1

Sentence comprehension before and after 1970: Topics, debates, and techniques

MONTSERRAT SANZ, ITZIAR LAKA, AND MICHAEL K. TANENHAUS

What is the Science of Linguistics a Science of?

(Bever 1970: 346)

1.1 Introduction

In "The cognitive basis for linguistic structures" (CBLS), published in 1970 and reprinted in this volume, Bever argued that specific properties of language reflect general cognitive laws. At the time, the competence vs performance dichotomy put forth by Chomsky (1965) was generally accepted, and hence linguistics developed on the assumption that the process of language acquisition was the attainment by children of the type of full competence adult speakers have. Given that linguistic theory accounted for grammar, the job of psycholinguists was to unveil how independent cognitive constraints, such as working memory, interact with grammar to result in linguistic performance.

In contrast to this view, Bever has argued that some formally possible, and therefore grammatically valid, structures never manifest themselves in natural language because children cannot understand, use, or learn them. In other words, Bever proposed that mechanisms of language learning and processing partially determine the form of grammar. This was a significant shift from the conception of grammar generally accepted in the Sciences of Language at the time, especially within generative linguistics. The consequences of this shift of perspective continue to unfold. This book presents a somewhat

impressionistic review of the many areas of research that have emerged since this new conception of the relationship between grammar and cognitive processes was hypothesized in 1970. The chapters show the evolution of some of these ideas in the last decades and the paths that researchers have explored, abandoned, or pursued. In this chapter we group chapters according to the major areas of inquiry that have emerged in the last four decades of research in the field of sentence processing. Reading them gives one an idea of the historical evolution after 1970. But, first we provide some historical context about some of the events that led Bever to the proposals presented in CBLS.

1.2 CBLS: A serendipitous beginning[1]

For some of our readers, knowing the history of what led to the philosophical change about the form of grammar and cognition may be as interesting as evaluating the current state of the field. The chain of events that led to the arguments in CBLS reminds us of how serendipitous events can lead to fundamental insights that open new avenues of scientific inquiry.

One of the pillars of the arguments in CBLS is children's perception of numerosity (see the reprint of CBLS in this book). A visit by Bever and Mehler as graduate students to Piaget's lab in Geneva in the mid-1960s inspired this research. Prior to that visit, the Harvard Center for Cognitive Studies had provided the nurturing environment for an emerging field to flourish, mainly under the guidance of George A. Miller. Harvard also hosted Lenneberg, a psychologist well versed in biology, who influenced generations of psychologists to explore the biological bases of language. As Mehler (personal communication) points out, at the Harvard Center of Cognitive Sciences, students and postdocs were starting the investigations that helped consolidate the nascent cognitive revolution. Bever and Mehler belonged to that group.

Nearby, the new Psychology Department at MIT, founded by Hans Lukas Teuber, promoted three areas of research: brain studies, including anatomy, neuroscience, and neuropsychology; development of visual perception and motor systems; and cognitive processes, with a strong emphasis on language processing. Teuber hired Jerry Fodor, Thomas Bever, and Merrill Garrett, who collaborated on pioneering research and later combined to write the first book that integrated ideas in philosophy, psychology, and linguistics to

[1] We have mainly taken the data and facts for this section from a piece written by Jacques Mehler originally for this book that was not included in the final version. Our deepest gratitude to him for sharing his personal memories with us.

explain what was called language performance (Fodor, Bever, and Garrett 1974).

It was in this atmosphere and after their experience in Geneva that Bever and Mehler decided to replicate some studies by Piaget. In particular, at Piaget's lab they had found that conservation of number, volume, and other parameters are mastered in children older than four-and-a-half years of age. After their return, at MIT and Harvard, Mehler and Bever decided to test children then considered too young for the techniques available: they tested one-and-a-half and two-year-olds for number conservation, as well as older children. Along with the classic technique, they utilized an innovative method that required little verbal response from the children: groups of M&Ms that the children could choose from, and later clay pellets.

Surprisingly, they found that young children's ability to make correct relative judgments actually declined between the ages of three and four. By the time children approach four years of age, they show the worst performance in making numerosity judgments, although they overcome this by using overt counting a few months later. These experiments were pioneering in testing very young infants and constituted a turning point for cognitive development: the first demonstration of the now well-established U-shaped curve in development. The U-shaped curve was explained by noting that younger infants estimate numerosity whereas older children use counting, a learned ability. This work helped usher in a new cognitive perspective in development and, importantly, linked cognitive development to language acquisition (Bever and Mehler 1967; Bever et al. 1968). The results also provided the foundation for many of the ideas that were later developed in CBLS, which can be considered the beginning of the current biolinguistic approach to the study of language.

In sum, the desire to expand Piaget's experiments to younger infants led Bever and Mehler to discover a surprising and unexpected pattern of learning. Moreover, an experiment on the perception of numerosity led to a new hypothesis about language acquisition. At the root of it all was an encounter between a legendary scientist and two graduate students, who were inspired and challenged by ideas from great teachers, and nurtured by innovative university departments.

In CBLS, Bever combined insights from experiments on numerosity perception in children with an analysis of several constructions in English and compelling examples like *The horse raced past the barn fell* to argue that psycholinguistic science would have a very important impact on linguistics. Mehler recalls that when that paper was published, the claim that certain ostensively grammatical structures might arise from domain-general

behavioral systems was widely viewed as an important challenge to both linguistic theory and methodology. He also notes that, whereas Bever's claim was initially resisted by most linguists, it is now endorsed by many cognitive scientists, including Chomsky himself (Hauser, Chomsky, and Fitch 2002) and, as we said, is now embedded in the biolinguistic approach to the study of human languages. Many of the chapters in this volume show how far the field has traveled along the path started by CBLS.

1.3 Setting the path for language-processing research

The main question that Bever addressed—the basic issue in the study of language processing at the time—was: how can children arrive at internal linguistic structures from external input sequences? Bever hypothesized several speech perception strategies that children adopt along their way to the discovery of grammar, among them segmentation, semantic labeling, and sequential labeling strategies, which would operate simultaneously in actual speech perception.

Segmentation strategies help the learner divide the external sequence into smaller and grammatically meaningful chunks. Strategies A and B are of this kind:

> **Strategy A**: Segment together any sequence X . . . Y, in which the members could be related by primary internal structural relations (actor–action–object . . . modifier).
>
> **Strategy B**: The first N . . . V (N) . . . clause (isolated by Strategy A) is the main clause, unless the verb is marked as subordinate.

Once these segmentation strategies have identified primary components of a sentence, labeling strategies assign semantic relations between them. Strategy C is of this type.

> **Strategy C**: Constituents are functionally related internally according to semantic constraints.

Sequential labeling strategies (like Strategy D, for instance) establish a mapping between the form and the semantics of an NVN sequence:

> **Strategy D**: Any noun–verb–noun sequence within an internal potential unit in the surface structure corresponds to the "actor–action–object."

By developing strategies of this kind, children learn the grammar corresponding to sentences they can understand, but not grammars corresponding to those they cannot.

Bever provides a series of examples of English grammatical structures whose form is consistent with these acquisition strategies: in particular,

main and subordinate clauses and the integrity of main clauses in external structure, relative clauses with deleted relative pronouns, and syntactic restrictions on pronominalization and on pronominal adjective ordering (see the reprint of CBLS in this volume for specific examples).

Bever also discussed some possible universal restrictions on linguistic complexity based on perceptual biases, such as restrictions on intervening constituents that interfere with the processing of a unit they are part of (**John called the not very well liked but quite pretty girl on the next block where Jack has lived for years up*), and in particular, restrictions on center-embedding sentences like *The dog the cat the fox was chasing was scratching was yelping.*

Relative clauses lacking a relative pronoun or complementizer, such as the now famous *The horse raced past the barn fell*, which would later come to be known as "garden path sentences," mislead comprehenders to initially pursue an analysis that turns out to be incorrect, as if taking the wrong turn while walking through a maze in a garden. A garden path sentence contains a temporary ambiguity (in this case, a verb root with an *-ed* suffix, which in English can correspond to either a past tense or a past participle). It becomes clear that something has gone wrong when the comprehender receives input that is incompatible with the past-tense analysis of the ambiguous form (in our example, this point is reached at the verb *fell*). The parser then must abandon its initial parse and attempt to reanalyze the input.

It was this example (*The horse raced past the barn fell*), mentioned in passing in CBLS to illustrate the power of perceptual and statistical strategies in shaping language, that caught the attention of researchers and helped spawn an entire subfield, sentence processing. In the sentence-processing community, language scientists from multiple disciplines explore the architecture of the language processor and the interaction between syntax and semantics during comprehension as a sentence unfolds in time. Garden path sentences have played a central role in motivating and testing alternative theories of language processing, typically under the assumption that syntax and semantics are distinct although interacting systems.

A series of processing models ensued, from those that proposed parsing preferences that prioritize syntax over semantics, to multiple constraint satisfaction models that establish the early influence of a variety of semantic factors in sentence processing (see the chapters by Altmann, MacDonald, and Fodor in this volume). In the four decades after CBLS was published, researchers have converged on the view that not only is semantics directly involved in sentence processing at early stages, but other language-independent cognitive factors also determine the structure initially assigned to the input when parsing. Among them, we can cite prediction based on statistical

properties of language or the cognitive prominence of actors in identifying goal-directed actions (see, for instance, Altmann, Dell and Kittredge, and Bornkessel-Schlesewsky and Schlesewsky, this volume). Although local syntactic ambiguity was the initial test bed for developing and evaluating different models of language comprehension, it has become clear throughout the years that ambiguity is an inherent feature of all linguistic forms. In language processing, even the simplest and superficially unambiguous sentence is temporarily ambiguous.

In the following sections, we summarize some of the main topics around which research in the field has progressed. As we will see, what becomes clear from the review of the chapters of this book is that some conclusions about the architecture of the sentence-processing mechanism are well-grounded in clear and uncontroversial empirical findings, among others: (a) previous linguistic experience modulates sentence-processing choices (i.e., prediction informed by distributional patterns modulates language processing); (b) animacy has a significant impact in syntactic processing (i.e., animate entities are more prominent than inanimates); and (c) the processes of production, comprehension, and acquisition interact in significant ways. We place the chapters in context as we examine how each one of them exemplifies one of the areas of research that have been developed in these decades.

1.4 Factors at play in processing

The chapters by Altmann, MacDonald, Lin, Bornkessel-Schlesewsky and Schlesewsky, Dell and Kittredge, and Townsend in this volume address some of the turns that the path of research has taken as language scientists have explored the specific role of syntax, semantics, and other cognitive factors in syntactic processing. In reading these chapters, it becomes clear that the mapping between syntax and semantics in processing is now viewed as an interaction shaped by cognitive factors including the properties of event change, a preference for goal-directed actions that involve agents, variation in the forms across languages, extralinguistic contextual information, properties of the neural system, and prediction based on previous experience, among others. Decades of research with different techniques, linguistic phenomena, and languages have converged on models of sentence comprehension in which constraint satisfaction plays a central role, although, as Altmann concludes, the exact implementation of these models is still hotly debated.[2]

[2] For this, connectionist and Bayesian models are proposed. Dell and Kittredge present evidence from connectionist models, whereas Dunbar, Dillon, and Idsardi, in a study of phonology, as we will see later, defend Bayesian models of prediction.

Gerry Altmann reviews the research trends that resulted from *The horse raced past the barn fell* and points out how research on these models expanded from ambiguity resolution to the general mechanisms for sentence processing. It becomes clear through his review that the boundaries between syntax and semantics became more and more blurry as extralinguistic factors made their way into models of sentence comprehension.[3]

Initially, at a time in which the strict modularity of cognitive processes was widely assumed, Altmann recounts how, despite the availability of a grammatical formalism in which syntax and semantics worked as one (namely, Combinatory Categorial Grammar), researchers shied away from accounts of syntactic ambiguity resolution where syntax and semantics acted simultaneously. Most of them adhered to an architecture where syntax proposed alternative structural interpretations which semantics could choose from on the basis of the best contextual fit. Furthermore, according to Altmann, researchers did not realize at the time that the really interesting action happened before the point of ambiguity, and not only in the sentence or in the text/discourse, but in the experience of the language user. In other words, accumulated experience with language leads to prediction, one of the mechanisms at play in sentence comprehension. Prediction applies to any kind of sentence, even unambiguous ones, and does not only refer to the upcoming linguistic input, but to the conceptual representations of the concomitant changes in the real world that would constitute the event described by the sentence at issue.

In sum, the view that the structure of language may not be totally independent from cognitive mechanisms, originally put forth in CBLS, has led researchers to explore how basic cognitive mechanisms underlie language processing. This has resulted in models that incorporate factors such as prediction based on previous experience. Researchers like Altmann turned from addressing only temporally ambiguous sentences to analyzing apparently unambiguous simple sentences: due to the role of prediction, ambiguity appears at every point in comprehension.

Maryellen MacDonald agrees that there is no need to divide the field into two different research avenues that separate the study of ambiguity resolution from other aspects of syntactic processing. She supports her claims with studies of the processing of relative clauses with and without a relative

[3] Current Minimalist approaches to language share one of the basic insights from CBLS: that perhaps the analyzable part of language is that which has to do with the constraints imposed by the cognitive system; that is to say, that the form of language is partially determined by conditions imposed by processing constraints.

pronominal/complementizer (garden path sentences vs regular relative clauses). The latter can be of two kinds, subject or object, depending on the grammatical function corresponding to the gap in the relative sentence. The observed generalization is that subject relative clauses are faster and easier to process than object relative clauses. After a review of the different explanations that have been provided for the increased difficulty of object relative clauses, MacDonald proposes to unify ambiguous and unambiguous sentence-processing mechanisms under a probabilistic constraint model.

According to MacDonald, it is not only the case that object relative clauses are more difficult to process than subject relative clauses: this difficulty is also modulated by animacy. To explain this, the author adopts an approach that explores the connections between the production and the comprehension systems—what she calls the *Production Distribution Comprehension* (PDC) account of language processing. Speakers generate object relatives naturally when talking about an inanimate being, but rarely when referring to animates. Therefore, the difficulty of object relatives would stem from the fact that speakers do not expect this type of phrasing when an animate is involved. In MacDonald's view, this is an experience-dependent phenomenon, not the product of inherent limitations in working memory. In a similar vein to Altmann, MacDonald thinks that the probabilistic constraints for relative-clause processing emerge, among other things, from prior experiences with form–meaning pairings, acquired via statistical learning by speakers.

In addressing the question of why object relatives should be more frequent with inanimates than with animates, MacDonald draws upon Bever's insight that language must have a form that speakers can comprehend, that children can learn, and that people can utter. The PDC account of language processing works in the following way: first, the computational difficulty of mapping between meaning and utterance form leads speakers to favor utterance forms that reduce the difficulty of the language production process. These production choices result in linguistic distributional patterns that are learned. Finally, this knowledge is used to guide comprehension of subsequent linguistic input. Thus, the processing asymmetries found in relative clauses in English and other languages can be traced back to production constraints. Since animate beings are conceptually more salient, this leads to their being less preferred in object position in general, and in relative clauses in particular.

To summarize so far, after years of research that treated ambiguity resolution as a specific phenomenon, distinct from general mechanisms of sentence processing, both Altmann and MacDonald remind us in their contributions to this volume that the current status of the processing

literature favors multiple constraint models for all types of sentences (ambiguous and unambiguous), incorporating previous experience with the world and prediction about upcoming linguistic material and world events as a determinant of linguistic form.

Chien-Jer (Charles) Lin's chapter addresses the issue of relative clauses in languages where the relative clause appears before its "antecedent." He reviews previous approaches to empty category processing, such as filler-gap models based on syntactic structures containing traces of movement. Gap-filler strategies are processing strategies derived from linguistic structures. Lin instead proposes that the processor adopts a perceptual strategy based on thematic templates that he calls the *thematic template mapping*. Thematic templates are made of dominant thematic orders between a verb and its arguments in a particular language, the canonical structures that Bever discussed in CBLS and that are recurrent in the literature (see Townsend and Bever 2001 for a review), as we see in other chapters. Lin bases his claims on the hypothesis that the comprehension of relative clauses (and sentence comprehension in general) involves two processes—syntactic parsing and thematic mapping—that operate in parallel. The strengths of these thematic templates are based on one's previous linguistic experience with thematic orders. The dominance of a thematic template is continuously adjusted, a claim that mirrors discussions by Altmann and MacDonald in their respective chapters. Furthermore, consistent with the *pseudosyntax* approach developed by Townsend and Bever (discussed in Townsend's chapter in this volume), Lin claims that content words in a sentence are matched with thematic templates for quick and easy thematic interpretations, whereas function words are temporarily ignored. In this way, these templates produce impressions about *who did what to whom*.

The chapter further discusses the strong convergence in the field of syntactic processing in the view that nonlinguistic factors significantly determine real-time sentence comprehension. As noted by MacDonald, in most languages, subject relatives are easier to process than object relatives. Lin states that, instead of focusing on "accessibility" of the antecedent, it is the "extractability" of sentence position in sentence planning that underlies this typology. For him, it is the ease of production that underlies the hierarchy, not the ease of comprehension. This is also in line with the view put forth by Bornkessel-Schlesewsky and Schlesewsky, to whom we will soon turn.

The chapter by **Gibson, Tily, and Fedorenko** also analyzes the nonlocal dependencies that occur in subject and object relative clauses with reference to the types of theories that have been proposed to explain their processing difficulties. However, they adopt a different approach. According to the

authors, most proposals fall into three general categories: (1) reanalysis-based theories; (2) experience-/surprisal-based theories; and (3) working-memory-based theories. The previous chapters have adopted approaches (1) and (2). In particular, MacDonald (this volume) argues against working-memory-based explanations. However, Gibson and collaborators, after a careful review of the predictions and evidence in favor of each of the theories, point out that the evidence for both the locus of the processing difficulty effects and for the effects of NP-type manipulations is mixed. In contrast, they argue that much of the empirical data can be explained by appeal to limitations in working memory.

Memory-based accounts, they explain, predict that most of the difficulty will be encountered in the embedded verb. However, some difficulty has been observed at the subject position, which could be explained by interference-based working-memory theories. Animacy effects, they admit, are hard to explain under memory-based accounts, although they are consistent with expectation-based theories. The authors reason that one possibility is that there are multiple sources of difficulty in processing an object relative clause: one source is retrieval from working memory, which surfaces at the embedded verb, and an independent source is associated with infrequent lexico-semantic configurations that appear at an earlier point or throughout the structure. Gibson and collaborators proceed to discuss two experiments examining nested relative clauses, such as *The vandal that the thief that the policeman wounded on the leg accused with some certainty was known to the authorities* or *The jewels that the thief that the policeman arrested on the weekend stole from the vault were worth a lot.* In both experiments, they obtain results that are most consistent with the working-memory accounts of nonlocal dependencies. More on the properties of working memory can be found in the chapter by McElree and Dyer, which we discuss in section 1.6 below.

In sum, the reader of this volume has the chance to contrast the different approaches that are currently being discussed in the literature on relative clauses, a cornerstone in the history of research on sentence processing: the chapters provide a rich source of historical reviews, ideas, and data, which should help the reader evaluate each approach.

1.5 More on the role of prediction and different sources of knowledge

As we have seen, a fundamental insight of CBLS was that prediction based on previous language experience affects processing in major ways. Altmann and

MacDonald discuss and defend this hypothesis. **Gary S. Dell and Audrey K. Kittredge** take it a step further. After a period in which research in language production, comprehension, and acquisition followed separate paths, addressing distinct questions, Dell and Kittredge suggest that modern psycholinguistics is unified by its search for general computational mechanisms that identify how comprehension, production, and acquisition support one another. Current research is converging on the view that there is continuous interplay between language production, comprehension, and acquisition, which is united by a refinement of the concept of prediction incorporated into CBLS. Dell and Kittredge claim that we can analyze the connections between production, comprehension, and acquisition on the basis of the influence they exert on each other. This is called the *psycholinguistic chain* or *P-chain*. The P-chain involves prediction from processing leading to production. Prediction, which stems from processing, leads to prediction error when the input mismatches what is predicted. Error minimization leads to revised predictions. Within this framework, updating predictions is implicit learning. Therefore, acquisition is part of the P-chain.

Prediction, like production, is a top-down process. Hence, most predictions are wrong; because many unpredictable sequences could be formed, our ability to predict is imperfect. This leads to a novel theory of priming. The authors point out that some priming phenomena (syntactic priming or orthographic-to-phonological mappings such as the fact that experiencing OU in 'couch' makes it difficult to read aloud 'touch') can be attributed to prediction error. Priming, the authors contend, is the result of prediction error. Priming in this way leads to implicit learning, and therefore syntactic priming and acquisition are achieved by the same mechanism: prediction error leading to the strengthening of certain connections.

David Townsend also discusses predictability and its effect on language comprehension, as he reviews evidence that led Bever and him to posit the strategy known as LAST (Late Assignment of Syntax Theory). This mechanism is based on a hypothesis-testing model of sentence comprehension. Townsend states that comprehension is highly structured, incremental, and interactive and that linguistic habits project representations at various levels simultaneously. As an enduring theme for pyscholinguists, he considers the fact that comprehenders *do* form linguistic structures, although, in his opinion, the relationship between semantic properties and whether they force structural commitments is still unresolved. Linguistic elements project structure, as evidenced by the fact that the nature of a verb (for instance, whether it is bounded or unbounded), determines its thematic grid in a way that guides parsing. According to Townsend, the projection of structure may arise either

from semantic or from structural information. Thus, semantic and syntactic representations interact, but how exactly they do so continues to be an ill-understood issue, although the data seem to point in the direction of the "multiple representation hypothesis." This hypothesis states that semantic information does not actually facilitate structural processing, but rather draws attention away from structural processing by eliminating some structural options.

For Townsend, increased predictability at one level (the sentence level or the discourse level) facilitates processing at another one when the levels share common representations; furthermore, projected structures are checked against grammar in the course of language comprehension. This checking is the reason for the existence of grammatical rules, he claims: grammatical rules are needed to contrast the initial rough parse done on the basis of semantic statistical patterns against a fully formed structure generated by grammar in order for incremental parsing to proceed.

In the final chapter of this section, **Robert Berwick** presents proposals about how different sources of knowledge can be used by the parser, while still maintaining the independence of grammar as a separate system. Addressing the question of the tension between modeling external language behavior and internal knowledge highlighted by CBLS, he states that the notions that guide linguistic analyses are not necessarily those that illuminate the goals of models of, for example, corpus linguistics that try to predict the upcoming material in a sentence. Whereas current models of corpus linguistics try to predict what a speaker will say next, the goal of traditional generative grammar is to capture law-like generalizations over some representation of knowledge of language with the smallest possible grammar. An expression like *walk on* is analyzed differently by a statistical method that attempts to describe language in terms of bigram properties (which would chunk *walk* together with *on*), and by a linguistically oriented representation that keeps *walk* and *on* in separate phrases.

Berwick points out that it is in such situations where one can best elucidate interactions between different knowledge sources that conspire to yield the distribution of actual sentences. The interaction of the basic constraints in a particular language yields *derived* regularities of that language. *Compilation* in computer science serves as a metaphor for the relationship between grammar and the parser. The programmer writes the instructions for some algorithm in a higher-level language. Then, through a series of intermediate steps, those statements are mapped into the actual step-by-step instructions that the computer must follow to arrive at the desired result. The end result does not resemble the original instructions. Hence, we can think of knowledge of

language as the "higher-level language" and the resulting machine instructions as the "knowledge put to use." In order to parse or produce sentences efficiently, the grammar could look quite different from the "actual" parser and the operations it uses to analyze language, since the parser's actions could include optimizations tailored to the particular language, extra-language contextual information, and properties of the neural system. In sum, different information sources can be combined, and "there is nothing principled that bars the infiltration of such information sources into one's model of language use, while retaining the advantages of the linguist's conventional notion of knowledge of language represented as a grammar." This perspective is a novel extension of the classic distinction between competence and performance and is, in effect, an argument against extending the prediction framework to linguistic knowledge.

1.6 Prosodic and working memory constraints in sentence parsing

The hypothesis that there is a distinction between "narrow" and "broad" syntax has become a central topic of research in current linguistic theorizing within the Minimalist tradition. We discuss this idea in more detail shortly. For now, it suffices to say that narrow syntax is proposed to form the limited core of syntax, with the explanatory burden of many aspects of broad syntax being assigned to cognitive and perceptual interfaces. Two of the chapters (Valian's and Mancini and collaborators') explore the contents of narrow syntax, with both contributing new proposals. Valian concludes that a general schema for determiners is part of narrow syntax, but the details of determiner behavior may fall in the broader syntax area, whereas Mancini and colleagues follow the Minimalist hypothesis that the "Agree" mechanism is part of narrow syntax. However, they argue that some features have interpreting anchors that fall outside of narrow syntax. **Janet Fodor**'s chapter in this volume also illustrates the difference between narrow and broad syntax, utilizing another of the sentence types analyzed in CBLS.

Bever discussed center-embedded sentences and attributed their processing difficulty to a syntax-independent perceptual strategy that rules out a constituent being perceived as holding two incompatible positions at the same time. In other words, a perceptual mechanism (Principle I in CBLS) explains why sentences such as *The dog the cat the fox was chasing was scratching was yelping* are nearly impossible to process. Janet Fodor resurrects an alternative explanation initially proposed by Frazier and Fodor (1978), based on a phrasal

packaging subcomponent of the parsing mechanism. She argues that the linear sequence of phrasal packages constructed in online processing is incompatible with the deeply hierarchical structure required by the syntax. In her own words, "Where the syntax wants NP1 [NP2 [NP3 VP1] VP2] VP3, the parser could most likely at best create [NP1] [NP2 NP3 VP1] [VP2] [VP3]. This is only partially helpful; an example like [*The beautiful young woman*] [*that the man the girl loved*] [*met on a cruise ship in Maine*] [*died of cholera in 1962*] remains awkward, especially in the transition from VP2 to VP3 (which suggests an explanation of why VP2 is often overlooked by the parser)."

This proposal is a reinstantiation of phrasal packaging, based on the limits of working memory, in terms of prosodic phrasing. While discussing these issues, Fodor reviews evidence of attachment preferences, another line of research within the field of sentence processing that led to a prolific literature. The packaging mechanism refers to the fact that, depending on category (relative clause vs prepositional phrase, for instance) and length (a long vs a short relative clause), the parser chooses either attachment to the constituent currently being processed or high attachment to a previous constituent. Thus, Fodor can convincingly account for the crosslinguistic differences found in the literature concerning attachment preferences by assuming that the packaging mechanism is not a memory-saving device, but a result of the prosodic component of the grammar of the language at issue: prosody divides strings of words into phrases for pronunciation. This explanation depends, of course, on the assumption that prosody is projected in silent reading as well as in pronunciation of sentences, an assumption that is supported by a growing body of experimental research.

Fodor claims that, while syntax thrives on recursion, prosodic phrasing does not. However, a sentence cannot be parsed without being assigned a supportive prosodic contour. In contrast to Bever's general account, this explanation is language-specific. It falls within the broad and not the narrow faculty of language, since it concerns the interface between prosody and syntax.

The chapter by **McElree and Dyer** focuses on the role of working memory in processing complex structures. At the time that CBLS was published, the limited capacity of working memory was generally considered to be the primary determinant of performance limits. Bever himself makes this assumption in his 1970 paper. McElree and Dyer review studies in comprehension that have been motivated by principles and procedures derived from memory research. Three fundamental questions have been addressed in these studies: (1) What is the nature of the memory representations formed during real-time comprehension, and what operations are used to access them? (2)

What factors determine the success of those operations? (3) When are memory operations required in comprehension? Following Bever (1970), the authors believe that understanding how memory functions in real-time comprehension will provide insights into the overall architecture of the comprehension system.

The authors argue that the evidence for a limited-capacity working memory is weak and, as a consequence, approaches based on this construct are not likely to provide a principled account of the limitations of comprehension. McElree and Dyer suggest that a more fruitful alternative is to look at the nature of the memory operations involved in comprehension. Rather than viewing comprehension problems as failures that arise when a limited-capacity working memory is overtaxed, they propose that difficulties in comprehension are due to failure to retrieve the product of past analyses, in a similar way in which memory loss can rather be considered a failure to access an existing representation. They argue that a linguistically dependent constituent is accessed via the same direct operations involved in access to long-term memory representations. More specifically, they propose an account of comprehension errors based on retrieval interference. According to this proposal, even a small amount of intervening material between the elements of a dependency (subject and verb, for instance) requires a retrieval operation, which could be prone to interference.

1.7 Why is the agent-initial pattern preferred?

One of the most salient claims in CBLS is that an NVN sequence is overwhelmingly interpreted as an Agent–Action–Patient thematic structure, at least in languages with an SVO canonical order. **Ina Bornkessel-Schlesewsky and Matthias Schlesewsky**'s chapter proposes an explanation of this observed universal of processing, which they frame in terms of an agent-initial preference.

The chapter illustrates a relatively new conception of how the structure of human languages is intimately tied to the functions of the brain. The authors point out that exceptionless universals (either absolute or implicational) are difficult to find, but some structural patterns clearly occur more often than others in the languages of the world. This justifies the quest for "statistical universals" and quantitative typology. Within the new field of neurotypology, which assumes tight connections between the structure of human languages and brain functions, the authors develop an account of the high frequency with which canonical forms respond to the template Agent–Action–Patient.

Data show that seemingly identical conflicts between form and meaning lead to different electrophysiological responses in different languages. These differences can be derived from the relevant cues to determine actorhood in a particular language or from the different ways in which the properties of languages affect processes of categorization and decision-making. Thus, both dimensions of variation can be explained via the interaction of language-specific properties and more general cognitive mechanisms. However, there is a crosslinguistic generalization concerning the identification of the actor. Bornkessel-Schlesewsky and Schlesewsky claim that the processing system attempts to identify the participant primarily responsible for the state of affairs under discussion as quickly and unambiguously as possible, and this would explain the prominence of the actor role. The result is that all arguments encountered by the processing system that have nominal properties within a sentence compete for the actor role. This is postulated as a universal of language processing. The reason for this is that the actor role is a cognitive and neural attractor category, that is, a language-independent category. It is a universal because of the general human ability to recognize goal-directed action and to differentiate between self and other. This claim is in line with MacDonald's conclusions about the prominence of animate agents and with the extractability condition discussed by Lin.

1.8 Universals, the syntax/semantics interface, and narrow syntax

The previous chapters present a historical overview and new evidence bearing on Bever's original hypothesis that cognitive factors shape grammar. Parallel to the psycholinguistic advances that are mentioned in those chapters, linguistic theorizing stemming from Chomsky (1993) evolved towards minimizing the contents of the "narrow faculty of language" (Hauser, Chomsky, and Fitch 2002) by placing more of the descriptive and explanatory burden on the cognitive interfaces of grammar. This perspective on the nature of linguistic structure is known as Minimalism or also biolinguistics. It is probably not an exaggeration to say that this new research agenda in linguistic theory developed in part because of advances in psycholinguistics, which were in turn strongly guided by the questions and central hypotheses put forth in CBLS.

Many language researchers—primarily, but not exclusively—coming from the generative tradition in linguistics and psycholinguistics represented by the Minimalist Program, now focus on exploring those aspects of language that would seem most resistant to explanations that are derived from extralinguistic (domain-general), constraints and principles. The construct of narrow

syntax is, of course, not uniformly accepted by all researchers; there is an active ongoing debate among language scientists on whether there are indeed such language-specific properties. This line of inquiry is represented in our volume in the contributions by Piatelli-Palmarini, Fodor, Valian, Mancini, Molinaro, and Carreiras, and Grodzinsky. These authors examine linguistic properties such as the determiner category, recursion, and agreement and attempt to elucidate which of these phenomena are candidates for inclusion among the core computational realm of grammar or "narrow syntax." The general question which provides the background for this research is: what is left of innate and domain-specific constraints on the form of grammar? The fact that this question is even being asked can be viewed in part as a response to the success of explanations for many phenomena that, as Bever suggested, might be rooted in more general perceptual and cognitive constraints and principles.

The chapter by **Montserrat Sanz** helps place these papers in context by reminding readers of the changes in the view of the syntax/semantics interface that Minimalism brought about. Focusing on research on events, she illustrates the gradual evolution of the syntax/semantics interface towards analyses based on features of functional categories. She also discusses how research on parsing proceeds on the basis of constructs that differ from those used by theoretical linguists (see also the chapter by Robert Berwick). In particular, reference to thematic properties and to argument positions might be assumed by researchers in sentence processing but is not part of the syntax/semantics mapping as conceived of by many linguists. Syntactic operations are now viewed as driven by features of functional projections, rather than by a certain mapping between thematic roles and structural positions. In line with Massimo Piatelli-Palmarini below, Sanz hints at the links between linguistics and physics, with the indeterminacies that physicists face: at this point in history, it is unclear what is pre-theoretically a syntactic, semantic, or lexical construct.

Massimo Piatelli-Palmarini reviews the "conservativity" property of determiners, a feature of human languages that he argues has no relation to the external world. According to Piatelli-Palmarini, determiner conservativity is therefore a true universal of language that must be explained with reference to syntactico-semantic structures and computations thereof. Determiners are two-place predicates whose arguments are ordered. The conservativity property can be exemplified as follows:

(1) All men are mortal = All men are mortal men
In which

$A \cap B = (A \cap B) \cap A$

In words: the overlap between A and B is exactly the same as the overlap between their overlap and the set A itself.

As Piatelli-Palmarini puts it, this property, which applies to all known languages, is a universal that cannot be explained by external factors, standard logical predicate-argument relations, elementary logical quantification, generic "laws of thought," or other kinds of language-processing constraints. He proposes that conservativity is a universal because no child could learn a nonconservative determiner. In this sense, constraints on internal computations in the domain of language explain both this universal and the impossibility of nonconservative determiners.

This property is abstract. Abstraction is a primitive and natural property of the system, because our nature makes it the only admissible generalization from impoverished stimuli. According to the author, the truths that are observed in linguistics are not necessary truths, but rather "the fallible outcome of a rational integration between empirical data and our science-forming faculty." In that sense, the work of the linguist resembles the job of physicists more than that of mathematicians or biologists, with their assumptions about the pre-existence of their object of study. This is a reference to an idea that Bever discussed in CBLS. Bever states that the concept of species or organ is a pre-theoretical assumption for the biological sciences. In order to define what a cow is, one can offer an exhaustive rendition of its physiological features and of its genetic material, but the fact that there is a bovine species is taken as a given. We can only describe the interactions between its isolable components but nothing like the "bovine essence." In contrast, Piatelli-Palmarini claims that the current predicament in linguistics is closer to that in physics.

Virginia Valian also focuses on issues that arise in investigating determiners. She argues that it is possible to prove that children have some innate abstract knowledge of this category, since, unlike any other category, its development can be traced from pre-verbal infancy to the age of two. Therefore, it is a likely candidate for narrow syntax membership. Valian argues that children's innate knowledge of determiners includes this: determiners head DPs and take as complements NPs with which they hold agreement relations. Therefore, a minimal hypothesis about determiners requires reference to other syntactic notions, such as head, complement, agreement, etc.

Even with this experience-independent knowledge, the child has to figure out the specific repertoire of determiners in her language, the contexts they can be used in, and so on. Reviewing a wealth of crosslinguistic data, Valian

concludes that the acquisition of determiners is a top-down process, given that children show continuity: they do not go from not having a representation to having one and they do not shift from one system of representation to another. Rather, they appear to have an abstract category (they flesh out a schema of what counts as a determiner) and they search for details about the members of that category (they learn about the language-specific particulars of each determiner). The part of grammar that specifies the behavior of a particular determiner is the result of acquisition mechanisms, in line with the hypothesis put forth by CBLS.

Agreement is another syntactic property that has been argued to belong in narrow syntax. This linguistic trait appears redundant from a semantic perspective, because it seems to encode the same information in two or more different elements of a sentence. **Simona Mancini, Nicola Molinaro, and Manuel Carreiras** explore a language-specific type of agreement. Many languages have morphological agreement, taken to be encoded as a bundle in a functional category which, in current Minimalist syntax, must be checked by a lexical item. In processing terms, this means that the parser would perform a unique agreement operation without regard to discourse or thematic functions. However, as the authors point out, person, number, and gender agreement hold different conditions and properties. In order to account for the experimental data that they discuss, the authors propose an agreement-anchoring operation.

Given that some nouns vary between a singular and a plural number depending on syntactic context, number cannot be identified either in the lexical or in the discourse representation of the sentence, but must be contained within the inflectional morphology of the nominal argument. Therefore, Mancini and collaborators argue, agreement does not expand beyond the inflectional layer of the sentence. Person features, on the other hand, express the status of an argument with respect to the participants in the speech act, which means that they reside in the participant representation of the sentence as related to the speech act. In other words, they consider that person and number have what they call different "anchoring points." They assume a processing correlate of the Agree syntactic operation postulated by Minimalist analyses, but their account of the different anchoring positions for the agreement features of person and number means that the current linguistic account in which features are bundled together in a functional head is untenable. In this way, instead of circumscribing the computation of the agreement dependency within the boundaries of narrow syntax, they identify an interplay between the purely formal character of feature consistency

checking and the semantic-pragmatic information that arguments carry, such as being a singular or plural entity.

The authors arrive at this conclusion through the analysis of data obtained by using event-related brain potentials (ERPs). Given some assumptions about the interpretation of different potentials, this technique allows researchers to infer whether the source of an anomaly lies in syntactic integration processes or in lexico-semantic processing. They find that person and number violations generate different negative effects. An agreement violation involving person may block the mapping between morphosyntactic properties and speech participant information, causing interpretation conflicts to arise. In contrast, an anomaly in number only affects the cardinality of the referent. This leads them to argue against a syncretic representation of features and a unique and strictly formal operation through which their content is checked. Furthermore, the results of this study lead the authors to claim that the directionality of agreement computation may not be as rigid as assumed in standard Minimalist analyses, because marked operations of the kind they call "unagreement" (a grammatical person mismatch) may reverse the Agree operation from verb to subject and shift the locus of person interpretation to verbal morphology. The contribution illustrates how some experimentalists use brain-imaging methods to examine specific syntactic hypotheses and take the results as source of evidence to evaluate linguistic hypotheses.

1.9 The role of grammar in language processing

One of the main questions put forth by CBLS concerned the place of grammar in language processing, as we have seen above (see discussions on chapters by Townsend and Berwick, for instance). **Colin Phillips** also takes up this question in his contribution, in which he challenges the assumption that the mental grammar is not directly recruited in language processing (Bever 1970; Townsend and Bever 2001). In doing so, he reviews the reasons that led to this view. Psycholinguistics in the 1960s considered transformational grammar as a model of the psychology of language and proceeded to test it as such. It is often concluded that these investigations found support for the syntactic representations argued for by generative linguists, but not for the transformational component that the model of grammar included in the 1960s and 1970s. Thus, it became accepted that transformations were not "psychologically real," because they were not supported by the linking hypothesis for grammar and processing known as the Derivational Theory of Complexity (DTC, Miller and Chomsky 1963). Phillips argues that this is an

oversimplification of what early psycholinguists found, because the DTC did not spell out a detailed linking hypothesis; it simply claimed that mental computations take time/effort, an assumption that remains standard in modern psycholinguistics. But discussions on the DTC at the time focused on the specifics of transformational length proposed by generative models. This in the end led to the conclusion that the grammar plays no significant role in language processing, a belief that continues to be widely held.

A second argument that made the place of grammar secondary in psycholinguistics, advanced by Fodor and colleagues, involved the impossibility of employing grammar directly as a sentence-processing device, because the derivational history of an expression could not be determined in "reverse," starting from the terminals up to the initial symbol: on the one hand, a bottom-up parser cannot incrementally assemble a right-branching tree of the type languages like English generate; on the other, transformational rules could generate output that made it impossible to determine what the input to the rule was. However, Phillips argues that subsequent work in computational parsing models has shown that phrase structure grammars can be used incrementally (Resnik 1992; Crocker 2010), that alternative formalisms can avoid the problems then raised (Pollard and Sag 1994; Steedman 2000; Kempson et al. 2001; Phillips 2003), and that current "Minimalist" descendants of 1960s transformational grammars can be associated with explicit parsers (Stabler 2011), all of which weakens the arguments against the hypothesis that grammar plays a direct role in parsing.

This issue is also taken up by **Edward Stabler** in his contribution. Stabler discusses how computational approaches can contribute to overcoming the problems faced by the notion in CBLS that grammar is "the epicenter of language." Stabler reviews and discusses three conceptual difficulties in granting the grammar a central role in language processing and argues that these difficulties have been largely overcome by advances in computational studies of language: (a) the determination of the common properties of human languages; (b) the quest for the assumptions that appropriately relate grammar and judgments about, or use of, particular expressions; and (c) how particular kinds of computations of those relations can be evidenced.

Regarding (a), one common property of human languages is that they are both strongly and weakly mildly context-sensitive (Joshi 1985) so that grammars can define the sentences of human languages (weak adequacy) and also provide the structures of those languages (strong adequacy). Computational methods hence provide tools for describing rather abstract similarities of structures and languages, allowing a perspective removed from concrete, typologically oriented universals, and moving into more

abstract, computational properties shared by all languages. Regarding the second difficulty, recent developments in the study of mildly context-sensitive grammars reveal a consistent "two-step" character: derivation-plus-mapping to derived and pronounced forms, where the details of derived structures are less important than the derivations themselves and their connections to pronounced forms. The simpler structure obtained when derivations are isolated catalyzes the study of how simple the mechanisms of analysis might really be, and of how those mechanisms could extend to, or even across, interfaces. The third problem of knowing how to look for reasonable implementations is reduced by comparing alternatives that really differ significantly, a task in which great progress is currently being made.

Phillips also discusses the role of heuristics and strategies in parsing, first proposed in CBLS, and generally assumed in psycholinguistics to be central to language processing and only indirectly related to the grammar. Phillips argues that these phenomena, on close examination, are less pervasive than generally assumed, and that they often result from the interplay of grammatical constraints in a noisy cognitive architecture. Phillips presents the alternative view that grammar is directly involved in language processing, both in perception and production. In such a view, there is no division of labor between grammar (knowing that) and processing (knowing how) in language, and processing is essentially a process of incrementally constructing a linguistic representation determined by grammatical constraints.

1.10 Uniquely linguistic? The neurocognitive perspective

As we have seen, some researchers continue to hold the position that there are aspects of language that are distinctively linguistic, even if the inventory of these aspects is shrinking, whereas others have taken the view that the faculty of language does not contain unique properties that do not have homologues in other cognitive domains. Two opposite stances in this debate are illustrated in the contributions by **Luciano Fadiga and Alessandro D'Ausilio** and **Yosef Grodzinsky.** Each ground their arguments in results from cognitive neuroscience. Fadiga and D'Ausilio argue that well-documented features of mirror neurons suggest the existence of a basic and primitive mechanism to acquire the symbolic representations that underlie language as well as other symbol-based capacities. They propose that the social function of mirror neurons and the ability to transform common objects into meaningful tools make mirror neurons a likely precursor for the capability to attribute meaning to novel or meaningless entities, which can be words or gestures. These authors argue that the primitive hierarchical organization of the motor system

displays all the required features that language deploys, including recursion and constituency.

Fadiga and D'Ausilio begin by pointing out that human cognition is geared toward the performance of goal-directed actions, which are based on the synergic composition of simpler motor constituents chained together according to a precise "motor grammar." Actions are directed to solve a problem. In order to accomplish the goal, different motor elements are integrated into a single unit. This reduces cognitive demands and makes it possible for complex skills to become automatized. The motor system is recursive, in the sense that it has the ability to repeatedly retrieve previously learned motor elements composing an action. Even though this differs from the notion of recursion in language (recursion in language expresses nested structures, whereas repetitive motor behaviors depict only sequential structures), they claim that the motor system could be conceived as a goal-driven hierarchical structure to concatenate simple motor acts. This hierarchical goal structure, along with the rules which connect individual motor elements, might subserve the syntactic computation of language. In particular, "hierarchical syntactic-like structures fulfill the two properties required for motor goal representation: Goal representations can (a) be reactivated as single units whenever required, and (b) have their component movements reactivated one by one or reassembled to enable learning of novel behaviors."

Fadiga and D'Ausilio review evidence that monkey motor area F5 is activated in goal-directed actions and that the area PFG in the parietal cortex plays a role in organizing natural actions. In looking for the human equivalent to area F5, they examine Broca's area and point out that it could represent the hierarchy of action goals, whether seen or executed, rather than the basic motor program to execute those actions.

Mirror neurons exhibit two important properties that enable them to code the actions of others in a social and communicative framework. First, they have some functional plasticity (they may extend their visuomotor properties to tools). Second, they show a special status when presented with an interaction. This points to the existence of a basic and primitive mechanism to acquire symbolic representations. According to the authors, mirror neurons are precursors for our capacity to attribute meaning to novel entities or entities without a meaning; in other words, for language. This is possible because of the hierarchical nature of goal abstraction that permits us to predict sub-actions or extend behavior to new situations. In this way, the primitive hierarchical organization of behavior has the basic features, including recursion and sequence chunking, that language needs.

In contrast, **Grodzinsky** reviews tests for modularity and discusses the relation between language and the perception-action loop. Grodzinsky argues against what he considers a holistic view held by Fadiga and colleagues. He claims that neuroscientists cannot analyze language impairment without a "linguistic tool kit." Grodzinsky approaches the issue of modularity by reviewing Fodor's four properties of modular systems (1983) and applying the computational perspective (i.e., whether the properties that govern one system can be deduced from those that govern another) to center-embedded sentences. He asks whether embedding is a central property of natural language syntax, as assumed by some holistic views that claim that language, music, and action share the same computational system, and points out that there are invisible properties that constrain structures in different ways (restricting relative clauses more than other embedded clauses, for instance). Hence, a simplified property cannot be taken as determining whether language is modular or not. He reviews neurological evidence that explores whether Broca's area governs the sequencing of both linguistic and action-based perceptual representations. He objects to experiments that attempt to prove this by noting that language sequencing has different properties from visual sequencing of video snapshots. He also argues that the deficit in Broca's aphasia is not directly related to sequencing, to embedding, or to the contrast between human action and physical events.

1.11 Language acquisition and abstractness

Arguments for the abstract nature of linguistic properties are discussed by Piatelli-Palmarini, Valian, and others. The following two chapters develop this topic further in connection with language acquisition. **Jacques Mehler** reviews research in his lab that has refined our knowledge of the mind of the neonate by progressively reducing the age of the subjects under study. Throughout the decades, it had become clear that prosodic cues are essential for children to develop their knowledge of words, but research had not identified reliable acoustic characteristics for the different rhythmic classes of languages. Mehler's research leads to the conclusion that vowels and consonants are specialized for different tasks: vowels are mainly specialized for conveying information about grammar, whereas consonants are used to individuate previously learned words. The progressive refinement of techniques, when coupled with the ability to study even the youngest infants, makes it possible to explore the contents of innate human knowledge with increasing accuracy. The conclusion that Mehler draws from these four decades of research is that humans are born with a left-hemisphere

superiority to process species-specific properties of speech, such as the difference between vowels and consonants.

Ewan Dunbar, Brian Dillon, and William J. Idsardi provide a complementary perspective. They argue that children arrive at an analysis based on abstract elements by using domain-general reasoning. Dunbar and colleagues revisit abstractness in phonology from a Bayesian perspective, and do so by exploring a specific case of opacity in Kalaallisut, an Inuit language of Greenland. Their main argument is that, all other things being equal, a Bayesian learner will favor the simplest model; in this particular case, it must arrive at an analysis involving abstract elements through independently motivated domain-general reasoning strategies. The authors thus show how Bayesian reasoning applies to the problem of abstractness in language modeling by the learner. Bayesian approaches are based on probability theory, the most widely accepted formal theory of reasoning under uncertainty. This contribution illustrates the ideas behind those methods and how they can apply to problems of inference in linguistics when confronted with underdeterminacy. In the particular case considered in this contribution, the authors show that while linguistic analysis can provide two empirically adequate accounts, it cannot determine which one is chosen by the learner. Dunbar and collaborators claim that Bayesian methods make this final determination possible and do so in favor of abstractness in phonology.

1.12 Recapitulation

The contributions discussed above review the main arguments developed during the last decades of research into the relationship between language production, comprehension, and acquisition mechanisms and between grammar and nonlinguistic cognitive factors. They introduce the reader to advances in the methods and techniques currently employed and to refinements in theoretical arguments as these methods and techniques are improved and better understood. The chapters bear witness to the reality that some aspects of the controversy remain: linguists and psycholinguists are gradually converging on an understanding of what might be at the core of grammar, but the question of whether there is a grammar whose features are independent of processing demands or other external cognitive constraints is still very much alive, even if the issues have been sharpened during the last decades. It is becoming increasingly clear that there are aspects of the syntax of sentences that cannot be detached from semantics and other factors, both linguistic and extralinguistic. Moreover, these factors influence the way that humans process language, affecting its acquisition and the actual shape that constructions take.

As a result, the argument for a core grammar increasingly focuses on aspects of language that arguably cannot be explained by how they interface with nonlinguistic systems. These include agreement, recursion (center embedding), and determiners, which are argued to be part of the narrow faculty of language.

Thomas Bever's own research program after 1970 has taken up many of the issues that he raised in CBLS. For this volume, he has written a piece in which he reviews how this initial proposal developed into a research program. His contribution begins with a reminder of the essential proposal he made in 1970 which constituted an early version of the biolinguistic approach: attested languages are the result of genetic endowment, maturational processes, experience, and other constraints. Thus, language universals can reflect a wide range of sources and constraints that influence the formal architecture of grammars, affecting how it is neurologically implemented, how it is learned, how it is understood and produced, and how it interacts with social practices. Over the succeeding years, Bever explored a number of potential universals, always with the goal of interpreting them as flowing from some language-external system or property. He reviews a series of arguments in several domains.

Concerning the general distinction between inductive statistical processes and structural computations in cognitive behavior and development (see also the chapters by Townsend and Berwick), he claims that the early appearance of U-shaped developmental functions in which very young children become temporarily worse at various tasks reveals that they have shifted to dependence on statistically valid structures. This duality of processes applies also to language acquisition and to the analysis by synthesis model of adult sentence processing, based on both statistical strategies and structural derivations. Bever demonstrates that while statistical processes account for a lot of language behavior, there is still evidence for derivations as suggested by the behavioral role of empty categories, the residue of derivations.

Bever also provides an argument that modularity of language representation and processing is definitional, not architectural: that is, the computational languages of distinct levels of linguistic representation are immiscible, hence opaque to each other. The left-hemisphere priority for language may be based in a general computational superiority, as opposed to having a specific innate computational mechanism that creates language. Mammals may have cerebral asymmetries computationally similar to humans, suggesting further that human asymmetries are not uniquely causal of language. Attempts to train animals in language-like behaviors show mixed results: they do show evidence of using representations, but not regular hierarchical structures.

Bever argues that certain linguistic universals such as hierarchical structure or movement constraints may have extrinsic causes: either they call on "uncaused" (Platonic) formal constraints or on discoverable physical laws: either way, some structural properties of language may be the result of the interaction of the human capacity with externally imposed natural constraints.

The emergence and persistence of a "psychogrammar" may happen because of its role during language acquisition in providing a consistent representation between the systems of production and comprehension. It has to call on independently available linguistic computational devices to create the reconciling structure. The discovery of grammar by the child has an intrinsic motivation, if one considers it as an expression of human-style problem solving: humans are perhaps the only animal that enjoys solving problems. So on this view, first-language learning is a kind of exciting fun. Classic investigations of problem solving suggest that it involves reconciling conflicting or disparate representations of a situation by accessing a different level or kind of computation, often oscillating between a statistically valid generalization and a structural analysis. This is consistent with the view that the psychogrammar emerges out of its role as reconciling cognitive conflicts. There are empirical consequences of this model for language universals: (a) Every language should have a "canonical form," a surface construction type that is basic and most frequent, so children have an overwhelmingly clear statistically supported structure to start with. (b) The canonical form should have a preponderant surface to thematic mapping relation, but the mapping need not be the same across languages. This creates the basis for a generalization by the child who is learning language, with enough exceptions to stimulate and require some form of derivational analysis of distinct surface–theme relations. (c) The existence of sentences with varying degrees of canonical form can mitigate the poverty of the stimulus: the child can think of and generate sentences that it has not experienced, based on the statistical generalizations it has built up. (d) The canonical form can interact with other levels of representation to explain certain dynamics of language change. A flagship case is the effect in Old English when nominal inflections were lost: sequences that violated the canonical order were no longer disambiguated by the nominal inflections, leading to the required presence of a complementizer in subordinate clauses of various kinds.

Bever's chapter concludes with two sections on some future directions that he anticipates in psycho- and neurolinguistics. First, current research is showing that normal conversational speech deletes and distorts the signal so much that large portions of sentences cannot be understood without prior or

following context (itself also somewhat garbled). This motivates consideration of "the psychological moment" in which processing can proceed both forward and backward in time, while preserving the conscious percept of always moving forward. That is, a later portion of a sentence may instantly clarify an earlier portion, at the acoustic/phonetic level. This has implications for motherese and the "real" poverty of the stimulus. In particular, it challenges the assumption that children have clear representations of words, so their only problem is to figure out the rules that govern their sequencing. Rather, children face a severe problem in just discovering the words, presumably by using their emerging syntactic capacities simultaneously with acoustic/phonetic abilities.

The second new area involves inroads into the study of the genetics of language. For forty years, Bever has been differentiating language behavior in right-handers with and without familial left-handedness: those with left-handers in their family history characteristically access lexical items faster than those without left-handers. Recently, he and colleagues have created a genomic model of the genetic load for left-handedness, and are relating it to various neurological differences during language behavior. Since about 40 percent of the population is right-handed with left-handed family members, the differences in neurological organization have to be treated as "normal," not the result of some particular genetic anomaly. Bever suggests that their results combine with well-known cases of abnormal brain organization for normal language behavior to support the claim that language is not caused by any particular neurological localizable organization. On this view, the capacity for language is rooted in a combination of general internal capacities (e.g., the ability to form a very large number of labeled categories) with general structural principles, some caused by neurological principles, some uncaused by them. Language finds its best neurological representation for individual brains: what commonalities there are for language across individuals is the result of having similar localizations for certain kinds of general computational processes that language calls on.

In sum, the chapters of the book present a panoramic view on forty years of research since a new conception of grammar ontology was proposed. In spite of the fruitful years of inquiry involving thousands of talented scientists in hundreds of labs around the world, there are still many unsolved mysteries about the relationship between grammar, language acquisition, comprehension, and production. We conclude this section by quoting Roger Brown who, at a tribute for his years of achievements at the Boston Conference on Child Development, after receiving much praise and eulogy, stood up and said: "Yes, but we still know so little."

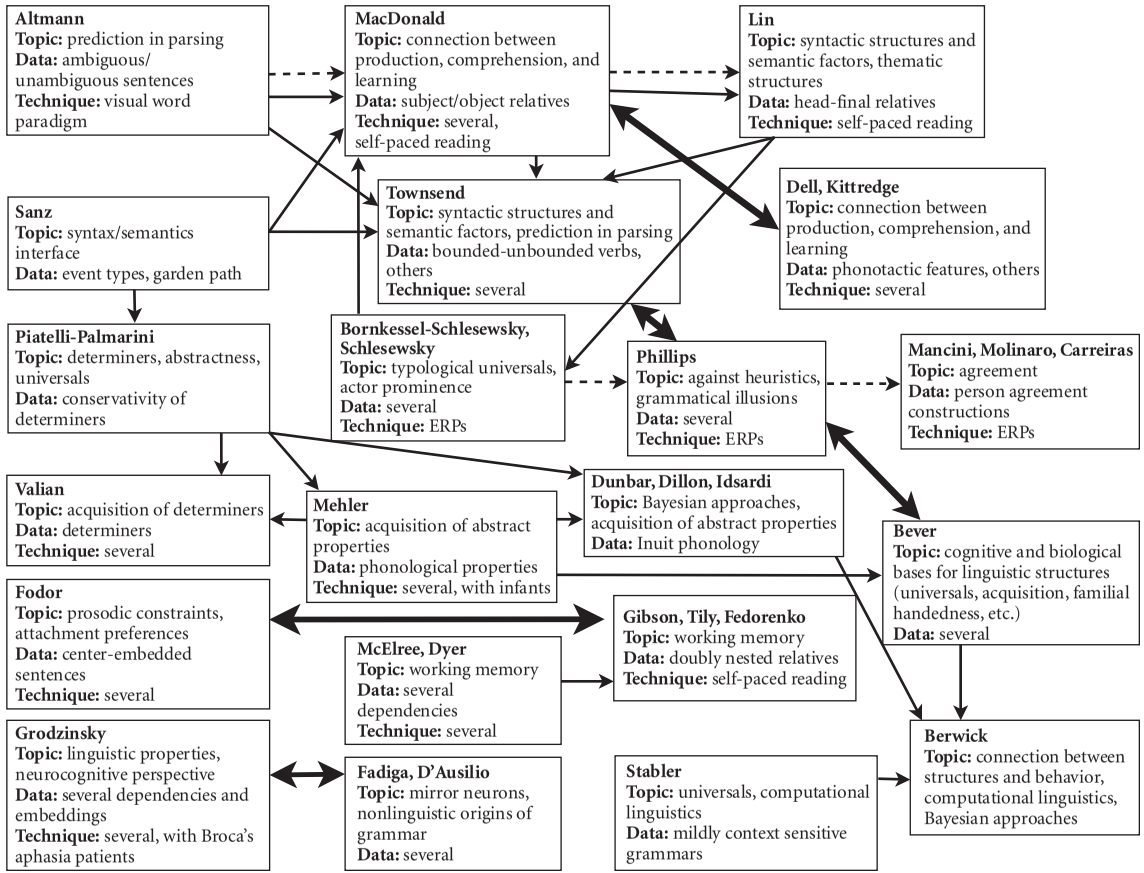

Figure 1.1. Connections between the chapters of this book

1.13 Using this book for a course

The chapters of this volume are interconnected in many ways that are not directly reflected in our grouping. Therefore, the divisions we have proposed should be considered only one, perhaps, imperfect attempt to guide the reader through the vast and varied content of the chapters. Several other groupings are possible. Some chapters focus on similar sets of data (MacDonald, Lin, and Gibson, Tily, and Fedorenko, for instance, each examine the processing of sentences with relative clauses). Some focus on the same mechanism, but apply it to different sets of data, considered from different angles (Gibson, Tily, and Fedorenko explore working memory, as do McElree and Dyer). Thus, the chapter by Gibson, Tily, and Fedorenko could serve as the center of discussions on different topics. It could be contrasted with MacDonald's and Lin's on the one hand, and with McElree and Dyer's, on the other.

The flow of the discussion in this chapter has followed a certain path, but we can also suggest alternative paths. For example, it would be possible to compare and contrast the chapter by Townsend with the chapter by Phillips, although in this summary we have placed them in different sections. Likewise, Mehler's and Valian's are complementary, in that they both present the latest research on child language acquisition; Piatelli-Palmarini discusses abstractness, and so do Dunbar, Dillon, and Idsardi, as well as Mehler. Figure 1.1 highlights other possible but not exhaustive links among chapters that we hope will provide helpful guidance in using the chapters of this book for undergraduate and graduate courses and seminars. Solid arrows indicate that the chapters share a topic or that they analyze similar sets of data. Doubly pointed arrows indicate that the two chapters can be used to exemplify contrasting arguments and ideas. Dashed arrows signal that the experimental techniques being reported are similar. Thus, the chapters can be packaged around research topics, data, and experimental techniques.