

Goethe's Leaf, Computing Galaxies, and the Human Mind: An Exploration of Modern Linguistic Theory



Joshua Bowles

To my wife—for hours of patience and endless support. MAYA!
Thank you also to Jade May Bowles for use of the beautiful pictures (most photos used here were taken by her).



Figure 1: Magenta Flower Courtesy Jade May Bowles

Contents

1	Introduction	1
1.1	What is Language?	3
1.2	What is a ‘linguist’ and where do I find one?	4
1.2.1	Brief Detour: Material and Immaterial	5
1.2.2	Back to the Question	7
1.2.3	A Brief Note on Evolutionary Adaptation and Exaptation	8
I	Exploration	11
2	World in a Leaf	13
3	Nature is a Computer	15
II	Functionalism and its Adherents	17
4	The Functionalist Program	19
4.1	20
4.2	20
4.3	20
5	Why Communicate?	21
5.1	Why do we communicate?	22
5.2	The World of Typology	22
5.3	Structuralism	22
5.3.1	Bloomfield, Sapir, and Harris	22
5.4	Documenting Languages	22
5.4.1	Boas and Sapir	22
5.4.2	The Ethical Imperative: Endangerment and Revitalization	22

III	Formal Considerations	23
6	The Minimalist Program	25
7	Optimality Theory	27
8	Other Approaches	29
8.1	Head-driven Phrase Structure Grammar(HPSG)	30
8.2	Generalized Phrase Structure Grammar (GPSG)	30
8.3	Lexical Functional Grammar (LFG)	30
8.4	Combinatory Categorical Grammar (CCG)	30
8.5	30
8.6	30
8.7	30
9	Infinity to Biology	31
9.1	Why Logic Matters to Linguistics	32
9.2	Cantor: Infinity Comes in Different Sizes	32
9.3	Search for Boole	33
9.4	Frege	33
9.5	Carnap and the Logical Syntax	33
9.6	Tell them Gödel Sent You	33
9.7	Tarski, Metamathematics, and Truth	34
9.8	Montague	34
10	Modern Semantics	35
10.1	Models and Functions	36
10.2	Events Really Happen	37
10.3	Concepts are Crucial	37
10.4	Cognition's the Thing—Cognitive Linguistic Semantics	37
10.5	Is that My Lexicon Over There?	37
10.6	Innate Ideas, Innate Meaning	37
10.7	The Neural Theory of Language	37
10.8	37
10.9	37
10.10	37
IV	Cross-disciplinary Inspiration	39
11	Galaxies, Flowers, and Snowflakes	41

12 Formal Conjecture	45
12.1 Human Language and Mathematical Systems	46
12.1.1 Ambiguity and Redundancy	47
12.1.2 Competence-Performance, Internal-External	49
12.1.3 Mathematicization	52
12.2 On With the Formalities	54
12.3 α Merge and symmetry	56
12.3.1 Ambiguous Merge	56
12.3.2 Fib(n) levels and the X-ratio	59
12.4 Strong Minimalist Thesis	67
12.4.1 Merge is not a formal object	68
12.4.2 Normative use of the term “computability”	71
12.4.3 GT and syntax	74
12.5 Final Remarks	77
 V Conclusion	 81
 13 Conclusion	 83

Preface

Before I say anything let me apologize to all linguists of every theoretical persuasion everywhere. I can never do justice to the richness and subtlety of your research. I know at the outset that I have ignored certain persons—this is not out of lack of interest but lack of knowledge. To those linguists who are mentioned in the following pages, I owe an even greater apology: I have condensed your life's work into a few pages of prose. While this is unforgivable, I hope that you are not offended in light the goal of this book—to present the field of linguistics as one of the most vital areas of academic work to the general reader. For this latter intention I do not apologize. A general reader of the almost unthinkable diversity of linguistic theory is, I feel, long overdue. Nonetheless, I apologize for superficial generalizations, omission of crucial details, etc., etc.

This book is an exploration of fundamentally different theories of grammar. Overall, I do not go into great detail about the specific datum of such theories, opting instead to paint with a broad brush. You will meet the researchers, scientists, and thinkers who compose the field of linguistics, and you will see that they are a very diverse group of people with different goals and backgrounds. Because of the general scope of such a work, many practitioners of the various theories may feel that I have only superficially covered their interests. This is the drawback of a broad exploration and I accept these consequences in return for the special view. It is the scientist's job to deal in details, but for the most part these details cannot be appreciated unless one has the appropriate vision of the larger picture. Just as the painter labors over every stroke and spot of color, the appreciation of such detail comes only in the collected whole.

There are numerous scientific endeavors that have come to be admired by people who do not practice them. This pop-culture admiration comes not from an appreciation for details, but an appreciation for the scope, magnitude, and daring of the endeavor itself. Fields such as mathematics, physics, and literature are age-

old classics. Chemistry, biology, and computer science are new additions. One gets a sense that modern culture feels a certain respect and awe for what chemistry or literature have given us; and what they may potentially offer. The respect and awe, I guess, come from a collective innocence of how exactly the chemist or poet brings their results to fruition. Nonetheless, a general sense of these of endeavors—more than likely harvested from a remembrance of High School or University classes and supported by popular media—promotes an appreciation of them. I hope that by the end of this book you will have gained a general appreciation for what the linguist does.

In this book you will find a discussion of language death and the attempts to document endangered languages. I will introduce the main tenets of the Functionalist research program, which views language as emerging from, and being directed by, the communication needs of humans and human groups. Closely related to the Functionalist project is the area of Typology—the study of similarity and difference in the world’s various languages. Along the way I will introduce the Minimalist Program—currently the dominant theoretical school of thought—and its focus on the idealized concept of “perfection” as a guiding tool to future research. I will also introduce the other dominant school of thought, Optimality Theory. This particular theory of grammar views language as arising from competing constraints on conditions for well-formed, “grammatical,” units.

Underlying discussion of the various research programs and theories in linguistics will be some issues close to my heart. These include mathematical logic, the study of biology and nature, concepts associated with “symmetry of form” in nature, computability theory, and even a little about fractal geometry and complexity theory. It may seem odd that a linguist would be interested in such things. It may even seem unrealistic that one person could study so many different and complicated sciences without resorting to superficial and overly simplistic generalizations. But with the appropriate focus one can find what is useful for an emerging science of language. I hope to show that one does not need to study the whole of science in order to find useful tools in cross-disciplinary research. Think of a threaded sewing needle: once one has decided on a particular thread of interest, the needle can weave its way through various issues in order to tie together what once appeared to be disparate phenomena. As long as that thread of interest is woven together by various strands that are consistently related it will be strong enough to hold disparate pieces together. More concretely: by showing that there exists a reasonable assumption that natural language is a product of the natural world, then certain types of patterns found in nature can be compared against human language. The assumption of human language being part of the natural world, which can be weakened to the point of non-controversy (i.e., to a fact), practically forces us to accept that there is a high probability that some kind of non-trivial symmetry is

encoded into the structure of whatever system allows us to have language in the first place. Given this conclusion, it should not be surprising to speculate that an alien race whose biology happens to be very different from ours might have a much different system of “language.” Contrarily, a race of aliens whose biology was similar to ours could very well speak a “natural” language as we know it. In this latter case, it would only be a matter of learning the alien language, just like learning any human language such as Czech, Japanese, or Xinka.

I hope to convince you, if you need convincing, that the human language capacity is worth studying in a scientific way. Painting with a broad scope, it does not matter if we consider the science of language a “social,” “formal,” or “hard” science. The mystery of human language is complicated and could benefit from a “united front” of scientific approaches. What does matter is that we recognize that human language capacity is a unique and interesting phenomena within the domain of humans and nature. Of course, internal to research on language the question of using social science, formal science, or hard science is a very important question. It will determine, in some cases, the assumptions you must make about the fundamental nature of human language. In other respects, the decision of which domain of science you study language in will clearly delimit certain options you have available in terms of the kinds of evidence allowed—and what importance such evidence will in your theory. It will also effect the kind so methods and testing you can do, dependent on what goals have been set for your investigation. These matters are important when it comes to the details. But this is not a book that will dive into many detailed studies of language or Language—it cannot. Instead, I am more concerned with painting a broad picture of the field of linguistics so that we can get a picture of how diverse this field has become. I hope to impart to you an appreciation for the complexity and diversity of the field of linguistics, and perhaps, a respect for those who dedicate their lives to it. Nonetheless, there is a lot I will be leaving out and this, while regrettable, is also inevitable.

Joshua Bowles February 11, 2010 Utah Valley University
--

Chapter 1

Introduction: What is Language?

Before asking the question “What is Language” we need to make a distinction between ‘language’ (with a lowercase ‘l’) and ‘Language’ (with a capital ‘L’). Let us agree, as is common in the field of linguistics, that the term ‘language’ refers to any one of the actual human languages spoken



on this planet. It may be a dead language like Latin, or a world language like Spanish or English. It can also refer to any signed language like American Sign Language. In general, the term ‘language’ may refer to any language that has been or will be spoken (or signed) by humans; this includes potential languages of the future, for example English in 300 hundred years. A word of caution though, since nobody knows what English (or Spanish, Chinese, Japanese, etc...) will look like in 300 hundred years—or if the particular language will even be used—linguists don’t typically consider potential, probable, or possible languages. While in theory the term refers to any human language, in practice it is restricted to the ones we know about for sure. Another note of caution, the created languages like Klingon,

Esperanto, Basic English, and the Tolkien languages¹ do not count as *natural* human languages. In other words, they are not a product of natural forces that were then shaped further by social, cultural, and historical forces. These latter languages were constructed consciously, and therefore, are only a product of social, cultural, and historical forces—but not by natural means.

By contrast, the term ‘Language’ refers to the unique human capacity. Depending on theoretical persuasion, this unique capacity will be explained in different ways. But no linguist—at least generally—will disagree with the notion that humans have a human capacity to learn, acquire, or develop a native language. Furthermore, most linguists agree that little to no instruction is actually required for a child to develop their native tongue. Of course, reinforcement strategies by parents, early reading activities, and early grammar education enhance the child’s capacity—but these do not *cause* a child to learn the language.

In regards to the uniqueness of this human capacity, there are no examples of any kind of non-human animal *naturally*² acquiring a language like Spanish or Mayan. With appropriate training, as everyone knows, a parrot can repeat human words, but no parrot has ever just simply acquired Spanish on its own. Or, as shown by the children’s cartoon *Martha Speaks*, which involves a dog who eats alphabet soup and is miraculously able to converse in English, no animal has ever just started talking after spending sufficient time around human speakers. Some scientists will argue that the capacity for animals to communicate has been grossly underestimated by us humans—and I agree—but it seems pretty clear after about 40 years of study that animals do not speak human languages. Even under all kinds of training and learning techniques, animals simply will not learn human languages. As far as the parrot goes, most linguists agree that the parrot, because of its unique physiology, obviously has the capacity to reproduce meaningful human sounds in sequences that sound like words or even sentences. But parrots don’t actually “know” English—not like a two year old child knows English. Despite whatever difference of opinion people may have about human and animal communication, the fact remains. Animals do not *naturally* acquire human languages in the same

¹Languages created by the Oxford linguist and fantasy author J. R. R. Tolkien for his series of popular books.

²What is meant by the term “naturally” is that little to no instruction is given within a critical period. In other words, within a very short amount of time the basic grammar and sound system is learned. If animals could naturally acquire human languages, then we would expect that spending a sufficient amount of time immersed in a human speaking community would result in fluency. Of course, everyone has pets who have learned certain commands. But this is not learning the language. Flip this example on its head: just because humans have the ability to distinguish an angry and friendly bark from a dog, does not mean we talk ‘dog’. Or more appropriately, just because we can tell the difference between the smell of one dog over another, does not mean we have fluency in the world of dog scent.

way humans do. Simply put, if you speak German to a human child for a couple years, that child will start to talk back in German. If you do the same thing to a gorilla, chimpanzee, dolphin, whale, mouse, or kangaroo you will not get the same result.

Throughout the rest of the book I will keep the terminology established here. That is, a distinction exists between ‘language’ and ‘Language.’

1.1 What is Language?

The question WHAT IS LANGUAGE? has a long history; at least 2,000 years, but one can speculate that it goes back much further. Throughout most of this history there have been two conflicting opinions that can be summarized simply by the contrasts in Table 1.1.

Table 1.1: Simple Explanatory Contrasts for Language

LEFT	RIGHT
Spirit	Bestial
Immaterial	Material
Designed by Nature	Designed by Society

The LEFT side of the contrast includes explanations such as

- (1) Spirit, Immaterial, and Natural Explanations
 - a. Language is a product of God.
 - b. Language is an ephemeral property of the mind. , Language is some kind of system that is *supervenied* onto the materiality of the human brain.
 - c. Language somehow arose from the general physical forces that shape all of nature.

The RIGHT of the contrast includes explanations such as

- (2) Bestial, Material, and Social Explanations
 - a. Language is a product of earthly (and sinful, perhaps) forces which try in vain to express the perfection of God.
 - b. Language is the product of physical mechanisms that express thought. , Language is some kind of system that emerges from the complexity of the brain and body; i.e., the *embodied brain*.

- c. Language is an evolutionary adaptation that arose from some kind of communicative need in social settings.

I want to highlight here that the simplistic contrasts given above in Table 1.1 are just that: simplistic examples. The explanations of Language in (1) and (2) are complex and cannot (in fact *should not*) be placed into over-simplistic dualisms. The truth is that a proper and complete explanation for Language and languages has to account for **both** sides. Specifically, I want to point out possible similarities between (1-c) on the one hand and (2-b), (2-c) on the other.

1.2 What is a ‘linguist’ and where do I find one?

The answer to the question in the title depends on what kind of linguist you are talking about. In other words, there are many different kinds of linguists who employ various tools to study or process various properties of natural language. You might find a linguist in



the jungles of Peru or Brazil, the deserts of Mexico, Australia, or Africa. Perhaps in the departments of philosophy, cognitive science, or computer science at any modern university. Or in special labs with soundproof booths asking undergraduate students to decide between certain sounds. Perhaps in other labs conducting some other kind of strange experiment such as a computer simulation based on an algorithm for resolving semantic ambiguity, parsing syllables, running statistical tests on a corpus of data, or looking at geographical spread patterns for of populations for some genetic variant as it relates to spread patterns for some language feature like tone. A linguist might even be found at various intelligence organizations such as the CIA, NSA, or FBI, doing some kind of translation or analytic work on Chinese and Farsi. They can even be found working at Google, Microsoft, Attensity, AT&T or any number of corporations that need lexicographers or syntacticians. There may even be a linguist around the corner from your house or apartment.

The most obvious place to find a linguist—if, for example, that linguist is me—

is at the library. The term linguist has been, and still is, used to refer to a person who speaks many languages. For example, the U.S. Army hires ‘linguists’ who speak Farsi, Poshtun, or Chinese. This concept of the linguist does not really contradict with what the *scientific* linguist is, but only limits it. In other words, scientific linguists are very aware of the world’s languages—and in fact may speak several of them or specialize in studying many of them—but our interest in these languages is not so that we can, for example, speak Italian when vacationing in Rome or translate a Iranian engineering manual from Farsi to English. Instead, we use languages like Spanish, Mayan, Navajo, or Irish as *data*. But data for what?

1.2.1 Brief Detour: Material and Immaterial

Different kinds of linguists use natural language data in different ways. This really boils down to different perspectives of what natural language is and the multiple interpretations of the language data. There are two dominant perspectives of the nature, cause, and origin of natural human languages and they have been around for a very long time. Summarizing a bit, they can be boiled down to the difference in focus between “material” and “immaterial” effects of language. For example, it is obvious that we don’t communicate telepathically—even if telepathic communication were real, it is definitely not used as the major mode of communication. Instead, it is a fact that movement of parts of the body are necessary to produce language. So when someone says, either verbally or in sign language “I would like to visit space.” or “This flower is so pretty!” they are manipulating the physiology of their lungs, throat, mouth, and tongue in the one case (spoken), and their arms and hands in the other (signed, or even written³ language). Additionally, a speaker also manipulates the physics of acoustics (for speaking); and recent evidence has shown that visual perception of spoken speech plays a crucial role in both understanding what is being said between fluent speakers, and in developing the language as a child and acquiring the language as an adult. All of this falls under the umbrella of the “material” effect of languages. Consequently, one can use language data to “get at the material” causes and effects of language and Language.

The “immaterial” effect comes in our understanding the meaning, structure, and sound of the intonation or level of voice—even in sign languages one can communicate intonation or emphasis. The interpretation of utterances on these three

³Though it should be noted, written language is an artificial symbol system that has been invented and exploited for practical purposes in a society. Writing is not a natural system; therefore, many languages past and present never had writing systems, and this makes no difference in their status as a language. However, lack of a writing system clearly has an impact on its preservation for the future. Latin is still studied today because not only was it written down (and also codified into legal and medical jargon), but most importantly, there were grammars of Latin written by native or fluent speakers.

levels—meaning, structure, sound—takes place within our heads. It is a mental or cognitive process that includes thoughts and concepts that nobody can actually see. We may communicate some of this mental activity in our body language or facial expression; but this does not make our thoughts visible. An interesting analogy can be made here between this internal “immaterial” sense of language and the gravitational force. I often like to begin a lecture by writing on the marker board the following question

(3) What do Language and gravity have in common?

I anticipate a mild state of confusion in the students. Then I pick up an eraser, holding it in the air, and let go. Predictably, the eraser falls to the ground. Then I ask “What just happened.” The usual response is varied but consistent. After some playful comments someone will usually say “You showed us gravity.” To which I can respond with “Where?” Some more chattering, and finally we all agree that we can’t actually see gravity—we are only able to see its effect. The same is true with Language—and even language. I challenge anyone to show me Language/language (Remember, writing systems are artifacts of languages, not languages themselves. Also, watching frequency waves of spoken language is also an artifact). Even though we cannot see the meaning of a sentence, or a sentence itself, or even a meaningful unit of speech like the word ‘the,’ we know without a doubt (or at least are justified to believe) that such things are real because we observe the effect all the time. The study of language data can also try to account for the causes and effects of this reality of meaning.

Tensions between the “seen” and “unseen” aspects of natural language were usually put into the terms of religion or philosophy by thinkers going back to as little as a few hundred years to as much as thousands. A common view was that the ability to think and reason like rational beings made us separate from the beasts of the world. This rational capacity came from God, or Allah, or the creative force of the universe. As rational beings designed by the omniscient creative God of the Universe we were left to ourselves to communicate and work through our reasoning ability. The only way to do this was through earthly means: sound. We could hear sound, and this kind of made a “seen” or “material” thing that belonged to the bestial world. Evidence to support this came directly from the fact that animals appeared to make all kinds of sounds. Some even seemed to communicate in some way or other. Consequently, the difference between “us” rational humans and “those” earthly beasts was that we had some kind of “immaterial” presence within us that could organize our sounds into systems that could communicate our divinely inspired gift of reason—though our ability to represent divinely inspired reason through bestial sounds was not perfect and varied between individuals. The

results of centuries of this kind of thinking have prevailed in some form or another even to this day. Furthermore, many people over the centuries have tried to combine the conflict between the “material” and “immaterial.” This attempt at combination is still underway, though not really in terms of God and beasts.

1.2.2 Back to the Question

Generally, there are three modern approaches to language data and all three can be associated with what might be called *flagship* departments at universities in the United States (though I am sure to offend many linguists either by not mentioning specific approaches or by associating certain approaches with certain American universities—but such are the sorrows of generalizing). These are departments that appear to have made a conscious effort at focusing intensively on a specific approach to language data in the work of their faculty and graduate students. The first approach is called typological-functional and is well-represented by linguistics departments at the University of California at Santa Barbara (UCSB) and the University of Oregon (UO). The second approach is called cognitive linguistics and can be loosely associated with the linguistics departments at University of California at Berkeley (UCB) and University of California at San Diego (UCSD), as well as the cognitive science department at Case Western Reserve University. It is difficult to choose *the* flagship American university department that represents the functionalist program, mostly because this approach was never centralized at any department. But I feel confident that the universities listed above are representative. Cognitive linguistics and typological-functional approaches to language data are closely related and share some common ideas—especially when viewed in relation to what is generally known as the generative linguistics approach. This approach is strongly rooted in the linguistics department at the Massachusetts Institute of Technology (MIT). The generative approach has changed considerably over the years, and is currently viewed as being more of a research program (versus a theory) called the Minimalist Program (MP). Because the generative approach, and especially the MP, is the dominant abstract theoretical approach it would be difficult to name other flagship university departments. MIT, for historical reasons, is the beginning of it all. Clearly, departments change by way of new faculty and new foci. But for those curious to know more about certain approaches, a good place to start is by looking at the the linguistics departments mentioned here.

So, what is a linguist? There is no easy answer, but I hope that by the end of this book you have some clear notion of what a linguist is. Specifically, I want to show that field is widely diverse and constitutes a multiplicity of approaches to natural Language and natural language data. A linguist is much more than a person who speaks many languages.

Before moving on, let me clarify some assumptions. There is no more room to go into how human language is unique within the animal kingdom. Instead of giving superficial examples based on a complex web of empirical data and theoretical inquiry, it is more convenient to agree to assume that natural human languages, such as Japanese, Kakchiquel, Winebego, and Portuguese, represent particular examples of a general human system of sound-symbol matching that creatures like dogs, monkeys, parrots, and even mammals like elephants, mice, whales, and dolphins, do not seem to have (though they may have particular aspects, or subsets of the human system—particularly in reference to mammals' capacity to have a syntax). That is, animals may be able to communicate specific messages in the context a present danger or a food source, but they do not seem able to produce messages in the way and of the kind the humans do. So, assumption number one is that natural human languages are unique to humans. Embedded in this assumption is another assumption: that different specific human languages are variations of and elaborations on a basic human system. Let us also agree on this with appeal to an intuitive proof of its reasonableness: Any child who grows up anywhere on the planet will speak the language they are exposed to. If you were born in the United States, but your parents moved to Japan when you were two, then it's fairly obvious that you would speak fluent Japanese had you grown to adulthood there. (In fact, you would probably be fluent in both English and Japanese if your parents always spoke English to you at home). One does not have to look far for immigrant families where this has occurred. Now that these two assumptions have been given, we can move on.

1.2.3 A Brief Note on Evolutionary Adaptation and Exaptation

The following question underlies the main debate between the major camps of linguistic theory: Do human beings really speak in order to get messages across to one another? Clearly the answer is Yes. But is it the FUNCTION of language to communicate messages? Put another way, did the uniquely human ability to use language arise to meet some pressure to communicate messages between each other? The answer to this is more difficult. On first glance many people feel compelled to answer Yes: language communicates messages, that is its function, and it came to be used by humans because of the external pressure to communicate messages between each other. But there is a complexity here that is often overlooked.

It is a matter of fact that language is used to communicate—that is the primary use to which language is put. The problem lies in thinking that just because something has been used for a specific purpose then it must have been created for that purpose. Take wings on a bird, for example. The evolution of winged animals provides an excellent counter-example to argument that nature is always an *adaptive*

force. If we applied the same reasoning used for explaining language to explain the appearance of winged creatures we would be wrong. Such reasoning would lead to us conclude that wings were developed through evolution so that birds could fly—because this is what wings are primarily used for. Based on this, we might speculate that wings developed through evolution in order to meet some need that certain creatures had, either to get away from prey or to cover large distances in order to find food (or both). However, the evidence is clear that this is not what happened. We might call this the FUNCTION-ORIGIN FALLACY, which generally says that it is mistaken to assume that the origin of a thing can be reduced to its current primary function. Such things that fall under the function-origin fallacy are referred to as the products of Exaptation, versus Adaptation. Now drawing the parallel closer to an explanation for language, the function-origin fallacy can only go so far: it only shows us that in some cases when one tries to pin down the origin of a natural object, physiological property, or organ, to its current primary function we come up short. Nothing in the fallacy will tell us what kinds of objects it applies to, including natural objects like language and Language. In other words, with specific reference to language, an argument that appeals to the function-origin fallacy does not by itself give us any reason to think that Language is a product of Exaptation; more is needed. I bring this up at the end of the introduction because it will haunt the rest of book; though I will not go into detail about the evolution of language. I briefly talk about exaptation in xxxxxxx, but that is the only place I mention these problems explicitly. They are very important and are the subject of much current debate. It would be lacking to not mention such issues—if only to point out that they will not be discussed, they are there nonetheless.

Part I

Exploration

Chapter 2

World in a Leaf

Goethe, Thompson, Gould. . .

Chapter 3

Nature is a Computer

Turing, Post, Deutsch, Lloyd, Kornai. . .

Part II

Functionalism and its Adherents

Chapter 4

The Functionalist Program

Functionalist theories of language do not typically call themselves a unified research program; or at least do not refer to themselves in that manner¹ But just because they do not use the term “research program” does not mean it is irrelevant. In fact, I do not think functionalist theorists would deny that Functionalism is a “research program.” Nor would they deny that the various theories under the functionalist umbrella share unifying principles. One can easily see in functionalist writings that they are very aware of their unifying principles.

4.1

4.2

4.3

¹This is in comparison with the practitioners of Minimalist Program theories and their acute sense of the unity of various theories as a “program.” See Chapter 6.

Chapter 5

Why Communicate?

5.1 Why do we communicate?

Talmy Givon

5.2 The World of Typology

Martin Haspelmath

Bernard Comrie

5.3 Structuralism

5.3.1 Bloomfield, Sapir, and Harris

5.4 Documenting Languages

5.4.1 Boas and Sapir

5.4.2 The Ethical Imperative: Endangerment and Revitalization

Lyle Campbell

Marianne Mithun

Greg Anderson, David Harrison, and Popular Media

The Dynamic Duo: Aikhenvald and Dixon

Part III

Formal Considerations

Chapter 6

The Minimalist Program

The term generative has come to be associated with a vibrant, if not controversial, movement in the latter half of twentieth century linguistics. If you do not agree with generative theories of linguistics, you must still study them at most university graduate programs. The movement has not only defined its own general methods and philosophy of science, but many different approaches have defined themselves specifically in opposition to generative approaches. Consequently, the movement has had a large influence on those invested in it *and* those who believe it is the wrong approach.

Chapter 7

Optimality Theory

Chapter 8

Other Approaches

Although the theories and methods described in this chapter are not the most prominent, this does not mean they are not crucially important to the overall enterprise of investigating Language and languages.

8.1 Head-driven Phrase Structure Grammar(HPSG)

8.2 Generalized Phrase Structure Grammar (GPSG)

8.3 Lexical Functional Grammar (LFG)

8.4 Combinatory Categorical Grammar (CCG)

8.5

8.6

8.7

Chapter 9

From Infinity to Biology: The Logical Influence of Formal Systems

9.1 Why Logic Matters to Linguistics

This chapter contains brief historical information on many of the leading figures of logic, meta-mathematics, and formal systems in the last 150 years. It is not a complete list and does not go into great detail of the specific contributions of these logicians and philosophers. There are many great sources that detail or collect the specific advances in logic and meta-mathematics. References are provided at the end of the book. In order to understand where theoretical linguistics comes from it is necessary (and interesting) to trace the logical lineage of the people who inadvertently influenced the direction of the discipline. A thorough knowledge of the work and life of those discussed here is not necessary to appreciate the work of contemporary linguistics. It is not even necessary to *do* contemporary linguistic work, but it can be very helpful in fully appreciating potential directions of research—not to mention the need for and possibility of mathematicization of grammatical theories. On a last note, as a teacher of scientific writing I have always stressed to my students that the lives of research academics—physicists, biologists, mathematicians, and even linguists—do not fit into any stereotypical mold. Such a mold would force us to see the lives of such scientists as spent twittering away in libraries and labs. The converse seems almost always to be the case. Certainly, a lot of time is spent reading, learning, writing, and performing experiments. But science is also about having an active life and following your intuitions and passions—and enjoying what you do. As I tell my students, you have to enjoy what you do in order to get through all the boring literature and the hours spent learning particular formalisms. Nobody claims that these are exciting activities; what is exciting is what one can do after reading the literature and learning the formalisms. In other words, the life of a scientist is a rich life—both internally and externally. Besides providing a small tapestry of the lives (and names) of logicians who have influenced parts of linguistic research, I hope also to give a brief glimpse into the rich and varied life that the scientist and academic enjoys.

9.2 Cantor: Infinity Comes in Different Sizes

Infinity comes in different sizes! For a short while I taught courses on science writing and would often sneak into my lectures on argumentation surprising or interesting information. When I would tell students that infinity comes in different sizes and that there is a multitude of infinities (in fact, an infinite number of infinities) their reactions would range from disbelief to wonder. Rarely would a student show boredom upon receiving such information—probably because the idea of different sizes of infinity is so counter-intuitive. Nonetheless, the German mathematician

Georg Cantor was able to show that this really is the case, though it arguably lead to his descent into madness. Georg Cantor was born xxxxx

9.3 Search for Boole

9.4 Frege

Gottlob Frege is sometimes called the Father of Modern Logic. Ironically, the last sentence contains an example of one of Frege's major insights into language. We know that 'Gottlob Frege' and 'Father of Modern Logic' refer to (or name) the same human being (or object), but they each have a very different "sense" in which they *name* the same object.

9.5 Carnap and the Logical Syntax

xxxx student of Frege xxxxx

9.6 Tell them Gödel Sent You

Kurt Gödel has not been very influential in the world of linguistics, at least not directly. But he has privileged place in any list of influential logicians. His work is regarded as *the* most important in the twentieth-century. Some consider his work in logic second only to the work of Aristotle. Aristotle, of course, founded the study of logic by providing explicit definitions and schemata for the analysis of propositions. His logical systems remained virutally unchanged for 2,000 years. Nobody can replace Aristotle as the greatest logician, but only because he came first. When it comes to profound insight, however, Gödel definitely has my vote for the greatest logician in all of world history. Most people would agree with this sentiment. Certainly, anyone with a reputation like this deserves to be mentioned, even if their work is not specially relevant to the topic at hand. Of course, you never know what the future has in store: Gödel's work may find a relevancy it had not had before. Nonetheless, current linguistic science has not found much use for Gödel. Gödel was infamously shy and reclusive. He died from complications due self-imposed starvation, the result of which was caused by his paranoia of being poisoned. He was born xxxxxxx

9.7 Tarski, Metamathematics, and Truth

Alfred Teitelbaum was born January 14, 1901 in Warsaw Poland. By the time he was 34 years old, in 1935, he had changed his name to Tarski and established himself as one of rising stars in the world of logic. Certainly, at this time he was seen as one of the best, if not *the* best, logicians in Poland—which was one of the best places to study logic in the world. By 1935, Tarski had formed relationships, and in some cases given lectures to, many of the people who would also become world famous for their work in logic: Kurt Gödel, Rudolf Carnap, W. V. O. Quine, Karl Popper, and many others associated with the Vienna Circle and the movement of logical positivism. For an excellent background on Tarski see the biography by Anita Burdman Feferman and Solomon Feferman titled *Alfred Tarski: Life and Logic*.

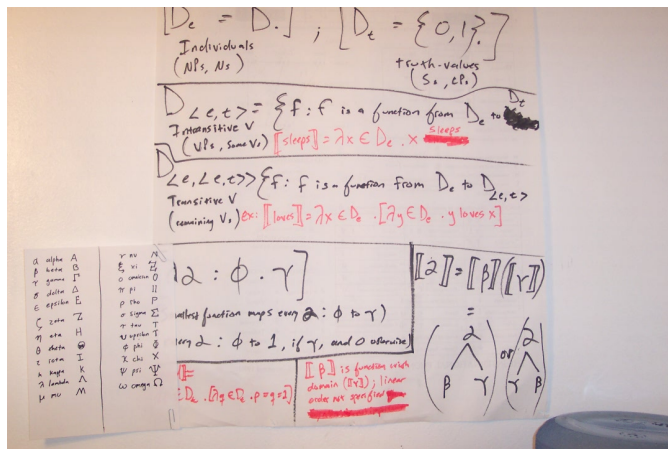
Besides being one of the greatest logicians of the twentieth-century (most consider his contributions second only to Gödel), Tarski is particularly important to the early development of theoretical linguistics because of his theory of truth. I'm not aware of Tarski having any influence on syntactic theories, but he did have an impact in the field of semantics. His two most direct influences on natural language semantics came through (i) his metamathematical definition of truth, and (ii) the work of one of his students, Richard Montague. I will discuss (i) here, and (ii) in the next section.

9.8 Montague

xxxx Montague semantics, formal grammar xxx xxxx student of Tarski xxxx

Chapter 10

Modern Semantics



10.1 Models and Functions

Barbara Partee

Barbara Partee is widely acknowledged as singularly reviving Montague Grammar, elaborating on it, and extending it for applications to natural language semantics.

Irene Heim and Angelika Kratzer

Carlson and Cheirchia

Horwich

10.2 Events Really Happen

Terrance Parsons

10.3 Concepts are Crucial

Ray Jackendoff \equiv Conceptual Semantics

10.4 Cognition's the Thing—Cognitive Linguistic Semantics

Gilles Fauconnier, Mark Turner, and Conceptual Blending

Mark Johnson, George Lakoff and the Embodied Mind

10.5 Is that My Lexicon Over There?

Pustejovsky's Lexical Frontier

John Lyons

Beth Levin

10.6 Innate Ideas, Innate Meaning

Wolfram Hinzen

Paul Pietroski

Paul Hagstrom

10.7 The Neural Theory of Language

Jerome Feldman

Srinivas Narayanan

10.8

10.9

10.10

Part IV

Cross-disciplinary Inspiration

Chapter 11

Galaxies, Flowers, and Snowflakes

There are general principles of organizational force in the universe: from the smallness of ice crystals forming into snowflakes, the shape of petal design in flowers, to the massive roundness of galaxies. The general

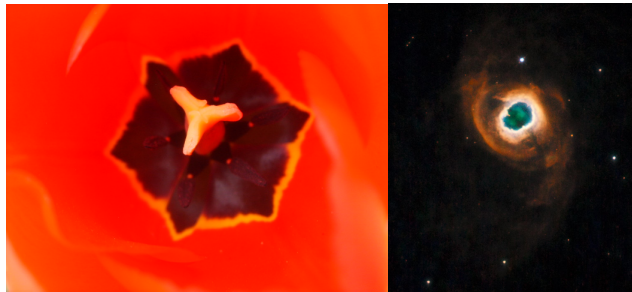


Figure 11.1: Spiral Galaxy Courtesy NASA

interaction of atomic, molecular, magnetic, and electric forces converge. This convergence results in certain shapes. It may be a funny coincidence that we find these shapes — hexagonal and pentagonal geometries of snowflakes and beehives, the shape of crystals, the shape of trees and flowers — pleasing. But it might also be a mixture of some interesting capacity for the human brain to ‘appreciate’ things and the forces of nature to produce optimally efficient repeating patterns. That is, the patterns we find in nature are not beautiful in themselves, but can only be beautiful to an audience ready to appreciate them. Long-term exposure to such patterns in nature would have most certainly affected a human mind. Such interaction between natural environment patterns and a human mind capable of such appreciation is most likely a factor in our finding images of spiral galaxies pleasing. Of course, much of this is speculation, but I don’t think many people would deny its intuitive force. There is something between nature, the patterns it produces, and the human mind that is unique and special. Granted, at some point in evolutionary

history we may not have been totally predisposed to appreciate patterns, and it may in fact be the case that our ability to appreciate patterns came directly from selective pressures in the natural world. Nonetheless, no one would argue that we find such patterns pleasing in a way that is very different from the way other species are attracted to patterns in nature. Whether or not our brains were already predisposed to find shapes and patterns pleasing is a moot point for us here. What is important to realize here is that such patterns are not 'beautiful' in themselves, but only in the context of a mind ready to be pleased. As the Astro-biologist Carl Sagan once said, the human mind is a way for the universe to study and appreciate itself. As a part of nature we should not find it surprising that we are fascinated with its visible geometry and expressive coloration.

What does this have to do with Language? When we study the abstract structure of Language, through investigation of language data, we have to keep in mind that Language itself is not 'beautiful.' That is, from one very abstract perspective, the patterns and asymmetries in language systems imply patterns in the overall structure of Language, but this structure is not due to our sense of what is pleasing, beautiful, or efficient; though nobody denies that aesthetic intentions can shape the form of produced language. The claim, then, is that some abstract biological information (pattern) encoded in the basic building blocks of the human organism is responsible for the general structure of Language, which is in turn responsible for drawing the boundaries possible structures in human languages. This is usually called the language faculty, or Universal Grammar. No form of poetry, no advertising jingle, no cultural movement, no government mandate can change the basic core structure of the language faculty. On the other hand, a more grounded view of Language sees the patterns in language data as coming directly from a long history of refinement of complexity for the purposes of communication. That is, a large part of the structure of languages is due to some sense that humans have for evaluating complexity, efficiency, and in some regards beauty of communicating messages and thoughts. The interplay between the need to communicate and the inherent complexity of language processes converge to shape the overall structure of Language. This structure can be gotten at by analyzing language data, but there is no reasonable expectation that we should find ONE overarching structure for ALL languages that have ever existed, ever will exist, and that exist presently. There in fact might be a few different templates that have arisen over the millennia and different languages probably exploit different templates. Both of these approaches are reasonable approaches to the scientific study of human languages and human language capacity.

What is interesting is that these two approaches do share quite a bit in common, even though they have radically different methods of evaluation and analysis and can come to some very different results. The most obvious similarity — besides

the fact that we have to view languages within the context of humanity, and humans have to be viewed, partly, in the context of nature — is that both approaches admit of some organizing principles in human languages and Language. However, they have radically different ways explaining these organizing principles: formalists hypothesize one general ‘cooker cutter’ template, while functionalists hypothesize that there really is no ‘cooker cutter.’ This is of course a gross simplification of the very rich theories that exist, but it will do for our general purposes. The similarity here, though, is not really that impressive. If you simply thought about these issues for a few hours you would have to come to the same conclusion. The only choice you really have is which side of the investigation you fall on. In other words, to deny that there are organizing principles in all languages — and that closely related languages follow some general schema (this is why they are classified as ‘closely related’) — would be tantamount to denying that there are organizing principles in the world. You simply would not have a leg to stand on: there would be nothing substantial you could say about languages, and in the end would have to abandon its study. But only a moment's reflection will reveal to anyone that languages are organized and do not need our help to stay organized. Most linguistic changes happen over generations and in the presence of strong resistance — like the loss of the who/whom distinction, which older people will insist on, but younger speakers will only use the distinction (if they memorize it) to please the older generation. Secondly, a native English speaker cannot deny that the typical word order of English is Subject, Verb, Object and not Subject Object Verb as in

- (1) Jon kicked the ball
- (2) Jon the ball kicked
- (3) Kicked the ball Jon
- (4) The ball kicked Jon

Whereas other languages have the typical orders that vary from SVO. Anyone who denies such organizational principles in language must be forced then to deny the study of language; and even worse, will have to deny organizational principles in nature and society.

Chapter 12

A Formal Conjecture: Symmetry, Syntax, and the Theory of Groups

12.1 Human Language and Mathematical Systems

Language is not math and languages are not numbers—this is intuitive enough. It may not be so intuitive to say that natural human language cannot be exhaustively modeled by a formal mathematical system. What do I mean by “modeled exhaustively?” The answer to this lies in how we define the crucial properties of human language and what our scientific goals are.

If I want to measure how many times the word “bedeviled” occurs in the modern English usage, then there are statistical techniques designed to calculate and measure such patterns based on a certain corpus or databank. This kind of study deals with the external distribution patterns of words and sentences and is largely dependent on cultural and historical accidents. For example, the word ‘rad’ may have had a larger, or more frequent, distribution in the 1980’s than it does now. It is simply a historical and cultural “accident” that ‘rad’ has the overall distribution that it does. The fields that study these kinds of matters are named lexicography and corpus linguistics, and the simplified example given here does not do justice their full complexity and richness, but it works to illustrate my point that the external output of human languages are subject to various kinds of measuring techniques with varying degrees of mathematical precision in showing language patterns. Importantly, these patterns (for example, how many times the word ‘rad’ is used) are not due to properties of Language itself, but to other sociological and historical-cultural properties. In other words, how many times a word is used in a language is not a crucial property of either that language itself, or of Language in general. Lastly, Frequency is not all that can be studied in lexicography and corpus linguistics, and in fact, some characteristics of human language can be revealed through these approaches (such as the size of a vocabulary). This still does not change the fact that, while these kinds of studies are important and can be quite complicated, they are not useful for those who wish to discover exactly what Language is and why it is a unique property of humans.

There is another mathematical way to study human language that might be relevant to explaining the *what* and *why* of human language. The study of what are called logical or formal languages has been used to explain human language. Examples include computer programming languages such as C++, or what is called the first order predicate calculus. In these languages efficiency is the golden rule. One can consciously design them to filter out redundancies and ambiguities. Focusing on the logical system of a first order predicate calculus, we might try to package natural human languages into this artificial logical language. Or, we might try to *reduce* natural language to the sleek form of the logical system. This is what is meant by ‘exhaustively model’ natural language by a formal system, and it cannot be done as I will show.

12.1.1 Ambiguity and Redundancy

Formal or logical languages are very different from human languages in at least two major ways: redundancy and ambiguity.

- (1) Jon runs and we like to run too.
- (2)
 - a. Jon paints cars.
 - b. Visiting relatives can be lame. , new drug technology

In (1) there are two kinds of redundancy. The first is in the redundancy of the sound of the distinct words ‘to’ and ‘too.’ In speaking, this kind of redundancy can actually lead to ambiguity—think also of the word ‘two.’ The second kind of redundancy in (1) is a more complicated kind known as agreement. In English, we have subject-verb agreement where the the same kind of information shows up on both the verb and the the subject. This information includes the number (singular or plural), gender (masculine, feminine, and neutral), and person (first, second, third). In the example given, ‘Jon’ is a singular, masculine, third person noun and so the verb that describes Jon’s action must also have singular, masculine, third person features. For this reason, we know that it must be the verb ‘runs’ that describes Jon’s action and ‘run’ that describes the action of ‘we.’ In languages other than English the agreement system can be much more complex—coding dual and various kinds of plural numbers, as well as a wide variety of different kinds of grammatical gender such as long, thin, animate, and also different kinds of person such as first-inclusive and first-exclusive. No matter what kinds of features or coded by agreement in natural languages, it seems pretty clear that natural language have to have agreement. Differences arise in the form of agreement, with some systems being observable and others not. Despite these differences, agreement in languages is a form of redundancy not found in formal languages—it is not efficient in an operational way in mathematical, logical, or computational languages. But in natural human language, there are strong arguments based on pretty good evidence that the redundancy of agreement represents an optimal solution to the different systems needed to run the language generator. Example (2) shows ambiguity. There are two types of general ambiguity, one is lexical and one is structural. In example (1) the words ‘to’ and ‘too’ sound alike but have different meanings—also ‘two.’ This is an example of lexical ambiguity. Think of other homonyms—words that sound alike but have different meanings—like the animal ‘bat,’ the baseball ‘bat,’ and the verb to ‘bat’ your eyelashes. Structural ambiguity is shown in the examples in (2). In (2-a) does Jon apply paint to cars or does he apply paint to a canvas while drawing cars? In (2-b) is it lame to go visit relatives or is it lame when visiting relatives come stay with you. Finally, in (2-b) are we talking about technology for

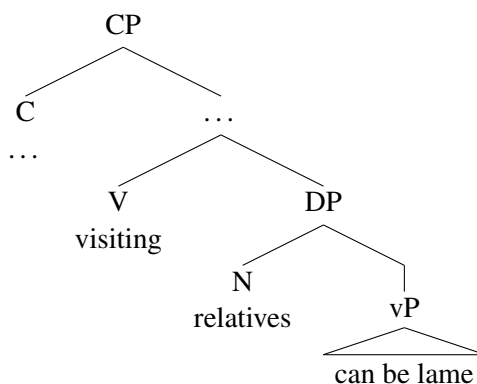


Figure 12.1: (Visiting (relatives (can be lame)))

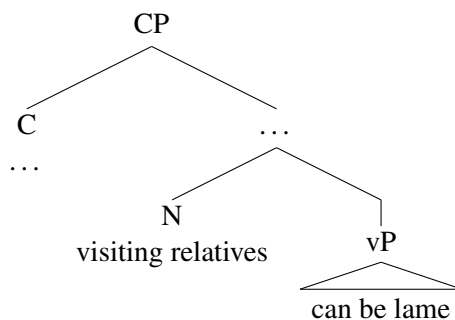


Figure 12.2: (Visiting relatives (can be lame))

new drugs, or new technology for old drugs? The answer to these questions can go either way, depending on your interpretation of the ambiguity. It is commonly understood in theoretical syntax that ambiguity provides robust support for the notion that there is an abstract hierarchical structure underlying natural language syntax. Simplified examples can be seen in the syntactic trees of Figures 12.1 through 12.4—category labels are not meant to be exact, I’m only trying to show different hierarchical groupings. Such tree structures are geometric models for the hierarchical relationships between constituents. There are no linguistic claims that such trees exist “in the head,” but the hierarchical relations between constituents that they model *are* assumed to be real properties of the the human mind.

Formal languages simply do not have the properties of redundancy and ambiguity because they make systems less efficient. We can build a logic or computer language with ambiguity in it but it might not be very useful. What linguists have found useful is to use tools from logical systems to model parts of natural language. This appears to work very well. When it comes to modeling all of natural

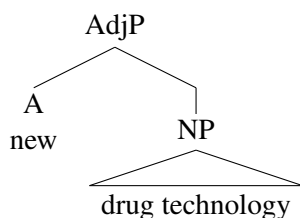


Figure 12.3: (new (drug technology))

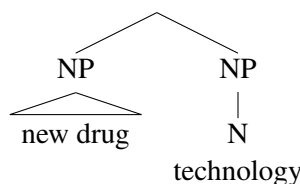


Figure 12.4: (new drug (technology))

language by a formal system, things quickly get out of hand. In fact, in computational linguistics—the field of linguistics that applies linguistic discoveries to computer languages—one of the biggest problems is ambiguity resolution, both structural and lexical.

12.1.2 Competence-Performance, Internal-External

Around the 1930's attempts to model natural language by an efficient and consistent model of grammar quickly showed that the model could only be gotten in one of two ways: (a) a synthetic grammar contrived from a first order predicate calculus, or (b) accounting for the distribution of particular morphosyntactic elements in produced language data—think of the example of the word of 'rad. Neither (a) nor (b) gave answers to the *what* and *why* of natural language. In other words, they did not shed any explanatory light on the following observations: any human being of sufficiently young age can acquire any human language with only a few years of exposure. Statistical studies have shown that the amount of language data a normal child is exposed to is not enough for the child to learn that language by a simple method. Some had thought that children begin as blank slates simply mimic their environment, or that they learn certain basics of a language and then extend on the base. But the fact is that the amount and quality of language data the child is exposed to are not significant enough to allow a child to start from a blank slate and learn everything the need by the age of about five. Not only this, but a child can speak sentences they never heard before. Furthermore, a child can understand

a great deal more than they can speak. This helped support a growing sense of the distinction between language *performance* and language *competence*. The latter is the grammatical system that we all seem to have intuitively, and the former is how this grammatical system is used in the world. Here we will take a short detour in order to see an example showing the difference between these two distinctions.

A modern speaker of English will rarely employ the infamous ‘who/whom’ distinction. Most people do not remember how the ‘who/whom’ distinction is to be used and will avoid it whenever possible. On the other hand, nobody has to think about the ‘he/him’ distinction. We all know that the sentences

- (3) He is the soccer player.
- (4) He did not intend to kick him.

are perfectly fine, but that

- (5) *Him is the soccer player.
- (6) *Him did not intend to kick he.

are ill formed sentences. The ‘who/whom’ distinction codes exactly the same information: he-who, him-whom; notice the *-m* ending. So how do we account for the massive difference in performance here? How do we account for the fact that practically no one can remember grade school teachers and grammarians scolding us on the proper use of ‘he-him,’ but most everyone I know can certainly remember having to pay extra attention the ‘who/whom’ distinction? The answer is quite simple, really. The ‘who/whom’ distinction is not part of our competence for English anymore, while the ‘he/him’ distinction is. I will explain. The English language has been losing what is called its Case system. Case is a way for languages to, essentially, tell us who did what to whom. Case information is almost always coded on nouns and helps distinguish, among other things, subject and object. In current syntactic theory Case plays a very important role in forming all sentences, and all languages must have some form of Case even if it does not show up on the nouns—it is usually called structural case and is distinguished from the case properties that we can see, called morphological case. In English, we used to have a very rich morphological case system but have been losing it for reasons too complicated to explain here (but see Appendix xxxx). However, the case system in English pronouns is still preserved: he, she, him, her. The different forms of the pronouns signal to us whether or not they are subject or object, no matter where they may occur in the sentence. For some reason this preservation of the case system does not seem to apply to pronouns that do not specify grammatical gender: it, who/whom. For this reason, modern English speakers do not intuitively code

the ‘who/whom’ distinction because it no longer belongs to the actual grammatical system of English—or what used to be called our language competence. This is why we all have to struggle to remember, or memorize, that ‘who’ is for subjects and ‘whom’ for objects. Some grammarians decided that the ‘who/whom’ distinction was still important and so they tried to tell everyone that the distinction was important. Clearly, telling people how to speak does not work if what you are telling them goes against their native intuitions; and we all have native intuitions about the English language, despite what prescriptive grammarians may try to tell us about the right and wrong way to speak. The grammarian rightness or wrongness is a matter of socio-political and cultural standards, NOT an objective standard of the nature of the English language—or any language for that matter. The competence-performance distinction is very difficult to define and it is probably not as cut-and-dry as it appears to be. The terms are not usually used in current syntactic theory, at least not in a really technical way, mostly due to the problems in trying to specify what parts of language belong to competence and what parts to performance. I have always been more inclined to see it as a duality than as a complementarity that can only be elaborated after much more is known about Language—how it develops in children from birth, how it changes over generations, what kinds of social factors can influence sentence structure. In other words, I see Language as partly defined by a competence-performance duality such that the grammatical system of a language may be altered by speakers’ performance/use of that system. Another way to describe this may be the statement that Language is a system composed of some finite invariant principles and some finitely variable parameters that have a small range of variability. The competence-performance duality belongs to the parts of language that are parametric.

Returning now to the problem of mathematical models in language. It soon became obvious (by about the early 1960s) that no simple mathematical model—whether it was a statistical distribution or a logical calculus—could capture the structure of Language. Eventually, conceptual tensions between (a)—synthetic grammar of a first order predicate calculus and (b)—distribution of morphosyntactic elements in language data, as well as between the notions of *competence* and *performance* forced a distinction between Internal and External structure known as I-language and E-language (around early 1990s). The focus of most generative syntacticians is in describing and explaining the internal mechanisms responsible for initial language acquisition. This is the **what** and **why** of Language and belongs to the domain of I-language. In the current climate of research it is now more appropriate to talk of the “growth form” or development of language in terms of acquisition. Humans do not really acquire language, unless we want to talk about humans acquiring vision, hearing, walking, and other kinds of processes. These processes clearly must develop in response to the environment and need guidance

from the adult in order to reach maximum potential, but they can also develop on their own. The same is true of language—it is just that the *apparent* variety of particular development paths (that is, languages such as Hindi, Swahili, or Hopi) seems stunningly distinctive when compared to the variable paths that vision, hearing, or walking can develop toward. What makes studying Language more challenging than vision or hearing is that, for one, the results of development seem to be so drastically different (in the form of different languages), and two, there is no easily detected material organ to which we can attribute the cause, and thus, put under a microscope like we do with eyeballs for vision, ear canals with hearing, and general physiology of legs, hip, and balance for walking. It should also be pointed out that the processes responsible for vision, hearing, and walking are only basically understood—much is still unknown.

I-language, as a growth form unique to the human organism, is a natural object—or so the argument goes. What makes this conjecture so contentious, it seems, is that unlike vision, hearing, or walking, it appears to show not only a great variety of resulting development paths, but also it seems to be unique in the animal kingdom. However, some of this contention can be resolved by noticing that Language is employed as a communication system and the animal kingdom has an extremely wide variety of communication systems. Within this variety, we might also assume that at least some of these communication systems did not evolve as an adaptation for communication, but may have been co-opted as such. In this context, the uniqueness of human language starts to look more like a system unique only because human biology is unique within the animal kingdom. That is, the unique status of Language is dependent on the unique status of human biology—and in particular, neurobiology. As such, I-language as a natural object, just like any other object of nature, can be modeled by various formal or mathematical systems with varying degrees of success and relative to the specific goals of the research. However, I emphasize the point that natural language is not a *formal* object. Natural language is—at a specifically defined level of analysis—an object of nature that is subject to constraints, laws, and tendencies of the patterns of nature. Consequently, we should not expect any formal/mathematical system to exhaustively model natural language; but useful methods for mathematicization should always motivate us.

12.1.3 Mathematicization

In the rest of this chapter I will actually show how linguists go about deciding whether or not some mathematical tool or theory is relevant to natural language. However, I should point out now, as I will continue to do, that linguistics is its own field of study and has as its object of study a particularly unique object in all of nature. Arguments for the adoption of certain mathematical tools are not ar-

guments for the wholesale adoption of the mathematical enterprise to which those tools belong. The assumption here is that because of the uniqueness of human language AND because attempts to exhaustively reduce human language to known formal systems has failed, linguists are going to have to create their own kind of formality. To a large degree, this has already happened, but the formal system is by no means adequate and it needs to be mathematicized to whatever degree is possible. In the history of science this is not that strange—new mathematics have either been developed or resurrected to meet the demands of a new science. For example, Claude Shannon and others took George Boole’s binary algebraic logic—which had essentially been un-influential in the world of logic because Frege was able to construct a much more useful logic based on the use of quantifiers such as ‘All’ and ‘Some’—and applied to electronic circuits to get the rudimentary base of all computers. Another example includes Einstein’s use and elaboration on non-Euclidean geometries—which had been seen as an interesting but impractical geometry—for the exposition of his general theory of relativity. A last example is Benoit Mandelbrot’s fractal geometry—which was first viewed as a somewhat interesting, if not superficial, geometry that made cool looking computer images but had no real applications. Needless to say, applications of fractal geometry have already found wide acceptance from their use in shaping the antennae of cell phones and blackberries to the measurement of numerous natural phenomena such as lightening, trees and forests, coastlines, and even heartbeat rhythms. What such a new mathematics for linguistics would look like is not certain, if possible at all. But the reasoning that states “If Language is a natural object then it should be able to be modeled by a mathematicized system” is not very controversial. In fact, this could be stated as a corollary to what I have termed the Weak Minimalist Thesis in § refsmt:sec. What is controversial is the status of natural language as a natural object—but for the purposes of this book we have already agreed to assume that it is a natural object (for dissenting opinions see the general literature from cognitive or functionalist linguistics). In this way, then, it is not so much that Language cannot be modeled by a mathematicized system, but that a mathematicized system does not yet exist that could model Language. Also, notice that when I am speaking about a formal model for Language I use the term ‘mathematicized’ and not ‘mathematical.’ This implies that whatever rigorous system scientifically minded linguists find useful may not actually be a mathematical one by nature. If linguistic models seek to be appropriately scientific they need to be able to be translated into a mathematical system, that is, mathematicized. The issue of finding an ‘exhaustive’ model for language is probably more an issue of science in general—Is any scientific model exhaustive of the phenomena it represents? I hazard a guess that no scientific model can live up to these standards, but it is nice to try. By these standards, then, an exhaustive model of something so intricately connected to human consciousness and

cognition—not to mention things like free will—could probably never even be exhaustively **described**. But this is not the linguists job. The job of the linguist is to give the **what** and **why** of Language—to explain the abstract principles and parameters of the language faculty (the human ability to develop and use particular languages). The really interesting areas of languages, such as poetry and literature, are out of bounds. A strictly exhaustive model would have to include this out of bounds area and if we figure that a description covers less empirical ground than an explanation, then we have reason to believe that any goal of a strictly exhaustive model of human language is highly unlikely. On the other hand, an exhaustive model of parts of human language seems more reasonable—especially if we limit these parts to abstract principles and parameters of the human language faculty.

As it stands now, with recent formulations of syntactic theory and interest in notions of optimal form in nature—such as the Fibonacci pattern and types of symmetry—the conscious effort to view Language as a natural object has started seeing some immediate and interesting results. However, it needs to be stressed that all results are, or should be, viewed with a very critical eye. Many of the arguments are very abstract and so is the evidence used to justify such arguments. Nonetheless, if one is committed to a research program like the Minimalist Program and an idealistic approach to science like the Galilean approach, then the results one gets are suggestive of future progress. This progress includes, in fact demands, that one day we will find a way to empirically test all of our hypotheses. The rest of what follows in this chapter is based on [Bowles \(2009\)](#) and represents results along the lines just discussed. They are speculative, abstract, and do model any actual natural language data. Instead, they investigate the possible mathematical structure of the operation Merge in syntax. For those of you who do not like math feel free to skip the rest of the chapter. Another reason for skipping the rest of this chapter is that you do not know anything about theoretical syntax. While lack of knowledge should never hold someone back from trying to gain knowledge, sometimes it can be a frustrating experience when the lack of appropriate educational background and training gets in the way.

12.2 On With the Formalities

Intense research into Fibonacci (F) patterns arising from ideal iterations in both prosody and phrase structures has shown that there is a special role for rigorously modeling the computability requirements of producing such patterns. F patterns produce self-similar structures along repeated and computable iterations. The computable iterations of the F sequence can in turn produce idealized phrase structures, as [Medeiros \(2008\)](#) and [Soschen \(2008\)](#) have shown. These idealized phrase struc-

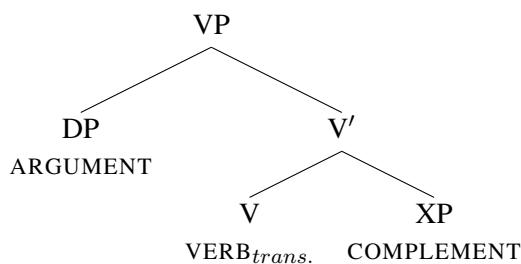


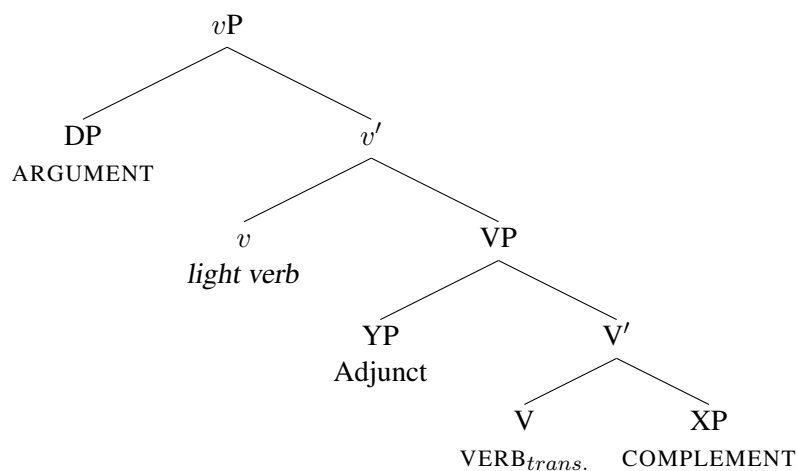
Figure 12.5: Simple argument structure

tures also map projections of phase domains that have implicational patterns, as [Piattelli-Palmarini and Uriagereka \(2008\)](#) discuss. Because F sequences are computable, iterated, self-similar generations they also produce symmetrical forms. If one takes the F sequence of phrase structures and compares two different levels, head (X^0) and phrase (XP), a ratio can be derived that shows symmetry between them; I call this the X-ratio in § 12.3.2. In the context of these developments it is only natural to look towards the formal study of symmetry for possible rigor.

It is well-understood that in any system that produces hierarchical embedding the key notion is not symmetry itself, but symmetry-breaking. In linguistic terms, a computable operation Merge will combine two objects $\{\alpha\}$, $\{\beta\}$ to produce a mirror-symmetric object if not restricted in any way. That is, no matter how you flip the two branches of the tree along a vertical axis, the relation will be the same—with the root node label unspecified. Such an object is ambiguous to the extent that the root node label is ambiguous or unspecified. When no specification for the root node label is given then no hierarchy can be established between the two syntactic objects, as [Boeckx \(2008\)](#) and [Chomsky \(1995c\)](#) discuss. If we iterate this process, unrestricted, then the projection¹ of label to the root node will be random, i.e., there is no deciding information for which of the two labels, α or β , projects. Real world syntax necessitates very specific projections—whether they are simple labels, lexical feature bundles, Probe with valued uF , or some other property.² In other words, empirical data show overwhelming evidence that syntactic relations between constituents are not altogether symmetrical and that hierarchical (asymmetric or antisymmetric) relations are a simple fact of human language; see Figures 12.1 to 12.6. The reason for focusing on symmetry, then, is to have some criteria by which to measure symmetry-breaking.

¹This is not the projection of mathematics.

²Of course, adjuncts and adjunction present problems of a different sort: such as a possible immediate asymmetry of Merge, represented by the ordered pair set notation $\langle a, b \rangle$ in contrast to typical set notation $\{a, b\}$; see [Boeckx \(2008\)](#), [Chametzky \(2000\)](#) [Hornstein and Nunes \(2008\)](#), and [Rubin \(2002, 2003\)](#) for examples.

Figure 12.6: Complex argument structure with *vP* shell

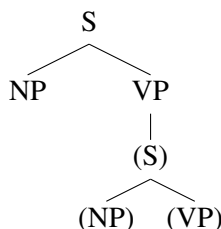
12.3 α Merge and symmetry

12.3.1 Ambiguous Merge

Given the naturalistic and rigorous framework for linguistic study briefly discussed in § ??, it is reasonable to experiment with the use of different kinds of rigor to explain Language. The search and justification for rigor in generative linguistics has a key issue since the beginning. Tomalin has an excellent introduction to the formal foundations of generative theories in his book *Formal Foundations of Generative Grammar*. In it he relates the following episode.

In 1954 Yehoshua Bar-Hillel published a paper, [Bar-Hillel \(1954\)](#) in the journal *Language* about the importance to linguistics of the then current practices of logical syntax and logical semantics. This paper was published within a growing context of linguistic interest in formal/logical methods applied to natural language data. A short while later the up-and-coming linguist philosopher Noam Chomsky published a response to Bar-Hillel in the same journal, [Chomsky \(1955\)](#). In this article Chomsky expressed doubts about the ability of formal systems to model language behavior. He particularly expressed doubt about the usefulness of certain kinds of *recursive definitions* that Bar-Hillel mentioned, particularly in reference to [Carnap \(1937\)](#).

(7) $S \rightarrow NP VP (S)$



- (8) a. $NP \rightarrow N (Det) (NP) (PP)$
 b. $PP \rightarrow P (NP)$

However, only a few years later Chomsky would make use of recursive definitions for defining hierarchical relationships of constituents. Examples can be found in (7) and (8).³ Such informal definitions by “recursion” proved to be problematic for describing the general environments of unique linguistic constituents—to the effect that they were understood to be too strong. Unique constituent definitions were quickly generalized under the X-bar schema first proposed in Chomsky (1970) and later extended by Jackendoff (1977). Years later, the X-bar model was modified and streamlined by Chomsky (1994) and Chomsky (1995a) in the form of a Bare Phrase Structure (BPS). Since that time, many proposals for a hyper-minimalist BPS have been made on the basis of both theoretical and empirical arguments; for example Carnie (2000), Citko (2005), Collins (2001), and Jayaseelan (2008) to name a few.

- (9) X-bar generalization for constituents
 a. $X'' \rightarrow \text{Specifier}, X'$
 b. $X' \rightarrow X, (\text{Complement})$

Such that:

- (i). X ranges over N, V, A, P, and S.
 (ii). Endocentricity applies:
 (ii'). X' is the head of X'' .
 (ii''). X is the head of X' .
 (iii). Heads share categorial properties with their projections.
 (iv). Complements are X'' categories.

In BPS, the basic operation Merge *Selects* two syntactic constituents or objects and “merges” them together, giving as its output one constituent or object. The

³Chomsky (1957) would come to recognize, specifically in light of analysis of finite-state grammars expressed in the form of Markov machines, that “If a grammar does not have recursive devices... it will be prohibitively complex. If it does have *recursive devices of some sort*, it will produce an infinitely many sentences” (Chomsky 1957:24, italics mine).

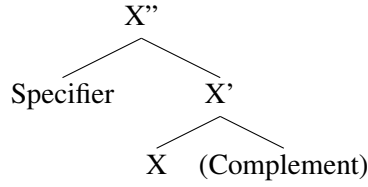


Figure 12.7: First X-bar structure

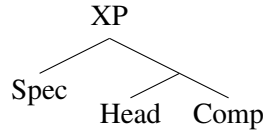


Figure 12.8: General BPS X-bar model

operation itself does not specify which of the two constituent labels projects—projection is assumed to be specified by the features of the objects being merged. For the purposes in this paper, no specification for projection will be given. It is reasonable to call this basic operation “ambiguous,” following Boeckx (2008), and I will denote its operation as αMerge .

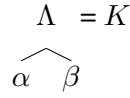
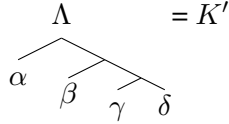
Definition 1. $\{\{\alpha\}, \{\beta\}\} \xrightarrow{\text{merge}} \{\Lambda, \{\alpha, \beta\}\} = K = \alpha\text{Merge}$ ⁴

The one thing that all, or most, current BPS models share in common is the general operation Merge defined as a physically computable operation. It has been hypothesized by Hauser et al. (2002) and Fitch et al. (2005) that Merge is operable in the domain of the Narrow Faculty of Language (FLN), specifically Narrow Syntax (NS), and is a unique property of human cognition.⁵ In tree form, αMerge produces a mirror symmetric object, seen in Figure 12.9.

With mirror-symmetric objects each side is the same in terms of their relative position and relation to the other side. Figure 12.9 has a bilateral mirror-symmetry: rotating the object along a vertical axis 180° or 360° will not change it. Under

⁴Where Λ = exclusively the label either $\{\alpha\}$ or $\{\beta\}$. Also, where α and β are two independent syntactic objects, K is the new object output of Merge: so $\{\alpha, \beta\} \in K$, but only $\{\{\alpha, \beta\}, \Lambda\} \subseteq K$; see Chomsky (1995c).

⁵It should be noted that αMerge , as presented here, may be operable in non-human cognition. See Pinker and Jackendoff (2005) for more on this, as well as arguments against Hauser et al. (2002). Also note, αMerge as presented here does not equate directly to adjunction defined in Boeckx (2008), Chapter 3, as a time-dependent insertion $\{\{\alpha\}_t, \{\beta\}_{t+1}\}$ with an inherent temporal asymmetry. Although, there doesn't appear to be any major incompatibility; especially in light of the fact that my narrowed treatment of Merge is fairly abstract and mathematical in order to get at a group theory definition of symmetry from which more realistic and empirically based definitions of symmetry-breaking can be applied. See § 12.4.3 for more detail.

Figure 12.9: α Merge tree produces object K Figure 12.10: Iterated α Merge tree produces object K'

these conditions it is understood that α Merge is an idealized operation and K is an idealized object. This operation can be iterated infinitely, where each label-node is ambiguous. A finite representation of the iteration is given in Figure 12.10—notice that K' is a compound of two mirror-symmetric trees but itself has no bilateral symmetry.⁶ Once the bilaterally symmetric object of Figure 12.9 is repeated to produce a more complicated object it loses its simple mirror-reflection property. By these standards, even if one wanted to argue that α Merge was necessary *and* sufficient to explain why human syntax is the way it is, some notion of symmetry-breaking must be introduced to account for strings containing more than two elements. For example, in Figure 12.10, $\{\alpha, \beta\}$ dominates the mirror-symmetric constituent $\{\gamma, \delta\}$, even though the label and/or category of $\{\gamma, \delta\}$ is ambiguous. Interestingly, once α Merge projects a label—no matter which label is, for the purposes of this paper, randomly projected from minimal/head level to maximal/phrase level—the number of phrase and head level constituents accumulated as the result of iterated Merge operations correlates with the Fibonacci pattern.

12.3.2 Fib(n) levels and the X-ratio

Rationale and justification

The idealized objects produced in Figures 12.9 and 12.10 are problematic because they do not allow *any* specification for label projection. The non-projective model of unrestricted Merge, or α Merge as I am calling it, does not seem to produce any interesting patterns suggestive of natural law. But when one allows α Merge to

⁶In fact, the tree in Figure 12.10 could be decomposed into $\{\alpha, \beta\}$, $\{\beta, \gamma\}$ and $\{\gamma, \delta\}$ with 3 ambiguous label slots, including the root node label, and 3×2 possible labels. Generalizing, for n syntactic objects, there are $(n-1)$ merge operations, or $(n-1)$ label slots, and $((n-1) \times 2)$ possible labels. Compare this to the observation in Medeiros (2008) of the relation between *string length* of n and *depth* of $(n-1)$.

project a random choice of label the result is an interesting mathematical F pattern, as [Medeiros \(2008\)](#) and [Soschen \(2008\)](#) have shown; see also [Boeckx et al. \(2006\)](#), [Carnie and Medeiros \(2005\)](#), [Idsardi \(2008\)](#), and [Piattelli-Palmarini and Uriagereka \(2008\)](#). However, Medeiros (2008:164) points out that

It is tempting to see the appearance of the Fibonacci sequence in the X-bar pattern as being deeply significant in itself. But the X-bar schema is after all a very simple mathematical object, and there may be nothing particularly magical about the appearance of the Fibonacci sequence in the structures it generates. Their appearance in this domain could be no more of a surprise than their appearance in the family trees of bees, or in Fibonacci's idealized rabbit populations, or in the number of metrical possibilities for a line of Sanskrit poetry, or any of the myriad situations these numbers describe. To put it another way, it could be that these properties are an accident of no real significance, or worse, merely a reflection of mathematical simplicity in linguists' description of language, rather than a property of language itself. Yet it is undeniable that patterns related to the Fibonacci sequence play an important role in nature, especially in optimal packing and optimal arboration.

Additionally, [Livio \(2002\)](#) provides analysis for many of the supposed F patterns that have been consciously encoded into art, architecture, and music, concluding that in many of the cases the investigators have been overzealous and mistaken—he even goes so far as to show how the dimensions of his television can be made to fit within the golden mean. Nonetheless, [Medeiros \(2008\)](#) is right, F patterns are undeniable and their investigation in human syntax is significant—if approached soberly.

The golden sequence (GS)

The golden sequence (GS) is derived by a simple algorithm that correlates to the F sequence; it is given in [Algorithm 1](#). A phrase structure equivalent is given in [Algorithm 2](#). The outputs of these algorithms are produced in [Appendix A](#) as [Table ??](#) and [Table ??](#)—the latter of which generates the tree in [Figure 12.14](#); see [Uriagereka \(1998\)](#) and [Piattelli-Palmarini and Uriagereka \(2008\)](#) for linguistic applications of the GS. In the algorithms and their outputs no attention has been given to bar-level (X') projections; only head (X^0) and phrase (XP) levels are of immediate concern—though [Medeiros \(2008\)](#) and [Soschen \(2008\)](#) include bar-levels. Lastly, the F level refers to the number of syntactic objects (heads and phrases), such that the number of syntactic objects follows the pattern of the F sequence.

Algorithm 1. GOLDEN SEQUENCE:

Start with 1, replace every 1 with 10, and every 0 with 1 in every iterated line.

Algorithm 2. GS FOR PHRASE STRUCTURE:

Start with GS; each 1 = Phrase ($XP = \alpha$); each 0 = Head (X^0).

Importantly, any specific parsing of elements by brackets in Table ?? is not determinate and can be changed. What is important is the fractional relation of heads to phrases, not how these head-phrases relations are parsed.⁷ In fact, the parsing, equivalent to the issue of determining the projecting head, is basically random in this approach. Also, it should be obvious that the name of the label doesn't matter and can be changed.

The GS applied to phrase structure embedding gives similar results to Medeiros (2008), who maps explicit Fibonacci correlates of phrase structure in his approach. His examples (1-5, 2008:153-154) are numbered Phrase Strings 1 to 5 in Appendix B. These Medeiros strings have equivalents when an alternative parsing of the strings from Table ?? is applied—these latter equivalents are numbered Phrase Strings 6 to 10 in Appendix B. Consequently, one can see that the two approaches converge on equivalent results—with obvious and trivial differences in label names; see also Figure 12.13.

By bracketing the phrase string produced from Fibonacci level 34 (F_{34}) differently than it is parsed in Table ??, a fully balanced binary tree, like the one in Medeiros (2008) and Soschen (2008) can be derived. Figure 12.11 is an example of a Medeiros tree and Figure 12.12 shows the same tree without specifiers. The Phrase Strings 11 and 12 represent the latter trees, respectively; 11 for Figure 12.11 and 12 for Figure 12.12. Phrase String 13 is derived from the GS algorithm for phrase structure and has been re-bracketed in order to correlate with the previous objects; it represents the embedding for Figure 12.13. One can easily see that by removing the X' projections from 11, both 12 and 13 are equivalent—and clearly, so are their tree projections.⁸

In Figure 12.14 I produce a binary right-branching alternative to Medeiros' balanced binary tree.⁹ This right-branching alternative is based on a well-known property of F patterns called the Golden Sequence; for example Livio (2002), but

⁷The F pattern is special because it codes a specific irrational value when two sequential numbers are put into a fraction such that the larger is the numerator; for example 13/8. As the numeral values get larger, and hence the fractions contain larger numeral values, their ratio gets closer to irrational limit. But this is an infinite limit, usually referred to as τ , which has the value 1.618033...

⁸Potential problems are associated not with the actual re-bracketing, which is not a problem under the assumption of virtually random projections, but of the redistribution of constituents to form the balanced binary tree of Figure 12.13; compare the right-branching tree of Figure 12.14 derived from a straightforward sequential embedding of F_{34} .

⁹Also work by Alana Soschen, for example Soschen (2008).

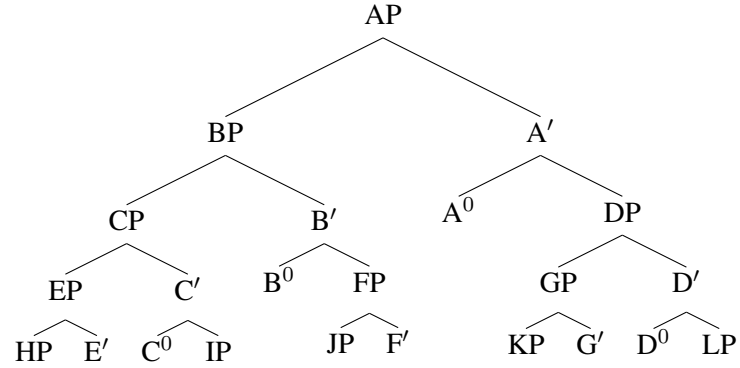


Figure 12.11: Medeiros tree

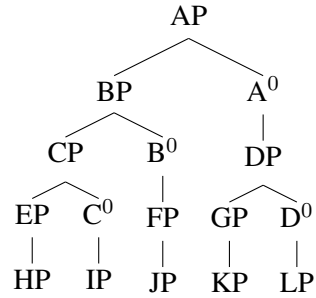
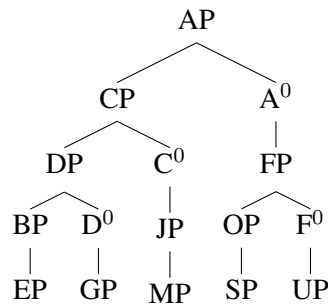
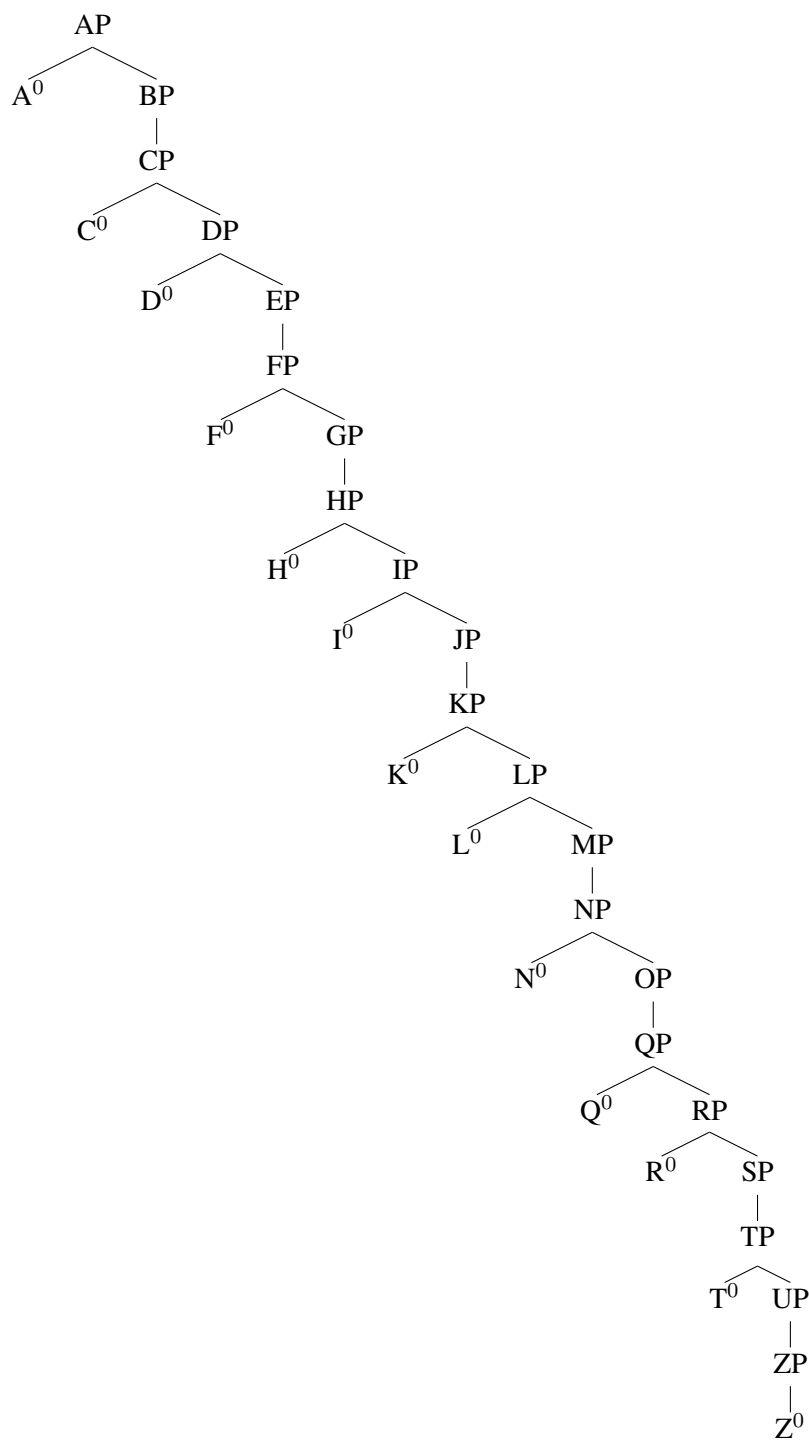


Figure 12.12: Medeiros tree with no Spec

Figure 12.13: Randomly projecting α Merge tree adapted to Medeiros string

Figure 12.14: Random label projection of α Merge

especially Uriagereka (1998). It has the same F pattern for syntactic objects that derivations in Medeiros (2008) do, and appears to display the same relations between maximal and minimal levels; though some redistribution of constituents is needed to reformulate the tree in Figure 12.14 as the Medeiros equivalent in Figure 12.13. Redistributing constituents in such abstract and idealized models—where label projections are random and no interface conditions or features play a role—does not seem to pose a problem. The placement and order of constituents—basic c-command relations—depends on the order in which they are merged, but in the abstract models used here it does not matter if, for example, BP is merged before AP or I^0 is merged before DP. In other words, if the label projection is random then so is the operation *Select*, which selects items to be merged. What is really at issue here is the fractional relation between the number of constituents at phrase and head levels. Such fractional relations appear to reveal, at a high level of abstraction and idealization, a sequence or pattern that is commonly found in nature: the F pattern/sequence. What is interesting about Medeiros’ treatment of the F pattern and its relation to idealized phrase structure patterns is that—abstracting even more—a $\text{Fib}(n)$ ratio between different comparative levels of phrase structure in a binary tree can be derived. For example, all maximal projections follow an initial sequence of the number pattern, $\text{Fib}(n)$, while terminal heads begin the pattern two levels ‘down’ at $\text{Fib}(n-2)$; (Medeiros 2008: 161). In other words, for maximal projections the sequence begins as shown in (10), but for terminal heads the sequence begins as shown in (11) (again, unlike Medeiros, I ignore intermediate (X') levels here and throughout). It does not matter how one derives the F pattern—whether by the golden sequence or other means—it still produces this ratio.

(10) Fib sequence for XP: [1, 1, 2, 3, 5, ...]

(11) Fib sequence for X^0 : [0, 0, 1, 1, 2, 3, 5, ...]

By this relationship between levels, then, the ideal ratio between maximal and minimal projections is $\text{Fib}(n : n-2)$ at any horizontal level—or between sisters—in a binary tree. For example, the F^3 string in Table ?? has one X^0 and two XPs (A^0 , α , and β) such that $2/1 = 2$. Looking at F^{34} , the ratio of XP to X^0 is $21/13 = 1.615384\dots$, an approximate to the limit ratio of the F sequence—usually denoted by ($\tau = 1.6180339\dots$)—for which the ratio of maxima and minima of projections in binary phrase structure trees would get closer as the number of projections reached the limit of infinity within the F pattern. I will refer to the ratio $\text{Fib}(n : n-2)$, defined over maximal and minimal projections in phrase structure, as the X-Ratio; see Carnie and Medeiros (2005), Boeckx et al. (2006), Medeiros (2008), and Soschen (2008). It is this ratio in Definition 2 that leads to symmetry and Group Theory.

Definition 2. X-RATIO:

The ideal distributional ratio of maximal projections to minimal projections is $\text{Fib}(n : n-2)$.

Symmetry in syntax

Group Theory is the formal application of study for the investigation of symmetry in nature; see [Livio \(2005\)](#) and [Stewart \(2007\)](#) for popularized accounts and [Rosen \(1995\)](#) and [Milne \(2008\)](#) for formal introductions—also Chapter 10 in [Partee et al. \(1990\)](#). Once the X-ratio has been established, then the symmetry between maxima and minima of constituents, i.e., between phrases and heads, can be defined. By symmetry, I refer to the conjunctive notion defined informally by [Rosen \(1995\)](#).

Definition 3. SYMMETRY:

(i) Possibility of change: It must be possible to perform a change, although the change does not actually have to be performed. (ii) Immunity: Some aspect of the situation would remain unchanged, if the change were performed.

Under this view of symmetry, XP and X^0 are symmetric because the interchange between them of the arithmetical operation $(n \pm 2)$ on the Fibonacci sequence defining them does not change their underlying Fibonacci sequence—it only shifts the X-bar level at which the sequence begins. We might call this a *translational symmetry* such that there is an immunity to change dependent on a certain displacement or shift along a specific boundary. Such *translational* symmetries produce repeating patterns. Another way to describe the X-ratio would be as a *level symmetry*: XP and X^0 are immune to change—in terms of their Fibonacci patterns—relative to minimal or maximal level; i.e., relative to their head or phrase level. The difference, then, between maximal and minimal projections in binary trees is $\text{Fib}(n \pm 2)$, which determines the X-bar level at which the F pattern begins, as in (10) and (11). This is generally consistent with the empirical and theoretical claims of the hyper-minimalist BPS of, for example, [Carnie \(2000\)](#), where lexical features can be projected as XP or X^0 depending on specifics of the syntax of predication.

However, the symmetry noted here is not the symmetry of αMerge in § 12.3.1. The X-ratio suggests a symmetry developed from constraints that determine the label projection; albeit the constraints used here are practically random. Needless to say, empirical data is more complicated and provides non-random constraints for the specification of projection. Importantly, the symmetries noted here and in § 12.3.1 do share one thing in common: they are both symmetries of syntactic objects and not of, say, syntactic rules of the sort discussed in [Boeckx \(2008\)](#).¹⁰ Thus,

¹⁰For example, the “Merge is source-independent” rule which applies to a symmetry between

heads and phrases are symmetrical in an important way that differs from traditional notions of symmetry in syntax.¹¹ The current status of the notion of symmetry in syntax and prosody—based on empirical and theoretical observations—is suggestive of possible new directions and certainly justifies a consideration of GT and anything this particular formalism might offer as useful tools to the linguist. I suspect that what is developing in current theoretical research is the groundwork for a potential discovery of a general principle (or group of principles) from which head and phrase levels can be derived. That is, head and phrase levels no longer need to be thought of as syntactic or linguistic primitives. Instead, they may be products of a general principle(s) of phrase structure based on some first principle(s) of natural law for efficient self-organization.¹² These laws or first principles may be coupled with a combination of interface constraints and/or syntactic operations, such as Agree in Boeckx (2008), that essentially work to break the initial symmetry of ambiguously merged elements or elements that have the potential for head-phrase (X-ratio) ambiguity.¹³

Of some immediate relevance here is the role of the formalisms of GT—the mathematical system used to study symmetry. It may not be necessary to incorporate much of the formal mechanism of GT, certainly not all of it, just as it is not necessary to incorporate most or all of Computability Theory simply because the notion of computability is crucial to theoretical syntax. However, having said this, it is at the very least interesting to flesh out some basic observations of symmetry in syntax in terms of the GT formalism; but I will not go into great detail.

A conceptual rationalization for GT and symmetry

The benefit of GT is that it can model practically any kind of real world object (animate bodies, the Rubik's cube) and it can model abstract laws or rules. Boeckx (2008) has shown how a rotational symmetry can be used to model symmetrical

External and Internal Merge such that under the operation *Select*, both types of Merge are immune to change; i.e., Merge is the same operation in all contexts. As a sidenote, Boeckx (2008) highlights how this observation can be traced back to the *Structure Preservation Hypothesis* of Emonds. This seems like a significant advance in syntactic theory.

¹¹Such as subject-object asymmetries.

¹²This kind of thinking meshes very well with Baker (2003) and his attempt to derive category information from more basic syntactic structural relations. Baker defines two such relations in the following way: *X is a Verb (defined in 1) or a Noun (defined in 2) if and only if X is a lexical category and X*

1. has a specifier.
2. bears a referential index, expressed as an ordered pair of integers.

¹³For example, clitics or nominal predicates.

relations between projections and chains (pg. 46), see also Uriagereka (1998), and he has also modeled symmetry relations between specific linguistic rules to point out the essential fractal (i.e., self-repeating patterns at different levels of scale) nature of syntax (pg. 159); see also Boeckx and Uriagereka (2007) for discussion of fractals in nature and the Minimalist Program. Medeiros (2008), Soschen (2008), Idsardi (2008), and Piattelli-Palmarini and Uriagereka (2008) have pointed out relevant patterns of F sequences in idealistic models of prosody, X-bar structure, and phase alternations. F sequences can be defined as a type of self-similar iteration that has a fractional limit of the value τ ($= 1.618033\dots$).¹⁴ This value is also the ratio responsible for making symmetrical spirals and other types of symmetrical objects. If one takes seriously the notion that the human ability to acquire and use a linguistic system—a system unique in the animal kingdom—is really more of a reflex to a natural process based on a “growth form,” then the idea that symmetry (and more importantly symmetry-breaking) is a necessary property of such a “growth form” is a tantalizing possibility. Symmetry is one of the most profound properties of nature—and in fact, it is the slight deviation from symmetry that is so profound. It means that multiple types of natural objects constructed of multiple kinds of material are at some abstract level following a similar blueprint. From spiral galaxies to human faces, this abstract blueprint is a defining characteristic of the laws of nature (not to mention fractals, which symmetrical forms can be related to). There is no reason to deny that, at an appropriate scale of measurement and abstraction, the “growth form” of the human language system must reveal the property of symmetry—and in fact does show deviance from symmetry in its real-world form.

12.4 Strong Minimalist Thesis

This section covers three areas that form part of the backdrop for evaluating directions and formalisms that the MP can utilize—they are indirectly related to the “three factors” of Chomsky (2005). The first area deals with the biological nature of natural language. This means that although operations like Merge can be defined by rigorous formal and mathematical tools, Merge itself is not a formal object. Furthermore, the fact that an interesting mathematical pattern found throughout nature can also be found in prosodic and phrase structures does not mean that these structures are themselves mathematical or formal objects. (In fact, it does not even mean that the mathematical patterns are *inherent* to prosody or X-bar; the pattern

¹⁴When one makes a fraction of any two sequential values in the F sequence, say $[\dots, 5, 8, \dots]$ as $8/5 = 1.6$ then take the next two values $[\dots, 8, 13, \dots]$ as $13/8 = 1.625$, one always gets closer to the limit value of τ but can never reach the limit because it is infinite.

may simply be a matter of our perception). The Strong Minimalist Thesis (SMT) proposes a general methodological backdrop for this and states, generally, that language is an optimal form meeting interface conditions. This implies that natural language is not a formal object, but an object of nature. The next area deals with the general notion that we expect, in leui of strong contrary empirical evidence, operations in language to be in fact *operations that are computable*, i.e., specifiable by an algorithm that can be discovered or constructed. These computable operations must be physically tractable. That is, operations are bounded by human processing limits. Therefore, MP theories about syntax must meet certain natural limits. The third area is about the nature and concept of computablity itself and is intimately related to the second area. It deals with what linguists mean by the terms “computable” and “recursive.” The following subsections briefly outline these areas in order to give an explicit context for evaluating the usefulness of GT and formal notions of symmetry in syntax.

12.4.1 Merge is not a formal object

Natural language is not a formal object. In this regard, neither is the operation Merge—though it, and its historical antecedents, are regularly defined through formal mathematical means. As early as Chomsky (1955) and Chomsky (1957) the status of natural language as a formal object has been viewed critically within the context of twentieth-century advances in logical sytnax and semantics of the type found in Carnap (1937), Ayer (1936), Reichenbach (1947), Quine (1953), and Tarski (1956). This does not mean that useful approaches from formal theories cannot be used productively, only that natural language probably cannot be modeled *exhaustively* by any formal/mathematical system that yet exists—the implication here is that it is up to linguists to construct such a system from the formalisms that do exist, coupled with empirical evidence from natural langauges; for example Langendoen and Postal (1984), Partee et al. (1990), Hornstein (1992), Prince and Smolensky (1993), Uriagereka (1998), Chierchia and McConnell-Ginet (2000), Tesar and Smolensky (2000), Bod et al. (2003), Niyogi (2006), and of particular relevance here Fortuny and Corominas-Mutra (to appear). No matter the formal methods that are drawn on (usually set theory in semantics and syntax and probability theory in phonology and phonetics) it has generally been understood that, as Fortuny and Corominas-Mutra (to appear) state,

As any abstract theoretical background, it is not reasonable to ask about the reality of the operations and objects defined. For example, although the algorithm runs through a time step indicator, such time step is only given for operational purposes and does not imply—in our

field of study—any temporal evolution. What is reasonable to ask is whether from the defined mathematical framework we can derive the core properties that we observe in the studied object (pp. 7-8).

Chomsky (1955) expressed a similar attitude in response to arguments by Bar-Hillel (1954) for the use of formal models in natural language—specifically the use of recursive definitions in a formal syntax of the type found in Carnap (1937):¹⁵

[t]he correct way to use the insights and techniques of logic is in formulating a general theory of linguistic structure. But this fact does not tell us what sort of systems form the subject matter for linguistics, or how the linguist may find it profitable to describe them. *To apply logic in constructing a clear and rigorous theory is different from expecting logic or any other formal system to be a model for linguistic behavior* (pg. 45; italics mine).

Additionally, a realistic theory of linguistics that takes natural language as a scientific object of inquiry must subject the brain to the known laws of biology, physics, and nature in general. In computational complexity theory, for example Deutsch (1985), Deutsch (1997), Lloyd (2000), Lloyd (2006), and also Kauffman (1995), all physical systems are understood to be computable and subject to the known laws of physics and computation, specifically entropy (the second law of thermodynamics) and universal computability (the Church-Turing Thesis). By normative scientific standards the brain, and by implication FL, cannot contradict laws of nature. I explicitly state this as a corollary of the *weak* minimalist thesis, which is an uncontroversial weakening of the strong minimalist thesis proposed in Chomsky (2008).¹⁶

Strong Minimalist Thesis. Human language is an optimal solution to interface conditions that the faculty of language (FL) must satisfy.

Proposition 1. *Weak Minimalist Thesis:* *Human language is the natural product of linking sound (or gesture) and meaning.*

Corollary to Weak Minimalist Thesis. Human language is a product of the natural state of the organism-as-a-whole, and therefore, is subject to the same laws and constants that all of nature is.

¹⁵Interestingly, as Tomalin (2006) points out, Bar-Hillel's linguistic work served as the foundation for a specific approach to natural language, Combinatorial Cateogrical Grammar (CCG), used in computational linguistics; for example Clark and Curran (2007).

¹⁶The SMT is widely regarded as not being plausible—though possible—and is thus used as a heuristic standard. For this reason, it seems strange to derive a corollary from a heuristic. This is the motivation for putting forward a weakened version of the SMT.

Notice that the Corollary to the WMT does not say that human language, or more precisely FL, is explained by or reduced to all natural laws; nor does it say that all natural laws apply to FL—only that FL is subject to the laws and constants of nature (and by implication only a subset of natural laws). The Corollary is simply a compression of arguments found in many sources, including [Boeckx and Piattelli-Palmarini \(2005\)](#), [Chomsky \(1986, 1995b,c\)](#), [Piattelli-Palmarini and Uriagereka \(2008\)](#), [Uriagereka \(1998\)](#) and many others past and present; see also [Formigari \(2004\)](#) for historical consideration of issues relevant to the Corollary—particularly historical tensions between assigning natural language a “spiritual” or “bestial” cause relative to “immaterial” and “material” effects.

Assumptions and the Church-Turing Principle

Assume the faculty of language (FL) is a finitely realizable physical system with an initial state Σ_{UG} of Universal Grammar followed by subsequent computational states (C_1, \dots, C_n) . These states have a finite limit simply because they emerge from and are ‘computed’ by a finitely organic brain—a product of biological nature. But a finite realistic limit does not negate the notion of *discrete infinity*. The latter is a property of the computability of FL, not its real-world limit. In theory the initial state of FL, Σ_{UG} , has the computable potential of reaching infinity, Σ_∞ , but because of the inherent constraints of natural law applied to biological systems no real (i.e., physically computable) FL actually reaches the limit. Further assume, as any physical system that can be realized by computable functions, the FL can be modeled by a universal Turing machine, (uTM), as determined by the informal definition of the Church-Turing Thesis (CTT) given by [Deutsch \(1985\)](#). However, [Deutsch \(1985\)](#) rejects the CTT as too vague and opts for a redefinition he calls the Church-Turing Principle (CTP). Assume the CTP applies to the MP.

Church-Turing Thesis. Every function which would naturally be regarded as computable can be computed by the universal Turing machine.

Church-Turing Principle. Every finitely realizable physical system can be perfectly simulated by a universal model computing machine operating by finite means.

[Deutsch \(1985\)](#) elaborates:

I propose to reinterpret Turing’s “function which would naturally be regarded as computable” as the functions which may in principle be computed by a real physical system. For it would surely be hard to regard a function ‘naturally’ as computable if it could not be computed in Nature, and conversely. To this end I shall define the notion of ‘perfect simulation’. A computing machine M is capable of perfectly

simulating a physical system S , under a given labelling of their inputs and outputs, if there exists a program $\pi(S)$ for M that renders M computationally equivalent to S under that labelling. In other words, $\pi(S)$ converts M into a 'black box' functionally indistinguishable from S (pg. 100)

A syntax utilizing the formalization of a u TM in the linguistic domain has been extremely successful, especially if viewed from the perspective of the CTP. According to the MP, the physical system S of the CTP is the brain—and there is an expectation that the computability of it be physically tractable, at least in terms of the computability of generating hierarchically nested structures that can be iterated (in)finately to produce various forms (i.e., discrete infinity). For example, [Chomsky \(2005\)](#) states the third factor of language design as “Principles not specific to the faculty of language. . . including principles of efficient computation” (pg. 6). Here, efficiency is relative to the naturally finite biological limits of the human brain. There is a strong (conceptual) parallelism between Deutsch’s redefining CTT as the physically bounded CTP and MP interests in the physical tractability of computable operations like Merge. Of course, David Deutsch employed his redefinition in terms of quantum phenomena and the computability of the universe—theoretical linguist’s concerns for physical computability are dramatically narrower and more deterministic. However, this should not restrain linguistic interest in the physical computability of nature as long as one can see that such parallels here are, at first glance, simply analogous, conceptual, and intended to inspire—not direct a path of research.

12.4.2 Normative use of the term “computability”

Although the normative use of term “computability” may seem like a mundane topic, it in fact strikes to the heart of the issue of what linguists mean when they use the term “recursive.”¹⁷ [Soare \(1996\)](#), and more recently, [Soare \(2007, 2008\)](#),

¹⁷One example of “recursive” can be found in [Partee et al. \(1990\)](#), who show that a recursive definition can be gotten by considering

the set M of all mirror-image string on $\{a,b\}$. A mirror-image string is one that can be divided into halves, the right half consisting of the same sequence of symbols as the left half but in the reverse order. For example, *aaaa*, *abba*, *babbab*, and *bbabbabb*. . . . The following is a possible recursive definition of M .

1. $aa \in M$ & $bb \in M$
2. $(\forall x)(x \in M \rightarrow (axa \in M \text{ \& } bxb \in M))$
3. M contains nothing but those members it has by virtue of lines 1 and 2 (pp. 181-82).

have pointed out a historical confusion in the use of the two terms that may lead to unfortunate misunderstandings relative to the field of study and the definitions assumed. He points out that there are both technical and now normative differences between the terms (i) “(primitive) recursive function” and (ii) “recursively enumerable” (r.e.). The latter term is closer to what linguists know as “recursion,” and as Soare suggests, there are important reasons for clearly delimiting the difference in terminology. Substituting (ii) with the term “computable” and making it clearly different from “recursion” will have positive results in the sciences. Soare (1996) states that anyone using the term “recursive” always risks the chance that a mathematician or computer scientist will confuse it with induction: “when we use ‘recursive’ to mean ‘computable,’ we are using it in a way that is not in any dictionary and which an educated scholar or scientist cannot reasonably be expected to know. Indeed, there is a danger that a computer scientist or mathematician might mistake it for ‘inductive’” (pg. 34). I use the term Turing computable (TC) or “computability” to designate r.e., or the linguistic notion of “recursion.” Additionally, while discussing the fact that Emil Post’s work focused directly on computably enumerable sets (which are usually called recursively enumerable sets) instead of computable functions—as found in the work of Church (1936) and Turing (1936), although Post (1936)—Soare (2008) says that

concentration on c.e. [computably enumerable] sets rather than partial computable functions may be even more fundamental than the thesis of Church and Turing characterizing computable functions because... often in higher computability theory it is more convenient to

Line 1 is the base, line 2 is the recursion step, and line 3 is the restriction. They go on to define the Principle of Mathematical Induction:

For any predicate Q , if the following statements [1 and 2] are both true of Q then the following statement [3] is also true of Q :

1. $Q0$
2. $(\forall x)(Qx \rightarrow Q(S(x)))$
3. $(\forall x)Qx$
(pg. 196)

Where $S(x)$ is the successor function—denotes the successor of (x) —and $Q0$ denotes the property Q of zero. They note that the similarity between lines 1 and 2 in both demonstrations is “readily apparent.” But also say that the “Principle of Mathematical Induction is not a definition, however, but a rule of inference to be applied to statements about the integers” (pg. 196). Whatever similarities exist, recursion and induction as given here are not, arguably, equivalent in the strong sense. Furthermore, even if they are equivalent in the weakest sense of the term, the applications to which mathematical induction is relevant are not the kinds of applications relevant to the physically tractable computability concerns of the linguist.

take the notion of a generalized c.e. set as basic and to derive generalized computable functions as those whose graphs are generalized computably enumerable (pg. 25).

Recall that [Chomsky \(1965\)](#) pg. 9, partly attributes the use of the term “generative” to the work of Emil Post (Chomsky is defending the normative and technical use of “generative”). A relevant quote can be taken from [Post \(1944\)](#), also in [Soare \(2008\)](#): “every *generated set* is effectively enumerable, every effectively enumerable set of positive integers is recursively enumerable” (pg. 286; italics mine). Chomsky attributes part of the motivation for the term “generative” to Post’s work on—what Soare argues should now be called—computably enumerable sets. That is, “generative” seems to refer more to functions defined by (or sets enumerated by) Turing machines, algorithms, and concepts generally associated with computation than to definition by induction, general recursive functions defined by Herbrand-Gödel, or fixed points in the Kleene Recursion Theorem. These latter notions do not generally appear to be utilized by linguists and are not familiar as part of the formal methods used to investigate I-language. Of the latter terms specific to recursion there is one possible exception: an apparent equivalence between **definition by induction** and **definition by recursion**, but this equivalence is only arguably apparent—though I give no demonstration. Additionally, even if a significantly relevant equivalence between definition by recursion and definition by induction is real, this does not change the fact that linguists do not use the latter because it is in fact only useful for investigating the natural numbers \mathbb{N} and arithmetic in general—not natural language syntax and semantics. Nonetheless, the concept of definition by recursion has been used widely; example ?? in this paper, [Partee et al. \(1990\)](#), and [Chierchia and McConnell-Ginet \(2000\)](#).¹⁸ But this kind of definition, used within the context of linguistics, is useful only in order to construct algorithms that are (finitely) computably efficient and physically tractable within the general domain of human cognition—and in terms of the operation Merge, within the domain of narrow syntax (NS) and the narrow faculty of language (FLN). Lastly, [Soare \(1996\)](#) proposes that a historically normative convention has largely directed the use, and sometimes misuse, of the term “recursive.”

¹⁸[Chierchia and McConnell-Ginet \(2000\)](#) state for the bracket $[_A B C]$ that

We have to specify the value of the tree whose root is A in terms of the values of the subtrees rooted in B and C . This means that the semantic value for the terminal string dominated by A is determined in terms of the values of the substrings dominated by B and C and the way these substrings are put together. If we do this for every syntactic rule in the grammar, we can interpret any tree admitted by it. A definition of this kind (with a finite number of base clauses and a finite number of clauses that build on the base clauses) is called *recursive* (pg. 76).

The Recursion Convention has brought ‘recursive’ to have at least four different meanings.... This leads to some ambiguity. When a speaker uses the word ‘recursive’ before a general audience, does he mean ‘defined by induction,’ ‘related to fixed points and reflexive program calls,’ or does he mean ‘computable?’ [...] Worse still, the Convention leads to imprecise thinking about the basic concepts of the subject; the term ‘recursion’ is often used when the concept of ‘computability’ is meant. (By the term ‘recursive function’ does the writer mean ‘inductively defined function’ or ‘computable function?’) Furthermore, ambiguous and little recognized terms and imprecise thinking lead to poor communication both within the subject and to outsiders, which leads to isolation and lack of progress within the subject, since progress in science depends on the collaboration of many minds (pg. 29).

Group Theory is not so much explicitly employed in computability theory but set theory is, and since GT presupposes some set theory it is an important, albeit minor, issue to note that research into how the operation Merge works is research into its computable nature—NOT its recursive nature. As with the parallel to the CTP, linguistic science also has parallels with computability theory and shares the latter’s distinction between (i) an operation M being computable by finite means that are describable by algorithms—for example [Fortuny and Corominas-Mutra \(to appear\)](#) are developing a nesting machine based on a set-theoretically defined algorithm within the general theory of order—or trying to determine computable operations for certain functions of M , or even generating sets of objects that are countable by M , and (ii) an operation M' that is defined by induction or uses primitive recursive functions. It is within the general backdrop of computability, not recursive function theory, that any formalism of GT will be successful in theoretical syntax. An analogy can be made to the linguistic use of Turing machines (TM): theoretical syntacticians are more concerned with a specific type of universal TM¹⁹ as it relates to the operation Merge within the larger backdrop of the concern for algorithms that are both physically tractable and computably efficient.

12.4.3 GT and syntax

Group Theory, or GT, presupposes some knowledge of set theory because, technically, any group is also a set. That is, any collection of objects, including a set or

¹⁹This is the Turing *automatic*-Machine, compared to a stochastic/probabilistic TM, or even a quantum TM that [Deutsch \(1985\)](#) formalized.

collection of sets, can be a group if it can be defined by Definition 4; the following is loosely based on [Rosen \(1995\)](#).²⁰

Definition 4. GROUP:

G is a group iff G , under a law of composition, meets Criterion 1 to 4.

Criterion 1. Closure:

For all $a, b : a, b \in G$, then $ab, ba \in G$.

Criterion 2. Associativity:

For all $a, b, c : a, b, c \in G$, then $a(bc) = (ab)c$.

Criterion 3. Existence of Identity:

G contains identity element $e : ae = ea = a$.

Criterion 4. Existence of Inverses:

For every $a \in G$, then $a^{-1} \in G : aa^{-1} = a^{-1}a = e$.

A “law of composition” is broadly defined as any procedure that combines any two elements in any way (i.e., a binary operation). This can include the combination of functions $f(n)$, relations R , integer values in addition or multiplication ($t + y, t \times y$), or the combination of collections of objects that are subsets of a set $s \in S$. Compositions also include permutations, rotations, and translations by displacement of a certain distance along a certain line. (For group theory, see especially Chapter 10 in [Partee et al. \(1990\)](#); also [Milne \(2008\)](#), section 39 in [Kleene \(2002 \[1967\]\)](#); also [Korfhage \(1974\)](#), [Dornhoff and Hohn \(1978\)](#)—or any abstract or modern algebra book; also [Livio \(2005\)](#) and [Stewart \(2007\)](#) for historical and popularized introductions to symmetry and groups.) The genius of GT is that it is so broadly defined that it can describe a wide variety of concrete or abstract objects, but is rigorous enough to derive very precise relations between those objects. I begin with a general notion for a law of composition.

Law of Composition. (informal):

Any two possible elements, a, b , may be combined in any possible way, where ‘ \circ ’ is any combination, iff $a \circ b$ equals a set S .

I now briefly sketch how α Merge meets the criteria in Definition 4, giving a proof. I then discuss some issues arising from the conjecture that syntactic operations can be defined using criteria from GT.

²⁰Notice and interesting parallel with the criterion given below and some examples from [Boeckx \(2008\)](#) when discussing the “symmetry problem” for Merge (pp. 79-80). In discussing the empirical inadequacy of, what I am calling α Merge, he stresses that it cannot produce *asymmetric* relations such as precedence ($a \prec b \neq b \prec a$), and also prosody defined over prominence ($[[A, B] C] \neq [A [B, C]]$). These are equivalent to Criterion 1 and Criterion 2, respectively.

Conjecture 1. The operation αMerge forms a group.

Proof. Assume no other operation Merge and no interface constraints, then the following conditions hold:

Closure: If $\{\{\alpha\}, \{\beta\}\} \xrightarrow{\text{merge}} \{\Lambda, \{\alpha, \beta\}\} = K$, (where $\Lambda = \alpha$ or β), then $\{\alpha, \beta\} \in K$ and $\{\beta, \alpha\} \in K$. This satisfies closure and is analogous to the c-command relation R in the following way: if $\alpha R \beta$, then $\beta R \alpha$.

Associativity: If $[\Lambda \alpha [\beta [\gamma \delta]]] = K'$ and all label projections are ambiguous, then K' equals the decomposition $\{\alpha, \beta\}$, $\{\beta, \gamma\}$ and $\{\gamma, \delta\}$. By associativity, then, K' can also be decomposed into $\alpha\{\beta, \gamma\}$ and $\{\alpha, \beta\}\gamma$. This is analogous to the c-command relation R in following way: if, for example, $\alpha R \{\beta, \gamma\}$, then $\{\alpha, \beta\} R \gamma$.

Identity: By analogy, where rotation by 360° is equivalent to the absence of rotation—which is equivalent to an identity for rotation—the absence of the operation αMerge occurring is equivalent to the identity function. Then, by stipulation $\exists(\neg\alpha\text{Merge})$, and by identity, $\{\alpha\}, \{Y\} \xrightarrow{-\text{merge}} \{\alpha\}$, where Y is any constituent or set of constituents.

Inverse: By stipulation, if $\exists\{\alpha\}$ then $\exists\{\alpha^{-1}\}$.

□

The last stipulation for the existence of inverses is the biggest problem. Compared to the stipulation for the existence of identity—which is simply the non-occurrence of the operation αMerge —the existence of inverses is harder to rationalize. Because it is specific for actual constituents in the derivation it might be interpreted as saying that lexical or functional items have inverses—or that lexical entries must encode some kind of information for its inverse. This is not so; in fact, the inverse can be part of the mechanism of a numeration N and not of actual constituents. For now, I assume the legitimacy of this reasoning based on the following line of argument. Assume a numeration N consists of a finite set of elements that must all be selected, merged, and meet further requirements specific to Case, Phase, Agree, and other mechanisms or features belonging to a real-world syntactic derivation. Within a particular N , all elements consist of a set of pairs (LI, i) where LI is the lexical item and i is the index, [Chomsky \(1995c\)](#). Once the index i is reduced to zero, such that it can no longer be selected to merge by External Merge (an extension or equivalent of αMerge), then it seems rational to conclude that such a zero index specification may act as the inverse for whatever LI it was indexing. For items with an integer value greater than 1 for its index—so that it must be selected multiple times in order to reduce to zero—the inverse is formed from its correlational index. That is, $(\text{DP}_1, \text{DP}_2, \dots, \text{DP}_n)$ correlates to $(\text{DP}_{-1}, \text{DP}_{-2}, \dots, \text{DP}_{-n})$. In this way no problem arises for having to propose an

inverse for lexical items. Instead, the inverse belongs to the abstract structure of the numeration. In other words, inverses are constructed from the indices of N so that an item in the numeration with (LI, i) as its pair will also have (i^{-1}) and would look like this: (LI, ii^{-1}) . This is an explicit representation of the idea that an index *must* reduce to zero (actually its inverse). Under this interpretation then, indices do not actually reduce to a numerical zero, but instead, they reduce to their prespecified inverse. Consequently, since the index is really just a placeholder within the numeration and does not belong to the feature content of LI's, the inverse of this placeholder is also not specified as a feature of LI, but of the abstract structure of N generally.

What GT has to offer syntax generally deals with the nature of derivations. For example, trees derived by iterations of α Merge, such as Figure 12.10, can be shown to form an Abelian group because any element can be *commuted* with any other element. This is certainly not the case with real-world trees derived from a more complicated and natural operation Merge. However, some elements in trees can commute with each other. Or rather, some *positions* can commute based on feature specification and probe-goal relations. These form chains CH , which, under GT considerations, might be Abelian subgroups g belonging to a non-Abelian group G . Certainly, in GT these kinds of things are possible. And if the group G is redefined as the derivation D , then a chain CH of D may possibly form such a subgroup relation.

Applications from a GT perspective, namely one that defines α Merge by Criterion 1 through 4, remain to prove their worth—but the possibility of rigor for defining a base notion of symmetry from which Merge originates and acts as *symmetry breaking agent* is promising. Such possibilities for syntax make certain aspects of GT a welcome convenience. It may provide formal rigor in establishing interesting relations of general symmetry, and especially symmetry-breaking, in natural language.

12.5 Final Remarks

This paper has given a very brief historical background to the operation Merge, showing how its current conception is the result of a narrowed focus on the role of the logical tool of *definition by recursion* for constituents. Early transformational-generative grammars specified unique recursive definitions for constituents, example ??, while the X-bar model that followed generalized this definition in the form of a variable defined over lexical and functional categories, example (9). The current MP notion of Merge is that it is an abstract operation occurring only at local intervals, blind to future operations, and based strongly on a Bare Phrase model of

the X-bar type. Given this context, [Medeiros \(2008\)](#) has constructed an idealized balanced binary branching tree that exhibits a common mathematical pattern found in nature, the Fibonacci pattern, that correlates the number of X-bar objects/levels with the mathematical pattern that defines a Fibonacci level in a downward sequence; see also [Uriagereka \(1998\)](#), [Carnie and Medeiros \(2005\)](#), [Boeckx et al. \(2006\)](#), [Idsardi \(2008\)](#), [Soschen \(2008\)](#), and [Piattelli-Palmarini and Uriagereka \(2008\)](#) for use of the F-pattern. I show that if one follows the algorithm of the Golden Sequence, which derives the F-pattern, and correlates the substitution criteria (1, and 0) with XP and X^0 levels, a right-branching equivalent to [Medeiros \(2008\)](#) and [Soschen \(2008\)](#) results. By abstracting further, and looking only at head and phrase levels, a *translational symmetry* can be seen between the two levels and defined as a ratio between their co-ordinate F-levels as $F(n : n - 2)$; also called the X-ratio. This is one measure of symmetry that can be found in syntax, but there are other non-trivial kinds, as [Boeckx \(2008\)](#) shows.

This paper also discusses issues relevant to methodological concerns stemming from the Strong Minimalist Thesis as they relate generally to the potential use of formal tools in syntax, namely tools from Group Theory. These issues deal mainly with (i) the biological nature of natural language, (ii) that operations in natural language are in fact *operations that are computable*, and (iii) the nature and concept of computability itself. An argument for the biological nature of Merge is given, contrasted to a view of Merge—and human language—as a formal object. Additionally, a weakened version of the SMT is proposed and a corollary derived from it—both focusing on the view that human language is at least partially a product of nature. Working from the “Corollary to the Weak Minimalist Thesis,” I highlight some parallels between the MP and research into concepts of the computability of natural systems—though such parallels are meant to be taken in a conceptual and analogous light. Most useful in this context is the Church-Turing Principle, which is a naturalistic redefinition of the Church-Turing Thesis, given in [Deutsch \(1985\)](#). Following this, I focus on arguments made by the mathematician and computability theorist [Soare \(1996, 2007, 2008\)](#) about the normative use of the term “computability,” in contrast to the sometimes misused term “recursive,” as it applies to the linguistic field. It seems reasonable that, given linguists are more concerned with algorithms computable through discretely infinite means and processed by finite machines (brains), the use of the term “recursive” should be dropped in favor of “computability;” the latter of which already implies the kind of recursivity linguists are interested in.

Finally, the fact that various kinds of non-trivial natural symmetries can be defined in syntax is strong support for the investigation of the formal study of symmetry, Group Theory. Additionally, given the far-reaching goal for more rigorous methods of measurement and analysis in syntax, it is reasonable to look at Group

Theory as possibly providing some useful tools to the theoretical linguist. For this reason, I provide a group theoretic definition of the idealized operation ambiguous Merge, α Merge, and briefly discuss some implications of this definition—specifically that a numeration N may have an inverse i^{-1} for any index i that exists. Now, instead of an index reducing to zero, it simply reduces to its inverse. There also appear to be, on the face of it, possible uses of Group Theory in defining, measuring, or analyzing the behavior of Chains and their formal relations to derivations.

Part V

Conclusion

Chapter 13

Conclusion

Bibliography

- Ayer, A. J. 1936. *Truth, logic, and language*. New York: Basic Books.
- Baker, Mark. 2003. *Lexical categories*. Cambridge, UK: Cambridge University Press.
- Bar-Hillel, Yehoshua. 1954. Logical syntax and semantics. *Language* 30:230–237.
- Bod, Rens, Jennifer Hay, and Stefanie Jannedy (eds.). 2003. *Probabilistic linguistics*. Cambridge, MA.: MIT Press.
- Boeckx, Cedric. 2008. *Bare syntax*. Oxford: Oxford University Press.
- Boeckx, Cedric, Andrew Carnie, and David Medeiros. 2006. Some reflexes of natural law in syntactic structure. Ms., Harvard University and University of Arizona.
- Boeckx, Cedric, and Massimo Piattelli-Palmarini. 2005. Language as a natural object—linguistics as a natural science. *The Linguistic Review* 22:447–466.
- Boeckx, Cedric, and Juan Uriagereka. 2007. Minimalism. In *Oxford handbook of linguistic interfaces*, ed. Gillian Ramchand and C. Reiss, 541–573. Oxford: Oxford University Press.
- Bowles, Joshua. 2009. α Merge, symmetry, and group theory. Ms., University of Utah and Utah Valley University.
- Carnap, Rudolf. 1937. *The logical syntax of language*. Chicago: Open Court Classics.
- Carnie, Andrew. 2000. On the definition of X^0 and XP. *Syntax* 3:59–106.
- Carnie, Andrew, and David Medeiros. 2005. Tree maximization and the extended projection principle. *Coyote Working Papers in Linguistics* 14:51–55.

- Chamtezky, Robert. 2000. *Phrase structure: From GB to Minimalism*. Oxford: Blackwell.
- Chierchia, Gennaro, and Sally McConnell-Ginet. 2000. *Meaning and grammar: An introduction to semantics*. Cambridge, MA.: MIT Press, second edition.
- Chomsky, Noam. 1955. Logical syntax and semantics: Their linguistic relevance. *Language* 31:36–45.
- Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.
- Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA.: MIT Press.
- Chomsky, Noam. 1970. Remarks on nominalization. In *Readings in Transformational Grammar*, ed. Roderick Jacobs and Peter Rosenbaum. Waltham, MA.: Ginn.
- Chomsky, Noam. 1986. *Knowledge of language: Its nature, origin, and use*. New York: Praeger.
- Chomsky, Noam. 1994. Bare phrase structure. *MIT Occasional Papers in Linguistics* 5.
- Chomsky, Noam. 1995a. Bare phrase structure. In *Government and Binding theory and the Minimalist Program*, ed. Geert Webelhuth. Oxford: Blackwell.
- Chomsky, Noam. 1995b. Language and nature. *Mind* 104:1–61.
- Chomsky, Noam. 1995c. *The minimalist program*. Cambridge, MA.: MIT Press.
- Chomsky, Noam. 2005. Three factors of language design. *Linguistic Inquiry* 36:1–22.
- Chomsky, Noam. 2008. On phases. In *Foundational issues in linguistic theory*, ed. Robert Freidin, Carlos P. Otero, and Maria Luisa Zubizarreta. Cambridge, MA.: MIT Press.
- Church, Alonzo. 1936. An unsolvable problem of elementary number theory. *American Journal of Mathematics* 58:345–363.
- Citko, Barbara. 2005. On the nature of merge: External merge, internal merge, and parallel merge. *Linguistic Inquiry* 36:475–496.
- Clark, Stephen, and James R. Curran. 2007. Wide-coverage efficient statistical parsing with ccg and log-linear models. *Computational Linguistics* 33:493–552.

- Collins, Chris. 2001. Eliminating labels. In *Derivation and explanation in the Minimalist Program*, ed. Samuel D. Epstein and T. Daniel Seeley. Oxford: Blackwell.
- Deutsch, David. 1985. Quantum theory, the church-turing principle and the universal quantum computer. *Proceedings of the Royal Society of London A* 400:97–117.
- Deutsch, David. 1997. *The fabric of reality: The science of parallel universes and its implications*. New York: Penguin Books.
- Dornhoff, Larry L., and Franz E. Hohn. 1978. *Applied modern algebra*. New York: Macmillan Publishing Co., Inc.
- Fitch, Tecumseh, Marc D. Hauser, and Noam Chomsky. 2005. The evolution of the language faculty: Clarifications and implications. *Cognition* 97:179–210.
- Formigari, Lia. 2004. *A history of language philosophies*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Fortuny, Jordi, and Bernat Corominas-Mutra. to appear. Some formal considerations on the generation of hierarchically structured expressions. Ms., Center for Language and Cognition Groningen. Rijksuniversiteit Groningen and ICREA-Complex Systems Lab. Parc de Recerca Biomdica de Barcelona-Universitat Pompeu Fabra; to appear in *Catalan Journal of Linguistics*.
- Hauser, Marc D., Noam Chomsky, and Tecumseh Fitch. 2002. The language faculty: What is it, who has it, and how did it evolve? *Science* 298:1569–1579.
- Hornstein, Norbert. 1992. *As time goes by: Tense and universal grammar*. Cambridge, MA.: MIT Press.
- Hornstein, Norbert, and Jairo Nunes. 2008. Adjunction, labeling, and bare phrase structure. *Biolinguistics* 2:57–86.
- Idsardi, William J. 2008. Combinatorics for metrical feet. *Biolinguistics* 2:233–236.
- Jackendoff, Ray. 1977. *X'-syntax*. Cambridge, MA.: MIT Press.
- Jayaseelan, K.A. 2008. Bare phrase structure and specifier-less syntax. *Biolinguistics* 2:87–106.
- Kauffman, Stuart. 1995. *At home in the universe: The search for laws of self-organization and complexity*. Oxford: Oxford University Press.

- Kleene, Stephen. 2002 [1967]. *Mathematical logic*. New York: Dover.
- Korfhage, Robert R. 1974. *Discrete computational structures*. New York: Academic Press.
- Langendoen, Terence, and Paul Postal. 1984. *The vastness of natural languages*. Oxford: Basil Blackwell.
- Livio, Mario. 2002. *The golden ratio: The story of phi, the world's most astonishing number*. New York: Broadway Books.
- Livio, Mario. 2005. *The equation that couldn't be solved: How mathematic genius discovered the language of symmetry*. New York: Simon and Schuster Paperbacks.
- Lloyd, Seth. 2000. Ultimate physical limits to computation. *Nature* 406:1047–1054. URL <http://arxiv.org/>, cornell arXive; arXiv:quant-ph/9908043v3.
- Lloyd, Seth. 2006. *Programming the universe: A quantum computer scientist takes on the cosmos*. New York: Random House.
- Medeiros, David. 2008. Optimal growth in phrase structure. *Biolinguistics* 2:152–195.
- Milne, James S. 2008. Group theory. Lecture Notes, University of Michigan.
- Niyogi, Partha. 2006. *The computational nature of language learning and evolution*. Cambridge, MA.: MIT Press.
- Partee, Barbara, Alice ter Meulen, and Robert E. Wall. 1990. *Mathematical methods in linguistics*. Dordrecht: Kluwer Academic Publishers, second edition.
- Piattelli-Palmarini, Massimo, and Juan Uriagereka. 2008. Still a bridge too far? bi-olinguistic questions for grounding language on brains. *Physics of Life Reviews* 5:207–224.
- Pinker, Steven, and Ray Jackendoff. 2005. The faculty of language: What's special about it? *Cognition* 95:201–236.
- Post, Emil L. 1936. Finite combinatory processes-formulation 1. *Journal of Symbolic Logic* 1:103–105.
- Post, Emil L. 1944. Recursively enumerable sets of positive integers and their decision problem. *Bulletin of the American Mathematical Society* 50:284–316.

- Prince, Alan, and Paul Smolensky. 1993. Optimality theory: Constraint interaction in generative grammar. Technical report, Rutgers University Center for Cognitive Science, New Brunswick, NJ.
- Quine, Willard. 1953. *From a logical point of view*. Cambridge, MA.: Harvard University Press.
- Reichenbach, H. 1947. *Elements of symbolic logic*. New York: MacMillan.
- Rosen, Joe. 1995. *Symmetry in science: An introduction to the general theory*. New York: Springer-Verlag.
- Rubin, Edward J. 2002. The structure of modifiers. Ms., University of Utah.
- Rubin, Edward J. 2003. Determining pair-merge. *Linguistic Inquiry* 34:660–668.
- Soare, Robert I. 1996. Computability and recursion. *Bulletin of Symbolic Logic* 2:284–321.
- Soare, Robert I. 2007. Computability and incomputability: Computation and logic in the real world. In *Proceedings of the Third Conference on Computability in Europe, CiE 2007, Siena, Italy: Lecture notes in computer science No. 4497*, ed. S. B. Cooper, B. Lowe, and Andrea Sorbi. Heidelberg: Springer-Verlag.
- Soare, Robert I. 2008. Incomputability, turing functionals, and open computing. Ms., Department of Mathematics, University of Chicago.
- Soschen, Alona. 2008. On the nature of syntax. *Biolinguistics* 2:196–224.
- Stewart, Ian. 2007. *Why beauty is truth: a history of symmetry*. New York: Basic Books.
- Tarski, Alfred. 1956. *Logic, semantics, metamathematics*. Oxford: Oxford University Press.
- Tesar, Bruce, and Paul Smolensky. 2000. *Learnability in optimality theory*. Cambridge, MA.: MIT Press.
- Tomalin, Marcus. 2006. *Formal foundations of generative grammar*. Cambridge: Cambridge University Press.
- Turing, Alan M. 1936. On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society* 42:230–265.
- Uriagereka, Juan. 1998. *Rhyme and reason: An introduction to minimalist syntax*. Cambridge, MA.: MIT Press.

Index

evolutionary, 4

Exaptation, 9

function of language, 8

Language, 1, 3

linguist, 4