Studies in Second Language Acquisition, 2016, **38**, 341–351.
doi:10.1017/S027226311600005X

SALIENCE, COGNITION, LANGUAGE COMPLEXITY, AND COMPLEX ADAPTIVE SYSTEMS

Nick C. Ellis
University of Michigan

CONCEPTUALIZING COMPLEXITY

This special issue of *SSLA* focuses on the interactions among linguistic complexity, cognitive difficulty, individual differences, and instruction and development. We are asked: "How are learning outcomes in instructed SLA determined or mediated by the cognitive mechanisms that L2 learners bring to bear in the acquisition of complex/difficult L2 structures?"

The last thirty years have evinced remarkable activity in researching complexity in SLA. Yet recently, there have been calls to step back, to reflect and review. In their introduction here, Housen and Simoens identify the need for conceptual analysis and uniformity of measurement. Research quality is predicated on conceptual clarity. A variety of operationalizations results in a variety of research outcomes. Hence there is too little conformity to allow sensible patterns to emerge.

Bulté and Housen (2012) provided a detailed taxonomy of constructs of complexity relevant to SLA research with, at the top level, a clear distinction between linguistic complexity (defined in terms of the structural properties of a language feature or [sub]system) and L2 cognitive complexity (as psycholinguistic difficulty indexed in language processing and language learning). A major goal was to identify the "major components of L2 complexity, each of which can be independently analyzed or measured" (p. 21).

My brief commentary on this special issue, *Cognitive Approaches to Complexity and Instruction in Second Language Acquisition*, first discusses the measurement of one aspect of cognitive difficulty—salience—and describes how, although different aspects of salience might be independently measured in the stimulus and in the learner, the essence of salience and its effects on learning are emergent properties

© Cambridge University Press 2016

of the stimulus-learner-context complex. It then reviews the multifarious ways in which salience appears in each of the articles of this special issue. The moral of this analysis is that the study of complexity should be informed by relevant theoretical frameworks relating to emergence and complex adaptive systems.

MEASURING COGNITIVE DIFFICULTY: THE CASE OF SALIENCE

In their introduction to this special issue, Housen and Simoens differentiate between objective and subjective difficulty factors: Objective difficulty factors are learner-independent properties of the L2 features themselves and potentially make a given feature more or less difficult for all learners (e.g., its input frequency, perceptual saliency, L1-L2 similarity, and markedness, as well as its structural complexity). Of these, perceptual salience has long been considered important. In regard to first language acquisition, Brown (1973) concluded that "some role for salience is guaranteed; the child will not learn what he cannot hear" (p. 463), and that, as a determinant of learning, salience is thus more important than frequency of experience. In the field of SLA, 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*, and *the*; third-person singular present *-s*; and regular past *-ed*. They investigated the effects of five potential objective difficulty factors (perceptual salience, semantic complexity, morphophonological regularity, syntactic category, and frequency). 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. Salience proved to be the most significant influence on acquisition order ($r = 0.63$). Linguistic forms of low psychophysical salience are more difficult both to perceive and to learn (Ellis, 2006).

But let us consider the "objective" measurement of perceptual salience in more detail. Psychological research uses the term *salience* to refer to the ability of a stimulus to stand out from the rest (Ellis, in press-b). 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. It is useful to think of three aspects of salience, one relating to psychophysics and the other two to what we have learned: (a) 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). (b) As we experience the world, we learn from it, and our resultant knowledge values some associations higher than others. We know that some stimulus cues have affordances—they are associated with outcomes or possibilities that are important to us—whereas others are negligible (Gibson, 1977; James, 1890a, Chapter 11). (c) 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). Each of the three phenomena is explained in more detail subsequently.

Three Aspects of Salience

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 (e.g., *O*s in a field of *T*s or a red poppy in a field of yellow) "pop out" from the scene, but this is not the case when such stimuli appear in a shared feature context (*O*s among *Q*s) (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 ($500000.0 vs. $0.000005, however similar the amount of pixels, characters, or ink in their sensation) or one that 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, 1890b, p. 82). Psychological salience is hugely 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 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, given that 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, when prediction goes wrong, it is surprisal that maximally drives learning from a single trial. Otherwise, the regularities of the usual course of our experiences add up little by little, trial after trial, to drive our expectations. Cognition is probabilistic, its expectations a conspiracy tuned from statistical learning from our experiences (Ellis, 2002).

By these psychological accounts, then, "objective" salience is a property of the stimulus, and of the learner, and of his or her learning history, and of the context. It is not so objective or simple after all, and we will see that all these aspects drive learning, in interactive ways.

Salience and Learning

Rescorla and Wagner (1972) presented a formal model of conditioning that 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 salience, and surprisal. The role of the US surprise and of the 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 that 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 US. Thus 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, Act II). Alternatively, with novel associations in which V is close to zero, there is much surprisal, and consequently much learning: first impressions, first love, first time, and so on.

This is arguably the most influential formula in the history of learning theory. Physical salience, psychological salience, and surprisal all affect what we learn from our experiences of the world. But they do so interactively. It is not enough to measure them individually and expect additive effects. The essence of salience and its effects on learning are emergent properties of the stimulus-learner-context complex (Ellis, in press-b). The articles in this special issue illustrate these interactions and their implications for research and measurement.

SALIENT FEATURES OF THE ARTICLES IN THIS SPECIAL ISSUE

Silva and Roehr-Brackin (this issue) present a rich conceptual analysis of the factors that affect learning difficulty, including frequency, perceptual salience, communicative redundancy, the opacity of form-meaning mapping, and the relative schematicity of a linguistic construction. These are separately evaluated as determinants of difficulty in implicit learning and as characteristics of metalinguistic descriptions (schematicity, conceptual complexity, the technicality of metalanguage, and truth value) as they contribute to difficulty in explicit learning. They assess these factors for 13 points of English grammar by having three applied linguists rate each grammar point on each dimension and having 30 learners and 11 teachers rate each grammar point for difficulty, and then they use these measures to predict learners' performance on measures of implicit and explicit knowledge of these L2 English grammar points. Learners' holistic difficulty rankings correlated significantly with their performance on the measure of explicit knowledge. Although correlations based on teachers' holistic difficulty rankings did not reach statistical significance, the judgments of this group were the only ones that showed trends toward successful prediction of learners' performance on both the implicit and the explicit L2 measure.

The experiment is as rich as the conceptual analysis on which it is based, though of course it is limited by the low *n*s for teachers and applied linguists. Such ratings are very relevant to the measurement of *subjective* factors, and they give us an important way forward. But further research is needed to investigate how the dimensions of learning difficulty as assessed by the applied linguists affect student and teacher difficulty assessments, and clearly many more respondents are required.

Ideally, student perceptions of difficulty should additionally be gathered and compared across different ability levels. Furthermore, the *objective* concepts that Silva Rodríguez and Roehr-Brackin begin with (frequency, perceptual salience, communicative redundancy, the opacity of form-meaning mapping, the relative schematicity of a linguistic construction, etc.) need to be assessed in the input. Physical salience can be operationalized in terms of sonority, syllabicity, morpheme boundedness, and so forth. Corpus studies, in turn, can assess such physical salience factors, as well as frequency and the opacity/reliability of form-function mapping, among other factors (Collins & Ellis, 2009; Collins, Trofimovich, White, Cardoso, & Horst, 2009; Kempe & MacWhinney, 1998). We must address the rich research program ahead of us before we can understand the interactions among physical salience, learner perceptions of salience, and learner level, never mind the full range of factors in the language input and usage that contribute to learner difficulty as it varies with proficiency.

Godfroid's study (this issue) presented upper-intermediate L2 German learners with an input flood of spoken exemplars of a difficult morphological structure involving strong, vowel-changing verbs. Toward the end of exposure, the mandatory vowel change was omitted, yielding ungrammatical verb forms. Word monitoring and oral production were used to gauge the development of learners' implicit and explicit knowledge, respectively, through reaction time (RT) and accuracy data. Although 33 out of 38 L2 learners remained unaware of the ungrammatical verbs in the input flood, they showed significant sensitivity during listening, as evidenced by a RT slowdown on ungrammatical trials. The unaware learners also improved significantly from pretest to posttest on the word monitoring and oral production measures, supporting the simultaneous development of implicit and explicit knowledge under incidental learning conditions. This study adds to the experimental demonstrations of implicit SLA, and in so doing it reminds us that conscious subjective factors are not the only relevant determinants of cognitive complexity.

Della Putta (this issue) manipulates the context of learning in order to modulate salience to instructional advantage. Textual enhancement (TE) such as color-coding, boldfacing, and underlining can be used to enhance forms in written input and thus to prompt learners' further processing of these cues. His study investigates the learning of two syntactic structures in Italian (the prepossessive determiner and differential object marking), which each pose specific learning difficulties for Spanish-speaking learners of L2 Italian (SSLI) because they have to learn to add the possessive determiner (APD) to their interlanguage and they have to learn not to add the prepositional accusative (PA), a property of their L1. Group A read five texts in which the *absence* of the PA was textually enhanced; group B read the same texts, but in these the *presence* of the

APD was enhanced. The findings that *unlearning* the PA is harder than learning the APD present a wonderful paradox for the measurement of salience. Objective measures focus on the physical salience of the relevant features, and TE studies are concerned with the necessary range of physical manipulations (for example, here, "a two-cue TE, bold and color, bold being one of the most commonly used cues to enhance forms"; Della Putta, 2016, p. 226). Della Putta's dilemma (2016, p. 234) dealt with how to enhance "the presence of an absence" (i.e., indirect negative evidence) of a form that is grammatical and realized in the L1 but ungrammatical and absent in the L2.

Cerezo, Caras, and Leow (this issue) illustrate salience as surprisal. Their study evaluates the effectiveness of a psycholinguistically motivated educational video game, as compared to a teacher-centered classroom lesson, in helping learners develop the complex Spanish *gustar* constructions. Leow's research has always admirably emphasized the role of awareness and attention in learning, and this study uses think-aloud (TA) protocols to track learners' depths of processing, levels of awareness (hypothesis testing and rule formation), and activation of recent prior knowledge. The think-aloud that they cite illustrates the role of *surprisal* rather well:

> Ohhh! It is an object! Because, that's why *gustan* agrees with the subject. The class is pleasing to her, making her the object and the class the subject. That's why it doesn't follow a literal English translation! There we go. *I just had a breakthrough*. Thank God. And that's why I'm also doing the Spanish lab now while we're learning all about this. (Cerezo, Caras, & Leow, 2016, Online Appendix A)

Yalçın and Spada (this issue) present an individual differences study showing that the aptitude of grammatical inferencing contributed to learners' gains on *passive* but not *past progressive* constructions on a written measure, whereas another component of aptitude (i.e., memory) contributed to gains on the past progressive on an oral measure. They argue that different components of aptitude may be involved in learning easy and difficult structures, and that salience is a key determinant of difficulty. The past progressive is an "easy" structure because of its transparent form-meaning relationship and salience in the input. It is physically salient, as it is realized by a free morpheme (*was/were*) and a syllabic bound morpheme (*-ing*), all progressive markers appear as intact syllables, and there are no allomorphs for the *-ing* marker that would reduce its salience (Yalçın and Spada, 2016, p. 247). Thus they argue that the progressive is more phonetically and perceptually salient and hence that the same type of instruction provided for both structures was enough for learners to significantly progress in their knowledge of the past progressive but not the passive without language aptitude as a resource: Salience and aptitude are in interaction.

Finally, Tagarelli, Ruiz, Vega, and Rebuschat (this issue) show that adult learners can acquire L2 syntax under incidental learning conditions while processing sentences for meaning, and additionally that the inclusion of an explicit component (metalinguistic rule presentation) prior to meaning-oriented exposure results in a greater learning effect. There were also effects of structural salience, as well as interactions of salience and individual difference factors whereby participants with poor procedural learning abilities in an incidental condition performed well on complex but salient items. These sentences consisted of a subordinate clause followed by a main clause; the verb in the subordinate clause was clause-final, and the verb in the main clause was clause-initial. Debriefing questionnaires suggest that these sentences were particularly salient to learners, perhaps because they contained two sequential verbs. Performance was also higher on simple sentences and salient complex sentences than on less salient complex sentences for both groups, though the difference between the complex sentence types was marginal. In sum, this study demonstrates interactions between salience and the effectiveness of the instructional condition, between salience and procedural learning ability as an aptitude, and among salience, instruction, and explicit/implicit knowledge.

LANGUAGE AS A COMPLEX ADAPTIVE SYSTEM

Salience is complex. It pervades the studies here, but in diverse ways. Second language acquisition research catalogues a rich history of such demonstrations: For example, the efficacy of recasts depends on the salience of the linguistic structure and the learner-level/target gap (Long, 2006). Psychological research shows the same: For example, novice learners are often aware of the cues they are using in problem solving, whereas experts have automatized long ago and so, having forgotten their first steps, are more likely to give post hoc rationalizations (Ericsson & Simon, 1993; Nisbett & Wilson, 1977).

Salience pervades language change as well. The salience of linguistic forms emerges diachronically as a result of their frequency of usage. Frequently used words become shorter with use. Considerable practice results in automaticity of a word's production along with sound reduction, assimilation, and lenition—the loss and overlap of spoken gestures (Bybee, 2003, 2006). Zipf's law describes the law relating frequency and length, which occurs in all languages (Ellis, in press-a; Zipf, 1935). Grammatical functors are the more frequently used parts of language, which is why they are of low salience. In the linguistic cycle, eventually some grammatical markers wear away entirely and disappear. It is not that fluent native speakers don't perceive them: They know they are there, despite the minimal sensation, but they perceive them from top-down knowledge. But second-language speakers don't perceive them. This is why high-contact languages—languages with a large proportion of second language speakers—tend to be morphosyntactically and grammatically less complex than more isolated languages with a preponderance of L1 speakers (Ellis, 2008). The mix of speakers in usage changes the nature of language. An elegant and persuasive recent illustration of this adaptive dance can be found in Bentz, Verkerk, Kiela, Hill, and Buttery (in press). By analyzing hundreds of languages within and across language families, regions, and text types, they show that languages with greater levels of contact have lower lexical diversity—that is, they typically employ fewer word forms to encode the same information content. Bentz et al. argue that language evolution and change should be modeled as the coevolution of multiple intertwined adaptive systems: on the one hand, the structure of human societies and human learning capabilities and, on the other, the structure of language.

Salience is adaptively complex, involving multiple agents at multiple levels in interaction. We need to acknowledge this complexity and to adopt theoretical perspectives concerning language emergence, dynamic systems theory, and language as a complex adaptive system (Beckner et al., 2009; de Bot, Lowie, & Verspoor, 2007; Ellis, 1998; Ellis & Larsen-Freeman, 2006a, 2006b, 2009; Holland, 2014; Larsen-Freeman, 1997; Larsen-Freeman & Cameron, 2008; MacWhinney & O'Grady, 2015).[1] Usage, too, is adaptively complex, involving multiple agents at multiple levels in interaction. Language is learned from such usage (Ellis, Römer, & O'Donnell, 2016). Emergentist principles apply across language, no more so than when the focus, as here, is the complexity of language itself.

W. B. Yeats's poem "Among School Children" (Yeats, 1989/1928) asks the same question in two different ways: "O chestnut tree, great rooted blossomer, Are you the leaf, the blossom, or the bole?" and "How can we know the dancer from the dance?" In language—as in biology, as in art—however much we might desire it for the sake of simplifying research, it is hard to localize complexity.

Received 12 January 2016
Accepted 22 January 2016

NOTE

1. There is no better 25-minute introduction to complex adaptive systems than the TED talk by John Holland, one of the pioneers of the field: https://www.youtube.com/watch?v=nzHVGd22vak.

REFERENCES

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.

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., & Schoenemann, T. (2009). Language is a complex adaptive system: Position paper. *Language Learning*, *59*(Suppl. 1), 1–26.

Bentz, C., Verkerk, A., Kiela, D., Hill, F., & Buttery, P. (in press). Adaptive communication: Languages with more non-native speakers tend to have fewer word forms.

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

Bulté, B., & Housen, A. (2012). Defining and operationalising L2 complexity. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Investigating complexity, accuracy and fluency in SLA* (pp. 21–46). Amsterdam/Philadelphia: Benjamins.

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

Bybee, J. (2006). From usage to grammar: The mind's response to repetition. *Language*, *82*, 711–733.

Cerezo, L., Caras, A., & Leow, R. P. (2016). Effectiveness of guided induction versus deductive instruction on the development of complex Spanish "gustar" structures: An analysis of learning outcomes and processes. *Studies in Second Language Acquisition*, *38*(2), 265–291.

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

Collins, L., & Ellis, N. C. (Eds.). (2009). Input and second language construction learning: Frequency, form, and function [Special issue]. *Modern Language Journal*, *93*(2).

Collins, L., Trofimovich, P., White, J., Cardoso, W., & Horst, M. (2009). Some input on the easy/difficult grammar question: An empirical study. *The Modern Language Journal*, *93*(3), 336–353.

De Bot, K., Lowie, W., & Verspoor, M. (2007). A dynamic systems theory to second language acquisition. *Bilingualism: Language and Cognition*, *10*, 7–21.

Della Putta, P. (2016). The effects of Textual Enhancement on the acquisition of two non-parallel grammatical features by Spanish-speaking learners of Italian. *Studies in Second Language Acquisition*, *38*(2), 217–238.

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. (2006). 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. (in press-a). Chunking. In M. Hundt, S. Pfenninger, & S. Mollin (Eds.), *The changing English language*. Cambridge: Cambridge University Press.

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

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

Ellis, N. C., & Larsen-Freeman, D. (2006b). Language emergence: Implications for applied linguistics (Introduction to the special issue). *Applied Linguistics*, *27*(4), 558–589.

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

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.
Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*, 179–211.
Ericsson, K., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (2nd ed.). Boston: MIT Press.
Gescheider, G. A. (2013). *Psychophysics: The fundamentals*. Hove Sussex: Psychology Press.
Gibson, J. J. (1977). The theory of affordances. In R. Shaw & J. Bransford (Eds.), *Perceiving, Acting, and Knowing*. Hillsdale, N.J.: Erlbaum.
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.
Holland, J. H. (2014). *Complexity: A very short introduction*. Oxford: Oxford University Press.
James, W. (1890a). *The principles of psychology* (Vol. 1). New York: Dover.
James, W. (1890b). *The principles of psychology* (Vol. 2). New York: Holt.
Kempe, V., & MacWhinney, B. (1998). The acquisition of case-marking by adult learners of Russian and German. *Studies in Second Language Acquisition*, *20*, 543–587.
Larsen-Freeman, D. (1997). Chaos/complexity science and second language acquisition. *Applied Linguistics*, *18*, 141–165.
Larsen-Freeman, D., & Cameron, L. (2008). *Complex systems and applied linguistics*. Oxford: Oxford University Press.
Long, M. H. (2006). Recasts in SLA: The story so far. In M. Long (Ed.), *Problems in SLA* (pp. 75–116). Mahwah, NJ: Lawrence Erlbaum.
MacWhinney, B., & O'Grady, W. (Eds.). (2015). *The handbook of language emergence*. Oxford: Wiley-Blackwell.
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.
Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*, 231–259.
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.
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.
Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136.
Yalçin, S., & Spada, N. (2016). Language aptitude and grammatical difficulty. *Studies in Second Language Acquisition*, *38*(2), 239–263.
Yeats, W. B. (1989). *Yeats's poems*. London: Macmillan. (Originally published 1928).
Zipf, G. K. (1935). *The psycho-biology of language: An introduction to dynamic philology*. Cambridge, MA: M.I.T. Press.