——(2006). Why not fossilization. In Z.-H. Han and T. Odlin (eds), *Studies of Fossilization in Second Language Acquisition*, pp. 173–88. Clevedon: Multilingual Matters.

Birdsong, D. and Molis, M. (2001). On the evidence for maturational constraints in second language acquisition. *Journal of Memory and Language*, *44*, 235–49.

Bley-Vroman, R. (1990). The logical problem of foreign language learning. *Linguistic Analysis*, *20*, 3–49.

DeKeyser, R.M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, *22*, 499–533.

Ellis, R. (1985). *Understanding Second Language Acquisition*. Oxford: Oxford University Press.

Hawkins, R. (2000). Persistent selective fossilization in second language acquisition and the optimal design of the language faculty. *Essex Research Reports in Linguistics*, *34*, 75–90.

——(2001). The theoretical significance of Universal Grammar in second language acquisition. *Second Language Research*, *17*, 345–67.

Hopp, H. (2010). Ultimate attainment in L2 inflection: Performance similarities between non-native and native speakers. *Lingua*, *120*, 901–31.

Johnson, J. and Newport, E.L. (1989). Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*, 60–99.

Lardiere, D. (1998). Case and tense in the 'fossilized' steady state. *Second Language Research*, *14*, 1–26.

——(2006). Establishing ultimate attainment in a particular second language grammar. In Z.-H. Han and T. Odlin (eds), *Studies of Fossilization in Second Language Acquisition*, pp. 35–55. Clevedon: Multilingual Matters.

——(2007). *Ultimate Attainment in Second Language Acquisition: A Case Study*. Mahwah, NJ: Lawrence Erlbaum.

Long, M.H. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition*, *12*, 251–85.

——(2003). Stabilization and fossilization in interlanguage development. In C.J. Doughty and M.H. Long (eds), *Handbook of Second Language Acquisition*, pp. 487–535. Malden, MA: Blackwell.

Marinova-Todd, S.H. (2003). Comprehensive analysis of ultimate attainment in adult second language acquisition. Unpublished doctoral dissertation, Harvard University.

Schumann, J.H. (1997). *The Neurobiology of Affect in Language*. Malden, MA: Blackwell.

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics*, *10*, 209–31.

Selinker, L. and Lamendella, J.T. (1978). Two perspectives on fossilization in interlanguage learning. *Interlanguage Studies Bulletin*, *3*, 143–91.

Sorace, A. (2003). Near-nativeness. In C.J. Doughty and M.H. Long (eds), *Handbook of Second Language Acquisition*, pp. 130–61. Malden, MA: Blackwell.

——(2011) Pinning down the concept of "interface" in bilingualism. *Linguistic Approaches to Bilingualism*, *1*(1), 1–33.

White, L. (2003). Fossilization in steady state L2 grammars: Persistent problems with inflectional morphology. *Bilingualism: Language and Cognition*, *6*, 129–41.

Frequency effects

Nick C. Ellis
University of Michigan

Usage-based approaches to SLA hold that we learn linguistic constructions while engaging in communication (Collins and Ellis, 2009; Ellis and Cadierno, 2009; Robinson and Ellis, 2008). The last 50 years of psycholinguistic research provides 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). 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. Psycholinguistic perspectives thus hold that language learning is the associative learning of representations that reflect the probabilities of occurrence of form-function mappings. Frequency is a key determinant of acquisition because "rules" of language, at all levels of analysis from phonology, through syntax, to discourse, are structural regularities which emerge from learners' lifetime unconscious analysis of the distributional characteristics of the language input.

Frequency determinants of construction learning

If constructions as form-function mappings are the units of language, then language acquisition involves inducing these associations from experience of language usage. Constructionist accounts of language acquisition thus 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 (Williams and Rebuschat, forthcoming) relating input and learner cognition.

Psychological analyses of the learning of constructions as form-meaning pairs is informed by the literature on the associative learning of cue-outcome contingencies where the usual determinants include such factors as: construction frequency, type-token frequency, Zipfian distribution, salience of form, prototypicality of meaning, and contingency of form-function mapping (Ellis and Cadierno, 2009). Consider each in turn:

Construction frequency

Frequency of exposure promotes learning and entrenchment. Learning, memory and perception are all affected by frequency of usage: the more times we experience something, the stronger our memory for it, and the more fluently it is accessed. The more recently we have experienced something, the stronger our memory for it, and the more fluently it is accessed (hence your reading this sentence more fluently than the preceding one). The more times we experience conjunctions of features, the more they become associated in our minds and the more these subsequently affect perception and categorization; so a stimulus becomes associated to a context and we become more likely to perceive it in that context. The power law of learning (Anderson, 2000) describes the relationships between practice and performance in the acquisition of a wide range of cognitive skills – the greater the practice, the greater the performance, although effects of practice are largest at early stages of learning, thereafter diminishing and eventually reaching asymptote.

Frequency of exposure also underpins statistical learning of categories. Human categorization ability provides the most persuasive testament to our incessant unconscious figuring or "tallying." We know that natural categories are fuzzy rather than monothetic. Wittgenstein's (1953) consideration of the concept *game* showed that no set of features that we can list covers all the things that we call games, ranging as the exemplars variously do from soccer, through chess, bridge, and poker, to solitaire. Instead, what organizes these exemplars into the *game* category is a set of family resemblances among these members – son may be like mother, and mother like sister, but in a very different way. And we learn about these families, like our own, from experience. Exemplars are similar if they have many features in common and few distinctive attributes (features belonging to one but not the other); the more similar are two objects on these quantitative grounds, the faster are people at judging them to be similar (Tversky, 1977). The greater the token frequency of an exemplar, the more it contributes to defining the category, and the greater the likelihood it will be considered the prototype. The operationalization of this criterion predicts the speed of human categorization performance – people more quickly classify as *birds* sparrows (or other average-sized, average-colored, average-beaked, average-featured specimens) than they do birds with less common features or feature combinations like geese or albatrosses (Rosch and Mervis, 1975; Rosch *et al.* 1976). Prototypes are judged faster and more accurately, even if they themselves have never been seen before – someone who has never seen a sparrow, yet who has experienced the rest of the run of the avian mill, will still be fast and accurate in judging it to be a bird. Such effects make it very clear that although people don't go around consciously counting features, they nevertheless have very accurate knowledge of the underlying frequency distributions and their central tendencies. Cognitive theories of categorization and generalization show how schematic constructions are abstracted over less schematic ones that are inferred inductively by the learner in acquisition.

Ellis' (2002) review illustrates how frequency effects the processing of phonology and phonotactics, reading, spelling, lexis, morphosyntax, formulaic language, language comprehension, grammaticality, sentence production, and syntax. 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.

Type and token frequency

Token frequency counts how often a particular form appears in the input. Type frequency, on the other hand, refers to the number of distinct lexical items that can be substituted in a given slot in a construction, whether it is a word-level construction for inflection or a syntactic construction specifying the relation among words. For example, the 'regular' English past tense *-ed* has a very high type frequency because it applies to thousands of different types of verbs, whereas the vowel change exemplified in *swam* and *rang* has much lower type frequency. The productivity of phonological, morphological, and syntactic patterns is a function of type rather than token frequency (Bybee and Hopper, 2001). This is because: (a) the more lexical items that are heard in a certain position in a construction, the less likely it is that the construction is associated with a particular lexical item and the more likely it is that a general category is formed over the items that occur in that position; (b) the more items the category must cover, the more general are its criterial features and the more likely it is to extend to new items; and (c) high type frequency ensures that a construction is used frequently, thus strengthening its representational schema and making it more accessible for further use with new items. In contrast, high token frequency promotes the entrenchment or conservation of irregular forms and idioms; the irregular forms only survive because they are high frequency. These findings support language's place at the center of cognitive research into human categorization, which also emphasizes the importance of type frequency in classification.

Zipfian distribution

In natural language, Zipf's law (Zipf, 1935) describes how the highest frequency words account for the most linguistic tokens. Zipf's law states that the frequency of words decreases as a power function of their rank in the frequency table (the most frequent word occurring approximately twice as often as the second most frequent word, which occurs twice as often as the fourth most frequent word, etc.). If p_f is the proportion of words whose frequency in a given language sample is f, then $p_f \sim f^{-\gamma}$, with $\gamma \approx 1$. Zipf showed this scaling law holds across a wide variety of language samples. Subsequent research provides support for this law as a linguistic universal. Many language events across scales of analysis follow this power law: phoneme and letter strings, words, grammatical constructs, formulaic phrases, etc. Scale-free laws also pervade language structures, such as scale-free networks in collocation, in morphosyntactic productivity, in grammatical dependencies, and in networks of speakers, and language dynamics such as in speech perception and production, in language processing, in language acquisition, and in language change (Solé *et al.*, 2005). Scale-free laws pervade both language structure and usage. And not just language structure and use. Power law behavior like this has since been shown to apply to a wide variety of structures, networks, and dynamic processes in physical, biological, technological, social, cognitive, and psychological systems of various kinds (e.g. magnitudes of earthquakes, sizes of meteor craters, populations of cities, citations of scientific papers, number of hits received by websites, perceptual psychophysics, memory, categorization, etc.) (Kello *et al.*, 2010). It has become a hallmark of Complex Systems theory. Zipfian scale-free laws are universal. Complexity theorists suspect them to be fundamental, and are beginning to investigate how they might underlie language processing, learnability, acquisition, usage and change (Ellis and Larsen-Freeman, 2009). Usage-based theories argue that it is the coming together of these distributions across linguistic form and linguistic function that makes language robustly learnable despite learners' idiosyncratic experience and the "poverty of the stimulus" (Ellis and O'Donnell, in press).

Form (salience)

The general perceived strength of stimuli is commonly referred to as their salience. Low salience cues tend to be less readily learned. Research in associative learning demonstrates that selective attention, salience, expectation, and surprise are key elements in the analysis of all learning, animal and human alike. The amount of learning induced from an experience of a cue-outcome association depends crucially upon the salience of the cue and the importance of the outcome. Many grammatical meaning-form relationships, particularly those that are notoriously difficult for second language learners like grammatical particles and inflections such as the third person singular *–s* of English, are of low salience in the language stream. This can lead to blocking effects. For example, '*today*' is a stronger psychophysical form in the input than is the morpheme '*-s*' marking third person singular present tense, thus while both provide cues to present time, *today* is much more likely to be perceived, and *-s* can thus become overshadowed and blocked, making it difficult for second language learners of English to acquire (Ellis, 2006b).

Prototypicality of meaning

Categories have graded structure, with some members being better exemplars than others. In the prototype theory of concepts (Rosch and Mervis, 1975), the prototype as an idealized central description is the best example of the category, appropriately summarizing the most representative attributes of a category. As the typical instance of a category, it serves as the benchmark against which surrounding, less representative instances are classified. Ellis and Ferreira-Junior (2009) show that the verbs that naturalistic second language learners first use in particular verb-argument constructions are prototypical and generic in function (*go* for Verb Locative, *put* for Verb Object Locative, and *give* for Verb Object Object ditransitive).

Contingency of form-function mapping

Psychological research into associative learning has long recognized that while frequency of form is important, so too is contingency of mapping. 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; eyes are neither. Raw frequency of occurrence is less important than the contingency between cue and interpretation. Contingency, and its associated aspects of predictive value, information gain, and statistical association, is a strong determinant of construction learning too (Ellis, 2006a; MacWhinney, 1987).

Conclusions

The primary motivation of usage-based approaches is that we must bring together linguistic form, learner cognition, and usage. Saussure (1916) said, "To speak of a 'linguistic law' in general is like trying to lay hands on a ghost … Synchronic laws are general, but not imperative … [they] are imposed upon speakers by the constraints of common usage … In short, when one speaks of a synchronic law, one is speaking of an arrangement, or a principle of regularity" (pp. 90–91). The frequencies of common usage count in the emergence of regularity in SLA. Usage is rich in latent linguistic structure.

See also: cognitive linguistics and SLA, construction learning, learned inattention and blocking, prototypes, statistical learning, type and token frequency

References

Anderson, J.R. (2000). *Cognitive Psychology and its Implications*, 5th edn. New York: W.H. Freeman.

Bybee, J. and Hopper, P. (eds) (2001). *Frequency and the Emergence of Linguistic Structure*. Amsterdam: Benjamins.

Collins, L. and Ellis, N.C. (2009). Input and second language construction learning: frequency, form, and function. *Modern Language Journal*, *93*, Whole issue.

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*, 143–88.

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

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

Ellis, N.C. and Cadierno, T. (2009). Constructing a second language. *Annual Review of Cognitive Linguistics*, *7*, 111–290.

Ellis, N.C. and Ferreira-Junior, F. (2009). Construction Learning as a function of frequency, frequency distribution, and function. *Modern Language Journal*, *93*, 370–386.

Ellis, N.C. and Larsen-Freeman, D. (2009). Language as a complex adaptive system (Special Issue). *Language Learning*, *59*, Supplement 1.

Ellis, N.C. and O'Donnell, M. (in press). Statistical construction learning: Does a Zipfian problem space ensure robust language learning? In Williams, J. and Rebuschat, J. (eds), *Statistical Learning and Language Acquisition*. Berlin: Mouton de Gruyter.

Kello, C.T., Brown, G.D. A., Ferrer-I-Cancho, R., Holden, J.G., Linkenkaer-Hansen, K., Rhodes, T. and Van Orden, G.C. (2010). Scaling laws in cognitive sciences. *Trends in Cognitive Science*, *14*, 223–232.

Macwhinney, B. (1987). The Competition Model. In Macwhinney, B. (ed.), *Mechanisms of Language Acquisition*. Hillsdale, NJ: Erlbaum.

Robinson, P. and Ellis, N.C. (eds) (2008). *A Handbook of Cognitive Linguistics and Second Language Acquisition*. London: Routledge.

Rosch, E. and Mervis, C.B. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, *104*, 192–233.

Saussure, F.D. (1916). *Cours de linguistique générale*. London: Duckworth.

Solé, R.V., Murtra, B., Valverde, S. and Steels, L. (2005). Language networks: their structure, function and evolution. *Trends in Cognitive Sciences*, *12*.

Williams, J. and Rebuschat, J. (eds) (forthcoming). *Statistical Learning and Language Acquisition*. Berlin: Mouton de Gruyter.

Wittgenstein, L. (1953). *Philosophical Investigations*. Oxford: Blackwell.

Zipf, G.K. (1935). *The Psycho-biology of Language: An Introduction to Dynamic Philology*. Cambridge, MA: The MIT Press.

Further reading

Ellis, N.C. (2002). Frequency effects in language acquisition. Special issue. *Studies in Second Language Acquisition*, *24* (2). (A special issue of SSLA featuring a target article on frequency effects in SLA, commentaries by leading researchers in the field, and a response.)

Collins, L. and Ellis, N.C. (eds) (2009). Input and second language construction learning: frequency, form, and function. Special issue. *Modern Language Journal*, *93* (3). (SLA research has always emphasised input. This special issue presents new ways of analyzing its latent properties.)

Ellis, N.C. and Cadierno, T. (eds) (2009). Constructing a second language. Special section. *Annual Review of Cognitive Linguistics, 7,* 111–29. (Cognitive Linguistics is a branch of linguistics which analyzes language in terms of the concepts which underlie its forms. It denies any autonomous linguistic faculty or language acquisition device. Instead it holds that language is learned by general cognitive processes, that grammar can be understood in terms of conceptualization, and that language is learned

from usage. The papers in this special section consider SLA from this perspective.)

Ellis, N.C. and Larsen-Freeman, D. (eds) (2009). *Language as a complex adaptive system*. Oxford: Wiley-Blackwell. (A special issue of *Language Learning* gathering experts from various language domains who share a CAS / Emergentist perspective.)

Robinson, P. and Ellis, N.C. (eds) (2008). *Handbook of cognitive linguistics and second language acquisition*. London: Routledge. (An edited collection bringing together leading researchers in Cognitive Linguistics and SLA to consider usage-based SLA.)

Functional categories

Pieter Muysken
Radboud University Nijmegen

Traditionally, grammarians have distinguished between function words and content words in the lexicon. Take a New York Times headline, "*Clandestine operatives have been sent into Libya to gather intelligence for military airstrikes and make contacts with rebels battling Col. Muammar el-Qaddafi's forces*." This sentence contains a number of content words, including proper names, nouns, adjectives, and verbs. There are also function words, such as auxiliaries, conjunctions, and prepositions. So far, so good. However, there are also some words, such as the "light verb" *make* and the locative preposition *into*, which have an inbetween status. These elements have a fairly abstract meaning, but not really a grammatical function. Finally, there are elements which do not occur as independent words but which clearly have a grammatical function, and in this sense resemble function words.

To avoid the need to distinguish between function words and grammatical endings linguists use the term "Functional Categories." These are used both for morpho-syntactic categories:

tense past / future / …
number singular / dual / …
person first / second / third / …

and for the words or word endings instantiating these categories:

pronouns *I* [first] / *you* [second], …
conjunctions *that* [finite] / *for-to* [non-finite], …
prepositions *of* [genitive] / *to* [dative], …
past participle *-ed,-en,-t*

Table 6 Content and function words

Content words	*In between content words and function words*	*Function words*	*Grammatical endings*
Muammar el-Qaddafi-'s	make	have	genitive *'s*
Libya	into	be-en	plural *–s*
col.(onel)		to	past participle *–en, -t*
intelligence		and	gerund *-ing*
airstrike-s		for	
contact-s		with	
operative-s			
rebel-s			
force-s			
military			
clandestine			
battl-ing			
sen-t			
gather			