corpus Linguistics, 9*, 97–129.

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.

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

Hopper, P. J., & Traugott, E. C. (2003). *Grammaticalization* (2nd ed.). Cambridge: Cambridge University Press.

Hulstijn, J. H., Young, R. F., Ortega, L., Bigelow, M. H., DeKeyser, R., Ellis, N. C., Lantolf, J., Alison, M., & 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), 361–421.

James, W. (1890a). *The principles of psychology* (Vol. 2). New York: Holt.

James, W. (1890b). *The principles of psychology* (Vol. 1). New York: Dover.

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

Jurafsky, D., & Martin, J. H. (2009). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

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: Appleton-Century-Crofts.

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

Klein, W., & Perdue, C. (1992). *Utterance structure: Developing grammars again.* Amsterdam: John Benjamins.

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

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

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

Larsen-Freeman, D. (1997). Chaos/complexity science and second language acquisition. *Applied Linguistics*, *18*, 141–165.

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.

Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. d. Bot, R. Ginsberg, & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39–52). Amsterdam: Benjamins.

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

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

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.

McWhorter, J. (2001). The world's simplest grammars are creole grammars. *Language Typology*, *5*, 125–166.

McWhorter, J. (2002). *The power of Babel: A natural history of language*. San Franciso, CA: W. H. Freeman & Co.

McWhorter, J. (2004). The story of human language lecture 24: Language interrupted. *The Great Courses*. The Teaching Company Limited Partnership. Retrieved May 12, 2017 from http://www.thegreatcourses.com/

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.

Mufwene, S. S. (2001). *The ecology of language evolution*. Cambridge: Cambridge University Press.

Mufwene, S. S. (2008). *Language evolution: Contact, competition and change*. London: Continuum International Publishing Group.

Norris, J., & 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*. New York: Cambridge University Press.

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

Rebuschat, P. (Ed.). (2015). *Implicit and explicit learning of language*. Amsterdam: John Benjamins.

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.

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.

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.

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.

Schumann, J. H. (1978). *The pidginisation process: A model for second language acquisition.* Rowley, MA: Newbury House.

Seidlhofer, B. (2004). Research perspectives on teaching English as a Lingua Franca. *Annual Review of Applied Linguistics, 24*, 209–239.

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

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing II: Perceptual learning, automatic attending and a general theory. *Psychological Review, 84*(2), 127.

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

Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning, 60*, 263–308.

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

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition.* Boston, MA: Harvard University Press.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*(1), 97–136.

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

Trudgill, P. (1983). *On dialect: Social and geographical perspectives.* Oxford: Basil Blackwell.

Trudgill, P. (2002a). Linguistic and social typology. In J. K. Chambers, P. Trudgill, & N. Schilling-Estes (Eds.), *The handbook of language variation and change* (pp. 707–728). Oxford: Blackwell.

Trudgill, P. (2002b). *Sociolinguistic variation and change.* Edinburgh, United Kingdom: Edinburgh University Press.

Van Gelderen, E. (2011). *The linguistic cycle: Language change and the language faculty.* Oxford: Oxford University Press.

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

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

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

Veronique, D. (1999). L'émergence de catégories grammaticales dans les langues créoles: grammaticalisation et réanalyse. In J. Lang & Neumann-Holzschuh (Eds.), *Reanaylse*

und Grammatikaliesierung in den romanischen Sprachen (pp. 187–209). Tübingen: Max Niemeyer Verlag.

Veronique, D. (2001). Genèse(s) et changement(s) grammaticaux: quelques modestes leçons tirés de l'émergence des créoles et de l'acquisition des langues étrangères. In M. Matthey (Ed.), *Le changement linguistique: Évolution, variation, hétérogénéité: Travaux neufchâtelois de linguistique 34/35* (pp. 273–303). Neuchâtel, Switzerland.

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

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: MIT Press.