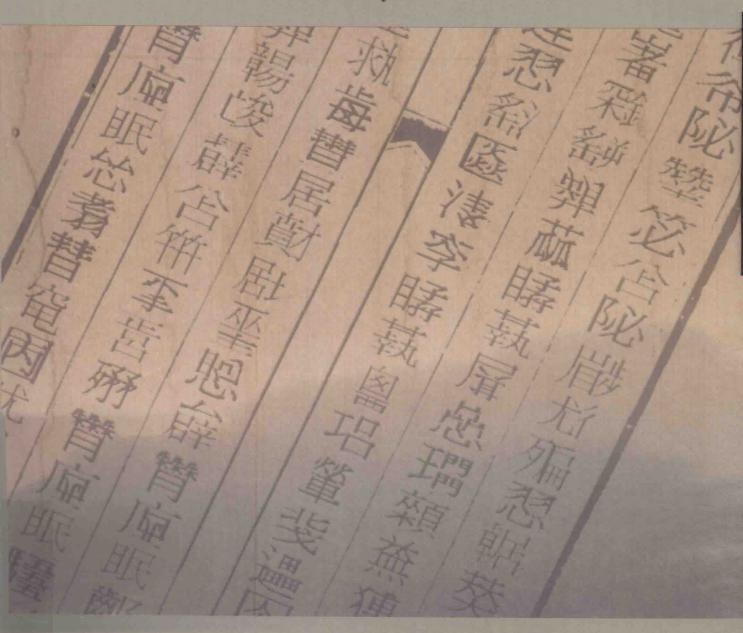# In Search of Grammar:
Experimental and Corpus-based Studies

Edited by
James Myers

In Search of Grammar:
Experimental and Corpus-based Studies

Edited by

James Myers

LANGUAGE AND LINGUISTICS MONOGRAPH SERIES 48

Edited by James Myers

In Search of Grammar:
Experimental and Corpus-based Studies

《語言暨語言學》

《語言暨語言學》隸屬於中央研究院語言學研究所，以出版語言學及相關領域之最新研究成果為宗旨。所有出版品均經過嚴格學術審查。

《語言暨語言學》專刊系列之四十八

語海尋規：語言學的實證方法

麥　傑　主編

LANGUAGE AND LINGUISTICS MONOGRAPH SERIES

Preface

This is a book for linguists who study grammar, that is, phonology, morphology, and syntax, and who are curious about new ways of studying it. Corpus linguistics can tell us not only how people actually use language, but also about the grammar that underlies language use. Experiments reveal not only the psychological processes controlling speech and language comprehension, but also the grammatical structures that these processes work with. Computer models and statistics can bring a greater degree of precision to the study of grammar, just as they have in many other fields of science.

Corpus analysis, experimentation, and computer modeling can at first seem intimidating to linguists more familiar with the traditional low-tech methods of grammatical research. Yet as this book hopes to demonstrate, it does not require much extra effort to expand one's methodological repertoire. The authors in this book, expert linguists from both Taiwan and abroad, provide easy-to-follow tutorials and case studies on a wide variety of grammatical issues. As befits a book published in Taiwan, many of the examples relate to challenges posed by the grammar of Chinese, but other languages are discussed as well. It is hoped that grammarians of all stripes, from syntacticians to phonologists, from formalists to functionalists, from students to tenured professors, will find inspiration in this book for their own research.

This book started life (many years ago now) in the International Workshop on Grammar and Evidence, held at National Chung Cheng University, in Chiayi, Taiwan, on April 13-15, 2007. This workshop was hosted by the Graduate Institute of Linguistics (Chung Cheng), cohosted by the Center for Research in Cognitive Science (Chung Cheng), and sponsored by the Center of International Affairs and Exchange (Chung Cheng) and the National Science Council (Taiwan). The website for the workshop, in both English and Chinese versions, remains available at www.ccunix.ccu.edu.tw/~lngproc/IWGE.htm. All of the chapters in this book saw their first incarnation at this workshop, but since not all of the workshop presentations saw their way into this book, the website remains a valuable resource, with slides and audio recordings of all of the presentations and the panel discussion too.

This book would not have been possible without the active participation of all of the workshop speakers. In addition to myself, they were Wayne Cowart from the University of Southern Maine, Michael Hammond from the University of Arizona, Mary Beckman from Ohio State University, Chao-Jan Chen from National Chi Nan University, Thomas

Lee from the Chinese University of Hong Kong, Jane Tsay from National Chung Cheng University, H. Samuel Wang from Yuan Ze University, and Chien-Jer Charles Lin, now at Indiana University.

The path from workshop to book required the hard work of yet more people, particularly the anonymous reviewers for each of the chapters and for the book as a whole, as well as Elizabeth Zeitoun, executive editor of *Language and Linguistics*. The book has been greatly improved by their insightful critiques and suggestions. I must also thank the *Language and Linguistics* editorial assistant Vicky Y-K. Chen for helping to coordinate the many strands of the editorial process, Chun-yu Kuo for shepherding the book through its final stages, and of course the publisher itself, the Institute of Linguistics, Academia Sinica.

Finally, for those readers puzzling over the text on the cover, yes, it is indeed a forgery containing nothing but fake characters. The frame was scanned from a facsimile of the Tang Dynasty *Book of Northern Qi* (北齊書) by Li Baiyao (李百藥), but all of the characters were created by my lab assistants (Guo-Ming Hong, Zi-Ping Hsu, Pei-Fen Du, Yu-Ting Su, and Chiung-Wen Hsu) for a series of experiments that will hopefully see publication some day. The images were put together by Yu-Guang Ko (who also suggested the book's Chinese title).

The point of the cover picture is to start the reader thinking about the theoretical and methodological issues addressed in this book. If all of the characters are fake, why do some look more plausible than others? Does this phenomenon demonstrate the "psychological reality" of formal character "rules"? Can we consider such rules to be part of a "grammar", or are they just the side-effects of superficial analogy? How could we ever find out?

I leave these questions, and an infinity of others about the nature of grammar and its empirical testing, to the interested reader.

James Myers
February 2012

Table of Contents

Chapter 1
Methods in search of grammar, grammar in search of methods[*]

James Myers
National Chung Cheng University

1. Introduction

Linguistics is the scientific study of language. Its core subject matter is mental grammar, that is, knowledge of linguistic structure. Mental grammar is shaped by universal human biology. The syntactic component of grammar is studied through native-speaker judgments of acceptability. The phonological component of grammar is studied through the analysis of lists of words and, occasionally, phrases.

At least these are the standard aphorisms in theoretical linguistics, explicitly stated or latent in mainstream practice. Eventually every budding linguist learns that all of these claims remain quite controversial, both inside linguistics and among researchers in the allied fields of cognitive psychology and neuroscience. Is grammar really mental, or just a convenient description of linguistic behavior? If there is a mental component, is it knowledge of linguistic structure ("knowledge that") or of language use ("knowledge how")? Is it in fact grammar that biology shapes, rather than other aspects of the language learning and processing system? Are native-speaker judgments consistent and unbiased enough to reveal reliable information about linguistic knowledge? Do patterns in word lists provide information about productive grammatical knowledge in contemporary speakers, or only about the history of the words themselves? Above all, can we honestly call linguistics scientific if there is so much controversy about such basic issues?

The confusion facing linguistic theory is reminiscent of the state of pre-Newtonian optics, vividly sketched by Kuhn (1970:12-13):

Each of the corresponding schools derived strength from its relation to some particular metaphysic, and each emphasized, as paradigmatic observations, the particular cluster of optical phenomena that its own theory could do most to explain. Other observations were dealt with by ad hoc elaborations, or they remained as outstanding problems for further research.... Being able to take no common body of belief for granted, each writer on physical optics felt forced to build his field anew from its foundations. In doing so, his choice of supporting observation and experiment was relatively free, for there was no standard set of methods or of phenomena that every optical writer felt forced to employ and explain.

For "optics" read "linguistics". Oversimplifying for rhetorical effect (as Kuhn presumably was doing as well), here we find one prominent school that takes native speaker judgments as paradigmatic evidence, while other schools challenge this school (and each other) with their own favored data sources (language processing, corpus data). Each school is satisfied to build its own version of linguistics in relative isolation from all of the others, with no generally agreed-upon methods for choosing among them.

Rather than bemoaning this sorry situation yet again (for a recent example of hand-wringing over the state of modern linguistics, see Carlson 2003), this book will instead ask how linguists might go about getting out of it. The guiding idea here is that scientific consensus relies on the one thing that all scientists undeniably have in common: observable reality. Newton did not have a unifying effect on optics solely through the sheer force of his ideas (some of which were immediately questioned; Shapin 1996). Rather, what won Newton's contemporaries over to his optical theory was his empirical argument for it, wherein he demonstrated certain undeniable facts through a series of experiments—cleverly designed, elegantly analyzed, and above all, replicable.[1] The best way to encourage greater cohesiveness in linguistic theory, then, is to encourage linguists to feel some responsibility towards the whole wide range of linguistic data, not just their own favorite sorts.

Though theoretical linguists have never entirely neglected "non-traditional" data sources, only in the past couple of decades has the trend begun to enter the mainstream, turning the non-traditional into the traditions of the future. The movement in this direction seems to have started to gain momentum a little over twenty years ago, in phonology, when Ohala & Jaeger (1986) and Kingston & Beckman (1990) demonstrated

---

[1] Historians and philosophers of science will recognize this version of the story as rather simplistic. Among other things, scientists surely share not just observable reality, but also strong biases about what makes a theory convincing. Nevertheless, while scientists often disagree about theory and even methodology, replicable observation seems to be a universally agreed upon virtue.

the advantages of adopting experimental, corpus-based, and quantitative methods from psycholinguistics and phonetics to the testing of hypotheses in theoretical phonology. The syntacticians entered the game a few years later, when the publication of Schütze (1996) and Cowart (1997) sparked new interest in improving the collection and analysis of native-speaker judgments in syntax via stricter experimental standards adapted from cognitive psychology. Since then many anthologies (including journal special issues) have been published on empirical methods in theoretical linguistics, including Borsley (2005), Coetzee, Kager & Pater (2009), Cole & Hualde (2007), Everaert, Lentz, De Mulder, Nilsen & Zondervan (2010), Everaert, Musgrave & Dimitriadis (2009), Fanselow, Fery, Schlesewsky & Vogel (2006), Featherston & Sternefeld (2007), Featherston & Winkler (2009), Gonzalez-Marquez, Mittelberg, Coulson & Spivey (2007), Kepser & Reis (2005), Krifka (2007), Meibauer & Steinbach (2011), Penke & Rosenbach (2007a), Rice & Newman (2010), Ritter (2002), Solé, Beddor & Ohala (2007), Stefanowitsch (2007), Stefanowitsch & Gries (2007), and Winkler & Featherston (2009). Even more importantly, perhaps, it is becoming ever more commonplace for experimental, corpus-based, and/or computational studies to appear in the core linguistics journals alongside the more "traditional" articles.

Yet this revolution in methodology has yet to have much effect on how linguistics is taught. The dedicated student can certainly learn a lot by carefully examining, and then attempting to emulate, the many recent research studies applying non-traditional data sources to grammatical questions, but this may not be the easiest way to start expanding one's methodological toolbox. After all, most of these studies are not intended to serve a pedagogical purpose, and many already assume a certain level of familiarity with experimental or quantitative methods. Some specialized tutorials do exist, including introductions to quantitative corpus linguistics (e.g. Kennedy 1998, Manning & Schütze 1999, Gries 2009) and Cowart's (1997) introduction to syntactic judgment experiments, but as yet there is no general resource providing gentle introductions to the wide range of empirical tools available for testing grammatical hypotheses.

This book provides a step in this direction. In one concise volume, it collects not only applications of non-traditional empirical methods in a variety of original linguistic studies, but also step-by-step tutorials, rich in examples, on how linguists may apply such methods in their own research. The contributions are selected to provide a balanced overview of the methodological landscape. Thus the book covers experimental, corpus-based, and computational approaches, addresses issues in syntax, morphology, and phonology, and includes data from English, Chinese, and a smattering of other languages.

Before introducing these contributions, in §2 we first revisit the question of why linguists should worry about methods in the study of grammar. Section 3 asks what grammar is anyway, why it cannot be observed directly, and why the standard methods

we use to study it have remained standard for so long. Finally in §4, we take a peek into the future of linguistic methodology hinted at by the contributions in this book.

2. Methods in search of grammar

Arguing that linguists should improve their methods may seem as irrelevant as arguing that the sun should rise; of course linguists will improve their methods in the course of time, since that's just the nature of science. After all, nobody could deny that the range of methods available to linguists today is a vast improvement over those available one hundred years ago. Yet arguments over methodology are also part of science. Methodological disagreements are particularly lively in social sciences like linguistics, presumably because human behavior is less well-behaved than the systems studied by physicists, chemists, and biologists.

Nevertheless, there are still many linguists who don't see what all the fuss is about. The traditional methods (traditional at least since the generative revolution of the 1950s) are working just fine, they say; supplementing or replacing them with more rigorous methods is more work than it's worth. Newmeyer (1983, 2007), Phillips & Lasnik (2003), and Phillips (2009) make this point explicitly with respect to informal acceptability judgments in syntax; Chomsky & Halle (1968) and Bromberger & Halle (1989) justify traditional phonological methods (i.e. studying word patterns in dictionaries) on philosophical grounds; Pullum (1983[1991]) argues more generally that linguists shouldn't waste time on metascientific issues at the expense of doing actual linguistic research.

This section attempts to pop this bubble of complacency with a two-pronged fork.[2] First, in §2.1, I examine the apparent source of the neglect of methodological issues in modern linguistics, namely the confusion between two distinct concepts: "empirical" and "empiricism". Then in §2.2, I show how the lack of methodological rigor has repeatedly got theoretical linguists into trouble.

2.1 A rational approach to empiricism

Calls to improve, or at least to think more deeply about, the methodology of linguistics have always been with us, but they tend to come with a certain amount of controversy. One possible reason for this is that they frequently get mixed up in an

---

[2] Featherston (2007) takes a similar tack, albeit using a different metaphor, urging empirical reform in syntax by offering both a carrot (the new theoretical insights made possible by improved methods) and a stick (the empirical mistakes that can arise when relying solely on traditional methods).

ancient debate over the source of human knowledge.

On the one side is empiricism, the view that knowledge derives primarily from our senses, advocated in contemporary linguistics most strongly by corpus linguists like Geoffrey Sampson (2001:1), who states that "if we want to deepen our understanding of language, our best way forward is to apply the same empirical techniques which have deepened our understanding of other observable aspects of the universe during the four centuries since Galileo." Sampson (p.1) defines the "empirical scientific method" as a bottom-up process: "Listen, look. Summarize what you hear and see in hypotheses which are general enough to lead to predictions about future observations." For Sampson, this means that the only scientifically appropriate linguistic data are corpus data, given that "[l]anguage ... is a concrete, tangible aspect of human behaviour," that is, directly observable via the senses, unlike abstract, intangible native-speaker intuitions.

On the other side is rationalism, the view that knowledge derives from deductive reasoning rather than solely from experience, exemplified in contemporary linguistics by Noam Chomsky, who applies rationalism not only to babies but to scientists as well. To take just one recent quotation (see Botha 1989 for others), Chomsky (2002:102) describes the development of modern science in terms of "the Galilean move towards discarding recalcitrant phenomena if you're achieving insights by doing so, the post-Newtonian concern for intelligibility of theories rather than of the world, and so on. That's all part of the methodology of science." Thus according to Chomsky, science became science precisely by becoming less dogmatic about the primacy of observation.

This view may seem peculiar at first; as Chalmers (1999) points out (in a highly accessible introduction to the philosophy of science), the empiricist view, in which tangible, real entities force their way unbidden into our attention, is the naive view most people have about science works. Problems with this naive view become clear, however, when one recognizes the logical impossibility of defining "evidence" completely independently of "theory." Far from facts speaking for themselves, the recognition of evidence as real and relevant is a complex psychological and social process (as Kuhn and others have documented). Even raw observations require some interpretation; in the late nineteenth century Giovanni Schiaparelli observed channels on Mars, but they later proved to be optical illusions (Yule 1978). Another problem is that history's greatest scientists, from Galileo on down, have tended to be a bit more stubborn in their theoretical positions than a strictly bottom-up, ugly-facts-kill-beautiful-theory empiricist view would lead us to expect, and this stubbornness is at least partly responsible for their greatest achievements. Still another problem is that the link between observation and hypothesis always involves subsidiary hypotheses, which require their own observations for support, which then require further subsidiary hypotheses, and so on (this is sometimes called the Duhem-Quine thesis; Gillies 1998). Thus even data in corpus linguistics do not

speak for themselves; interpreting them requires, among other things, subsidiary beliefs about the representativeness of corpus data and the link between language production and mental knowledge.

The most productive way out of the empiricist/rationalist debate is not to take sides but to try to get both sides to work together. Note the peculiar fact that Sampson invokes Galileo as the paradigmatic objective observer, while Chomsky praises him for playing fast and loose with facts. Historically, both views about Galileo are correct (see Botha 1982), and they are true more generally as well: modern science is neither pure rationalism nor pure empiricism, but a novel (if sometimes awkward) synthesis of these two ancient traditions. It's hard to maintain the balance, and linguistics is hardly alone in having scholars who tilt more towards one side than the other. Shapin (1996) contrasts the mathematically elegant Newton (another of Chomsky's heroes) with his contemporary Boyle, who was a far more careful experimentalist, yet so cautious about theorizing that he never even formalized the gas law that today bears his name. Shapin calls their contrasting approaches simply "different games that natural philosophers might wish to play" (p.117).

Thus in advocating that linguists pay more attention to evidence, I emphatically do not favor empiricism over rationalism: one without the other just isn't science. Instead, my goals are purely pragmatic: observations are public, concepts are private, so consensus in science ultimately depends just a bit more on the former than the latter. Consensus in science thus requires some sort of general agreement about methodology, and just a bit more respect to the data favored by competing schools.

2.2 What could possibly go wrong?

In a defense of traditional syntactic acceptability judgments, Phillips & Wagers (2007:740) are careful to limit their discussion to the "best practices" in the discipline. Yet this leaves as open questions how well linguists keep up these high standards in general, and whether the risk of empirical blunders could be reduced by adopting more stringent methodological protocols. Phillips & Wagers (2007:740) themselves go on to note that "[i]t is not difficult to find instances of careless misrepresentation of linguistic data." In this section I cite several examples.

Starting with syntax, consider the Chinese sentence in (1). Huang (1982) claims that for syntactic reasons, *shenme* ("what") can have wide scope here, so that the sentence can mean "What was bought, such that you wonder who bought it?" Xu (1990) denies that this meaning is possible for this sentence. Who is right? How could we be sure?

(1) Ni   xiang   zhidao   shei   mai-le   shenme?
     you   want   know   who   buy-ASPECT   what

Such disagreements over acceptability (see Schütze 1996, Featherston 2007, and Myers 2009a, 2009b for further examples) may arise due to different idiolects (though Labov 1975, 1996 argues that the empirical evidence for idiolects, divorced from speech communities, is very weak), differences in speakers' thresholds for "acceptable" versus "unacceptable" (see Sprouse 2009 for data relating to this notion), differences in the theoretical interpretation of clear judgments (as argued in Newmeyer 1983, 2007), differences in sensitivity to imagined pragmatic context (Carroll, Bever & Pollack 1981 even show how judgments can be affected by the presence vs. absence of a mirror, supposedly influencing how speakers see themselves as social beings), or random variations across speakers within a speech community, or indeed from one judgment to the next by the same speaker (a central issue in Cowart 1997).

The fact that judgment data and their interpretation can be so ambiguous has led some critics to dismiss them entirely (e.g. Sampson 2001, 2007). A more pragmatic approach, however, is to recognize their enormous value in generating hypotheses, and even in resolving empirical debates when they are not ambiguous, while still working to understand and reduce their limitations. As argued by Schütze (1996, 2011), Cowart (1997), Featherston (2007), and Myers (2009a, 2009b), this can be done by collecting and analyzing judgments in accordance with protocols building on the almost two centuries of experience that experimental psychologists have developed for dealing with their similarly variable and multifaceted behavioral response measures.

A key part of these protocols is experimental design, a point emphasized in Cowart (1997) and Myers (2009b). To illustrate what can go wrong when linguists fail to design informal judgment experiments carefully, consider the English judgment contrast in (2) claimed by Di Sciullo & Williams (1987:33). These examples are intended to show that affixes like *-er* are transparent to the assignment of semantic roles (*bread* is still the patient of *bake*), whereas roots like *man* are not.

(2) a.   a baker of bread
     b.   *a bake-man of bread

However, the claim actually involves a relationship between two elements, *bake* and *bread*, and therefore the implicit design of the judgment experiment should involve two factors, not just *-er* versus *man*, but also *bread* versus no *bread*. Adding the [±*bread*] factor results in the two additional forms in (3), where (3b) is already worse than (3a), even without the presence of a patient argument. Moreover, the unacceptability of (3b)

is no accident; as Spencer (1991:333) points out in his commentary on this failed argument, *man* generally doesn't attach to verbs, especially not transitive ones (see Myers 2007 for further discussion).

(3) a.   a baker
     b. *a bake-man

Phonologists also sometimes employ native speaker judgments, as when dealing with phrasal patterns. Thus a whole stratum in Halle & Mohanan's (1985) model of English lexical phonology is based on a rule of /l/ resyllabification which they justify using the phonetic intuitions given in (4) (based on their figure (20), p.65).

(4) a.   palatalized /l/:   a whale edition
                           the seal office
     b.   velarized /l/:   the whale and the shark
                           the seal offered a doughnut

Unfortunately for their analysis, Sproat & Fujimura (1993) showed, using instrumental phonetics, that /l/ resyllabification actually seems to involve not a binary palatalized/velarized /l/ contrast but a complex interplay between prosody and the two subgestures involved in forming the lateral (lowering the tongue body and raising the tongue blade). Unsurprisingly, then, Hayes (2000) found that there was a considerable degree of variability and gradience in native-speaker acceptability judgments for palatalized versus velarized /l/. Given that these later studies were based on a larger amount of data, collected in a more systematic fashion that those cited by Halle & Mohanan, and then analyzed quantitatively, they are inherently more convincing.

The dominant traditional method in phonology involves not judgments but the analysis of word sets, taken from a dictionary or elicited from a native speaker. Like acceptability judgments in syntax, dictionary data have earned their central place in phonology by providing systematic arguments for a whole host of fundamental theoretical concepts and are unlikely to disappear from the phonologist's toolbox any time soon. Yet detecting patterns in a data set is essentially a statistical problem, and phonologists who neglect to quantify can get stuck.

A striking example of this is the confusion over tone spreading in Mende. On one side, Gussenhoven & Jacobs (2011:159) want to show that in this language, tone consistently spreads left to right, and so they choose examples from Leben (1978) like those in (5), with surface high-low-low and low-high-high patterns. By contrast, Zoll (2003:230-231) wants to show that the direction of tone spread in Mende depends on

tone quality (high vs. low), and so she chooses examples from Leben (1978) like those in (6), with surface high-low-low and low-low-high patterns. Presumably the examples chosen by these authors are all genuine Mende words, but only Gussenhoven & Jacobs (2011) claim that the pattern in their sample represents "the majority of Mende words" (p.159). This essentially statistical point seems quite relevant if we want to know which of these two analyses may be a more accurate description of Mende mental grammar.

(5) a. [félàmà] "junction"
     b. [ndàvúlá] "sling"

(6) a. [félàmà] "junction"
     b. [lèlèmá] "mantis"

Linguists are only human. Nevertheless, it seems likely that the risk of such problems could be reduced by paying more attention to methodology, both in the training of linguists and in the evaluation of their research by their peers.

3. Grammar in search of methods

Advocating better methods makes obvious sense if we are searching for generally recognized entities that merely happen to be hard to reach, such as underground oil reserves. Yet not only is grammar hidden, but its very nature is subject to much controversy, with some doubting that it even exists.

Not quite knowing what one is looking for is hardly a unique situation in science. As I point out in §3.1, history suggests that the most promising strategy is to swallow one's doubts and start somewhere anyway. Surprisingly perhaps, in the case of grammar, linguists of competing schools already tend to agree about several of its key features. One of these features, as discussed in §3.2, is that grammar can be inferred only indirectly via observed linguistic behavior. The methodological consequence of this is that we must proceed cautiously when building chains of inferences from observation to hypothesized grammar.

3.1 Where to start the search for grammar

It might seem natural to start the search for grammar by defining explicitly what we mean by grammar. However, practicing scientists rarely, if ever, begin this way. Instead, scientists generally start out with a vague notion of what they're looking for, and then as they work, they gradually shift their focus until they may end up somewhere

quite different from where they were first headed.

Churchland (2002), who is interested in the study of consciousness, a concept even more elusive than grammar, illustrates this point with a number of historical examples. Take "fire" for instance. Initially (and still naively today), this concept seemed to apply both to a candle and to the sun, since both give off heat and light. Today we know, however, that the sun's heat and light are the result of nuclear reactions whereas the candle's result from the chemical reaction of oxidation. Thus the candle's flame actually classes with rust, also caused by oxidation, rather than with the sun. If scientists had insisted all along that, by definition, the candle and the sun must be identical, they may never have discovered that they actually weren't. In the same way, we don't know exactly what "grammar" will turn out to be in reality, and so it would be a bad idea to fix a definition for it before starting the search.

But how can it be possible to study something without defining what it is that we are studying? Again, history can be our guide. The notion of fire began with relatively intuitive and pre-theoretic observations about the world. By analogy, then, our quest for grammar should start with views about it that are arguably most "intuitive" (at least among linguists).

Despite the rancor across the various schools of linguistics, it is surprisingly easy to come up with a list of features that most linguists would agree to hold of "grammar." First, grammar is that which is covered in a traditional descriptive grammar book, namely pronunciation (phonology) and the formation of words and sentences (morphosyntax). Grammatical analyses have existed for centuries and in many cultures, from Babylonian descriptions of Sumerian (Civil 1994) to Pāṇini's analysis of Sanskrit (Cardona 1994) to traditional Chinese phonology (Downer 1963).

Grammar is also seen (by definition perhaps) as systematic (i.e. regular, lawful, productive, rule-governed, constrained, structured), which means that in principle it goes beyond rote memorization and ad hoc analogy (though the borderline between grammar and analogy is sometimes seen as fuzzy; see e.g. Albright & Hayes 2003). Seuren (1998) traces the notion of grammar as system surprisingly far back in the linguistic tradition in the West.

Another central aspect of the notion of grammar is that it is mental. This has been the explicitly stated mainstream view since the cognitive revolution of the 1950s, but its roots are older. Sapir (1933[1949]) made the phrase "psychological reality" famous, Bloomfield (1914) assumed linguistics was a "mental science" (though in later work like Bloomfield 1933 he made a sharp division between the work done by linguists and by psychologists), and so on back through the centuries (again, see Seuren 1998). Long before the scholars got involved, folk psychology already saw a language as something that people "know." Mentalism is also embraced by the large contingent of Chomsky's