those assumed by Wang and Spelke illustrates this well. Therefore we need not only to raise hypotheses about ‘the mechanisms by which humans overcome the limits of our primitive navigational systems’ [1], but also to extend these same hypotheses to a vast spectrum of animals, in the study of which the lack of conceptual and technical tools might so far have prevented us from appreciating the true nature of their spatial representations.

References

Lasnik’s review of the Minimalist program in syntax [1] offers cognitive scientists help in navigating some of the arcana of the current theoretical thinking in transformational generative grammar. One might observe, however, that this journey is more like a taxi ride gone bad than a free tour: it is the driver who decides on the itinerary, and questioning his choice may get you kicked out. Meanwhile, the meter in the cab of the generative theory of grammar is running, and has been since the publication of Chomsky’s Syntactic Structures in 1957. The fare that it ran up is none the less daunting for the detours made in his Aspects of Theory of Syntax (1965), Government and Binding (1981), and now The Minimalist Program (1995). Paraphrasing Winston Churchill, it seems that never in the field of cognitive science was so much owed by so many of us to so few (the generative linguists).

For most of us in the cognitive sciences this situation will appear quite benign (that is, if we don’t hold a grudge for having been taken for a longer than necessary ride), if we realize that it is the generative linguists who should by rights be paying this bill. The reason for that is simple and is well-known in the philosophy of science: putting forward a theory is like taking out a loan, which must be repayed by gleaning an empirical basis for it; theories that fail to do so (or their successors that might have bought their debts) are declared bankrupt. In the sciences of the mind, this maxim translates into the need to demonstrate the psychological (behavioral), and, eventually, the neurobiological, reality of the theoretical constructs. Many examples of this process can be found in the study of human vision, where, as in language, direct observation of the underlying mechanisms is difficult; for instance, the concept of multiple parallel spatial-frequency channels, introduced in the late 1960s, was completely vindicated by purely behavioral means over the following decade (see, for example, [2]).

In linguistics, the nature of the requisite evidence is well described by Townsend and Bever: ‘What do we test today if we want to explore the behavioral implications of syntax? …the psychological basis for the two primary and ever-present operations, merge and move.’ (Ref. [3], p.82).

Unfortunately, to our knowledge, no experimental evidence has been offered to date that suggests that merge and move are real (in the same sense that the spatial-frequency channels in human vision are). Generative linguists typically respond to calls for evidence for the reality of their theoretical constructs by claiming that no evidence is needed over and above the theory’s ability to account for patterns of grammaticality judgments elicited from native speakers. This response is unsatisfactory, on two accounts. First, such judgments are inherently unreliable because of their unavoidable meta-cognitive overtones, because grammaticality is better described as a graded quantity, and for a host of other reasons [4]. Second, the outcome of a judgment (or the analysis of an elicited utterance) is invariably brought to bear on some distinction between variants of the current generative theory, never on its foundational assumptions. Of the latter, the reality of merge and move is but one example; the full list includes assumptions about language being a ‘computationally perfect’ system, the copy theory of traces, the existence of Logical Form (LF) structures, and ‘innate general principles of economy’. Unfortunately, these foundational issues have not been subjected to psychological investigations, in part because it is not clear how to turn the assumptions into testable hypotheses.

Lasnik is optimistic that Minimalism, which is “as yet still just an ‘approach’, a conjecture about how human language works (‘perfectly’)” (Ref. [1], p. 436), can be developed into an ‘articulated theory of human linguistic ability.’ Such optimism would seem to require that the foundational issues be thoroughly addressed, but to our
Linguistics and empirical evidence
Reply to Edelman and Christiansen

Colin Phillips and Howard Lasnik
Department of Linguistics, 1401 Marie Mount Hall, University of Maryland, College Park, MD 20742, USA

The main claims of Edelman and Christiansen’s (E&C’s) comment [1] on Lasnik’s article [2] are that generative grammar is built upon empirically weak (perhaps non-existent) foundations, and that generative grammarians aggressively resist experimental testing of their assumptions. Neither of these claims survives even brief scrutiny.

There is a long history of laboratory-based work that investigates foundational constructs of generative grammar. The operation merge, which produces recursive hierarchical structures, was already a focus of interest among psycholinguists 30 years ago [3]. Under the heading of the ‘binding problem’, the same issue is a leading problem in neuroscience research today [4]. The nature of long-distance dependencies, which some versions of generative grammar capture via movement operations, has been widely investigated using reaction-time methodologies since the 1960s [5]. The inconclusiveness of earlier results has led to progressive sharpening of the issues [6–8], particularly concerning traces. Developmental considerations have been decisive in areas such as binding theory [9], and argument structure [10]. More recently, techniques from cognitive neuroscience have been added to the inventory of tools [11,12]. This listing could be continued for pages.

Moreover, the list grows exponentially once we move beyond the methodological imperialism of E&C’s letter.

Gathering of native-speaker judgments is a trivially simple kind of experiment, one that makes it possible to obtain large numbers of highly robust empirical results in a short period of time, from a vast array of languages. Any good linguistics study involves carefully constructed materials, appropriate control items, and robust and replicable results. It is only because the technique is so easy and requires no more than a notebook that it is not usually described as an ‘experiment’. Outsiders would surely be puzzled by the attitude that seeks to deny the psychological relevance of easy, robust results, while insisting on other, far more subtle measures, such as 20-ms differences in reaction times, or 1-s changes in how quickly babies get bored, or 2% changes in regional cerebral blood flow. The variability that one observes in native-speaker judgments is real, but very small relative both to the variability in other measures (we have observed this repeatedly in our own studies). Furthermore, it is a truism in linguistics, widely acknowledged and taken into account, that acceptability ratings can vary for many reasons independent of grammaticality [15].

In language, as in any other area of inquiry, decisive evidence can – and does – come from a variety of sources, and it is hard to know in advance where the key evidence will come from. It is at best misleading for E&C to...