Opinion

Three factors in the design and acquisition of language

William O'Grady*

Recent advances in linguistic theory offer new proposals about the factors that are crucial to understanding the design and acquisition of language—the genetic endowment, experience, and principles not specific to the language faculty. Of particular interest is the third of these factors, whose importance is now widely recognized, raising questions about its character, its role in shaping the language faculty, and its impact on the future of linguistic research.

How to cite this article:
WIREs Cogn Sci 2012, 3:493–499. doi: 10.1002/wcs.1188

INTRODUCTION

Contemporary linguistics has two major objectives, one descriptive and the other explanatory. On the one hand, there is the challenge of documenting how individual languages employ form to express meaning—e.g., how they use case, agreement and word order to distinguish among the participants in an event, how they encode contrasts involving time and space, how they convey new and old information, and so forth. On the other hand, there is the challenge of explaining why language has the particular properties that it does (the problem of language design) and how those properties emerge so reliably in the course of early childhood (the problem of language acquisition). It is the search for answers to these two problems that make work in linguistics central to the larger enterprise of cognitive science.

A signature thesis of linguistic theory for the last half century is the 'innateness hypothesis'. First put forward in the 1960s by Noam Chomsky, it posits two separate inborn mechanisms: a sensory system for the preliminary analysis of input and a Universal Grammar (UG) (Refs 1, p. 12, and 2, p. 269). The idea of an innate sensory system is widely accepted, but the UG thesis has always been deeply divisive. Indeed, several branches of linguistics (syntax, language acquisition, and typology, to name three) have parallel research programs, one committed to UG and the other opposed.

*Correspondence to: ogrady@hawaii.edu
Department of Linguistics, University of Hawai'i, Honolulu, HI, USA

This schism notwithstanding, the playing field for explanatory initiatives is well bounded. As Chomsky[3] observes, recapitulating the long-standing consensus, there are really just three factors that might be responsible for the character of language and for the ease with which it is acquired.

1. A possible genetic endowment specifically for language (UG)
2. Experience
3. Principles not specific to the faculty of language[a]

I will consider each of these factors in turn, with a view to summarizing recent developments and to assessing their contribution to our understanding of language. A concluding section raises the possibility that recent work on the 'third factor' opens the door for a reunification of the discipline around a common research question.

UNIVERSAL GRAMMAR

The central thesis of UG, clearly and consistently stated from the 1960s onward, is that the human capacity for language is rooted in a faculty-specific 'system of [grammatical] principles, conditions, and rules that are elements or properties of all human languages not merely by accident but by [biological] necessity' (Ref 1, p. 29).

In early work, principles of UG played a direct and crucial role in the explanation of a broad range of phenomena—phrase structure, agreement, case, *wh* movement, quantifier scope,

pronoun interpretation, and so on. As the field progressed, however, the abstractness and faculty-specific character of the principles that were being put forward generated increasing skepticism among opponents and proponents alike. The Empty Category Principle (ECP), a key component of the Government-and-Binding instantiation of UG, is a case in point.

(1) *The Empty Category Principle* (Ref. 4, p. 8ff).
An empty category must be properly governed.

Deceptively simple in its formulation, the success of the ECP is built on an extraordinarily complicated definition of government, which entails reference to m-command, barriers, blocking categories, domination, exclusion, and L-marking, among other technical notions. Such intricacies eventually gave pause even to Chomsky (Ref 5, p. 233), who observed that many explanations had come to have 'roughly the order of complexity of what is to be explained'. Corrective action was initiated in the form of the Minimalist Program, which rejected many of the key claims and assumptions of traditional generative grammar.

A particularly radical version of Minimalism focuses on 'the faculty of language in the narrow sense' (FLN), the part of our linguistic capacity that is unique both to humans and to language. Upending the long-standing view that it includes a rich UG, Hauser et al. (Ref 6, p. 1573) propose that FLN can be reduced to 'mechanisms of recursion' that are manifested most clearly in Merge, the operation that builds structure by combining and rearranging words and other formative elements. For a discussion of recursion and its significance for language, see Ref 7 and the recent special issue of *Biolinguistics* (Ref 8, 5.1-2).

This view of FLN forces a fundamental reconceptualization of language universals, as Boeckx (Ref 9, p. 197) observes.

> If by [linguistic universals] we mean specific universals for language, then we are going to be looking at a very narrow field, a very narrow set . . . what we expect to find will be basically abstract general principles such as minimal search, or various refinements of relativized minimality, cyclicity, etc.
>
> . . . those universals are not specifically linguistic, but might be generic or general principles of efficient computations belonging to third-factor properties, for example. But these would be the kind of [universals] that may actually be at the core of FLN.

The potential advantages of this proposal can be appreciated even by those opposed to the UG-based research program.

For one thing, the new view brings UG into better alignment with current thinking about the evolution of language. The plethora of complex grammatical principles that had been proposed in the Government and Binding framework raised seemingly intractable questions about how the language faculty evolved (Refs 10, p. 2, and 11). Strong versions of Minimalism avoid this problem by focusing on general computational constraints rather than narrow grammatical principles. (For further discussion, see Refs 12, p. 369ff, 13, and 8, p. 264.)

Another problem that can be dispensed with involves challenges to the descriptive adequacy of classic grammatical universals—a common occurrence as field work makes available data from an ever-growing number of languages.[14–16] As Chomsky (Ref 10, p. 2) acknowledged, the sorts of principles associated with earlier versions of UG 'pose serious problems for dealing with the diversity of languages'. On the Minimalist view, such principles no longer exist.

But now new difficulties arise. A first challenge involves the issue of empirical coverage: can all the core properties of language really be explained in terms of recursion? Pinker and Jackendoff (Ref 17, p. 220) suggest that they cannot, noting the challenges presented by numerous phenomena that have long been considered central to the study of language, including case, agreement, co-reference, topic, focus, scope, and the like. On this point, see also Ref 18, p. 588).

A second challenge touches upon the original motivation for UG—language acquisition. Development is underdetermined by experience, Berwick et al. (Ref 19, p. 1207–1209) insist, so by the argument from poverty of stimulus, there is still a significant explanatory role for 'innate domain-specific factors'. The question of how this can be made to work without the traditional rich set of inborn grammatical principles is barely addressed. Chomsky (Ref 3, p. 8) suggests only that language acquisition is now 'a matter of parameter setting and therefore divorced from the [remaining] principles of UG' (see also Ref 10, p. 3). The suggestion is accompanied by a favorable reference to Baker[20] and Yang,[21] but does little to help, as both authors draw on traditional principles of UG. For instance, Baker's parameters include a polysynthesis option, a verb-raising option, a serial verb option, and a null subject option, all hierarchically arranged in an intricate network of implicational relationships.

It is important in this regard to distinguish between where parameters come from (if not from UG), and the manner in which the appropriate values

are selected. The latter task may well be a matter of 'data processing', as Chomsky (Ref 3, p. 7) suggests, citing Yang's proposal that experience increases or decreases the statistical viability of particular parameter settings. But this does not explain how children know which parametric options are available in the first place. Longa and Lorenzo[22] offer an insightful discussion of the general tension between the study of syntax and the study of language acquisition that has resulted from the Minimalist conception of UG.

In sum, there is little hope that UG, in whatever form it still exists, can contribute much to our understanding of the many typological and developmental phenomena for which it once accounted. Almost all of its former explanatory burden must be shifted to the other two factors at our disposal—experience and principles not specific to the language faculty.

EXPERIENCE

It has always been understood that experience plays a major role in shaping linguistic development, but there is a long-standing dispute over whether it contains sufficient information to bypass the need for an innately specified UG. Advocates of experience-based approaches to language acquisition hold that there is no poverty-of-stimulus problem; the properties of language are learned, they claim, not given. Research of this type follows several different but related lines.

One strategy focuses on the manner in which language use by caregivers shapes development. Often referred to as the 'usage-based' theory[23,24], it holds that language acquisition occurs as children, relying on their ability to intuit the intentions of others, make sense of what they hear in particular situations. In most versions of this theory, the early stages of development are dominated by item-specific learning, with a focus on high-frequency utterances (e.g., *What's that?*, *I want that*). With time and additional experience, general patterns are discerned, leading to the formation of more abstract grammatical constructions, such as *wh* questions, transitive sentences, relative clauses, and so on.

Another research strategy draws on experimental work on statistical learning, usually with a focus on artificial 'mini-languages'. Based on the idea that learners are highly sensitive to patterns of co-occurrence, early research in this area examined infants' remarkable ability to segment strings of nonsense syllables into 'words' by taking note of recurring sequences, such as *tutibu* in *bupadatutibubabupupatutibupadababupupidabututibu*.[25] Subsequent work has extended the scope of the enterprise to include phrase boundaries, hierarchical structure, and other higher-level phenomena.[26,27]

A third strategy makes use of computational modeling, often with the help of simple recurrent networks (SRNs), which attend to the same sorts of transitional probabilities exploited in the work on statistical learning. Computational simulations using SRNs have explored a variety of syntactic phenomena, including category creation, verbal agreement, and subject–verb inversion.[28,29]

Experience-based approaches to language acquisition face two fundamental challenges. First, they have thus far focused their attention on phenomena that even proponents of UG agree must be learned largely from experience—inflection, the argument alternations permitted by particular verbs (*give Max a book*, *give a book to Max*), and so on. There has been no comprehensive effort to address 'UG-type' phenomena, such as the intricate constraints on co-reference, contraction, and scope interpretation whose effects are instantiated far less systematically in the input, if at all.[30]

Second, as even their proponents acknowledge (Ref 31, p. 110), experience-based models of acquisition offer no account for why the input has the particular properties that it does. It is one thing to explain how children learning English come to know that verbs agree only with subjects; it is an entirely different matter to explain why there are no languages in which verbs agree only with direct objects. For questions such as these, a largely promissory note must be issued: the properties of human language are shaped not only by experience, but also by yet-to-be-discovered constraints on processing, perception, cognition, and interaction.[31–33] In an important respect, then, UG-based and experience-based approaches to language find themselves in a similar situation. Each must look to third-factor effects to fill in key pieces of the language puzzle. Herein may well lie the future of linguistics.

THE THIRD FACTOR

Even as rival frameworks look to a third factor to supplement their respective accounts of how language works, a new controversy looms on the horizon: what is the nature of that factor? At least in the case of syntax, there seem to be two opposing views, one based on the notion of computational efficiency linked to the Minimalist Program and the other based on the idea of processing cost associated with psycholinguistics. It is worthwhile to consider each in turn.

Computational Efficiency

Appeals to computational efficiency in generative grammar first came to the fore in the early 1990s, with the emergence of principles whose name and content evoked a concern for locality and economy—notions that invite a more general computational interpretation. One such principle is the Minimal Link Condition (Ref 5, p. 264).

(2) *The Minimal Link Condition*
Make moves as short as possible.

A consequence of (2) is that the direct object *wh* word in a sentence such as *What should they discuss?* moves first to the left edge of vP ('small VP') and from there to the left edge of CP.[b]

(3) [$_{CP}$ What should [$_{TP}$ they [$_{vP}$ _ [$_{VP}$ discuss _]]]]
↑_____|↑_____|

CPs and vPs count as *phases*—pieces of structure that, once assembled, immediately undergo the relevant phonological and semantic operations, thereby becoming inaccessible to further syntactic intervention. Chomsky (Ref 3, p. 9) suggests that phases exist for reasons of computational efficiency, since they allow completed portions of a derivation to be 'forgotten', leading to 'a substantial saving in memory' (p. 16–17).

Computational efficiency also plays a crucial role in deriving another staple of the Minimalist Program—the so-called copy theory of movement. On this view, a fuller representation of (3) would be (4), with a copy of the direct object *wh* word and the auxiliary verb in their default positions (where, supposedly, they are interpreted) and in any intermediate positions that they occupy in the course of movement.

(4) [$_{CP}$ WHAT *should* [$_{TP}$ they *should* [$_{vP}$ WHAT [$_{VP}$ discuss WHAT]]]]

These copies exist, so the story goes, because it would be 'inefficient' to delete them (by the 'No Tampering Condition'; see Refs 3, p. 13, and 19, p. 1219). And they go unpronounced for another efficiency-related reason—phonological computation is costly.

Despite the allusions to memory and forgetting, Chomsky is not proposing a model of how sentences are produced and comprehended: computational efficiency is not the same thing as processing cost. Indeed, Chomsky (Ref 34, p. 146) draws an explicit distinction between the two, asserting that processing would be easier in patterns such as (4) if all the copies were *retained*, thereby eliminating empty positions, which are known to be difficult to process. The copies are deleted, he suggests, only because computational efficiency trumps processing, forcing 'erasure of all but one copy, so that the phonological component can forget about the others'.

Such examples illustrate the extent to which the viability of computational efficiency in the Minimalist sense depends on a network of theory-internal assumptions about cost, copies, movement operations, and syntactic representations. Such assumptions raise questions about whether an independently verifiable third factor is really in play here at all.

No such doubt arises with respect to processing cost, a performance-based notion whose effects are measured and tested through psycholinguistic experimentation.

Processing cost

It has long been understood that processing plays an important role in shaping language (see Ref 35 for an overview), but typically with the understanding that its effects are in some sense secondary. As Ferreira et al. (Ref 36, p. 13) observe, the 'most basic assumption about the nature of the human sentence processor is that it obeys the fundamental principles of grammar'. This consensus has recently been challenged by work that treats processing considerations as *primary*. On this view, a simple processor, constrained by the need to minimize the burden on the limited resources of working memory, lies at the heart of the human language faculty. The grammar, to the extent that it exists, is simply a system of 'frozen processing preferences' (Ref 37, p. 280) or 'a processor that has become set in its ways' (Ref 38, p. 212).

An illustrative example from typology involves the sort of filler-gap dependencies found in certain types of *wh* questions and relative clauses, among other patterns.

(5) a. What did the dog find _?

b. the book [which Harry recommended _]

It is widely recognized that filler-gap dependencies place a special burden on working memory (Refs 39, 40, p. 102, and 41), and that their cost increases when they extend across clause boundaries (Refs 42, 43, p. 253, and 44, p. 383).

As Hawkins 45, p. 193ff) demonstrates, the cumulative effects of processing cost may explain a well-established typological asymmetry: languages that permit cross-clausal filler-gap dependencies also permit intra-clausal ones, but not vice versa. In a conservative language such as Russian, a filler-gap dependency can extend into an embedded infinitival VP, but not into an embedded clause.

(6) Russian

a. Filler-gap dependency extending into embedded VP:
Vot ogurcy [kotorye ja obeščal [$_{Inf}$ prinesti _]]
|_____|
Here.are cucumbers which I promised to.bring

b. Filler-gap dependency extending into embedded clause:
*Vot ogurcy [kotorye ja obeščal [$_{S}$ čto prinesu _]]
|_____|
Here.are cucumbers which I promised that I.bring

In contrast, English permits a filler-gap dependency to descend into either an infinitival or an embedded clause.

(7) English

a. Filler-gap dependency extending into embedded VP:
Here are the cucumbers [which I promised [$_{Inf}$ to bring _]]
|_____|

b. Filler-gap dependency extending into embedded clause:
Here are the cucumbers [which I promised [$_{S}$ that I'd bring _]]
|_____|

The general idea, developed in some detail by Hawkins 45, p. 192ff, 266) and O'Grady (Ref 38, p. 203ff, 214ff), is that processing factors create a continuum of difficulty along which there are certain natural 'break points' (clause boundaries, for instance). Different languages choose different break points, but always with the same consequence: if the processor tolerates a more demanding pattern, it must permit less demanding patterns as well. Thus, we find languages such as Russian and English, but no language that allows a filler-gap dependency to extend only into an embedded clause.

The processing-based approach also offers a possible solution to a major problem in language acquisition research, which is essentially this: Why do children formulate the particular hypotheses that they do? A particularly notorious example involves a simple question pattern that has been crucial to the decades-long debate over the existence of UG.

(8) Can [$_{NP}$ birds [$_{S}$ that sing]] fly?

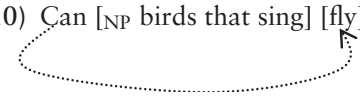
Without an *a priori* constraint, Berwick et al. (Ref 19, p. 1210ff) argue, a child could associate the auxiliary verb in (8) with either *sing* or *fly*. On their view, the right choice is made only because UG requires syntactic rules to be structure-dependent—i.e., to make reference to sentence structure rather than linear order. Thus, subject–verb inversion applies to the auxiliary verb in the main clause (a structural constraint) rather than to the first auxiliary verb in the sentence (a linear constraint). This gives the analysis in (9a), in which *can* is correctly associated with *fly* in the main clause—an interpretation that has been documented in children as young as age 2.[46]

(9) **a.** Structure-dependent—*can* is associated with the verb in the main clause:
Can [$_{NP}$ birds [$_{S}$ that sing]] _ fly?
(cf. Birds [that sing] *can* fly.)

b. Non-structure-dependent—*can* is associated with the first verb:
*Can [$_{NP}$ birds [$_{S}$ that _ sing]] fly?
(cf. Birds [that *can* sing] fly.)

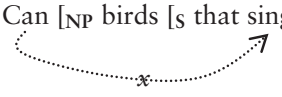
Processing considerations too favor this result, but in a different way. A processor committed to minimizing operational cost will opt for the analysis in (9a), in which *can* is associated with the verb in the same clause.

(10) Can [$_{NP}$ birds that sing] [fly]?



In contrast, the analysis in (9b) requires *can* to be associated with a verb in a lower clause that is itself further embedded inside an NP—at a cost that should deter language learners from treating it as a viable alternative to (9a).

(11) Can [$_{NP}$ birds [$_{S}$ that sing]] [fly]?



This gives the same result as the UG stipulation, but derives it from a commitment to parsimony that characterizes real-time processing in general.

CONCLUSIONS

In sum, we are left with something of a sea change in linguistics. There remains a significant explanatory role for 'innate domain-specific factors', Berwick et al. (Ref 19, p. 1207–1209) insist, but not as a matter of first recourse. There is a shared desire, they write, 'to reduce any language-specific innate endowment, ideally to a logical minimum', noting that 'responsible nativists try to account for the knowledge attained with the sparest plausible language-specific schematism'.

There are strong indications of crossing trend lines here: interest in the role of a third factor in the design and acquisition of language has risen sharply, while work on inborn grammatical principles has declined precipitously. Although these developments perhaps do not signal the 'end of history' for theoretical linguistics, the shift of focus to third-factor effects in generative grammar marks a milestone of sorts. Not because the idea is new, for it is not. Broadly speaking, the rest of the field has been committed to the primacy of third-factor explanations for decades. What is new is the opportunity—the first in half a century—for the discipline to focus on a common research question: What are the nongrammatical mechanisms and forces that shape language and contribute to its effortless acquisition?

No doubt, different perspectives will emerge. Indeed they already have, as can be seen in the contrast between explanations based on processing cost and those based on computational efficiency in the Minimalist sense. But at least there is now the realistic hope that these and other competing lines of research will be able to engage each other in productive ways, furthering the shared goal of understanding the mysteries of language.

NOTES

[a]This includes principles that may be part of 'extrabiological natural law' (Ref 2, p. 263).

[b]CP stands for 'complementizer phrase', an extended clausal projection that includes positions both for complementizers such as *that* and *whether* and for *wh* words. TP designates a 'tense phrase', corresponding roughly to the subject–predicate complex traditionally labeled 'S'. The vP projection provides a locus for the merger of agent arguments with verb phrases; thus a more precise representation would have the agent *they* originate within vP, before moving to TP.

ACKNOWLEDGMENT

I am grateful to Miho Choo, Peter Culicover, Lynn Nadel, Kevin Gregg, and two anonymous referees for comments on earlier versions of this essay.

REFERENCES

1. Chomsky N. *Reflections on Language*. New York: Pantheon; 1975.
2. Chomsky N. Language and other cognitive systems: What is special about language? *Lang Learn Dev* 2011, 7:263–278.
3. Chomsky N. Three factors in language design. *Linguist Inq* 2005, 36:1–22.
4. Chomsky N. *Barriers*. Cambridge, MA: MIT Press; 1986.
5. Chomsky N. *The Minimalist Program*. Cambridge, MA: MIT Press; 1995.
6. Hauser MW, Fitch T, Chomsky N. The language faculty: What it is, who has it, and how did it evolve? *Science* 2002, 298:1569–1579.
7. Coolidge F, Overmann K, Wynn T. Recursion: What is it, who has it, and how did it evolve? *WIRES: Cogn Sci* 2011, 2:547–554.
8. Boeckx C, Longa V. Lenneberg's views on language development and evolution and their relevance for modern biolinguistics. *Biolinguistics* 2011, 5:254–273.
9. Boeckx C. Round table: language universals: yesterday, today, and tomorrow. In: Piattelli-Palmarini M, Uriagereka J, Salaburu P, eds. *Of Minds and Language: A Dialogue with Noam Chomsky in the Basque Country*. Oxford: Oxford University Press; 2009, 195–220.
10. Chomsky N. Approaching UG from below. In: Sauerland U, Gärtner H, eds. *Interfaces + Recursion = Language?* New York: Mouton de Gruyter; 2007, 1–29.
11. Christiansen M, Chater N. Language as shaped by the brain. *Behav Brain Sci* 2008, 31:489–558.
12. Elman J, Bates E, Johnson M, Karmiloff-Smith A, Parisi D, Plunkett K. *Rethinking Innateness: A Connectionist Perspective on Development*. Cambridge, MA: MIT Press; 1996.

13. Benítez-Burraco A, Longa V. Evo-Devo—Of course, but which one? Some comments on Chomsky's anaologies between the biolinguistic approach and Evo-Devo. *Biolinguistics* 2010, 4:308–323.
14. Newmeyer F. Against a parameter-setting approach to typological variation. *Linguist Variation Yearbook* 2004, 4:181–234.
15. Evans N, Levinson S. The myth of language universals: language diversity and its importance for cognitive science. *Behav Brain Sci* 2009, 32:429–448.
16. Levinson S, Evans N. Time for a sea-change in linguistics: responses to comments on 'The myth of language universals.' *Lingua* 2010, 120:2733–2758.
17. Pinker S, Jackendoff R. The faculty of language: What's special about it? *Cognition* 2005, 97:201–236.
18. Newmeyer F. Review of *On nature and language*, by N. Chomsky. *Language* 2003, 79:583–599.
19. Berwick R, Pietroski P, Yankama B, Chomsky N. Poverty of the stimulus revisited. *Cogn Sci* 2011, 35:1207–1242.
20. Baker M. *The Atoms of Language: The Mind's Hidden Rules of Grammar*. New York: Basic Books; 2001.
21. Yang C. *Knowledge and Learning in Natural Language*. Oxford: Oxford University Press; 2002.
22. Longa V, Lorenzo G. What about a (really) minimalist theory of language acquisition? *Linguistics* 2008, 46:541–570.
23. Tomasello M. *Constructing a Language: A Usage-Based Theory of Language*. Cambridge, UK: Harvard University Press; 2003.
24. Lieven E. Input and first language acquisition: evaluating the role of frequency. *Lingua* 2010, 120:2546–2556.
25. Saffran J, Aslin R, Newport E. Statistical learning by 8-month old infants. *Science* 1996, 274:1926–1928.
26. Thompson S, Newport E. Statistical learning of syntax: the role of transitional probability. *Lang Learn Dev* 2007, 3:1–42.
27. Saffran J, Hauser M, Seibel R, Kapfhamer J, Tsao F, Cushman F. Grammatical pattern learning by human infants and cotton-top tamarin monkeys. *Cognition* 2008, 107:479–500.
28. Elman J. Computational approaches to language acquisition. In: Brown K, ed. *Encyclopedia of Language & Linguistics*, vol. 2. 2nd ed. Oxford: Elsevier; 2006, 726–732.
29. Reali F, Christiansen M. Uncovering the richness of the stimulus: structure dependence and indirect statistical evidence. *Cogn Sci* 2005, 29:1007–1028.
30. O'Grady W. Does emergentism have a chance? In: Chan H, Jacob H, Kapia E, eds. *Proceedings of the 32nd Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press; 2008, 16–35.
31. Saffran J. Statistical language learning: mechanisms and constraints. *Curr Dir Psychol Sci* 2003, 12:110–114.
32. Sebastián-Gallés N. Biased to learn language. *Dev Sci* 2007, 10:713–718.
33. Chater N, Christiansen M. Language acquisition meets language evolution. *Cogn Sci* 2010, 34:1131–1157.
34. Chomsky N. On phases. In: Freidin R, Otero C, Zubizarreta M, eds. *Foundational Issues in Linguistic Theory*, Oxford: Oxford University Press; 2008, 133–166.
35. Jaeger TF, Tily H. On language 'utility': processing complexity and communicative efficiency. *WIRES: Cogn Sci* 2011, 2:323–335.
36. Ferreira F, Christianson K, Hollingworth. A. Misinterpretations of garden-path sentences: Implications for models of reanalysis. *J Psycholinguist Res* 2001, 30:3–20.
37. Hawkins J. Processing complexity and filler-gap dependencies across grammars. *Language* 1999, 75:244–285.
38. O'Grady W. *Syntactic Carpentry: An Emergentist Approach to Syntax*. Mahwah, NJ: Erlbaum; 2005.
39. Gibson E. Linguistic complexity: locality of syntactic dependencies. *Cognition* 1998, 68:1–76.
40. Goodall G. On the syntax and processing of *wh*-questions in Spanish. *ProcWest Coast Conf Formal Linguist* 2004, 23:101–114.
41. Phillips C, Kazanina N, Abada S. ERP effects of the processing of syntactic long-distance dependencies. *Cogn Brain Res* 2005, 22:407–428.
42. Frazier L, Clifton C. Successive cyclicity in the grammar and the parser. *Lang Cogn Processes* 1989, 4:93–126.
43. Kluender R. On the distinction between strong and weak islands: a processing perspective. In: Culicover P, McNally L, eds. *The Limits of Syntax*, (Syntax and Semantics 29). San Diego: Academic Press; 1998, 241–279.
44. Hoffmeister P, Sag I. Cognitive constraints and island effects. *Language* 2010, 86:366–415.
45. Hawkins J. *Efficiency and Complexity in Grammars*. Oxford: Oxford University Press; 2004.
46. Crain S, Nakayama M. Structure dependence in grammar formation. *Language* 1987, 63:522–543.

