2

Cognitive Approaches to Second Language Acquisition

Nick C. Ellis and Stefanie Wulff

2.1 Introduction

Cognitive approaches to L2 acquisition minimally share these two assumptions:

- The primary source for both first (L1) and second language (L2) learning is the learner's participative, contextualized experience of language. Language learning is largely usage-based. Humans use language in order to communicate and make meaning.
- The cognitive mechanisms that learners employ in language learning are not exclusive to language learning, but are general cognitive mechanisms associated with learning of any kind.

In this chapter, we describe the constructs and working assumptions that characterize such approaches to language learning, with a particular focus on their cognitive underpinnings and how these explain differences between the linguistic forms that distinguish L1 and L2 speakers. We first define constructions as the targets of language learning and then describe the processes of construction learning in terms of exemplar-based, rational, associative learning. Not all constructions are equally learnable by all learners: naturalistic second language learners process open-class words more efficiently than grammatical cues even though the grammatical cues may be more frequent. We outline a usage-based account of this phenomenon in terms of salience, contingency, and redundancy, and explain how effects of learned attention and blocking further limit learning in adult L2 learners. We describe educational interventions taking these findings into consideration and conclude with further readings on usage-based approaches to L2 acquisition.

https://doi.org/10.1017/9781108333603.003 Published online by Cambridge University Press

There are other relevant and interesting aspects of usage-based second language acquisition (SLA) that we simply cannot deal with here. We have made the conscious choice to focus in this chapter upon L1–L2 differences in morphosyntax. In the final section, we provide pointers to more social, interactional, and meaning-based investigations of L2 cognition.

2.2 Constructions as the Targets of L2 Acquisition

Learning a language involves the learning of constructions, which are the conventionalized form–meaning mappings used in a speech community. Constructions include morphemes—the smallest pairing of form and meaning in language—as well as words, phrases, and syntactic frames (Goldberg, 2006; Trousdale & Hoffmann, 2013). Simple morphemes such as *–aholic* (meaning "being addicted to something") are constructions in the same way as simple words like *nut* (meaning "a fruit consisting of a hard or tough shell around an edible kernel"), idioms like *It is driving me nuts* (meaning "It is greatly frustrating me"), and abstract syntactic frames like Subject–Verb–Object–Object (meaning that something is being transferred to someone, as realized in sentences as diverse as *Max gave the squirrel a nut, Nick gave Max a hug*, or *Steffi baked Max a cake*, where nuts, hugs, and cakes are being transferred, respectively). Including abstract syntactic frames means that not all constructions carry meaning in the traditional sense, but rather serve a functional or meaningful purpose. For example, the passive construction serves the function of shifting the attentional focus in an utterance from the agent of the action to the patient undergoing the action (compare the passive *A cake was baked for Max* with its active counterpart *Steffi baked Max a cake*).

From this, it follows that constructions have to be stored in multiple forms simultaneously that differ in their level of complexity and abstraction. For instance, the word *nut* and the plural *–s* morpheme are both simple constructions that are also stored as constituent parts of the more complex construction *nuts* ("more than one nut"). Different levels of constructional abstraction (also referred to as schematization) are evident in the fully lexicalized formula *You're welcome* versus the partially schematized slot-and-frame greeting pattern [*Good* + (time of day)], which renders lexicalized phrases like *Good afternoon* and *Good evening*, and the completely schematic [Adjective + Noun Phrase] construction, which in turn could be lexically specified as *happy baby, delicious cake*, or *grand opening*, to give just three examples.

This wide definition of constructions entails a blurred dividing line between the lexicon and the grammar, or what traditional approaches have labelled words and rules: from a construction grammar perspective,

https://doi.org/10.1017/9781108333603.003 Published online by Cambridge University Press

a sentence is not the product of applying a rule that strings a number of words into a particular order, but the product of combining a number of constructions—some simple, some complex, some lexically specific, some abstract—in a particular way. A sentence like *What did Max give the squirrel*, for instance, combines the following constructions:

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

We can therefore equate an adult speaker's knowledge of their language(s) to a huge warehouse of constructions that vary in terms of complexity and abstraction. Constructions have properties that specify if and how they can combine with other constructions; these properties are mostly semantically and/or functionally motivated such that constructions can only be combined if their meanings or functions are compatible or can at least temporarily attain compatibility in a specific context or discourse situation (Goldberg, 2006). Constructional compatibility is crucially solidified by the frequency with which they are used (and therefore, heard) together: the more often they co-occur, the more entrenched that particular constructional arrangement becomes. Likewise, L2 learners will acquire constructions first in the contexts of the constructions with which they most often co-occur in the input before they gradually expand the repertoire of combinations to less frequent combinations and even acceptable novel combinations.

2.3 The Processes of L2 Acquisition: Exemplar-Based Rational Contingency Analysis

In other words, language learning means learning the associations within and between constructions. Constructionist accounts of language acquisition involve the distributional analysis of the language stream and the parallel analysis of meaning in terms of contingent perceptual experience, with abstract constructions being learned from the conspiracy of concrete exemplars of usage following statistical learning mechanisms relating input and learner cognition (see Rebuschat & Williams, 2012, on statistical learning mechanisms). Psychological analyses of this learning of constructions is informed by the literature on the associative learning of cue–outcome contingencies that hinge on both construction-related and learner-related factors. For constructions, their frequency of experience, salience of form, significance of meaning, prototypicality, redundancy versus surprise value, and the contingency of form and function seem to

https://doi.org/10.1017/9781108333603.003 Published online by Cambridge University Press

be relevant factors; for learners, cognitive factors like learned attention, automaticity, transfer, overshadowing, and blocking each play important roles (Ellis, 2008b). These various psycholinguistic factors conspire in the acquisition and use of any linguistic construction (see Ellis & Wulff, 2015a, 2015b, for a detailed discussion of each factor).

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

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

In the same fashion, the form–function mappings between a phonological form and its interpretation are strengthened through continued use: every encounter of /wʌn/ as *one* strengthens the association between the two; every encounter of /wʌn/ signalling *won* is tallied as well; as is the association between /wʌn/ when it is the initial part of *wonderland*.

After a first memory representation is formed, the language system compares each subsequent exemplar that the learner encounters in their language environment against that representation, and gradually modifies it to fit the accumulating experience of that construction, its properties, and its contexts. Since repeated encounters with exemplars of a construction manifest similar or identical properties time and again, prototypes emerge that then serve as the basis of comparison for future encounters. Prototypes are knowledge representations of the most typical properties of a construction. They are mental constructs in the sense that they are abstractions of a learner's accumulated encounters of sufficiently similar exemplars. Prototypes are the defining centrepieces of categories: they are maximally similar to other members of that category and maximally dissimilar to non-members of that category. For example, people are quicker to confirm that sparrows are birds than they are with other kinds of birds, like geese or albatrosses. This is because sparrows are more prototypical

https://doi.org/10.1017/9781108333603.003 Published online by Cambridge University Press

birds: they unite the most typical characteristics of birds in terms of size, beak shape, wing length, etc.

Importantly, this adaptive fine-tuning of a learner's language representations is not conscious and explicit in nature, but happens unconsciously and implicitly. As far as properties of categories are concerned (whether it is a conceptual category like *bird* or a linguistic category like *noun phrase*), learners do not consciously inventory frequencies in the cognitive and linguistic environment; instead, statistical learning happens unconsciously (Ellis, 1994; Rebuschat, 2015).

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

2.4 Usage Leads to an Emerging Language System

Through usage experience, form–function mappings are woven into a network of construction forms and their meanings. This language system is sometimes referred to as the "constructicon". Through this network, activation spreads as a function of the learned probabilities of the different form–meaning associations that a speaker has formed over his or her lifespan. The resulting mental model is, at any time in language development, a custom-tailored, adaptively fine-tuned reflection of the learner's summed language experience (Ellis, 2006a). In that sense, language learning is rational as defined in the field of rational cognition: a major impetus for human psychology is to adapt behaviour as best as possible to its environmental conditions (Anderson, 1989). Language learning is also emergent in the sense that the mechanisms that learners employ are few and simple, yet the knowledge networks that arise from employing these mechanisms over time are complex, dynamic, and adaptive (Ellis, 1998; Beckner et al., 2009; Ellis & Larsen-Freeman, 2009). Furthermore, language is a complex-adaptive system in the sense that it involves many agents (people communicating with each other) in many different configurations (individuals, groups, networks, and cultures), and it operates across many different levels of the system architecture (neurons, brains, and bodies; phonemes, constructions, interactions, and discourses), as well as on multiple time scales (evolution, epigenesis, ontogenesis, interactional, neuro-synchronic, and diachronic; Ellis, Römer, & O'Donnell, 2016; MacWhinney & O'Grady, 2015).

2.5 Lexical and Grammatical Constructions in L1 and L2 Acquisition

As stated above, the frequency of use of a construction drives its learnability. That said, not all frequent constructions are equally learnable to all learners. In early stages of acquisition (and for many learners, even after years of language immersion), learners process open-class words (nouns, verbs, adjectives, and adverbs) more efficiently than grammatical cues. The limited ultimate attainment of L2 learners who remain at that stage stabilizes at a "Basic Variety" of interlanguage that is less grammatically sophisticated than that of native-like L1 ability (Klein & Perdue, 1992; Bardovi-Harlig, 1992). Although naturalistic L2 learners are exposed to rich target language input, only a subset of that input turns into intake that promotes further language development (Corder, 1967). A classic case study demonstrating the limitations of intake is that of the naturalistic language learner Wes, who was described as being very fluent, with high levels of strategic competence, but low levels of grammatical accuracy: "using 90% correct in obligatory contexts as the criterion for acquisition, none of the grammatical morphemes counted have changed from unacquired to acquired status over a five year period" (Schmidt, 1984, p. 5).

The Basic Variety might be sufficient for everyday communicative purposes, but grammatical morphemes tend not to be put to full use (e.g., Bardovi-Harlig, 1992; Clahsen & Felser, 2006; Schmidt, 1984; VanPatten, 1996, 2007). For example, learners initially reference time 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 L2 acquisition (e.g., Bardovi-Harlig, 2000; Noyau, Klein, & Dietrich, 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; VanPatten, 2007).

A key challenge for L2 acquisition research is therefore to explain why closed-class constructions are more difficult to learn than open-class constructions, for some learners throughout their L2 career. Usage-based theories attribute this to three tenets of the psychology of learning: the learnability of a construction is affected by (i) salience, (ii) contingency of form–function association, and (iii) learned attention.

https://doi.org/10.1017/9781108333603.003 Published online by Cambridge University Press

2.6 Salience and Learning

Less salient cues are less readily learned than highly salient ones (Ellis, 2006c, 2017; Rescorla & Wagner, 1972). Salience refers to the property of a stimulus which makes it stand out from the rest and consequently more likely to be perceived, attended to, and entered into subsequent cognitive processing and learning. Salience can be determined independently by physics and the environment, and by our knowledge of the world:

- The physical world, our embodiment, and our sensory systems jointly render certain sensations to be more intense (louder, brighter, heavier, etc.) than others.
- As we experience the world, we learn from it, and our resultant knowledge values some associations higher than others (James, 1890, p. 82). A favoured stimulus stands out, either because of weighty associations ($500000.0 vs. $0.000005, however similar the number of pixels, characters, or ink in their sensation) or because it matches a motivational state (a meal when hungry but not when full).

Psychological salience is experience-dependent: hotdog, sushi, and 寿司 mean different things to people of different cultural and linguistic experience. This is why, unlike sensation, the units of perception are subjective and therefore cannot simply be measured in physical terms. This is reflected in 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).

Rescorla and Wagner (1972) presented a formal model of conditioning that describes 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 encounter of their pairing. The formula below is arguably the most influential formula in the history of learning theory, encapsulating over eighty years of research. It elegantly unites psychophysical salience, psychological salience, and surprisal:

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

The associative strength of the US to the CS is referred to by the letter V, and the change in this strength which occurs on each trial of conditioning is called dV. On the right-hand side, a is the salience of the US, b is the salience of the CS, and L is the amount of processing given to a completely unpredicted, surprising, US. Thus, both the salience of the cue (a) and the psychological importance of the outcome (b) are essential factors in any associative learning. As for ($L - V$), the more a CS is associated with a US,

https://doi.org/10.1017/9781108333603.003 Published online by Cambridge University Press

the less additional association the US can induce. As Beckett (1954) put it: "habit is a great deadener." Alternatively, with novel associations where *V* is close to zero, there is much surprisal, and consequently much learning: first impressions, first love, first time …

One factor determining the learning of construction form is psychophysical salience. In his landmark study of first language acquisition, Brown (1973) breaks down the measurement of perceptual salience, or "clarity of acoustical marking" (p. 343), into "such variables as amount of phonetic substance, stress level, usual serial position in a sentence, and so on" (p. 463). Temporal phrases (such as *on my birthday*, *at Christmas*, etc.), temporal adverbs (*yesterday*, *tomorrow*, *later*, etc.), and other lexical temporal cues (*morning*, *winter*, etc.) are salient and stressed in the speech stream, while verb inflections usually are not.

Many grammatical form–function relationships in English, such as grammatical particles and inflections like the third person singular *–s*, are of low salience in the language stream. This is a consequence of the well-documented effect of frequency and automatization in the evolution of language. The basic principles of automatization apply to all kinds of motor activities and skills (like playing a sport or a musical instrument) as well as to languages across the world: through repetition, previously independent sequences of units come to be processed as a single unit or chunk (Ellis, 1996). The more frequently they use a form, the more speakers reduce it. 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. Grammatical functors are the most frequent words of a language, and so they lose their emphasis and tend to become abbreviated and phonologically fused with surrounding material (Bybee, 2008; Jurafsky et al., 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.

Since grammatical function words and bound inflections tend to be short and unstressed, they are difficult to perceive from the input. When words are clipped out of connected speech and presented in isolation, adult native speakers perceive open-class words (*buy, four, know, ewe*, etc.) 90 to 100% correctly, but grammatical function words (*by, for, no, you*, etc.) only 40% to 50% of the time (Herron & Bates, 1997). Clitics (accentless words or particles that depend accentually on an adjacent accented word and form a prosodic unit together with it) are the most extreme examples of this: the /s/ of "he's", /l/ of "I'll" and /v/ of "I've" can never be

https://doi.org/10.1017/9781108333603.003 Published online by Cambridge University Press

pronounced in isolation. If native speakers have trouble correctly perceiving grammatical function words in isolation, the task for the L2 learners is even greater.

In sum, grammatical functors are extremely difficult to perceive based on bottom-up auditory evidence alone. Fluent language processors provide top-down schematic support to perceive these elements in continuous speech. However, this top-down knowledge is exactly what learners lack: they haven't had sufficient experience in the L2 and corresponding retuning of their L1 system to develop a sufficiently schematized knowledge system (or constructicon) that would afford them the same levels of L2 top-down support as in fluent L1 processing. Thus, the low psychophysical salience of grammatical functors contributes to L2 learners' difficulty in learning them (Ellis, 2006c; Goldschneider & DeKeyser, 2001).

Salience effects are compounded by redundancy. Grammatical morphemes often appear in redundant contexts where their interpretation is not essential for correct interpretation of the sentence (Schmidt, 2001; Terrell, 1991; VanPatten, 1996). For instance, tense markers often appear in contexts where other cues have already established the temporal reference (e.g. *yesterday he walked*), plural markers are accompanied by quantifiers or numerals (*10 nuts*), etc. Since their neglect does not result in communicative breakdown, they carry little psychological importance of the outcome (term *b* in the Rescorla–Wagner equation).

2.7 Contingency and Learning

The degree to which animals (including humans) learn associations between cues and outcomes also depends upon the contingency of the relationship. In classical conditioning, it is the reliability of the bell as a predictor of food that determines how easily this association is acquired (Rescorla, 1968). In language learning, it is the reliability of the form as a predictor of an interpretation that determines its acquisition and processing (Ellis, 2006b; Gries & Ellis, 2015; Gries & Stefanowitsch, 2004; MacWhinney, 1987). The last thirty years of psychological research into humans' sensitivity to cue–outcome contingencies (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). ΔP measures the directional association between a cue and an outcome, as illustrated in Table 2.1.

a, *b*, *c*, and *d* represent frequencies of occurrence. 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 showing 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 $P(O|\neg C)$, which can be calculated using this formula:

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

When the outcome is just as likely when the cue is present as when it is not, there is no covariation between the two events, and ΔP amounts to 0. As the presence of the cue increases the likelihood of the outcome, ΔP approaches 1.0. A learnable cue is one where the outcome is there whenever the cue is there, and where the outcome is not there when the cue is not there either, i.e., where a and d are large and b and c are small.

The less reliably a form is associated with a particular meaning or functional interpretation, the more difficult learning becomes (Ellis, 2006b; Shanks, 1995). Cues with multiple interpretations are ambiguous and thus hard to resolve; cue–outcome associations of high contingency are reliable and readily processed. Consider how, in the learning of the category of birds, while eyes and wings are equally frequently experienced features in the exemplars, it is wings which are distinctive in differentiating birds from other animals. Wings are important features to learning the category of birds because they are reliably associated with class membership while being absent from other categories of animals. Raw frequency of occurrence is therefore less important for construction learning than the contingency between cue and interpretation.

In language, there are rarely 1:1 mappings between forms and their interpretations. Cue–outcome reliability can be reduced in two directions: either forms have multiple interpretations (polysemy and homophony) or interpretations are realized by more than one form (synonymy). The same usage-phenomenon that promotes the reduction of frequently used words also drives grammatical functors towards homophony: different functions associated with forms that were originally distinct eventually merge into the same shortened form. An example is the –*s* suffix in English: in modern English, it has come to encode a plural form (*squirrels*), it indicates

https://doi.org/10.1017/9781108333603.003 Published online by Cambridge University Press

possession (*Max's toy*), and it marks third person singular present (*Nick sleeps*). The –*s* form is abundantly frequent in learners' input, but not reliably associated with any or just one of these meanings and 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 assess 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 short, the low cue-interpretation contingency of plurals and many other highly frequent grammatical constructions (see Gries, 2015) makes them difficult to learn (DeKeyser, 2005; Ellis, 2008a; Goldschneider & DeKeyser, 2001).

2.8 Learned Attention

L2 acquisition is vulnerable to attentional biases that stem from the L2 learners' knowledge of a prior language. For example, Ellis (2006a, 2006c) 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 that occurs in animals and humans alike. It shifts the learner's attention to selective aspects of the 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 paired with that same outcome is also a good predictor of it, and so the prior association effectively "blocks" further associations. For example, all languages have lexical and phrasal means of expressing temporality, so anyone with knowledge of any first language is aware 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*. Benati (2013) reviews a series of studies showing learners are better able to identify temporal reference when presented with temporal adverbs rather than verbal morphology.

A number of theories of L2 acquisition incorporate related notions of transfer and learned attention. The Competition Model (MacWhinney, 2001; MacWhinney & Bates, 1989), for one, focuses on dealing with competition between multiple linguistic cues to interpretation. Similarly,

https://doi.org/10.1017/9781108333603.003 Published online by Cambridge University Press

Input Processing (IP) theory (VanPatten, 1996) includes a *Lexical Preference Principle*: "Learners will process lexical items for meaning before grammatical forms when both encode the same semantic information" (VanPatten, 2007, p. 118) as well as a *Preference for Nonredundancy Principle*: "Learners are more likely to process nonredundant meaningful grammatical markers before they process redundant meaningful markers" (VanPatten, 2007, p. 119).

The basic mechanisms of learned attention in SLA have been examined through a series of experiments in which participants learned a small number of Latin expressions and their English translations. Ellis and Sagarra (2011) included three groups: an Adverb Pretraining, Verb Pretraining, and Control group. In Phase 1, the Adverb Pretraining group learned two adverbs and their temporal reference—*hodie* "today" and *heri* "yesterday"; the Verb Pretraining group learned verbs (shown in either first, second, or third person) and their temporal reference—e.g., *cogito* present or *cogitavisti* past; and the Control group had no pretraining. In Phase 2, all groups 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 both cues, their judgments in Phase 3 should be equally affected by them. If, in contrast, they paid more attention to adverb (or verb) cues, then their judgments would be biased towards them in Phase 3.

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

This experiment demonstrated how short-term instructional manipulations affect attention to language. Ellis and Sagarra (2010, 2011) furthermore illustrated long-term language transfer effects from the L1. They found that 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

https://doi.org/10.1017/9781108333603.003 Published online by Cambridge University Press

speakers of Spanish or Russian (rich morphology) to acquire inflectional cues from the same language experience where adverbial and verbal cues were equally available, with learned attention to tense morphology being in standardized ß coefficients: Chinese (–0.02) < English (0.17) < Russian (0.22) < Spanish (0.41) (Ellis & Sagarra, 2011, Table 4).

Ellis et al. (2014) replicated Ellis and Sagarra (2010), extending the investigation using eye-tracking measures to determine the extent to which short-term learned attention biases in the acquisition of temporal reference in L2 Latin in English as a foreign language (EFL) learners are overt or covert biases. The results indicated that prior experience of particular cue dimensions affected upon what participants overtly focused during subsequent language processing, which in turn resulted in covert attentional biases in comprehension and in productive knowledge. These learned attention effects include elements of both positive and negative transfer. Prior use of adverbial cues causes participants to pay more attention to adverbs, i.e., a positive effect of entrenchment of the practised cue. At the same time, however, 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) while that of blocking was moderate (–0.52).

Sagarra and Ellis (2013) showed the results of blocking over years of learning in intermediate and advanced learners of Spanish (as opposed to one hour of learning Latin). 120 English (a morphologically simple language) and Romanian (a morphologically rich language) learners of Spanish (also a morphologically rich language) 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 or incongruencies. Eye-tracking data revealed significant effects of incongruency (all participants were sensitive to tense incongruencies), cue location in the sentence (participants spent more time at their preferred cue), L1 experience (learners and monolinguals with morphologically rich L1s looked longer at verbs than learners and monolinguals with morphologically simple L1s), and L2 experience (intermediate learners read more slowly and regressed longer than advanced learners).

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

It is important to emphasize here that the limitations of L2 learning do not license the conclusion that L2 learning is qualitatively different from L1 learning. Second language learners employ the same statistical learning mechanisms that they employed when they acquired their first language. Rather, first language learning is (nearly always) so incredibly successful that it—somewhat ironically—hampers second language learning.

2.9 Implications for Language Teaching

The fact that L2 learners have to learn to adjust the attention biases they obtained through 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 conscious attention to what is being learned. In contrast, L2 acquisition is largely characterized by *explicit learning*. For reviews on implicit and explicit language learning, see Ellis (1994) and Rebuschat (2015).

Schmidt's (2001) *Noticing Hypothesis* holds that "people learn about the things they attend to and do not learn much about the things they do not attend to" (Schmidt, 2001, p. 30). That is, in order to successfully acquire specific aspects of their L2, learners must pay conscious and selective (i.e., focused) attention to the target structures. Given the bottleneck effect of input versus intake discussed above, even in dense-input, immersive environments, explicit learning and teaching gain even more relevance for the second language learners in foreign language environments with only limited L2 input. This holds in particular for aspects of form in the L2 that are redundant or polysemous and/or lack perceptual salience as discussed above. Form-focused instruction (FFI) thus attempts to encourage noticing, drawing learners' attention to linguistic forms that might otherwise be ignored (Ellis, 2012). Different 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 (Ellis, 2005).

Norris and Ortega (2000) compared the outcomes from studies that employed differing levels of explicitness of L2 input in a meta-analysis. Their results suggest that FFI instruction results in substantial target-oriented L2 gains, that explicit types of instruction are more effective than implicit types, and that the L2 instruction has durable effects. Similarly, more recent meta-analyses of effects of type of instruction by Spada and Tomita (2010) and Goo et al. (2015) 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 that future research should control to determine how robust effects of FFI are.

Cintrón-Valentín and Ellis (2015, 2016) used eye-tracking to investigate whether different types of FFI can aid learners in overcoming learned attention and blocking biases. English and Chinese native speakers viewed Latin utterances combining lexical and morphological cues to temporality under control conditions (CC) and three types of explicit focus on form (FonF): verb grammar instruction (VG); verb salience with textual enhancement (VS); and verb pretraining (VP). All participants completed an exposure phase, 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; and VP participants had an additional introductory phase where they were trained on solitary verb forms and their English translations (the rationale being that when the verb is presented in isolation rather than in potentially redundant combination with adverbial cues, there is less scope for blocking). CC participants were significantly more sensitive to the adverbs than verb morphology, while instructed participants showed greater sensitivity to morphological cues in comprehension and production.

Such results demonstrate that form-focused instruction recruits learners' explicit, conscious processing capacities and allows them to consolidate unitized form–function bindings of novel L2 constructions (Ellis, 2005). When a construction is learned this way, its use in subsequent implicit processing can update the statistical tallying of its frequency of usage and probabilities of form–function mapping.

2.10 Conclusion and Further Reading

This chapter has focused upon salience, contingency, and learned-attention in usage-based accounts of the L2 acquisition of morphology. In so doing, we have covered in detail just one facet of usage-based approaches to the L2 acquisition of linguistic form. Ellis and Wulff (2015a, 2015b) and Ortega et al. (2016) provide broader descriptions. Ellis et al. (2016) describe a large body of complementary work showing the joint effects of type-token frequency, contingency, and prototypicality in the usage-based L1 and L2 acquisition and processing of verb-argument constructions.

There are many other relevant and interesting aspects of usage-based SLA relating to social and cultural motivations:

> The target for many second language learners is not just "to speak another language", but to become part of the social and cultural environment in which the language is used. This entails frequent and rich participation in the second-language life worlds into which the learner "bricolages" his or her way. (Wagner, 2015, p. 75)

https://doi.org/10.1017/9781108333603.003 Published online by Cambridge University Press

Digital technologies give increasing opportunities for rich "rewilding" of education (Thorne, 2018), and such embodied, environmentally embedded, enacted, socially encultured, and situated environments support rich language learning (Eskildsen & Wagner, 2015). Beckner et al. (2009), Cadierno and Eskildsen (2015), Douglas Fir Group (2016), Hulstijn et al. (2014), and Ellis (2019) outline how the cognitive underpinnings of usage-based approaches can be integrated with a social perspective on SLA. Robinson and Ellis (2008), Littlemore (2009), and Tyler (2012) give broader overviews of applied cognitive-linguistic research in L2 learning and teaching.

References

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

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.

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.

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*(S1), 1–26.

Benati, A. (2013). The input processing theory in second language acquisition. In M. P. García Mayo, M. J. Gutierrez Mangado, M. Martínez Adrián, F. Myles, J. Rothman, & B. VanPatten (eds.), *Contemporary approaches to second language acquisition* (pp. 93–110). Amsterdam: John Benjamins.

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

Bybee, J. (2008). Mechanisms of change in grammaticization: the role of frequency. In B. D. Joseph & R. D. Janda (eds.), *The Handbook of Historical Linguistics* (pp. 602–623). Malden, MA: Wiley.

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

Chater, N., & Manning, C. (2006). Probabilistic models of language processing and acquisition. *Trends in Cognitive Science*, *10*, 335–344.

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.

https://doi.org/10.1017/9781108333603.003 Published online by Cambridge University Press

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

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

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

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.

Douglas Fir Group. (2016). A transdisciplinary framework for SLA in a multilingual world. *The 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). The Associative-Cognitive CREED. In B. V. Patten & J. Williams (eds.), *Theories in second language acquisition: An introduction* (pp. 77–96). Mahwah, NJ: Lawrence Erlbaum Associates.

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

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

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

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

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

Ellis, N. C. (2019). Essentials of a theory of second language cognition. *The Modern Language Journal*, 103, 39–60.

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

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., 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, N. C., & Wulff, S. (2015a). Second language acquisition. In E. Dąbrowska & D. Divjak (eds.), *Handbook of cognitive linguistics* (pp. 409–431). Berlin: DeGruyter Mouton.

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

Ellis, R. (2012). Form-Focused Instruction and Second Language Learning. In R. Ellis (ed.) *Language Teaching Research and Pedagogy* (pp. 217–306). Hoboken, NJ: Wiley-Blackwell.

Eskildsen, S. W., & Wagner, J. (2015). Embodied L2 construction learning. *Language Learning*, *65*, 268–297.

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: Wissenschaftlicher Verlag Trier.

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

Goldschneider, J. M., & DeKeyser, R. (2001). Explaining the "natural order of L2 morpheme acquisition" in English: A meta-analysis of multiple determinants. *Language Learning*, *51*(1), 1–50.

Goo, J., Granena, G., Yilmaz, Y., & Novella, M. (2015). Implicit and explicit instruction in L2 learning. In P. Rebuschat (ed.), *Implicit and explicit learning of languages* (pp. 443–482). Amsterdam: John Benjamins.

Gries, S. Th. (2015). Polysemy. In E. Dąbrowska & D. S. Divjak (eds.), *Handbook of Cognitive Linguistics* (pp. 472–490). Berlin: De Gruyter Mouton.

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.

Hulstijn, J. H., Young, R. F., Ortega, L., Bigelow, M. H., DeKeyser, R., Ellis, N. C., & Talmy, S. (2014). Bridging the gap: Cognitive and social approaches to research in second language learning and teaching. *Studies in Second Language Acquisition*, *33*(3), 361–421.

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

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: John Benjamins.

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). Learned Attention. Paper presented at the Fifth International Conference on Development and Learning, Indiana University.

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

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

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

Littlemore, J. (2009). *Applying cognitive linguistics to second language learning and teaching*. Basingstoke and New York: Palgrave Macmillan.

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.

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: 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.

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.

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

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

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

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

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.

Seidenberg, M. S., & Plaut, D. C. (2014). Quasiregularity and its discontents: The legacy of the past tense debate. *Cognitive Science*, *38*(6), 1190–1228.

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

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

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

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

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

Thorne, S. L. (2018). Technologies, morphologies of communicative action, and the rewilding of language education. Paper presented at the American Association for Applied Linguistics, Chicago, IL.

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

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

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

VanPatten, B. (2007). Input processing in adult second language acquisition. In B. VanPatten & J. Williams (eds.), *Theories in second language acquisition: An introduction* (pp. 115–135). Mahwah, NJ: Lawrence Erlbaum Associates.

Wagner, J. (2015). Designing for language learning in the wild: Creating social infrastructures for second language learning. In T. Cadierno & S. W. Eskildsen (eds.), *Usage-based perspectives on second language learning* (pp. 75–101). Berlin: de Gruyter Mouton.

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.