Natural Syntax

An Emergentist Primer

CMO

©**William O'Grady**
Department of Linguistics
University of Hawai'i at Mānoa
ogrady@hawaii.edu

2020
Version B.1.2

Contents

Prefatory Note

This short monograph draws on materials from my classes, talks and publications, as well as on new ideas that reflect my current thinking about language and emergence. It is not in its 'final form,' and probably never will be, as I intend to make continual revisions and additions as new ideas and information become available.

Natural Syntax is not written in an entirely conventional way. One obvious departure from the usual practice involves the very extensive use – especially in the first few chapters – of direct quotes to document the various views and proposals that have shaped linguistics. This may disrupt the flow of the text in places, but I found it helpful, as a writer at least, to work with the verbatim record.

As its subtitle suggests, *Natural Syntax* is intended for an audience with little or no background in the study of emergence or its possible relevance to the understanding of language. I have also tried to assume no more than a modest background in linguistics, favoring readability over technical details wherever possible. The chapters are divided into three main groups. Chapters 1 through 3 focus on matters of history and methodology and should be easy reading. The next six chapters (4 through 9) examine properties of fairly basic syntactic phenomena – word order, *wh* questions and coreference, all of which should be familiar to most readers, including those with a relatively modest background in linguistics. Chapters 10 through 12, which take us well beyond the familiar languages of Europe, examine the differences between accusative and ergative languages with regard to case, agreement and filler-gap dependencies. The book comes to a close with two chapters on language acquisition, followed by some brief concluding remarks on the roots of natural syntax. (An appendix seeks to shed light on the '*that*-trace effect,' one of the most baffling phenomena in all of language.)

Linguistics – and especially syntax – has been a hotbed of controversy for many decades, for reasons that await scrutiny and assessment by scholars in the history and philosophy of science. I do not believe that we should be looking forward to that day. In the meantime, however, I offer yet another perspective on what language might be like and how we might go about making sense of it.

Acknowledgments

I’ve benefitted a great deal from conversations and correspondence with a number of scholars on matters related to the topics on which this book touches, including Sharon Bulalang, Lyle Campbell, Victoria Chen, Kamil Deen, Patricia Donegan, Kevin Gregg, John Hawkins, Raina Heaton, Gary Holton, Brad McDonnell, Brian MacWhinney, Yuko Otsuka, Peter Schuelke and David Stampe. The editor for this project was Miho Choo, who uncovered many flaws and suggested many improvements, for which I am very grateful.

1

Emergence

Language is one of nature's most mysterious phenomena. Its workings have eluded scholars for thousands of years, and there is no reason to think that a major breakthrough is at hand. The danger of a deep irony is real: the quest to explain language may lie beyond the reach of the only creatures who are able to use it. But of course we must nonetheless try.

1.1 The emergentist idea

One avenue of inquiry that has shown some promise in recent years remains lightly explored. It takes as its starting point *emergence*, the process whereby the interaction of simple components, forces and events produces a system with its own novel properties.

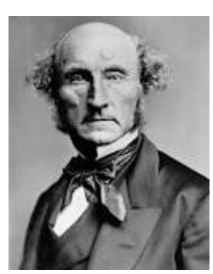
The term was coined by the nineteenth-century philosopher George Henry Lewes, who drew a distinction between 'resultants' and 'emergents.'

> *… every resultant is clearly traceable in its components, because these are homogenous and commensurable. It is otherwise with emergents... The emergent is unlike its components …, and it cannot be reduced to their sum or their difference.*
>
> (Lewes 1875:413)

Lewes's friend and colleague, the British philosopher and economist John Stuart Mill, offered an early concrete example.

> *The chemical combination of two substances produces, as is well known, a third substance with properties different from those of either of the two substances separately, or both of them taken together. Not a trace of the properties of hydrogen or oxygen is observable in those of their compound, water.*
>
> (Mill 1842, Book III, Ch. 6)

George Henry Lewes (1817-1878)



John Stuart Mill (1806-1873)

Emergentism reinforces the need for caution in scientific inquiry. Things are often not what they appear to be at first glance. There are lots of examples of this in nature, and it's worthwhile to consider a couple of them before going forward.

A first illustration comes from the striking architecture of bee hives, with their carefully arranged rows of perfect hexagons.

Inside a bee hive

Evidence of geometrical ability in bees? Alas, no; just evidence for emergence.

> *When [spherical] honey cells are packed together, the wax walls will undergo deformation. … the surface tension between two honey cells acts to create a flat plane surface. Since each cell contacts just six other cells, the surface tensions of the wax and the packing pressure of the bees will force each sphere into a hexagonal shape. The hexagonal shape maximizes the packing of the hive space and the volume of each cell and offers the most economical use of the wax resource… The bee doesn't need to "know" anything about hexagons.*
> (Elman, Bates, Johnson, Karmiloff-Smith Parisi & Plunkett 1996:111; see also Thompson 1917:608ff)

A second example involves the behavior of flocks of birds, whose seemingly coordinated dance back and forth across the sky is best appreciated by direct observation.

Watch the video: https://www.youtube.com/watch?v=V4f1r80RY

I recall watching such a display once and marveling at the complexity of the avian ballet playing out in the sky. In fact, I should have been marveling at its simplicity. It turns out that flocking behavior reflects the interaction of the three simple 'principles' paraphrased below. (The terms and their description vary somewhat from author to author.)

- Alignment: Go in the same direction as those around you.
- Cohesion: Don't get too far from those around you.
- Separation: Don't get too close to those around you.
  (Reynolds1987; see also https://www.red3d.com/cwr/boids/)

You can see a computer simulation of these three principles in action here:

> *How do boids work? A flocking simulation*
> https://www.youtube.com/watch?v=QbUPfMXXQIY;
> see also http://www.harmendeweerd.nl/boids/

Emergentism versus essentialism

A typical linguistic analysis starts with a grammatical phenomenon (or, at least, what *appears* to be a grammatical phenomenon) and ends up with a grammatical explanation of its properties. That's called *essentialism*.

> *An … "essentialist" … explanation generally appeals to one or more factors that are unique to the complex phenomenon itself.*
> (Tachihara & Goldberg 2019:237; see also Scholz, Pelletier & Pullum 2016)

The vast majority of work in linguistics is essentialist in character, as best exemplified by 'Principles-and-Parameters' theory, the influential variety of generative grammar outlined in Noam Chomsky's seminal book, *Lectures on Government and Binding* (Chomsky 1981).

Lectures on Government and Binding – the founding document of modern generative grammar

The heart of Principles-and-Parameters theory consists of Universal Grammar (UG), a system of inborn grammatical principles. Here's an example of one of the more famous principles, which occupied the attention of hundreds of linguists for decades. (I deliberately present the principle in its full technical detail to give readers a sense of what traditional UG actually looked like, outside of textbooks. I do not expect the average reader to actually understand the principle, nor is there any reason to do so at this point.)

> ***The Empty Category Principle (ECP)***
> An empty category must be properly governed.

α governs β iff α m-commands β and there is no barrier for β that excludes α.

α excludes β if no segment of α dominates β.

γ is a barrier for β iff (a) or (b):
(a) γ immediately dominates δ, and δ is a blocking category for β.
(b) γ is a blocking category for β, and γ is not IP [Inflectional Phrase].

γ is a blocking category for β iff γ is not L-marked and γ dominates β.

α L-marks β if α is a lexical category that θ-governs β.

α θ-governs β iff α is a zero-level category that θ-marks β, and α & β are sisters.

(Chomsky 1986a:8ff)

It has been repeatedly noted in the literature that principles like the ECP are 'domain-specific' – they are relevant only to language.

> *… the system of basic principles, UG, appears to be unique in significant measure to [language].*
>
> (Chomsky 1977:3)

> *It would be surprising indeed if we were to find that the principles governing [linguistic] phenomena are operative in other cognitive systems, although there may be certain loose analogies… there is good reason to think that the language faculty is guided by special principles specific to this domain…*
>
> (Chomsky 1980b:44)

> *There seems little reason to suppose, for the moment, that there are general principles of cognitive structure, or even of human cognition, expressible at some higher level, from which the particular properties of particular "mental organs," such as the language faculty, can be deduced, or even that there are illuminating analogies among these various systems.*
>
> (Chomsky 1980b:215)

> *My own belief is that the principles [of UG] do not generalize, that they are in crucial respects specific to the language faculty…*
>
> (Chomsky 1986b:xxvi)

> *… a Universal Grammar, not reducible to history or cognition, underlies the human language instinct.*
>
> (Pinker 1994:238)

None of the highly abstract and tightly knit principles and parameters of universal grammar bear any resemblance whatsoever to derivations from non-linguistic principles ...
…
There is no hope, not even the dimmest one, of translating these entities, these principles and these constraints into generic notions that apply to language as a "particular case."
(Piattelli-Palmarini 1998:334 & 339)

Human knowledge of natural language results from – and is made possible by – a richly structured and biologically determined capacity specific to this domain.
(Anderson & Lightfoot 2000:17)

Put simply, classic generative grammar explains grammatical phenomena by reference to grammatical principles. That's not the way emergentism works.[1]

The emergentist program

A defining feature of the emergentist program is a commitment to explaining the properties of language by reference to the interaction of deeper and more general properties of cognition. The point has been emphasized in a large and varied literature over a period of several decades, as the following representative assertions help illustrate.

Language is a new machine built out of old parts.
(Bates & MacWhinney 1988:147)

… in very significant ways, language is a radically new behavior. At a phenomenological level, it is quite unlike anything else that we (or any other species) do. It has features that are remarkable and unique. The crucial difference between this view and the view of language as a separable domain-specific module … is that the uniqueness emerges out of an interaction involving small differences in domain-nonspecific behaviors.
(Elman 1999:25)


Elizabeth Bates (1947-2003)


Brian MacWhinney


Jeff Elman (1948-2018)

[1] In fact, as we will see in the next chapter, it may not be the way generative grammar works any more either.

… while the grammar forms a separate cognitive system, the categories and principles needed for language are constructed from notions that are not specifically linguistic in character … none of the biologically determined structures involved in development are unique to the language faculty.
(O'Grady 1987:181)

The structures of language emerge from interrelated patterns of experience, social interaction, and cognitive mechanisms.
(Beckner, Blythe, Bybee, Christiansen, Croft et al. 2009:2)

The phenomena of language are best explained by reference to more basic non-linguistic (i.e., "non-grammatical") factors and their interaction – physiology, perception, processing, working memory, pragmatics, social interaction, properties of the input, the learning mechanisms, and so on.
(O'Grady 2008b:448)

… the structure of human language must inevitably be shaped around human learning and processing biases deriving from the structure of our thought processes, perceptuomotor factors, cognitive limitations, and pragmatic constraints.
(Christiansen & Chater 2008:490)

To the extent that there are striking similarities across languages, they have their origin in two sources: historical common origin or mutual influence, on the one hand, and on the other, from convergent selective pressures on what systems can evolve. The relevant selectors are the brain and speech apparatus, functional and cognitive constraints on communication systems, including conceptual constraints on the semantics, and internal organizational properties of viable semiotic systems.
(Evans & Levinson 2009:446)

In order to qualify as emergentist, an account of language functioning must tell us where a language behavior "comes from." In most cases, this involves accounting for a behavior in a target domain as emerging from some related external domain. For example, an account that shows how phonological structures emerge from physiological constraints on the vocal tract involves external determination …
(MacWhinney 1999b:xii; see also MacWhinney 1999a:362)

> *Emergentism refers to the idea that unique, complex phenomena can be explained by the development of more basic processes that interact in dynamic ways with each other and the environment.*
> …
> *The emergentist perspective predicts that languages will be constrained by communicative pressures and domain-general processes related to attention, memory, categorization, and cognitive control.*
> (Tachihara & Goldberg 2019:237 & 238)

1.2 Complex systems

An emergentist approach to language in the 21st century situates linguistics within the study of *complex systems* – systems composed of multiple parts that interact with each other in intricate ways. A defining feature of complex systems is the presence of emergence; they become more than the sum of their parts, taking on properties and manifesting effects that could not have been predicted in advance.[2]

- Complex systems are 'dynamic;' they are subject to flux and change rather than long-term stability. The weather is constantly changing, traffic increases and decreases throughout the day, the human body ages with the passage of time.
- Complex systems are subject to 'phase transitions' that bring about qualitative changes over time. Water can turn into ice, fall becomes winter, a tadpole can morph into a frog.
- Complex systems are 'non-linear.' Because of the way their parts interact with each other and with other factors, even a seemingly insignificant event can yield consequences that are far out of proportion to its own importance. A randomly selected lottery number can bring wealth, a dropped ball can determine a national championship, the cross-species transmission of a microscopic virus to a single human can result in a pandemic.

Complex systems are everywhere: the weather, evolution, society, the economy, politics, ecosystems, football games, rush-hour traffic, human history, your life … The possibilities are essentially endless, and some scenarios are deliberately fanciful. Might it be possible, for example, for a butterfly in Brazil to trigger a tornado in Texas by simply flapping its wings, thereby creating a cascade of ever more complex effects?[3]

---

[2] For a brief and very helpful on-line lecture on complex systems (made by Complexity Labs), see: https://www.youtube.com/watch?v=vp8v2UddPM. For a brief treatment of the non-linear effects, see: https://www.youtube.com/watch?v=1xn98WnMxMs. For phase transitions and various related notions: https://www.youtube.com/watch?v=xSWm8p5tSiw.

[3] E. N. Lorenz, 'Does the flap of a butterfly's wings in Brazil set off a tornado in Texas?,' American Association for the Advancement of Science, 139th Meeting, Washington DC. December 29, 1972. The title of the talk, chosen by the session chair, was inspired by a critical remark directed at Lorenz several years earlier to the effect that if his theory were right, 'one flap of a seagull's wings could change the climate forever.'

Other scenarios are more realistic and more related to our field of study. The most obvious question has to do with whether language could be a complex system, with its form and use reflecting the interaction of yet to be discovered forces and factors. The possibility is worth considering.

Language certainly looks like a complex system. It consists of many interacting parts: sounds, meanings, words, inflection, ordering, exceptions, metaphors, idioms and so forth. Moreover, it also seems to have those other properties of complex systems mentioned above:

- It is in a constant state of flux: language change, linguistic development, attrition.
- It manifests non-linear behavior: access to even a small vocabulary and a small set of combinatory operations results in the ability to produce and understand an unlimited number of sentences of potentially unlimited complexity.
- It exhibits signs of phase transition: languages with OV order transition to VO order; a child's halting production of two-word utterances quickly evolves into the ability to produce complex sentences of any length; the entire system can be lost as the result of neurological trauma or decline; and so on.

The big question is whether recognizing that language is a complex system sheds light on the sorts of problems that we need to confront as linguists:

- Why do languages have the properties that they do, as manifested in phenomena such as case, agreement, relativization, referential dependencies and the like?
- How are these properties acquired with such ease by children?
- What are the limits on variation across and within languages?

These are the questions with which this book deals. Before beginning to address them, however, it is useful to consider some developments in a more mainstream line of linguistic research. The next chapter examines the Minimalist version of generative grammar initiated by Noam Chomsky in the early 2000s.

2

The Minimalist Turn

By the mid 1990s, doubts had emerged about the viability of essentialism as a central tenet of the generative research program. Indeed, the doubters included Chomsky himself, who launched a new explanatory strategy that came to be known as 'Minimalism.'

A key feature of this approach involved a fundamental rethinking of Universal Grammar.

> *There is no longer a conceptual barrier to the hope that the UG might be reduced to a much simpler form, and that the basic properties of the computational systems of language might have a principled explanation instead of being stipulated in terms of a highly restrictive language-specific format for grammars.*
> (Chomsky 2005:8)

This line of thinking culminated in the Strong Minimalist Thesis, the focus of this chapter.

2.1 The Strong Minimalist Thesis

In a major reversal of the earlier view that Universal Grammar consists of universal grammatical principles (as its name implies), Chomsky began to move toward an entirely different conception of the language faculty.

> *The optimal situation would be that UG reduces to the simplest computational principles, which operate in accord with language-independent conditions of computational efficiency.*
> (Chomsky 2017:296)

Strictly speaking, the Strong Minimalist Thesis is an endorsement of emergentism, at least in the broadest sense of the term[1] – a development predicted years earlier by the emergentist pioneer Brian MacWhinney.

> *Recent developments within the theory of generative grammar, as well as psycholinguistic extension of the generative framework, have suggested that linguistic theory may eventually be compatible with an emergentist framework.*
> (MacWhinney 1999:xvi)

Indeed, the label 'emergentist' is now being used by some proponents of the Minimalist approach to describe their work.

> *… we will present a new view of the nature of parameters, one which represents a major departure from the 'classical' view, and which is compatible with minimalist assumptions, as well as being in certain respects more compatible with functionalist views on language acquisition and change. This is the 'emergentist' theory of parameters…*
> (Biberauer & Roberts 2017:142; see also Newmeyer 2017:560)

An even more extreme emergentist idea, reaching into the domain of animal cognition, has also been put forward. (Notice the name of the fourth author.)

> *A recent development, the variational learning model (Yang, 2002), proposes that the setting of syntactic parameters involves learning mechanisms first studied in the mathematical psychology of animal learning.... Parameters are associated with probabilities: parameter choices consistent with the input data are rewarded and those inconsistent with the input data are penalized. The computational mechanism is the same – for a rodent running a T-maze and for a child setting the head-directionality parameter – even though the domains of application cannot be more different.*
> (Yang, Crain, Berwick, Chomsky & Bolhuis 2017:116).

The problem of evolution

Minimalist thinking in contemporary generative grammar is driven by a concern that had previously received little attention – the need for a plausible theory of how language evolved in the human species, apparently quite recently.

> *… any linguistic theory is going to have to meet two conditions: the conditions of acquirability and evolvability. UG must permit acquisition of [language], and it must have evolved in the human lineage – and if current best guesses are correct, it must have evolved recently.*
> (Chomsky, Gallego & Ott 2019:25)

[1] Chomsky clearly has no objection to the idea of emergence in general; indeed, he has called the idea that the mind is an emergent property of the brain 'close to a truism' (2002:74).

> *… the emergence of language is a very recent evolutionary event [within the last 80,000 – 100,000 years]: it is thus highly unlikely that all components of human language, from phonology to morphology, from lexicon to syntax, evolved independently and incrementally during this very short span of time. Therefore, a major impetus of current linguistic research is to isolate the essential ingredient of language that is domain specific and can plausibly be attributed to the minimal evolutionary changes in the recent past.*
> (Yang, Crain, Berwick, Chomsky & Bolhuis 2017:114)

Obviously, 'less is more' in confronting this sort of challenge.

> *We can all agree that the simpler conception of language with a reduced innate component is evolutionarily more plausible – which is the impetus for the Minimalist Program of language.*
> (Yang, Crain, Berwick, Chomsky & Bolhuis 2017:115)

The recent attention to evolutionary considerations is intriguing. In 1975, Noam Chomsky participated in an exchange of views, billed as a 'debate,' with the celebrated Swiss psychologist Jean Piaget.

Jean Piaget (1896-1980)

Piaget, who was 80 at the time of the meeting, had devoted his career to the study of child cognitive development. Although he was clearly not familiar with the details of Chomsky's theory, he expressed skepticism about Universal Grammar based on considerations of evolutionary plausibility.[2]

> *I cannot accept the [UG] hypothesis … this mutation particular to the human species would be biologically inexplicable; it is already very difficult to see why the randomness of mutations renders a human being able to "learn" an articulate language, and if in addition one had to attribute to it the innateness of a rational linguistic structure, then this structure would itself be subject to a random origin…*
> (Piaget 1980:31; see also Bates 1994:141)

[2] Elizabeth Bates, a founder of modern emergentism, made a similar point a decade later: 'It is hard to imagine how we could have developed elaborate, innate and domain-specific mechanisms for language in a relatively short period of time' (Bates 1994:141).

Chomsky dismissed Piaget's concerns.

> *Although it is quite true that we have no idea how or why random mutations have endowed humans with the specific capacity to learn a human language, it is also true that we have no better idea how or why random mutations have led to the development of the particular structures of the mammalian eye or the cerebral cortex…. Little is known concerning evolutionary development but from ignorance, it is impossible to draw any conclusion.*
> (Chomsky 1980a: 36)

In retrospect, it appears that Piaget was prescient in his skepticism over how a cognitive system as elaborate as Universal Grammar, as then conceived, could have evolved (let alone in a very short period of time). Moreover, Chomsky's response was barely relevant, since it evoked biological structures (the mammalian eye and the cerebral cortex) that evolved over millions of years.

At the time, Piaget's position was not taken seriously, but today he is being hailed in some quarters as an early Minimalist!

> *… some aspects of Piaget's vision were on the right track, but premature. In some sense, Piaget was a minimalist, a modern biolinguist ante litteram. Chomsky's 'classical stance' (rich, overspecified UG) may have been a necessary step in the development of the field, but in the long run, at least some aspects of Piaget's vision may prove more productive …*
> (Boeckx 2014:90-91)

It is easy to see a non-linear effect here. Had Chomsky grasped the importance of Piaget's point at the time, UG would have been nipped in the bud. There would have been no Government-and-Binding theory or its various Principles-and-Parameters successors. Some scholars who went into other professions would have become linguists, and some who became linguists would have gone into other professions. Linguistics would have been a very different discipline, although, of course, we will never know how different. That's just the way emergence works.

2.2 Rebooting Universal Grammar

The term 'language faculty' is widely used within linguistics to refer to the ability to acquire and use language – a distinctly human trait. Proponents of Universal Grammar note the uniqueness of language at every opportunity as if it were a novel discovery or an issue of contention. It is neither. As Jackendoff (2002:94) notes, 'it is an ancient observation that only humans speak.' As the following broad sampling of remarks illustrate, this is the consensus view among modern scholars as well, regardless of their theoretical persuasion.

… man is the sole possessor of language.
(Whitney 1875:2)

Language is a purely human and noninstinctive method of communicating ideas, emotions, and desires by means of a system of voluntarily produced symbols.
(Sapir 1921:7)

From time immemorial, the animals and spirits of folklore have had human characteristics thrust upon them, including always the power of speech. But the cold facts are that Man is the only living creature with this power.
(Hockett 1958:569)

Clearly no one denies that human beings are the only creatures that are able to learn and use grammar in its full-blown form.
(Elman, Bates, Johnson, Karmiloff-Smith, Parisi & Plunkett 1996:372)

Certainly, humans are endowed with some sort of predisposition toward language learning.
(Pullum & Scholz 2002:10)

Language is one of the hallmarks of the human species – an important part of what makes us human.
(Christiansen & Kirby 2003:300)

Language is a unique hallmark of the human species.
(MacWhinney 2005:383)

… language has evolved only in humans …
(Ambridge & Lieven 2011:365)

Language is used by humans in a way that no animal can match.
(Ibbotson & Tomasello 2016:75-76)

Inside the language faculty

The question has always been (and remains) simple and straightforward: Of what does the language faculty consist? It is here that generative linguistics made its mark in the history of ideas by putting forward Universal Grammar as the centerpiece of a *theory* of the language faculty.

UG is the theory of the human faculty of language.
(Chomsky 1977:3; see also p. 74)

UG is a theory of the "initial state" of the language faculty, prior to any linguistic experience.
(Chomsky 1986:3)

UG is a theory of the initial state S_0 of the relevant component of the language faculty.
(Chomsky 1993:1)

The theory of the initial cognitive state is called Universal Grammar.
(Chomsky 2002:8)

It is amply clear what that theory proposed: an inborn system of grammatical categories and principles of the type exemplified in *Lectures on Government and Binding* and countless subsequent publications by hundreds of scholars, led by Chomsky himself, over a period of many years. (If you're having any doubts about what traditional Universal Grammar was like, have another look at the Empty Category Principle, outlined in §1.1.) The published literature is replete with assertions that confirm the character of UG as it developed over a period of several decades, beginning in the 1960s.

… the child has an innate theory of potential structural descriptions that is sufficiently rich and fully developed so that s/he is able to determine … which structural description may be appropriate [for any particular sentence].
(Chomsky 1965:32)

… universal grammar (UG) [is] the grammar underlying the grammars of all particular human languages…
(Piattelli-Palmarini 1989:23)

Children are preprogrammed to adhere to [certain linguistic principles] as part of the blueprint for their development. Just as a child cannot help but grow fingers and toes, and not wings or claws, so these linguistic principles, and not others, grow in the child.
(Crain & Thornton 1998:27).

Where does linguistic knowledge come from? … this knowledge is inborn.
(Guasti 2002:17)

UG articulates the essential architecture, the epistemological primitives necessary to each level of linguistic representation, and the linguistically specific computational system for combining multiple levels of representation.
(Lust 2006:265)

... children are born with a set of universal linguistic principles ... many aspects of adult grammar are innate.
(Crain, Goro & Thornton 2006:31)

Grammar, though abstract and remote, is just like our arms and legs – an apparatus that we have from birth, whose use we refine by experience.
(Roeper 2007:247)

… a great deal of grammatical information must already be present in the child's brain at birth.
(Snyder & Lillo-Martin 2011:670)

Modern nativists assumed that there exists a language-specific "genetically determined initial state" … for explaining language growth in the individual. That initial state, named Universal Grammar, contains the innate linguistic principles for language to develop.
(Boeckx & Longa 2011:258)

The UG hypothesis argues that at least some of the learning biases necessary to solve [the] induction problem take the form of innately specified, language-specific constraints …, which often correspond to specific linguistic phenomena.
(Pearl & Sprouse 2013:24-25)

Universal Grammar … is essentially a series of biases that allow the child to make sense of hugely ambiguous linguistic data in a uniform and efficient manner. But crucially, the UG-based approach assumes that all of these biases are specific to the domain of language and cannot be derived from other domains at all.
(Becker & Deen 2020:45)

Salvaging the name

The UG-based theory of the language faculty now appears to have been wrong, as even many of its former proponents, including Chomsky, have come to recognize (see §2.1). Yet, for some reason, there is an attempt afoot to salvage the term 'Universal Grammar' by making it the *name* for the language faculty.

The term UNIVERSAL GRAMMAR (UG), in its modern usage, [is] a name for the collection of factors that underlie the uniquely human capacity for language – whatever they may turn out to be…
(Nevins, Pesestsky & Rodrigues 2009:357)

UG is the technical name for the genetic component of the language faculty.
(Chomsky 2013 'After 60+ years of generative grammar: A personal perspective,' lecture at Princeton University: https://www.youtube.com/watch?v=Rgd8BnZ2-iw, around 21:45)

The term Universal Grammar (UG) is a label for [the] striking difference in cognitive capacity between "us and them [humans and animals]."
(Chomsky, Gallego & Ott 2019:230)

The strategy has not gone unnoticed.

> *UG comes in two forms: a definitional form and an empirical form. The definitional one is that UG consists of those attributes characteristic of the human mind that makes first language acquisition both possible and systematically successful in the absence of teaching.*
> (Rooryck, Smith, Lipták & Blakemore 2010:2651)

> *Chomsky's notion of Universal Grammar (UG) [is] the programmatic label for whatever it turns out to be that all children bring to learning a language.*
> (Evans & Levinson 2009:430)

> *We do not use the term [Universal Grammar] in its most general sense, in which it means simply "the ability to learn language". The claim that humans possess universal grammar in this sense is trivially true…*
> (Ambridge, Pine & Lieven 2014:e53)

> *Very strangely, what I observe is anti-Chomskyans rejecting universal grammar (e.g., Evans & Levinson 2009, Ibbotson & Tomasello 2016), and Chomskyans defending universal grammar in some unclear abstract sense – whereas Chomsky himself seems to largely agree with the anti-Chomskyan view. (I really pity newcomers to the field of linguistics – they must be terribly confused by what is going on.)*
> (Martin Haspelmath, https://dlc.hypotheses.org/1269)

The reduction of UG to a terminological convention has proven to be polemically useful.

> *The question of whether language exists is, basically, the question of whether UG exists.*
> (Chomsky 2011:270)

> *The controversy about the existence of universal grammar is a controversy over whether there is some genetic property that distinguishes human beings from everybody else.*
> (Chomsky 2012 question-and-answer session:
> https://www.youtube.com/watch?v=vbKO-9n5qmc, around 2:30)

> *No rational person can believe that there is no genetic basis for the fact that a human infant, but no other organism, instantaneously and effortlessly extricates from the environment language-relevant data, and in a rather comparable way quickly attains rich linguistic competence, again a feat beyond other organisms even in its rudimentary aspects. But that is what is entailed by denying the existence of UG.*
> (Chomsky 2015:96)

> *There is a huge literature that denies the existence of UG. What the literature is saying is that [language acquisition] is a miracle. There is no coherent alternative to UG.*
>
> (Chomsky 2013 'After 60+ years of generative grammar: A personal perspective,' lecture at Princeton University: https://www.youtube.com/watch?v=Rgd8BnZ2-iw, around 21:45)

These remarks are quite surprising. True, many scholars have expressed opposition to the idea of inborn grammatical principles, including now Chomsky himself. However, there is no literature that denies the existence of a uniquely human language faculty and there never has been, as the quotes at the beginning of §2.2 help show. The proposal that humans have an inborn ability to learn and use language is not a breakthrough for cognitive science. It is simply a statement of the position that *everyone* in the field has always held, including many opponents of generative grammar, as the following assertions confirm.

> *In one way or another, every modern approach to language acquisition deals with the fact that language is constructed anew by each child, making use of innate capacities of some sort.*
>
> (Slobin 1985:1158)

> *… a great deal has been learned … about the biological foundations for language development. Evidence of innateness is good, but evidence of a domain-special 'mental organ' is difficult to find.*
>
> (Bates 1994:148)

> *… there is a biological foundation for language, just not in the form of specific linguistic structures preformed in the human genome*
>
> …
>
> *The issue, then, is not whether human beings are biologically prepared for language acquisition; they are.*
>
> (Tomasello 1995:133 & 137)

> *There is little controversy in contemporary linguistics over the claim that the human mind includes an innately structured acquisition device. There is simply no other plausible explanation for why human beings – but not dogs or rabbits – can acquire and use language.*
>
> (O'Grady 1997:298)

> *… there are compelling reasons to believe that the possession of language by human has deep biological roots. We are the only species that has a communication system with the complexity and richness of language.*
>
> (Elman 1999:1)

> *Proponents of the emergentist view acknowledge that something is innate in the human brain that makes language possible, but that something may not be a special-purpose, domain-specific device…*
> (Bates & Goodman 1999:33)

> *All researchers agree that there is clearly some genetic involvement in speech and language…*
> (Ambridge & Lieven 2011:368)

The challenge is now the same for everyone – to offer a characterization of the human capacity for language without positing inborn grammatical principles. The evolving Minimalist Program has brought generative grammar 'back to the pack,' where it joins a crowded field of ideas and hypotheses, including many of emergentist inspiration. Perhaps too many! Recent interest in emergentism has created 'an embarrassment of riches,' notes MacWhinney (2015:9), leaving students and scholars to make sense of a landscape in which 'all of these approaches fall under the general category of Emergentism.'

In the next chapter, I will outline a deliberately narrow conception of what a contemporary emergentist approach to language might look like.

3

The Strict Emergentist Protocol

There are many varieties of emergentism and many ways to exploit its leading ideas in confronting the explanatory challenges presented by the study of language, including the two most central issues of all:

- Why do languages have the particular properties that they do?
- How are those properties acquired by children?

In developing an emergentist approach to these questions, I will adopt three restrictive theses, each of which runs directly counter to standard lines of thought within formal linguistics. The first thesis challenges the existence of syntactic structure, the second rejects the reality of grammar, and the third proposes that the operations required to bring together form and meaning in natural language are shaped primarily by processing pressures. Let us consider each in turn.

3.1 Direct mapping

It has long been a matter of consensus that language provides a way to map meaning onto form (sound or gesture), and vice versa.

> *A sentence is a spoken sound with meaning.*
> (Aristotle, cited by Everson 1994:91)

> *Language may be defined as the expression of thought by means of speech sounds. … Every sentence or word by which we express our ideas has a certain definite form of its own by virtue of the sounds of which it is made up, and has a more or less definite meaning.*
>
> *The first thing in the study of language is to realize clearly this duality of form and meaning, constituting respectively the formal and the logical (or psychological) side of language.*
> (Sweet 1900:1)

> *Grammars are complex behavioral solutions to the problem of mapping structured meaning onto a linear string of sounds.*
> (Elman, Bates, Johnson, Karmiloff-Smith, Parisi & Plunkett 1996:39)

> *… every approach to language presupposes, at least tacitly, that a language determines a specific sound-meaning correlation.*
> (Chomksy 2015:93)

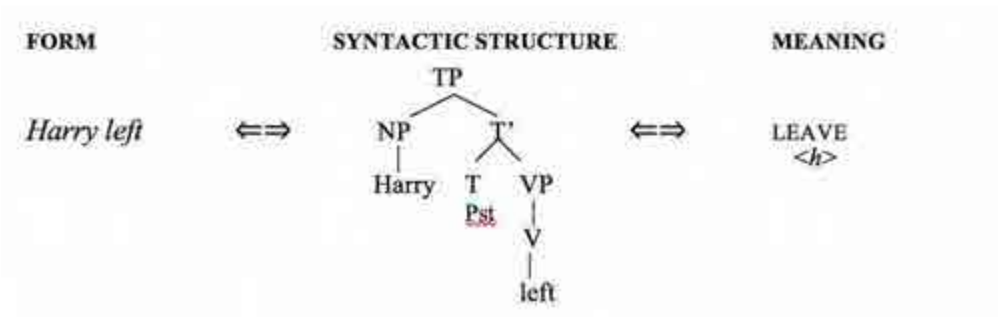
The challenge, of course, lies in determining the mechanisms by which sound and meaning are associated.

> *[The study of language is] concerned not with form and meaning separately, but with the connections between them, these being the real phenomena of language.*
> (Sweet 1891:7)

> *… a major task of linguistics should be to investigate … how this thought-to-sound pairing is possible, for thoughts and sounds themselves have little in common aside from the fact that both flow through time.*
> (Chafe 2018:28)

The dominant view in formal linguistics holds that the mapping between sound and meaning must be 'mediated' by syntactic structure. I will refer to this idea as 'mediated mapping.'

> ***Mediated Mapping***
> The relationship between form and meaning in natural language is mediated by hierarchical binary-branching syntactic representations (i.e., 'tree structures').



Mediated mapping is a central doctrine in the literature on generative grammar.

> *… the correlation of sound and meaning is mediated by syntactic structure ...*
> (Jackendoff 2007:3)

> *A crucial part of understanding a sentence is to construct its syntactic structure.*
> (van Gompel & Pickering 2007:289)

> *Structural relations such as c-command, expressed in hierarchical sentential representations, determine all sorts of formal and interpretive properties of sentences: agreement and other morphosyntactic properties, the binding of anaphors and other aspects of referential dependencies, etc.*
> (Rizzi 2016: 338)

> *… any theory of [generative grammar] must assume the existence of a computational system that constructs hierarchically structured expressions…*
> (Chomsky, Gallego & Ott 2019:232)

In contrast, I will seek to advance the case for a more direct relationship between form and meaning – the thesis of direct mapping.

> ***Direct Mapping***
> The mapping between form and meaning does not require the mediation of syntactic representations.

FORM		MEANING
Harry left	⇐⇒	LEAVE <*h*>

The principal consequence of this proposal is the elimination of syntactic structure as a level of representation.

> *Only two levels of processing are specified: a functional level (where all the meanings and intentions to be expressed in an utterance are represented) and a formal level (where the surface forms appropriate for a given meaning/intention are represented). Mappings between the formal and functional levels are … direct.*
> (MacWhinney, Bates & Kliegl 1984:128)

> *[Language] maps a string of words directly onto a semantic representation without the mediation of grammatical principles or syntactic structure.*
> (O'Grady 2015:102)

> *Syntactic representations are neither computed during comprehension nor in production.*
> (Christiansen & Chater 2016:17)

We will return to this claim in much more detail in the chapters that follow.

3.2 Algorithmic orientation

In influential work, Marr (1982) proposed that cognitive systems can be studied at three different levels of analysis.

Marr's three levels of analysis:

- *The computational level describes the goal(s) of the system, the information that it manipulates and the constraints that it must satisfy.*
- *The algorithmic/representational level is concerned with the representations and data structures involved and the algorithms that manipulate these representations.*
- *The implementational/physical level focuses on how the system is physically realized (e.g., what neural structures and neuronal activities implement the visual system).*

(Marr 1982; see also Johnson 2017)

David Marr (1945-1980)

Generative grammar offers a computational-level theory: it is concerned solely with the study of language as a system of knowledge, not with the question of how that knowledge is put to work in the course of speech and comprehension.

> *[F]inding algorithms by which Chomsky's theory may be implemented is a completely different endeavor from formulating the theory itself. In our terms, it is a study at a different level, and both tasks have to be done.*
> (Marr 1982:28)

> *Linguistic theories are computational-level theories of language, while psycholinguistic theories of comprehension or production are algorithmic-level descriptions of how knowledge of language can be put to use.*
> (Johnson 2017:172)

On Chomsky's view, here is what a generative grammar does:

> *To avoid what has been a continuing misunderstanding, it is perhaps worthwhile to reiterate that a generative grammar … attempts to characterize in the most neutral possible terms the knowledge of the language that provides the basis for actual use of language by a speaker-hearer. When we speak of a grammar as generating a sentence with a certain structural description, we mean simply that the grammar assigns this structural description to the sentence...*
> (Chomsky 1965:9)

To make the point clear, Chomsky proposes an analogy with arithmetic.

> *… the algorithm level doesn't exist for [arithmetic]. It's the same with language. Language is kind of like the arithmetical capacity. There's some system in there that determines the sound and meaning of an infinite array of possible sentences. But there's no question about what the algorithm is.*
> (Chomsky 2012a; see also Momma & Phillips 2018:235)

> *… a rule of multiplication is described as if it's an action but it isn't. The same for generative processes; there's no algorithm. There's an implementation, but no algorithm.*
> (Chomsky 2013, from a talk entitled 'After 60+ years of generative grammar: A personal perspective,' given at Princeton University. Available at: http://www.youtube.com/watch?v=Rgd8BnZ2-iw; the excerpt above occurs at 1:28:56)

> *When we say that a sentence has a certain derivation ..., we say nothing about how the speaker or hearer might proceed, in some practical or efficient way, to construct such a derivation. These questions belong to the theory of language use – the theory of performance.*
> (Chomsky 1965:9)

As traditionally conceived, the study of grammatical systems informs and guides a separate research program devoted to the investigation of 'performance' – the use of language to produce and understand speech.

> *[Parsing and perception] have their own mechanisms, and can access unbounded external resources, but in doing so they surely access the generative mechanisms ...*
> (Chomsky 2015:95-96)

This view is widely adopted in psycholinguistic work as well.

> *… the most basic assumption about the nature of the human sentence processor is that it obeys the fundamental principles of grammar.*
> (Ferreira, Christianson, & Hollingworth 2001:13)

> *A syntactic representation is developed essentially as each word is encountered…*
> (Frazier 1998:126; see also Hagoort, Brown & Osterhout 1999:275 and Pickering 1999:124)

> *Because the correlation of sound and meaning is mediated by syntactic structure, the processor must … develop enough syntactic structure in both perception and production to be able to make the relation of sound and meaning explicit.*
> (Jackendoff 2007:3)

> *When sentences are processed, speakers and listeners build syntactic structures…*
> (Traxler 2011:355)

> *Humans use their linguistic knowledge in at least two ways: on one hand, to convey what they mean to others or to themselves, and on the other hand, to understand what others say or what they themselves say. In either case, they must assemble the syntactic structures of sentences in a systematic fashion, in accordance with the grammar of their language.*
> (Momma & Phillips 2018:234)

In contrast, the Strict Emergentist Protocol focuses on the algorithmic level; there is no grammar.

> ***Algorithmic Orientation***
> The mapping between form and meaning is executed by processing algorithms operating in real time in the course of speaking and understanding.

Direct Mapping and Algorithm Orientation fit well together: the first thesis denies the existence of syntactic structure, and the second rejects the mechanism (generative grammar) that produces representations of this type in the first place.

3.3 Processing determinism

Taken together, direct mapping and algorithm-level analysis virtually force a third thesis: in the absence of grammar, the properties of language must be explained in some other way. I propose that the explanation lies in the effect of processing pressures.

> ***Processing Determinism***
> The properties of algorithms are shaped by processing considerations.

Two types of forces seem to be in play – one internal and the other external.

Internal forces

Internal forces are focused on minimizing the cost of processing operations.

> *… the [processor] should operate in the most efficient manner possible, promptly resolving dependencies so that they do not have to be held any longer than necessary. This is a standard assumption in work on processing, where it is universally recognized that sentences are built in real time under conditions that favor quickness.*
> (O'Grady 2005:7)

> *Speed in communicating the intended message from speaker to hearer and minimal processing effort in doing so are the two driving forces of efficiency…*
> (Hawkins 2014:48; see also Hawkins 2004:9)

As we will see in due course, speed and efficiency can be achieved in a variety of ways. One is to favor form-meaning mappings that minimize the burden on working memory.

> *… there is an advantage to reducing the burden on working memory, whatever its nature and whatever its capacity … the effects of this advantage can be discerned in the way that sentences are built.*
> (O'Grady 2005:6)

> *The conclusion I derive from much of the working memory literature, and from comparisons of different domain sizes within and across languages, is simply that the more items there are to process … the harder it is – i.e., … processing difficulty … increases where there are more forms and their properties to process and hold in working memory…*
> (Hawkins 2014:232)

Another way to improve efficiency is to enhance the prospects for predicting what lies ahead in the sentence.

> *[Another] big idea involves frequency, predictability, expectedness, and surprisal. There are different terms here, and differences in detail, but the basic intuition seems to be much the same: the more expected something is, the easier it is to process.*
> (Hawkins 2014:232)

> *Current psycholinguistic theories emphasize the importance of prediction for efficient language comprehension. This trend dovetails with the broader treatment of cognitive agents as "prediction engines" that anticipate the future and update knowledge databases when their prediction are disconfirmed.*
> (Ferreira & Chantavarin 2018:443)

External forces

In contrast, external pressures arise from factors manifested in experience, including the relative frequency of particular items and patterns in the speech of others. This too makes sense: the more frequently a word or pattern is heard and used, the stronger and more accessible the associated processing operations become.

> *The processes that create templates … automatically give greatest associative strength to those aspects of analysis that occur most frequently. This has the*

consequence that broad-ranging canonical sentence patterns will be the most strongly confirmed.
(Townsend & Bever 2001:175)

A word or … syntactic construction … must reach a certain activation threshold in order to become available … The activation … will fluctuate as a consequence of its recency and frequency of use…
(Paradis 2004:29)

Repeated exposure to a particular [linguistic] pattern … increases [the] speed and fluency of processing of the pattern.
(Bybee & McClelland 2005:396)

Every time a target is retrieved, its activation increases.
(Jaeger & Tily 2011:325)

… frequency information is an important source of predictive processing in native speakers.
(Kaan 2014:262)

… automatization is a gradual process driven by frequency.
(Diessel 2015:315)

… the more frequently a construction is used, the easier it becomes to process.
(Imamura, Sato & Koizumi 2016:2)

The drive for efficiency, whether it is the product of internal or external forces (or both), has the effect of directing language learners and language users down particular computational paths, creating processing routines that mediate the association between form and meaning and shape much of what we think of as 'syntax.' I will pursue this idea in some detail, starting in the next chapter. First, though, it is necessary to avert a potential confusion involving the goals of the minimalist and emergentist programs.

An overlap in terminology

A key claim of the Strong Minimalist Thesis is that language is shaped by the demands of 'efficient computation.' (In the generative literature, these demands are sometimes dubbed 'third-factor effects,' in contrast to the constraints imposed by UG and the information gleaned from experience.)

… the principles of language are determined by efficient computation.
(Berwick & Chomsky 2011:30)

A primary goal of linguistic theory … has been to try to reduce UG assumptions to a minimum ... There have been two approaches to this problem: one seeks to reduce or totally eliminate UG by reliance on other cognitive processes; the

> *second has approached the same goal by invoking more general principles that may well fall within extra-biological natural law, particularly considerations of minimal computation...*
>
> (Chomsky 2011:263; see also 2005:9.)

> *The optimal solution would be that UG reduces to the simplest computational principles, which operate in accord with language-independent conditions on computational efficiency.*
>
> (Chomsky 2017:296)

The mention in these passages of 'cognitive processes,' 'simplicity' and 'efficiency' overlaps with the terminology used in the emergentist literature to describe processing.

> *… efficiency is the driving force behind the design and operation of the computational system for human language.*
>
> (O'Grady 2005:2)

> *… efficiency relates to the most basic function of languages which is, as I see it, to communicate information from the speaker (S) to the hearer (H): Communication is efficient when the message … is delivered … in rapid time and with the most minimal processing effort that can achieve this communicative goal.*
>
> (Hawkins 2014:230-31)

At times, it *appears* as though Chomsky is referring to real-time processing when he discusses computation, as he makes reference to memory, forgetting, load, cost and the like.

> *We therefore hope to be able to establish a "Phase Impenetrability Condition," which guarantees that mappings to the two interfaces can forget about what they have already done, a substantial saving in memory.*
>
> *…*
>
> *What objects constitute phases? They should be as small as possible, to minimize computational load.*
>
> (Chomsky 2005:16-17)

> *[The mapping to the conceptual-intentional interface] is universal, hence in effect instantaneous and costless.*
>
> (Chomsky 2007, note 17)

> *Nothing more than phase-level memory is required to identify [certain] properties at the semantic interface, where the information is required.*
>
> (Chomsky 2008:145)

> *… minimization of computation calls for erasure of all but one copy, so that the phonological component can forget about the others.*
> (Chomsky 2008:146)

Despite the allusions to memory and forgetting, to cost and to speed, Chomsky is not proposing a model of how sentences are produced and comprehended; his notion of computational efficiency cannot be equated with processing cost, which he sometimes refers to as 'communicative efficiency' (e.g., 2011:275).[1] As previously noted (§3.2), Chomsky's proposals apply to a system of knowledge, viewed abstractly and sometimes described metaphorically, not to the algorithms that create form-meaning mappings in real time, as called for by the Strict Emergentist Protocol.

The three components of the Strict Emergentist Protocol (direct mapping, algorithmic orientation and processing determinism) define what might be called a 'natural syntax' for human language. Taken together, they ensure that the mapping between form and meaning – the essential activity of language – is shaped and constrained by factors such as memory, prediction and processing cost that have a natural and well established role in cognition. In this, they differ from the formal laws of grammar on which most analytic work in contemporary syntax is based. I will return often to this theme as we proceed through the chapters that lie ahead.

---

[1] Both uses of the term 'computational' are appropriate in their own right, since a computation is simply an operation on a representation.

4

Mapping

Of all the tenets of traditional linguistics, none is more widely accepted or fiercely defended than the idea that languages have a *grammar* – a system of 'rules' that determine the acceptability and structure of particular sentences. Indeed, this has long been the first goal of work in generative linguistics.

> *The fundamental aim in the linguistic analysis of a language L is to separate the grammatical sequences which are the sentences of L from the ungrammatical sequences which are not sentences of L and to study the structure of the grammatical sequences.*
> (Chomsky 1957:13)

This idea is inconsistent with the type of theory called for by the Strict Emergentist Protocol, which adopts an algorithmic orientation and a commitment to processing determinism within a framework that provides a direct mapping between form and meaning, without the mediation of syntactic structure.

> ***Direct Mapping***
> The mapping between form and meaning does not require the mediation of syntactic representations.
>
> ***Algorithmic Orientation***
> The mapping between form and meaning is executed by processing algorithms operating in real time in the course of speaking and understanding.
>
> ***Processing Determinism***
> The properties of algorithms are shaped by processing considerations.

My goal in this chapter is to outline the first of several steps needed to implement a system of form-meaning mapping that complies with the Emergentist Protocol.

4.1 The basics of mapping

There can be no doubt that the use of language requires a cognitive system that brings together form and meaning in a way that is compatible with the practices of other members of a speech community. It has even been suggested that this system is *by definition* a generative grammar.

> *To question the validity of generative grammar is to hold that there is no specific sound-meaning correlation … that differentiates, say, English and*

> *Finnish. Since no one believes that, there can be no serious question about the validity of generative grammar, though of course that leaves open the form that it might take.*
> (Chomsky 2015:93)

This unique conceptualization of generative grammar notwithstanding, there is a clear-cut distinction in the linguistic literature between a grammar and a processor which aligns roughly with the contrast between Marr's computational and algorithmic levels (see §3.2).

> *[Theories of grammar] are concerned with what the [sentence's] structure is for either speaker or hearer. Processing theories are concerned with how the structure is built in real time…*
> (Jackendoff 2002:57 & 31)

Moreover, methodologically, it makes perfect sense to call into question the need for a grammar.

> … *there must be psychological mechanisms for speaking and understanding, and simplicity considerations thus put the burden of proof on anyone who would claim that there is more than this. To defend the more traditional view, what is needed is some sign of life from the postulated mental grammar.*
> (J.D. Fodor 1978:470)

The challenge for a theory of natural syntax – and for linguistic emergentism in general – lies not so much in distinguishing between a grammar and a processor as it does in confronting the widely held view that grammar is essential to an explanatory theory of language. My plan is to start, in this chapter and the next two, with a very simple set of facts involving canonical word order – arguably the most fundamental typological feature of language.

Representing form and meaning

The sort of algorithms that are needed to satisfy the Strict Emergentist Protocol assume the existence of a representation of sound (gesture in the case of sign language) and a representation of meaning. I will have little to say here about the representation of sound, for which a string of written words will stand as a proxy. Moreover, for the most part, I will make use of very simple semantic representations that contain little more than information about predicates and their arguments. (In the examples below, predicates are represented in upper case and arguments in italicized lower case, along the lines illustrated below; see Kroeger 2018:67-68 for a similar notation.)

Form	Meaning	
Harry left.	LEAVE	
	<*h*>	(*h* = *Harry*)
Robin studies math.	STUDY	
	<*r m*>	(*r* = *Robin; m* = *math*)

A full semantic representation must of course include information about other things, including tense, aspect, modality, number, gender, definiteness and the like. As the need arises, I will occasionally add information of this type. Nonetheless, as we will see, the form-meaning mappings underlying a very substantial number of syntactic phenomena require reference to little more than the pared-down representations I'll be using here.

Semantic representations do not come ready-made of course; they must be built. My key proposal in this regard is that all mappings between form and meaning start with a simple template, called a *semantic base*, whose presence can be reliably predicted in essentially every sentence that is produced or encountered in any language. It consists of a predicate position and a single argument position, as depicted below, with PRED standing for *predicate* and the symbol β representing the predicted argument (henceforth the *base argument*).

The semantic base
PRED
<β>

The essential claim is simply this:

The Semantic Base Hypothesis
In both production and comprehension, the processor takes the semantic base as its starting point.

Proceeding in this way contributes to predictability, a powerful force in both production and comprehension, as noted in the previous chapter (§3.3).

> *… both production and comprehension involves extensive use of prediction.*
> (Pickering & Garrod 2013:332)

> *Predictive mechanisms are likely important for the robustness of language understanding ... Prediction may also play a key role in learning: … if a learner can use [his/her] current knowledge of the language and the context to predict how the sentence will unfold, then the comparison of what [s/he] expects with what actually occurs could provide valuable additional information.*
> (Phillips & Ehrenhofer 2015:414)

Put simply, in both speech and interpretation, the processor benefits from knowing what lies ahead. In the case at hand, this means being able to proceed with the confidence that every clause will have as its foundation a semantic representation consisting of a predicate and at least one argument.

In order to pre-empt possible misunderstandings, three additional points are in order:

- The base argument position is thematically and topically unspecified. Depending on the choice of predicate, it could ultimately house an agent, a patient or some other argument type. Moreover, depending on the context, the base argument could be used to convey new or old information; it could be definite or indefinite; and so on.
- The base argument need not be overtly expressed. In languages that allow null arguments (so-called *pro*), the base argument need not be phonetically realized but is nonetheless present in the semantic representation.
- There may be a small number of verbs in some languages that, contrary to the prediction of the Semantic Base Hypothesis, literally have no argument (so-called ‘zero valency’); Spanish *llover* ‘to rain’ is often mentioned as an example. However, this will be a rare occurrence, and the cost of correcting the failed prediction (by deleting the default argument position) is too minor to be disruptive.

With these ideas in mind, we can now consider some actual examples. I will initially focus on extremely simple syntactic phenomena – mostly word order in canonical intransitive and transitive sentences, for which there are innumerable alternative analyses. However, the stakes will be raised significantly as we progress through the next several chapters.

4.2 Simple intransitive patterns

I will begin by considering the simple syntax of clauses that consist of an intransitive verb and a sole argument, often informally called its ‘subject.’ (I use the terms ‘subject’ and ‘direct object’ for the purposes of descriptive convenience only; the use of grammatical relations in any other way would be incompatible with the thesis of Direct Mapping.)

Two algorithms

Let us take as our first example the English sentence *Harry left*, which has the (simplified) form and meaning repeated below.

Form:	Meaning:
Harry left.	LEAVE <*h*>

Two mapping algorithms are needed here, one to identify the nominal *Harry* as the predicate’s only argument and another to associate the verb *left* with the predicate function.

As a first approximation, these algorithms can be formulated as follows.[1] (The symbol ↦ stands for ‘is mapped onto.’)

- The First-Argument Algorithm:

  Map the referent of nominal$_x$ onto the first-argument position.

  N$_x$ ↦ PRED
  <**REF$_1$**>

- The Predicate Algorithm:

  Map the event denoted by the verb onto the predicate position.

  V ↦ **EVENT**
  <…>

Here is how these algorithms go about mapping the form *Harry left* onto the appropriate semantic representation.

Harry left.

Step 0: Projection of the semantic base
PRED
<β>

Step 1: First-Argument Algorithm
Harry… ↦ PRED
<***h***>
(The referent of the nominal is mapped onto the first-argument position.)

Step 2: Predicate Algorithm
Harry ***left*** ↦ **LEAVE**
<*h*>
(The event denoted by the verb is mapped onto the predicate position.)

Used in this way, the two algorithms make up a *processing routine* that helps define a language’s sentence-level syntax. Most obvious for now is the fact that the order in which the algorithms are activated reflects (and determines) a sentence’s word order. Thus, in the example above, activation of the First-Argument Algorithm before the Predicate Algorithm reflects the fact that English is a ‘subject–verb’ (SV) language.

[1] Talk of mapping ‘events’ onto positions is of course really shorthand for mapping *representations* of events onto those positions.

VS patterns

Now think about a verb-initial language such as Irish.

D'imigh Harry.
Left Harry.

The VS order exemplified here reflects the order in which the two algorithms are activated.

D'imigh Harry. Left Harry.

Step 0: Projection of the semantic base
PRED
<β>

Step 1: Predicate Algorithm
D'imigh… ↦ **LEAVE**
<β>

Step 2: First-Argument Algorithm
D'imigh ***Harry*** ↦ LEAVE
<***h***>

As these examples illustrate, the mapping operations for English and Irish are identical. They differ only in the order in which they apply, in accordance with the language's choice of linearization option – SV or VS.

	SV routine	VS routine
Step 1:	Map 1st argument	Map verb
Step 2:	Map verb	Map 1st argument

This idea offers a natural way to make sense of word order: it is simply a reflection of the order in which algorithms are activated.

Now consider Spanish, in which both SV and VS orders are permitted.

Harry partió.	Partió Harry.
Harry left.	Left Harry.

The choice of word order here reflects a pragmatic distinction: the VS order is reserved for situations that place the referent of the subject in focus, as in the answer to a question such as 'Who left?' (e.g., Bolinger 1954, Contreras 1978). This suggests that the choice of processing routine is influenced by context and intent, a point that has long been acknowledged in one form or another.

> *It is generally recognized that sentence-level constructions … are associated with their own information structure properties.*
>
> (Goldberg 2006:429)

4.3 Algorithms for production

Although I will focus in this book on the mapping of form to meaning rather than vice versa, it is worth noting that speech production also invites an algorithmic treatment. Indeed, in the simplest case, production and comprehension could be the product of the same algorithms, running in opposite directions.

Two production algorithms

- The First-Argument Algorithm:

  Map the first argument onto nominal$_x$.

  PRED $\mapsto$ $\mathbf{N_x}$
<REF1>

- The Predicate Algorithm:

  Map the predicate onto the verb$_e$.

  PRED $\mapsto$ $\mathbf{V_e}$
< …>

Here is an example of how this might work for the sentence *Harry left*.

Production routine for *Harry left.*

Step 0: Meaning to be encoded

LEAVE
<*h*>

Step 1: First-Argument Algorithm for Production

LEAVE $\mapsto$ ***Harry*** …
<***h***>
(The first argument is mapped onto a nominal.)

Step 2: Predicate Algorithm for Production

LEAVE $\mapsto$ *Harry* ***left.***
<*h*>
(The predicate is mapped onto a verb.)

The close parallels between comprehension and production are consistent with the view that the two uses of language draw on essentially the same processing mechanisms.

> *I take the computational system … to be identical in the relevant respects for both production and comprehension ...*
>
> (O'Grady 2005:7)

> *Traditional accounts of language assume separate processing "streams" for production and comprehension. They adopt … a perspective that is incompatible both with the demands of communication and with extensive data indicating that production and comprehension are tightly interwoven.*
> (Pickering & Garrod 2013:346)

> *… fundamental properties of parsing and generation mechanisms are the same and … differing behaviors reduce to differences in the available information in either task.*
> (Momma & Phillips 2018:248)

I leave open the question of the extent to which the mirror-like symmetry evident in the production and comprehension of simple intransitive patterns might hold more generally.

4.4 Algorithms and 'grammaticality'

The standard view of the relationship between grammar and processing within generative linguistics is that the processor draws on procedures that implement grammatical rules and principles.

> *Most theories of sentence processing today incorporate the claim that parsing is both fast and grammatically controlled… Each new word is incorporated into the current representation in accord with the grammar.*
> (Frazier 1998:126).

> *… the most basic assumption about the nature of the human sentence processor is that it obeys the fundamental principles of grammar.*
> (Ferreira, Christiansen, & Hollingworth 2001:13)

> *… humans have available to them two systems for interpreting natural language. One system … pairs a syntactic form with its interpretation using grammatical rules of composition… The other … uses the grammar together with knowledge of how the human production system works.*
> (Frazier 2015:7)

This division of labor allows the grammar to be responsible for well-formedness, leaving the processor with the job of expressing and interpreting sentences that satisfy the 'grammaticality filter.'

My view is substantially different. Consistent with the principles of direct mapping and algorithmic orientation, I treat 'ungrammaticality' simply as a failure of mapping. The key idea, to be discussed in more detail in chapter 13, is that the repeated activation and use of particular algorithms and sequences of algorithms creates processing routines, which come to serve as well-worn paths for mapping a form of a certain type onto a corresponding semantic representation. With time and experience, processing routines become entrenched and inflexible to the point where alternatives are no longer countenanced.

> *Once strengthened to the point where they are more or less fixed, … routines have the de facto effect of defining a set of acceptable sentences.*
> (O'Grady 2005:212)

Other options – such as a sentence with VS order in English – are rejected as unnatural, creating the impression of unacceptability or 'ungrammaticality' that helps distinguish English from, say, Irish or Spanish.

This leaves the grammar without a job to do in at least one of the areas (word order) for which it has traditionally been assigned responsibility. Returning to Fodor's criterion ('what is needed is some sign of life from the postulated mental grammar'), we must ask whether the grammar, as traditionally construed, will show signs of life in other domains. If the Strict Emergentist Protocol is on the right track, any such signs should turn out to be illusory.

In sum, at least for the cases considered in this chapter, the traditional functions of the grammar – building syntactic structure and adjudicating grammaticality – have been set to the side in favor of algorithms that directly map form onto meaning and vice versa. The point of proceeding in this way is not to eliminate syntax; it is to uncover deeper explanations for why syntactic phenomena have the particular properties that they do. In this regard, three claims lie at the heart of an algorithmically oriented approach to mapping.

- Algorithms provide a direct mapping between form and meaning, without the mediation of syntactic structure and without reference to grammatical principles.
- The order in which the algorithms are activated reflects and determines a sentence's word order.
- The entrenchment of particular algorithms has the effect of licensing certain form-meaning mappings to the exclusion of others, thereby creating a standard for assessing a sentence's acceptability.

The next chapter seeks to extend these ideas to transitive patterns.

5

Verb–Object Languages

A survey of basic transitive patterns in 1377 languages reveals that the two most common word orders by far are SOV and SVO, which together represent almost 90% of world's languages. (The estimates below exclude 189 languages for which no basic word order has yet been determined.)

Order of subject, verb and direct object

SOV	565	(47.5%)
SVO	488	(41.1%)
VSO	95	(8.0%)
all others	40	(3.4%)

(Dryer 2011, ch. 81)

I will focus in this chapter on SVO and VSO languages with a view to outlining how three simple algorithms interact to create the form-meaning mappings associated with these types of sentences. The next chapter will deal with SOV languages.

5.1 Extending the semantic base

The defining feature of transitive patterns is that they are built around a verb that takes *two* core arguments.[1]

PRED
<$\underline{1}$ $\underline{2}$>

Given its argument structure, transitivity clearly requires an extension to the semantic base, which consists of just a predicate and one argument.

PRED
<β>

There are two obvious ways to extend the semantic base. One option is to add a second-argument position, allowing the base-argument to maintain its status as first argument.

[1] For the sake of exposition, I will define a core argument as an argument whose relationship to the verb is not mediated by a preposition. The prototypical core arguments are agents and patients.

Argument Addition (Add a second-argument position)

PRED ⇒ PRED
<β> <β _>
↑
added argument position

The other option is to extend the semantic base in the opposite direction by adding a first-argument position.

Argument Addition (Add a first-argument position)

PRED ⇒ PRED
<β> <_ β>
↑
added argument position

Both options yield the same result – a predicate with two argument positions.

Two notes of caution are in order before proceeding.

- Argument addition has nothing to do with thematic roles. The operation that creates a first-argument position does not add an agent argument, and the operation that creates a second-argument position does not add a patient argument. These operations create *positions* only.
- The thematic role of the arguments that ultimately end up in those positions is determined by the semantics of the predicate. For example, *hunt* has the type of meaning that calls for an agent as its first argument, whereas *fear* implies an experiencer. (*ag* = agent, *exp* = experiencer)

HUNT	FEAR
<ag …>	<exp …>
Harry hunts tigers.	*Harry fears tigers.*

It is true, of course, that the first argument of a transitive verb tends to be an agent and the second argument is almost always a patient. However, this is a fact about transitive predicates in general, not about the operations that make room for an additional argument position.

As we will see as we proceed, the strategy used to extend the semantic base has important consequences for a language's morphosyntax. I will return to this matter in chapter 10. In the meantime, I will be focusing on languages, including English, that expand the semantic base by adding a second-argument position.

SVO patterns

Bringing together form and meaning in a transitive pattern requires two operations. The first is a procedure that expands the semantic base to create an additional argument position. The second is an algorithm that maps the second argument onto that newly added position.

The Second-Argument Algorithm

Map the referent of nominal$_y$ onto the second-argument position.

$N_y \mapsto$ PRED
<_ **REF$_2$**>

In an SVO language such as English, the Second-Argument Algorithm applies after the Predicate Algorithm, reflecting (and ensuring) that the second argument is post-verbal. Here is an example.

Harry plays football.

Step 0:	Projection of the semantic base PRED <β>
Step 1:	First-Argument Algorithm ***Harry***… ↦ PRED <***h***>
Step 2:	Predicate Algorithm + Argument Addition *Harry* ***plays***… ↦ **PLAY** <*h* **_**> *added argument position*
Step 3:	Second-Argument Algorithm *Harry plays* ***football*** ↦ PLAY <*h* ***f***>

As illustrated here, encountering the verb in Step 2 has two effects:

- It activates the Predicate Algorithm, which identifies the event denoted by the verb as the predicate PLAY.
- It triggers an operation ('Argument Addition') that creates a second-argument position to accommodate the argument structure called for by a transitive predicate.

Now let's consider how these same procedures might work in a language with VSO order.

VSO patterns

For a VSO language such as Irish, the order of the first two algorithms is reversed, ensuring that the verb is the first component of the sentence to be mapped onto the semantic representation.

	SVO routine (English)	**VSO routine (Irish)**
Step 1:	Map 1st argument	**Map verb**
Step 2:	Map verb	**Map 1st argument**
Step 3:	Map 2nd argument	Map 2nd argument

Here is an actual example from Irish.

Imríonn Harry peil. Plays Harry football.

Step 0: Projection of the semantic base
PRED
<β>

Step 1: Predicate Algorithm + Argument Addition
Play… $\mapsto$ **PLAY**
<β **_**>
added argument position

Step 2: First-Argument Algorithm
Play ***Harry…*** $\mapsto$ PRED
<***h*** _>

Step 3: Second-Argument Algorithm
Play Harry ***football*** $\mapsto$ PLAY
<*h* ***f***>

Here Step 1 is crucial, as it involves both the identification of the predicate (the Predicate Algorithm) and the creation of a second argument position to accommodate its transitivity (Argument Addition).

Some VSO languages permit a VOS variant, and languages such as Malagasy even have a preferred VOS order to begin with.[2]

[2] It has been suggested that there may be two types of VSO languages – those like Irish, in which just the verb occurs before the subject, and those like Niuean, in which the verb and at least some of its dependents appear in a pre-subject position. For discussion, see Clemens & Polinsky (2017).

Malagasy:
Manao mofo Harry.
bake bread Harry
'Harry baked bread.'

Even in these cases, however, the consensus view is that the agent still occupies the first position in argument structure (a matter to which I return in chapter 9).

BAKE
<**ag** pat>

This makes ample sense from a processing perspective. As the instigator of the action denoted by the verb, the agent is the starting point of the event, as widely recognized in the semantic literature.

> … *the agent is at the head of the causal chain that affects the patient.*
> (Kemmerer 2012:50)

Moreover, patienthood often entails prior agency: an entity cannot become a patient until an agent has acted upon it (Schlesewsky & Schlesewsky 2009:41). For instance, in the event described by the sentence *The cat scratched the dog*, the patienthood of the dog depends on prior action by the cat. The ubiquity of this asymmetry suggests that events are systematically conceptualized in a way that treats the agent as the first argument.

In sum, what makes VOS patterns unusual is that the order in which the argument algorithms are activated need not (and does not) align with the internal organization of argument structure.[3] Put simply, depending on the language, the Second-Argument Algorithm can be activated before or after the First-Argument Algorithm.

VOS order

Step 1: Map verb.
Step 2: Map 2nd argument.
Step 3: Map 1st argument.

5.2 Signs of processing determinism

So far, I have said nothing about the role of processing determinism in shaping the algorithms that I have been describing – other than to suggest that these procedures are strengthened and entrenched through repeated activation (see §4.3). However, there are internal pressures at work here too. One such effect shows up in the positioning of various 'non-nuclear' parts of sentences, such as prepositions, possessors (genitives) and modifiers.

[3] This does not rule out the possibility that an alignment match might be the default state of affairs within and across many languages.

It has been known for many decades that interesting correlations exist with regard to the placement of these items.

Preferred Word Orders in Verb–Object Languages

- Prepositions rather than postpositions
- N + Adjective
- N + Possessor
- N + Relative Clause

(Greenberg 1963:61ff)

John Hawkins has proposed that these preferences reflect a processing-driven propensity to minimize the distance between the verb and particular items to which it is semantically related.

> ***Minimize Domains***[4] (paraphrased)
> The human processor prefers to minimize the domain in which relations of dependency are processed.
> (Hawkins 2004:123)

Take for instance a verb of motion such as *go*, *walk* or *dash*, which calls for a preposition denoting direction such as *to* or *toward*. As illustrated below, domain minimization favors prepositions over postpositions in SVO languages so as to minimize the distance between the verb and the direction-marking element. (The statistical data cited below is from Hawkins 2004:224.)

Preferred	***Dispreferred***
Preposition pattern	Postposition pattern
walk ***to*** *the door*	*walk the door* ***to***
V P	V P
↑___↑	↑___________↑
(observed in 161 languages)	(observed in just 18 languages)

Another example involves the positioning of adjectives and possessors. In order to minimize the distance between a verb and the noun representing its patient argument, we would expect them to occur after the noun rather than between the verb and the noun. Malay works this way. (English is unusual in this regard; I'll talk about that shortly.)

[4] A very early version of this insight is evident in Behaghel's First Law: 'Conceptually related entities are placed close together' (Behaghel 1932). The calculus underlying Hawkins' principle is somewhat different from what I have in mind, since he focuses on calculating the number of words needed to recognize the phrases of which they are a part (2014:11ff). There are no phrases in the system of mapping that I propose.

Malay:
beli kereta ***lelaki itu***
buy car man the
'buy the man's car'

Preferred	***Dispreferred***
buy car man the	buy man the car
V N Possessor	V Possessor N
↑___↑	↑_____________↑
(observed in 134 languages)	(observed in just 14 languages)

These positioning facts are expected in a syntax shaped by processing pressures.

> *If a particular property can be derived from a smaller rather than a larger domain, then less effort will be expended in its processing, and processing can be faster… Simultaneous processing and working memory loads are reduced when domains are small.*
> (Hawkins 2004:27; see also 2014:13)

An unexpected pattern

Evolution does not produce perfect creatures, and natural syntax does not create perfect languages. It is therefore no surprise that some languages allow mappings that could have been less costly than they actually are. A striking example comes from Mandarin, which allows relative clauses to intervene between a verb and the noun representing its patient argument.

yudao **[zuotian hui jia de]** pengyou
run.into yesterday go home LNK friend
↑__________________________________↑
'encountered friends who went home yesterday'

The positioning of the relative clause in this pattern creates a very substantial distance between the verb and the noun. Given domain minimization, it would have made more sense to place the relative clause after the noun, as English does.

met friends **[who went home yesterday]**
↑____↑

And, in fact, 99% of all verb–object languages do place relative clauses after the noun that they modify (Hawkins 2014:102). But Mandarin, with its billion-plus speakers, does not.

How could this have happened? There is probably no way to know for sure. The more immediate question has to do with whether and how this typologically unusual pattern might fit into the language's overall syntax in a way that best accommodates the usual processing pressures. The compromise lies in a calculus of 'affordability' that creates an implicational scale for managing departures from frugality.

Affordability
If a language permits a more costly operation, it should also permit a less costly operation of the same type.
(Hawkins 2004:129-30, 256ff; O'Grady 2005:214, 2016:43)

The underlying logic is comparable to that used by consumers: if you're willing to buy a five-dollar cup of coffee, you'll also be willing to buy a three-dollar cup. Here's what this means for Mandarin.

Relative clause are on aggregate longer and more complex categories than possessive phrases, which are on aggregate longer than single-word adjectives. If a language permits the relative clause to intervene between [V] and N, it permits the less complex possessive phrase, and if it permits the latter it generally permits single-word adjectives.
(Hawkins 2004:129)

Calculus for Domain Minimization

Single-word adjective > Possessor > Relative clause

And, sure enough, that's just the way things work in Mandarin. Not only can relative clauses intervene between a verb and the associated noun, so can a possessor and a single-word adjective.

Intervening possessor in Mandarin:

yu dao	**nage**	**nuhai**	**de**	pengyou
visit	that	girl	LNK	friend

'visit that girl's friend'

Intervening single-word adjective in Mandarin:

yu dao	**hao**	pengyou
visit	good	friend

'visit a good friend'

A less spectacular example of the same sort of effect can be seen in English, where a possessor can intervene between a verb and the noun representing its patient argument.[5]

Intervening possessor in English:
visit **that girl's** friends
↑______________↑

[5] The complexity of the possessor is a factor here: long possessors like *the right honorable gentleman* are more likely to occur after the possessee (*the policy of the right honorable gentleman*) than before it, according to a corpus study conducted by Rosenbach (2005). The same is true for adjectives; compare *a tall man* with *a man taller than Jerry*.

This may not be optimal, but it does comply with the logic underlying Affordability since a single-word adjective is also allowed in the pre-nominal position.

> Intervening adjective in English:
> visit **good** friends
> ↑________↑

All these contrasts illustrate an important feature of natural syntax, which is that pressures and preferences – rather than absolute laws – shape the form of sentences. As the examples above illustrate, this strategy accommodates the observed facts of language while also helping uncover the forces that help fashion and constrain the mapping between form and meaning.

5.3 Sentences and structure

According to the ideas that I have been outlining, a sentence is just what it has always appeared to be – a string of words. It has no syntactic structure, and it is mapped directly onto a semantic representation by processing algorithms, consistent with the Strict Emergentist Protocol.

This does not mean that there is no structure of any sort in language. Words still consist of syllables and morphemes. And, of course, the semantic representation has a modest internal structure, since a distinction is made between predicates and arguments. Moreover, in the case of biclausal sentences, such as *Mary thinks Harry left*, we see embedding within the semantic representation. (There will be more about this in chapters 7 and 8.)

THINK
<*m* _>
|
LEAVE
<*h*>

The key point for now is simply that *syntactic structure*, in the sense of a hierarchical representation of a sentence's component words and phrases, does not exist – contrary to the standard view.

> *Sentences are composed of smaller expressions (words and morphemes). These smaller units are composed into units with hierarchical structure…*
> (Hornstein, Nunes & Grohman 2005:7)

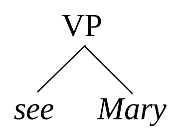
> *It is beyond dispute that hierarchical structure plays a key role in most descriptions of language.*
> (Frank, Bod & Christiansen 2012:4522)

> *Syntax is about the study of sentence STRUCTURE.*
> (Carnie 2013:72)

> *A fundamental property of syntactic structure is that words forming sentences are assembled into hierarchically organized units.*
> (Rizzi 2016:341)

If there is no syntactic structure, then of course there is also no need for combinatory operations of the sort proposed in generative grammar.

> *… any theory of generative grammar must assume the existence of a computational system that constructs hierarchically structured expressions ... The optimal course to follow … is to assume a basic compositional operation MERGE, which applies to two objects X and Y, yielding a new one, K = {X,Y}.*
> (Chomsky, Gallego & Ott 2019:232)

see, *Mary* ⇒ MERGE ⇒ 

In sum, consistent with the ideas developed in the two preceding chapters, there are no procedures to combine words into phrases and sentences. And there are no grammatical rules either. There are just algorithms that implement efficient operations for directly mapping form onto meaning and vice versa. What we think of as ‘grammatical’ sentences are just sentences for which a language’s processing algorithms can create form-meaning mappings.

It is important to acknowledge that we are still dealing with the most elementary of facts, largely related to canonical word order. I will extend the scope of our inquiry to a small degree in the next chapter by examining some basic properties of SOV languages.

6

Object–Verb Languages

Typologists generally draw a line between ‘verb-early’ SVO and VSO languages on the one hand and ‘verb-final’ SOV languages on the other.

> … *there are two, and only two, logically possible ways for a grammar to be optimally efficient …: namely when heads consistently precede or consistently follow their non-heads.*
>
> (Hawkins 2014:136)

At first glance, SOV word order appears to be dysfunctional since it postpones to the very end of the sentence its most important and informative word, the verb. This state of affairs has triggered some wry humor among speakers of SVO languages.

> … *after which [in an SOV language] comes the* VERB, *and you find out for the first time what the man has been talking about.*
>
> (Mark Twain, ‘The awful German language,’ an essay written in 1880)

Although targeted at German, Twain’s complaint could just as easily be applied to almost half the world’s languages – raising the obvious question of how SOV languages manage to work as well as they evidently do.

At the algorithmic level, the difference between verb-final languages and their SVO and VSO counterparts is minor. In fact, it can be reduced to a difference in the order in which three simple and familiar algorithms are activated.

The order of activation of algorithms by word order type

	SVO	VSO	**SOV**
Step 1:	Map 1st argument	Map verb	**Map 1st argument**
Step 2:	Map verb	Map 1st argument	**Map 2nd argument**
Step 3:	Map 2nd argument	Map 2nd argument	**Map verb**

Nonetheless, as we will see next, the position of the verb in SOV languages has a significant impact on how the burden of mapping form onto meaning is distributed in the course of syntactic computation.

6.1 The basic syntax of SOV languages

Let's begin by considering an example from Korean that illustrates how algorithms go about mapping an SOV pattern onto an appropriate semantic representation.

Harry-ka chwukkwu-lul hanta. Harry-NOM football-ACC plays.

Step 0: Projection of the semantic base

PRED
<β>

Step 1: First-Argument Algorithm

***Harry*-NOM…** ↦ PRED
<***h***>

Step 2: Argument Addition + Second-Argument Algorithm

Harry-NOM ***football*-ACC…** ↦ PRED
<*h* ***f***>
↰added argument position

Step 3: Predicate Algorithm

Harry-NOM *football*-ACC ***play*** ↦ **PLAY**
<*h f*>

A fundamental difference between SOV languages on the one hand and SVO and VSO languages on the other shows up early in this example. In Step 2, the appearance of nominal ('football') forces the addition of a second-argument position in a predicate that has not yet been encountered, thereby identifying it as transitive. In an SVO or VSO language, in contrast, the first signs of transitivity come directly from the verb itself, which occurs before its second argument. Put simply, the two language types differ with regard to how and what they predict in the course of incremental processing.

- In SVO and VSO languages, the verb triggers a prediction about the number and type of arguments that can be expected later in the sentence.
- In SOV languages, the number and type of arguments triggers a prediction about the type of verb that can be expected later in the sentence.

A role for case

It is commonly observed that SOV languages often use case marking to distinguish among a verb's arguments.

Universal 41
If the verb follows both the nominal subject and the nominal object as the dominant order, the language almost always has a case system.
(Greenberg 1963:75)

Although Greenberg's initial generalization was based on a mere 30 languages, its essential correctness has since been confirmed by data from more than 500 languages.

Proportion of languages with case distinctions (sample of 502 languages)

SOV	SVO	V-initial
72%	14%	47%
(181/253)	(26/190)	(28/59)

(Dryer 2002:152; see also Nichols 1986, Siewerska 1996, Comrie 2011, Hawkins 2014:43)

It's easy to see why case marking is so useful in SOV languages. As illustrated in the Korean example on the previous page, an encounter with a pre-verbal argument carrying an accusative marker both predicts an upcoming transitive verb and identifies the bearer of the accusative case as that verb's second argument, thereby triggering the Second-Argument Algorithm.

> *Case marking can signal differences in [thematic roles] prior to the verb, … reducing forward-looking dependencies …*
> (Hawkins 2004:247)

> *[Case-marking] allows for the early determination of the thematic role and grammatical function of each nominal argument before the parser reaches the clause-final verb position. … overt case-marking is a strategy that compensates for the late appearance of V in SOV languages.*
> (Ueno & Polinsky 2009:679)

Consistent with these considerations, there is a 'surprisal effect' when the expectations generated by the case marking on a preverbal argument cannot be accommodated by the later-occurring verb.

> *There is a clear surprisal effect [in Japanese] when the verb ATTA 'met' is received after the NP MARY-O, because [that] verb does not take an [accusative]-marked NP as its argument.*

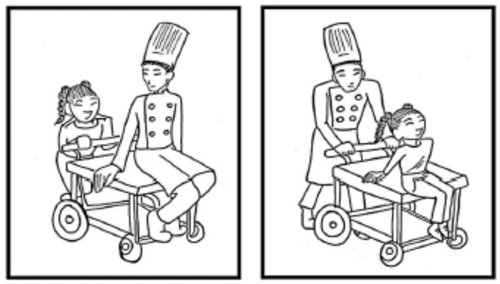
		!
Bob-ga	*Mary-o*	*atta …*
Bob-NOM	*Mary-ACC*	*meet.PST*

(Mazuka & Itoh 1995:317; see also Inoue & Fodor 1995:14 and Kamide & Mitchell 1999)

A signature feature of SOV languages with case marking is the flexibility they show in the ordering of arguments. Because case distinguishes the first argument from the second argument, the two can in principle occur in either linear order. Indeed, there is evidence from eye-tracking research that native speakers of verb-final languages can anticipate the structure of an event based just on information provided by case markers, prior to hearing the verb.

For both canonical (SOV) and scrambled (OSV) structures, [Korean] native speakers looked toward the correct image starting from the second noun.
(Frenck-Mestre 2019:399)

SOV:	*OSV:*
Sonye-ka **yolisa-lul** minta.	**Yolisa-lul** sonye-ka minta.
girl-NOM chef-ACC push	chef-ACC girl-NOM push
‘The girl pushes the chef.’	‘The girl pushes the chef.’



(Sentences and sample picture from Frenck-Mestre 2019:393)

Case has no meaning of its own and is therefore not mapped onto the semantic representation. Its function is to serve as a ‘signpost,’ helping ensure that each nominal in a sentence triggers the appropriate algorithm and ends up being associated with the proper argument position in the semantic representation. The presence of case therefore has the effect of diminishing the role of word order in the mapping process, with an important consequence that we will consider next.

6.2 The mechanics of scrambling in SOV languages

The OSV pattern in the example above – often referred to as the ‘scrambled’ option – can easily be mapped onto a semantic representation by simply reversing the activation order of the first two algorithms used for transitive clauses.

	SOV order	OSV order
Step 1:	Map 1st argument	**Map 2nd argument**
Step 2:	Map 2nd argument	**Map 1st argument**
Step 3:	Map verb	Map verb

Case plays a key role in this alternation, as illustrated in the following example involving one of the sentences from the Frenck-Mestre experiment.

Yolisa-lul sonye-ka minta. Chef-ACC girl-NOM pushes.

Step 0: Projection of the semantic base

PRED
<β>

Step 1: Argument Addition + Second-Argument Algorithm

***chef*ACC** … ↦ PRED
<β ***c***>
↰*added argument position*

Step 2: First-Argument Algorithm

*chef*ACC ***girl*NOM**… ↦ PRED
<***g*** *c*>

Step 3: Predicate Algorithm

*chef*ACC *girl*NOM ***push*** ↦ **PUSH**
<*g c*>

The key to accommodating OSV order lies in Step 1, where two things happen.

- The presence of the accusative case on the sentence-initial nominal triggers the creation of a second-argument position in the predicate (Argument Addition).
- Application of the Second-Argument Algorithm maps the sentence-initial nominal onto the newly created position.

As in the case of VOS patterns (discussed in §5.1), we once again see that the order in which the argument algorithms are activated does not have to align with their ordering in argument structure (even though that may well be the default arrangement).

A matter of usage

OSV patterns are *far* less common than their SOV counterparts in languages such as Korean and Japanese.

> … *only 1.2% of the sentences in the Korean Sejong corpus have OSV order.*
> (Kwon, Polinsky & Kluender 2006:3; see also Nam 1988)

> … *the ratio of frequencies of occurrences between the [SOV] pattern and the [OSV] pattern in Japanese is 17 to 1.*
> (Kuno 1973:353-54)

Indeed, use of the OSV pattern is largely restricted to situations in which the patient argument corresponds to the discourse topic.

> *Leftward movement [increases] the topicality of a moved phrase.*
> (Choi 1997:556; see also Choi 1999)

> *… given information is produced earlier and new information later.*
> (Ferreira & Yoshita 2002:670)

> *Speakers tend to spell out discourse-prominent items earlier than non-prominent items.*
> (Hwang-Jackson 2008:121)

As in the example of languages such as Spanish that permit either SV or VS order (discussed in §4.2), this pragmatic information is presumably ‘attached’ in some way to the processing routines that yield canonical and non-canonical order in verb-final languages.

A further consequence now follows. If it is true that frequency of activation contributes to the strength and accessibility of a processing routine (see §4.2), the less used OSV routine should be harder to access and therefore more costly than the more entrenched SOV routine. This seems to be correct.

> *In all of the experiments [participant questionnaires, eye-tracking, self-paced reading], we found that scrambled sentences cost more processing resources than canonical sentences.*
> (Mazuka, Itoh, and Kondo 2002:157)

> *… scrambled sentences should be more difficult to process than canonically-ordered sentences. And indeed, many studies have shown longer response latencies and/or higher error rates for scrambled structures in tasks involving comprehension questions or sensicality/acceptability judgments.*
> (Witzel & Witzel 2016:477)

> *… numerous studies on Japanese have observed that non-canonical [OSV] sentences … incur higher processing costs than canonical [SOV] sentences.*
> (Imamura, Sato & Koizumi 2016:2)

> *[Japanese scrambled sentences have] a non-canonical structure, which demands an additional cognitive load in a human parser.*
> …
> *[Our] result showed more activation at the left inferior frontal gyrus and the left dorsal prefrontal cortex during the comprehension of scrambled sentences than that of canonical sentences. This indicates, in accordance with previous findings on scrambling … that the parsing strategy … demands an additional cognitive activation in the brain.*
> (Kim, Koizumi, Ikuta et al. 2009:163 & 151-52)

A further observation sheds additional light on the functioning of the algorithms that are responsible for managing the expressive options permitted by verb-final languages.

> *… we found that the locus of the processing load is at the second argument of the scrambled sentence.*
> (Mazuka et al. 2002:158)

O S V
↑
locus of the processing load

Why does the spike in processing load not occur on the sentence-initial argument in OSV sentences? The answer lies in the possibility of 'subject drop' in these languages.

Transitive sentence with a null subject in Korean and Japanese

Chayk-ul	ilkessta. (Korean)	Hon-o	yonda. (Japanese)
book-ACC	read.PST.DECL	book-ACC	read.PST.DECL
'(S/he) read a book.'		'(S/he) read a book.'	

The null-subject option is widely exploited in Korean and Japanese, where it is far more commonly manifested than scrambling.

> *The average proportion of overt subjects across caregivers' speech [in Korean] … in adult-to-adult conversations … is 0.31.*[1]
> (Kim 2000:328, citing Hong 1985; see also Kwon & Sturt 2013)

Because subjects are so frequently unexpressed, many sentences require activation of the Second-Argument Algorithm as the *initial* step in the mapping process, with the identity of the first argument (represented here as '?') left to contextual inference.

Chayk-ul ilkessta. *[Kor]* ***/ Hon-o yonda.*** *[Jpn]* Book-ACC read.

Step 0: Projection of the semantic base
PRED
<β>

Step 1: Argument Addition + Second-Argument Algorithm

***book*-ACC** … ↦ PRED
*<? **b**>*
added argument position

Step 2: Predicate Algorithm

book-ACC ***read*** ↦ **READ**
<? b>

[1] Fry (2001) and Kim (2008) give even lower estimates for the use of overt subject in transitive patterns in Japanese and Korean, respectively.

For this reason, the processor is not able to discern the presence of an unusual word order until it encounters the nominative-marked argument in a position to the right of its accusative-marked counterpart – the very point at which the spike in processing difficulty is observed.

6.3 More signs of processing determinism

The previous chapter (§5.2) documented an important effect of processing determinism on the ordering of non-nuclear components of sentences (prepositions, possessors and adjectives) in verb–object languages. A parallel set of generalizations hold for object–verb languages.

Preferred word orders in Object–Verb Languages

- Postpositions rather than prepositions
- Adjective + N
- Possessor + N

(Greenberg 1963:61ff)

These facts help confirm Hawkins' observation that languages seek to minimize the distance between the verb and parts of the sentences to which it is semantically related.

Minimize Domains (paraphrased)
The human processor prefers to minimize the domain in which relations of dependency are processed.
(Hawkins 2004:123)

The relationship between a verb of motion and a morpheme denoting direction or location is once again informative. In contrast to their SVO counterparts, object–verb languages strongly favor postpositions over prepositions. There is a good processing-related reason for this, as the examples below help illustrate. (The statistical tendencies cited here are from Hawkins 2004:224.)

Preferred	***Dispreferred***
Postposition pattern	Preposition pattern
the door ***to*** *walk*	***to*** *the door walk*
P V	P V
↑___↑	↑_________↑
(observed in 204 languages)	(observed in just 6 languages)

By the same reasoning, pre-nominal adjectives and possessors should be preferred over their post-nominal counterparts in SOV languages. This too is true.

Preferred

the man's *car* *buy*
Possessor N V
↑____↑

(observed in 177 languages)

Dispreferred

car ***the man's*** *buy*
N Possessor V
↑______________↑

(observed in just 11 languages)

Preferred

good *food* *eat*
Adj N V
↑____↑

Dispreferred

food ***good*** *eat*
N Adj V
↑___________↑

These are just the sorts of asymmetries that are expected in a syntax shaped by processing pressures. To repeat Hawkins' observation from the previous chapter:

> *If a particular property can be derived from a smaller rather than a larger domain, then less effort will be expended in its processing, and processing can be faster… Simultaneous processing and working memory loads are reduced when domains are small.*
> (Hawkins 2004:27; see also 2014:13)

In sum, even in the case of the very simple examples that we have been considering, it is possible to see how a theory of natural syntax constrained by the Strict Emergentist Protocol can shed light on various important puzzles:

- Why SOV languages tend to have case.
- How case contributes to the possibility of scrambling.
- Why scrambling increases processing cost.
- Why SOV languages tend to have postpositions, pre-nominal adjectives and pre-nominal possessors.

I will turn next to a puzzle that takes us into the darker depths of syntax, where we find a phenomenon that seems almost deliberately designed to increase processing cost – but in a natural way and for natural reasons.

7

Wh Questions

There are two basic strategies for forming *wh* questions. The first, called the *in situ* strategy, places a *wh* argument in the position appropriate for its grammatical relation. In an SVO language such as Mandarin, for example, a 'subject' *wh* word occurs in the preverbal position and a 'direct object' *wh* word in the postverbal position.

Subject *wh* question:
Shei kanjian le ni?
who see ASP you
'Who sees you?'

Direct object *wh* question:
Ni kanjian le **shei**?
you see ASP who
'Who do you see?'

In contrast, English employs a 'movement' strategy that places the *wh* word at the left edge of the clause.

Subject *wh* question:
Who is helping Mary?

Direct object *wh* question:
Who is Mary helping?

The formation of patterns of this type creates a 'filler-gap dependency,' whose signature feature is most evident in direct object questions, in which the *wh* word is clearly not in the position where a second argument would normally appear. (That position is often represented by a gap for the sake of explicitness.)

Who is Mary helping _?
↑
'gap'

In patterns such as these, the second-argument status of the *wh* word cannot be ascertained until the processor comes upon the 'open' position right after the verb. This places a significant burden on the processing mechanisms.

> *A well-known finding from the processing literature is that filler-gap structures (such as wh questions) strain working memory capacity, because the filler (the wh phrase) must be held in working memory until it can be assigned to a gap.*
> (Goodall 2004:102)

The syntax of filler-gap dependencies in *wh* questions raises a series of important issues and challenges for a theory of natural syntax, some of which I will address in this chapter.

7.1 A *Wh* Algorithm

The mapping of *wh* questions onto a semantic representation in English requires two steps.

- The *wh* word must be temporarily held in a neutral position outside argument structure until there is an opportunity to associate it with an open argument position. I will refer to this as '*Wh* Storage.' (The symbol '#' stands for 'sentence-initial.')

Wh Storage

Store the wh argument in a neutral position.

$$\#wh \;\mapsto\; \begin{array}{c}\text{PRED}\\ {}^{\boldsymbol{wh}}\!\!<\ldots>\end{array}$$

- At some later point, the *wh* dependency must be resolved by associating the stored *wh* argument with an open position in argument structure. I will formulate the relevant algorithm as follows.

The Wh Algorithm

Associate the stored wh argument with an open position in argument structure.

$$\begin{array}{c}\text{PRED}\\ {}^{\boldsymbol{wh}}\!\!<\ldots>\end{array} \;\mapsto\; \begin{array}{c}\text{PRED}\\ <\ldots\boldsymbol{wh}\ldots>\end{array}$$

Here is an example of how these algorithms work in the case of the subject *wh* question *Who left?*

Who left?

Step 0: Projection of the semantic base
PRED
$\langle\beta\rangle$

Step 1: *Wh* Storage
Who… ↦ PRED
$^{\mathbf{wh}}\langle\beta\rangle$

Step 2: Predicate Algorithm
Who ***left***… ↦ **LEAVE**
$^{\mathbf{wh}}\langle\beta\rangle$

Step 3: *Wh* Algorithm: LEAVE $^{\mathbf{wh}}\langle\beta\rangle$ ↦ LEAVE $\langle \boldsymbol{wh} \rangle$

In Step 3, the *Wh* Algorithm seizes the opportunity to map the *wh* word onto the open 'subject' position in the verb's argument structure.

Now consider an example involving a direct object *wh* question.

Who did Harry see _?

Step 0: Projection of the semantic base
PRED
$\langle\beta\rangle$

Step 1: *Wh* Storage
Who… ↦ PRED
$^{\mathbf{wh}}\langle\beta\rangle$

Step 2: First-Argument Algorithm
Who (did) ***Harry***… ↦ PRED
$^{wh}\langle \boldsymbol{h} \rangle$

Step 3: Predicate Algorithm + Argument Addition
Who (did) Harry ***see***… ↦ **SEE**
$^{wh}\langle h\ \_ \rangle$
added argument position

Step 4: *Wh* Algorithm: SEE $^{\mathbf{wh}}\langle h\ \_ \rangle$ ↦ SEE $\langle h\ \boldsymbol{wh} \rangle$

As illustrated here, the *wh* argument is held in storage while the First-Argument Algorithm maps *Harry* onto the first-argument position (Step 2). An opportunity to activate the *Wh*

Algorithm arises after the Predicate Algorithm has mapped SEE into the semantic representation (Step 3), making available an open second-argument position.

The psycholinguistic literature supports exactly this scenario. Indeed, electrophysiological activity in the brain reflecting event-related potentials precisely traces the path that the processor follows in seeking to resolve a filler-gap dependency.

> *Holding an incomplete wh dependency in working memory [gives] rise to a sustained anterior negativity… Completion of the wh dependency elicit[s] a P600 response.*[1]
>
> (Phillips, Kazanina & Abada 2005:426; see also Kluender & Kutas 1993:620 and Felser, Clahsen & Münte 2003)

Who do the neighbors usually visit _?
|___*SUSTAINED ANTERIOR NEGATIVITY*___↑
P600

(For a review of further evidence in support of quick resolution of filler-gap dependencies, see Phillips 2006:797-98 & Phillips & Wagers 2007:747-58)

7.2 The cost of filler-gap dependencies

There is widespread agreement that the resolution of filler-gap dependencies comes with significant processing cost.

> *… there appears to be a consensus on the following basic point in the psycholinguistic literature. Filler-gap dependencies are hard to process, and they are characterized by a heightened processing load and a constant effort to relate the filler to the appropriate gap site …*
>
> (Hawkins 2004:173)

> *It has been clear for some time that a sentence with [a filler-gap dependency] incurs a relatively high degree of processing difficulty, compared to a minimally different sentence without the dependency.*
>
> (Hofmeister & Sag 2010:380)

> *There is good evidence that wh dependencies increase the difficulty of a sentence. Sentences with wh dependencies are rated as harder and less acceptable than sentences without wh dependencies, and sentences with longer wh dependencies are rated as harder and less acceptable than sentences with shorter wh dependencies.*
>
> (Phillips 2013a:90)

---

[1] Sustained negativity and P600 are commonly studied components of brain responses known as event-related potentials (ERPs).

So, why do languages use the ‘movement strategy’ for *wh* questions? Presumably because the front edge of a sentence is a highly salient position, which makes it a good place to put the most important word in a *wh* question, namely the *wh* argument. The more urgent question for now has to do with how the mapping system adjusts to the challenge of dealing with the filler-gap dependencies that this option calls for.

Various considerations suggest that the difficulty of filler-gap dependencies depends on a number of factors, including the position of the gap. I will focus here on the cost and consequences of *wh* dependencies that extend across clause boundaries.

Cross-clausal *wh* dependencies

A striking feature of filler-gap dependencies in many languages is that they can extend across a clause boundary. (Here and elsewhere, I use the term ‘clause’ as a convenient way to refer to what might better be called a ‘predicate-argument complex.’)

> ***Intra-clausal filler-gap dependency:***
> **What** did you put **_** on the table?
>
> ***Cross-clausal filler-gap dependency:***
> **What** do you think **[**I put **_** on the table**]**?
>
> **Who** did he say **[_** fixed the car**]**?

Not all languages permit cross-clausal filler-gap dependencies, and even those that do – like English – often restrict the phenomenon to the clausal arguments particular matrix verbs.

> *… the only long-distance filler-gap expression to occur with any regularity at all are specific formulas with the verb* THINK *or* SAY *[as in the examples above].*
> (Ambridge & Goldberg 2008:353; see also Dabrowska 2004:196-200 and Verhagen 2006)

Moreover, there is good evidence that cross-clausal filler-gap dependencies incur significant additional processing cost.

> *One very clear result was that [sentences with cross-clausal wh dependencies were read more slowly than monoclausal sentences containing wh dependencies] … carrying a filler across a clause boundary and assigning it to a gap … is the source of the difficulty.*
> (Frazier & Clifton 1989:104)

> *If the gap is located in an embedded clause, the filler must be transported across the embedded clause boundary… Carrying a filler across a clause boundary results in additional processing cost.*
> (Kluender 1998:253; see also Kluender & Kutas 1993)

> *Psycholinguistic research shows … that processing clause boundaries generally lowers acceptability ratings and causes an increase in processing time.*
>
> (Hoffmeister & Sag 2010:383)

This brings us to the question of why and how the extension of a filler-gap dependency into a lower clause incurs this extra processing cost.

A Transfer operation

A bi-clausal sentence such as *Mary said Harry met Sue* is mapped onto a semantic representation such as the one below, in which the predicate SAY takes a clausal second argument. (Put more technically, its second argument is a predicate-argument complex.)

```
 SAY
<m _>           m = Mary
   |
  MEET
  <h s>         h = Harry, s = Sue
```

A key issue in the study of bi-clausal sentences involves the occurrence of question patterns such as *Who did Mary say Harry met?*, in which the sentence-initial *wh* word must ultimately be associated with the second-argument position in the *lower* clause.

Let's start with the matrix clause, in which the *wh* word is encountered and stored, *Mary* is identified as first argument and SAY as the predicate in search of a clausal argument. (The details are in the footnote.[2])

The matrix clause

Who did Mary ***say***
^{wh}<*m* _>
↰*position for clausal argument*

---

[2] ***What happens in the matrix clause:***

Step 0: Projection of the semantic base
PRED
<β>

Step 1: *Wh* Storage
Who*…* ↦ PRED
^{wh}<β>

Step 2: First-Argument Algorithm
Who (did) ***Mary****…* ↦ PRED
^{wh}<*m*>

Step 3: Predicate Algorithm + Addition of clausal argument position
Who (did) Mary ***say****…* ↦ **SAY**
^{wh}<*m*_>
↰*position for clausal argument*

At this point, the processor runs into an apparent dead end: there is no place for the *wh* argument. Crucially, however, SAY takes a clausal argument, whose addition will create a new set of opportunities to resolve the *wh* dependency.

Who did Mary ***say*** *… (Harry met)?*

SAY
wh<*m* _>
|
PRED
<…>

In order to take advantage of those opportunities, the processor must pass the *wh* dependency to the lower clause and reactivate it there. That operation, which I will call the Transfer Algorithm, produces the result illustrated below in the case at hand. As you can see, the stored *wh* argument is relocated from its position just outside the argument grid in the higher clause to a similar position in the lower clause.

The Transfer Algorithm

Who did Mary ***say*** *… (Harry met)?*

SAY
~~***wh***~~<*m* _>
↓ |
PRED
wh <…>

Once the *wh* argument has been transferred to the embedded clause and reactivated there, the processor proceeds with its usual intra-clausal routine, identifying the first argument (*Harry*) and the predicate (MEET).

Who did Mary say … ***Harry met _****?*

MEET
wh<*h* _>

The *wh* dependency is then finally resolved by associating the *wh* argument with the second-argument position of MEET.

MEET
wh<*h* _> ↦ <*h* ***wh***>
↑

Some evidence

Evidence from a variety of sources supports the Transfer-based scenario, including the reactivation of the *wh* dependency in the embedded clause.

i. ERP evidence

Left Anterior Negativity, a sign that a *wh* dependency is being held in memory, is observed at the beginning of the embedded clause in cross-clausal *wh* questions.

What did you say **[**that the children bought **_]**?
↑
LAN continues into this region.

(Kluender & Kutas 1993:608-10; see also Phillips, Kanzina & Abada 2005:423.[3])

ii. Reading time evidence

Increased reading time – interpreted as a reflex of reactivating the *wh* word – has been observed at the beginning of the embedded clause in the first sentence below compared to the second one. (The test items used in the relevant study are relative clause patterns, but relative pronouns create the same sort of filler-gap dependencies as interrogative pronouns.)

Native speakers of English [N = 25, mean age = 24] 'showed elevated reading times' at the complementizer THAT in the first sentence below compared to the second one.[4]

(Marinis, Roberts, Felser & Clahsen 2005:69; see also Gibson & Warren 2003 and Keine 2020.)

Bi-clausal pattern containing a cross-clausal filler-gap dependency:

[FILLER] [GAP]
The nurse **who** the doctor argued **[that** the rude patient had angered **_]** went home.
↑
reactivation of the wh dependency around here

Bi-clausal pattern with no filler-gap dependency:

The nurse thought the doctor argued **[that** the rude patient had angered the staff**]**.
↑
no wh word in the sentence,
so no reactivation

---

[3] Phillips et al. (2005:423) note that although the negativity persists throughout the cross-clausal *wh* dependency, 'the growth of the negativity is confined to the first clause of the dependency plus the complementizer *that* at the beginning of the second clause.'

[4] Another measure, considered by Gibson & Warren and Marinis et al. but not discussed here, involves response times at the gap in the embedded clause, which appear to be shortened in cases where a long filler-gap dependency includes a clause boundary at which the *wh* phrase can be reactivated.

iii. Acquisition evidence

When young children produce *wh* questions containing a cross-clausal filler-gap dependency, they sometimes repeat the sentence-initial *wh* word at the beginning of the lower clause.

What do you think **[what** pigs eat **_]**?

Who did he say **[who _** is in the box**]**?

… adults can readily reactivate the wh phrase covertly and move on to the articulation of the embedded clause ... Some children … produce another instance of the sentence-initial wh phrase to strengthen their memory representation of the filler-gap dependency.

(Lutken, Legendre & Omaki 2020:37[ms]; see also Crain, Goro & Thornton 2006:33, Thornton 1990 and McDaniel, Chiu & Maxfield 1995)

iv. Typological evidence

Some languages place a copy of the *wh* word at the beginning of the embedded clause.

German (some Cologne dialects):

Wen	glaubt	Hans	**[wen**	Jakob	**_**	gesehen	hat**]**?
whom	thinks	Hans	whom	Jakob		seen	has

‘Whom does Hans think that Jakob has seen _?’

(McDaniel 1989:569n)

Wen	willst	du	**[wen**	Hans	**_**	anruft?**]**
who	want	you	who	Hans		call

‘Who do you want Hans to call _?’

(Thornton 1990:220)

Romany:

Kas	o	Demìri	mislinol	**[kas**	Arìfa	dikh̀la _ **]**?
whom	does	Demiri	think	whom	Arifa	saw

‘Who does Demiri think that Arifa saw _?’

(McDaniel 1989:569n)

Passamaquoddy:

Wenil	Mali	wewitahamacil	**[wenil**	kisniskamuk	**_]**?
who	Mary	remember	who	dance.with	

‘Who does Mary remember I danced with _?’

(Bruening 2006:28)

A somewhat different effect is found in languages such as Irish, where a special complementizer (*aL*) occurs at the left edge of clauses through which a *wh* dependency extends, presumably signaling reactivation of the *wh* argument. (The complementizer for clauses with no *wh* dependency is *goN*.)

Cé	**[aL**	dúradh	léithi	**[aL**	cheannódh	**_**	é **]]**?
who	COMP	was.said	with.her	COMP	would.buy		it

'Who was she told would buy it ?'
(McCloskey 2001:94n)

7.3 Islands on the horizon

A central issue in the study of filler-gap patterns involves the need for constraints on cross-clausal dependencies, as illustrated in the second of the two sentences below.

What do you know (for certain) **[**that he sent **_** to Jane**]**?

***What** do you know (for certain) **[how** he sent **_** to Jane**]**?

The embedded clause in the second example is an 'island' – a configuration that blocks the extension of a filler-gap dependency.

The search for an explanation of island effects has helped shape the history of syntactic theory since the 1960s.

> *I think that it is fair to say that within linguistics there is almost no controversy about the importance of island effects and the need for general and abstract explanations for them.*
> (Phillips 2013a:79)

Two general lines of inquiry can be identified. One posits the existence of formal grammatical constraints that are typically treated as principles of Universal Grammar.

> *By far the largest body of work on island effects has assumed that they are consequences of formal grammatical constraints that block displacement operations that remove phrases from island domains.*
> (Phillips 2013a:78)

The other approach, often dubbed the 'reductionist' account, holds that island effects can be traced to processing considerations.

> *… a competing line of thought [is] that the variation in acceptability judgments associated with [islands], both language-internally and crosslinguistically, can*

> *be better explained by appealing to cognitive constraints on language processing.*
> …
> *… if one of the goals of linguistic inquiry is to explain why we have the linguistic constraints that we do, the [processing]-based view accounts for islands as the byproduct of general principles of cognition. In contrast, the [grammar]-based theory offers no insight into* WHY *islands exist, at least at this time.*
> (Hofmeister & Sag 2010:367 & 403-04)

The debate is often heated:

> *There are many challenges facing resource-based reductionist accounts of island effects. What matters here is not the sheer number of challenges, but the fact that so many of the premises of such accounts are not met.*
> (Phillips 2013a:108)

> *The grammatical constraints that have been proposed to account for syntactic islands are almost uniformly complex, arbitrary, and ultimately either too strong, too weak, or both. They express intricate and highly specific limitations on just a subset of the linguistic dependencies possible in natural language. They are arbitrary in the sense that they bear no relationship to other constraints, emanate from no general principles of language, and have no relevance or parallel outside language.*
> (Hofmeister & Sag 2010:406)

Only the reductionist line of inquiry is compatible with the Strict Emergentist Protocol, and I will pursue it in the next chapter.

8

Constraints on Filler–Gap Dependencies

Many different types of constraints on filler-gap dependencies have been proposed over the past several decades. My purpose in this chapter is to outline a reductionist explanation for three of these constraints: a little-discussed island pattern in Russian, a widely studied restriction on multiple *wh* dependencies, and the curious interaction of filler-gap relations with phonological contraction in English.

Setting aside the possibility of cognitive pathologies, I assume that every normal human being has the processing capacity needed to manage the complexities of any language of which he or she is a native speaker. However, processing cost is real and it has a major effect on shaping the internal workings of language, including the syntax of filler-gap dependencies. The following two assumptions are crucial for this line of reasoning.

- Costly processing routines are resisted, unless they are encountered in the speech of others.
- If a costly routine is permitted, less costly routines of the same type must also be allowed.

The latter assumption follows directly from the notion of Affordability (see §5.2), restated here.

> ***Affordability***
> If a language permits a more costly operation, it should also permit a less costly operation of the same type.
> (Hawkins 2004:129-30, 256ff; O'Grady 2005:214, 2016:43)

An application of this idea to the syntax of filler-gap dependencies has been promoted, in one guise or another, at least since the mid 1970s, when it surfaced in work by Keenan & Comrie (1977) on the typology of relative clauses. (I will return to this matter in chapter 12.)

> *… if a given strategy is used to encode a fairly easy meaning and that strategy is "strong" enough to encode a rather difficult meaning, then it is surely strong enough to encode the meanings of intermediate difficulty.*
> (Keenan & Comrie 1977:88)

The idea was later quite widely adopted in processing-based approaches to typology.

> *If there is a preference ranking A > B > C > D among structures of a common type in performance, then there will be a corresponding hierarchy of grammatical conventions (with cut-off points and declining frequencies of languages).*
>
> (Hawkins 2004:256; see also p. 216ff.)

> *… the greater the demands that a particular pattern makes on the processor, the less likely that languages will develop a computational routine for dealing with it… This in turn explains the familiar implicational facts … if the 'harder' pattern is possible, then so is the comparable 'easier' one.*
>
> (O'Grady 2005:214)

As we will see next, Affordability also figures in a possible natural explanation for the syntax of islands.

8.1 An island constraint in Russian

There are substantial differences across languages in their tolerance for filler-gap dependencies. A very basic difference involves the contrast between intra-clausal and cross-clausal *wh* dependencies, first mentioned in the preceding chapter. (As previously noted, I use the term 'clause' as shorthand for what might better be called a 'predicate-argument complex.')

Intra-clausal dependency:
Who do you like **_**?

Cross-clausal dependency:
Who do you think **[**you like **_]**?

The latter pattern places a greater burden on the processor because of the need for a Transfer operation to carry the filler-gap dependency across a clause boundary, reactivating it in a lower clause.

Transfer

PRED
~~wh~~<… _>
PRED
wh<…>

This fits well with common observations in the processing literature relating to the cost of cross-clausal dependencies.

> *We consistently found that a [filler-gap dependency in English, Greek and German] with single, double, or triple embedding is less acceptable than an unembedded structure.*
> …
> *A main finding was that [extractions] from that-clauses are less acceptable than unembedded ones.*
> (Alexopoulou & Keller 2007:133 & 136)

> *Psycholinguistic research shows … that [the] processing [of] clause boundaries generally lowers acceptability ratings and causes an increase in processing time.*
> (Hofmeister & Sag 2010:383)

Moreover, as observed in the preceding chapter, there is direct evidence (ranging from ERP effects, to errors in child speech, to languages with ‘*wh* copying’) that confirms the reactivation of the *wh* argument in the lower clause.

Two predictions now follow.

- Because of the cost of the Transfer operation, there could be languages whose speakers uniformly reject cross-clausal *wh* dependencies.
- Consistent with Affordability, if a language permits cross-clausal *wh* dependencies, it should also allow intra-clausal *wh* dependencies.

As we will see next, both predictions turn out to be correct.

Russian versus English

Striking support for the first prediction comes from Russian, in which cross-clausal *wh* dependencies cannot extend into a finite clause.[1]

Intra-clausal dependency:

Kogo	oni	videli	**_** ?
who.ACC	they.NOM	see.PST.PL	

‘Who did they see _?’

Cross-clausal dependency:

***Kogo**	Olga	skazala	**[**čto	oni	videli	**_]**?
who.ACC	Olga.NOM	say.PST.FEM	that	they.NOM	see.PST.PL	

‘Who did Olga say that they saw _?’

(Dyakonova 2009:215)

---

[1] van Gelderen (2001:97) suggests that there may be individual variation with regard to the extent to which these patterns are unacceptable.

The most obvious way to capture the difference between Russian and English is to reduce it to the following simple contrast:

> ***Russian versus English***
> English has the Transfer operation; Russian doesn't.

This fits well with both of the predictions under consideration.

- Russian lacks the Transfer operation and therefore rejects cross-clausal *wh* dependencies.
- English has the Transfer operation and, consistent with Affordability, allows not only cross-clausal *wh* dependencies but also intra-clausal dependencies.

Infinitival clauses

Russian offers a further instructive fact: despite its aversion to cross-clausal dependencies, it permits patterns of this type when the lower clause is infinitival.[2]

Kogo	Olga	nadeyalas'	**[uvidet' _]**?
who.ACC	Olga.NOM	hope.PST.3SG	see.INF

'Who did Olga hope to see _?'
(based on Dyakonova 2009:216; see also Hawkins 2004:194)

Why should a *wh* dependency be able to extend into an infinitival clause, but not a finite clause?

One possibility is that infinitival clauses, which generally cannot stand alone as independent sentences, are more integrated into the argument structure of the matrix verb than a finite clause is. Indeed, an idea along these lines is frequently mentioned both in discussions of English and in the typological literature.

> *The more the two events coded in the main and complement clauses share [arguments],*[3] *the more likely they are to be semantically integrated as a single*

---

[2] An improvement also occurs when the embedded clause is subjunctive:

Komu	Ira	hočet	**[**čtoby	my	otdali	_	kotjat?**]**
who.DAT	Ira.NOM	want.PRS.3SG	that.SUBJ	we.NOM	give.away.PST.PL		kittens.ACC

'Who does Ira want that we give the kittens to _?'

Gibson (1998:12) was among the first to suggest that nonfinite verbs do not contribute to locality costs; see also Hofmeister, Casasanto & Sag (2013:44-45).

[3] Argument sharing is a defining feature of infinitival clauses. In the first example below, the subject arguments of *hope* and *stay* have the same referent; in the second example, the direct object of *persuade* and the subject of *leave* have the same referent.

Max hoped [. to stay].

They persuaded Mary [. to leave].

> *event; and the less likely is the complement clause to be coded as an independent finite clause.*
> …
> *The less independent a complement clause is, the higher is the probability of its being fully integrated into the main clause…*
> (Givón 1990:527 & 561)

I will represent this difference as follows, with the infinitival clause directly integrated into the argument structure of the matrix predicate. In contrast, the finite clause stands on its own, connected to the matrix predicate – but a short step removed.

I hope ***to see Mary****.*	*I hope that* ***I'll see Mary****.*
HOPE <*i* **SEE**> *i* = *I* < **.** *m*> (. = unexpressed argument)	HOPE <*i* _> *i* = *I* | **SEE** <*i m*>

Now consider the case of a *wh* dependency that extends into an embedded infinitival clause.

Who do you hope [to see _]?

If in fact an infinitival clause is tightly integrated into the argument structure of the matrix clause to begin with, it seems plausible to suggest that the filler-gap dependency might be resolved without transfer. Rather, as depicted below, the *wh* dependency might simply extend into the infinitival clause, where it can be associated with an open position in the argument grid of SEE. Depicted simply:[4]

Who do you hope [to see _]?

This proposal leads to a series of predictions that help confirm its plausibility.

- Since there is no Transfer operation, there should be no language in which a copy of a sentence-initial *wh* word appears at the infinitival-clause boundary.

  ***What** did you promise **[what** to bring **_]**?

[4] Here is a depiction of the actual mapping operation, in which the stored *wh* dependency extends directly into the infinitival clause, where it is associated with the second argument position of SEE by the *Wh* Algorithm.

HOPE ***wh***<*y* SEE> < . _>	↦	HOPE ***~~wh~~***<*y* SEE> < . ***wh***>

- Young language learners should never reproduce a sentence-initial *wh* word at the infinitival-clause boundary.

  ***Who** did you decide **[who** to visit **_]**?

- There should be no psycholinguistic signs of *wh* reactivation at an infinitival-clause boundary (e.g., no ERP effects or reading slowdown).

As far as I know, the first two predictions are correct. I am aware of only one psycholinguistic study (involving a timed self-paced reading task) that has looked for signs of a reactivation at an infinitival-clause boundary. As predicted, none was found.

> *The results indicate that [an infinitival clause] crossed by movement does not facilitate dependency completion [by reactivating the wh argument], in direct contrast to [full tensed clauses]. This asymmetry … is of course precisely what is predicted if intermediate landing sites are created in [the latter case but not in the former].*
> (Keine 2020:145)[5]

8.2 *Wh* islands in English

Although English permits the transfer of a filler-gap dependency across a clause boundary in some cases, a restriction on this practice is evident in the following contrast, as noted in the previous chapter.

No *wh* word in the embedded clause:
What do you know (for certain) **[**that he said **_** to Jane**]**?

Wh word in the embedded clause:
***What** do you know (for certain) **[who** he said **_** to **_]**?

The lower clause in the second sentence constitutes a '*wh* island.' Intuitively, the filler-gap dependency initiated by sentence-initial *what* is blocked by the *wh* word at the left edge of the lower clause (*who*).

> *When a filler is being held in working memory and a clause boundary is being crossed, which already taxes the parser, the occurrence at the clause boundary of a second expression whose [interpretation] must simultaneously be activated represents an additional processing load. This convergence of multiple processing tasks is perceived as ungrammaticality.*
> (Kluender & Kutas 1993:579)

[5] The key contrast involves sentences of the following sort:

Who did the lawyer claim **[**that the accusations had hurt _ **]**? (tensed embedded clause)
Who did the lawyer claim **[**the accusations to have hurt _ **]**? (infinitival embedded clause)

> *A dependency into a wh island … requires processing an embedded interrogative … while holding a filler phrase in working memory and actively searching for a suitable [gap].*
> (Hofmeister & Sag 2010:385; see also p. 383)

In other words, the problem can be traced to the need to hold two *wh* fillers in memory at the onset of the lower clause – the one being transferred from the higher clause and the one that is already present in the embedded clause.

***What** do you know (for certain) **[who** he said **_** to **_]**?
↑ the *wh* filler seeking transfer ↑ the *wh* filler already in the lower clause

At least for English, this is a 'break point.' The storage of a second *wh* argument requires an incremental increase in processing cost for which there is no precedent in experience. The mapping is therefore blocked.

Unattested Transfer

PRED
~~***wh***~~<… _>
PRED
****wh wh***<…>

The end result is an Affordability metric that exactly matches the typological facts.[6]

Affordability metric for extending wh dependencies

- no extension across a clause boundary
- extension across a clause boundary if there is no other *wh* dependency
- extension across a clause boundary even if there is another *wh* dependency

(based on O'Grady 2012:497; see also Hawkins 2004:192ff, 266 and O'Grady 2005:214ff)

---

[6] I would be remiss not to mention here a (welcome) parallel in the generative literature.

> ***The Subset Principle***
> *Learners must select the smallest possible language compatible with the input at each step of the learning procedure.*
> (Clark:1992:101; see Fodor & Sakas 2005 for extensive discussion.)

By 'smallest' language, proponents of the Subset Principle have in mind the language with the more restrictive grammar – hence a grammar that permits only intra-clausal *wh* dependencies versus one that also permits cross-clausal dependencies, for example. Although the Subset Principle is formulated in terms of grammatical mechanisms, the underlying intuition can easily be understood in terms of processing cost in most cases.

The first option is exemplified by Russian (see §8.1) and the second by English. Instances of the third option have been reported in Italian relative clauses.

tuo fratello,	**a cui**	mi domando	**[che storie**	abbiano	raccontato	**_ _]**
your brother	to whom	I wonder	which stories	they.have	told	

(Rizzi 1982:50)

In sum, there is a cost-based continuum consisting of natural 'break points' (clause boundaries and the number of stored fillers) beyond which language learners and users will not proceed without sufficient prior exposure. Different languages choose different break points, but always with the same consequence: if the processor tolerates a more demanding pattern, it must permit less demanding patterns as well.

Wh dependency options (with additional cost on the lower tiers)

Russian	English	Italian
intra-clausal	intra-clausal	intra-clausal
	cross-clausal [no other filler]	cross-clausal [no other filler]
		cross-clausal [plus another filler]

One final point is worth considering before moving on. In §8.1, I made a two-part suggestion:

- Infinitival arguments are more tightly integrated into the argument structure of the higher verb than a clausal argument that can be a stand-alone sentence.
- For this reason, a *wh* dependency can extend into an infinitival clause without the need for Transfer and reactivation (see the examples in §8.1).

If this is true, then a surprising result can be predicted: *wh* island effects should disappear, or at least be sharply diminished, when the clausal argument is infinitival. This seems to be correct.

Which engines does he know **[how** to repair **_ _]**?

Which clothes were you wondering **[what** to do **_** with **_]**?

(Pesetsky 1982, Richards 1997:40, O'Grady 2005:120-21)

8.3 A constraint on contraction

A very different type of constraint on filler-gap dependencies involves a notorious contrast in the contractibility of *want to* in sentences such as the following.

Contraction is possible:	Contraction is unnatural:
Do you **want to** stay?	Who do you **want to** stay?
↓	↓
wanna	**wanna*

It is a remarkable fact about both language and science that a single casually pronounced word might shed light on the workings of human cognition, but that does seem to be the case.

> *Syntacticians' enduring interest in the* WANNA *facts reflects the momentous role they have played in shaping modern linguistic theory.*
> (Getz 2019:120)

The search for an account of this contrast has yielded a number of spectacular proposals, including the idea that contraction is blocked by an invisible Case-marked trace.

> *A Case-marked trace blocks contraction.*
> (Jaeggli 1980:242)

> Who do you want **[$_{NP}$ t_{OBJ}]** to stay?

> *If it is true, as a matter of principle, that movement rules leave a trace in mental representations, then contraction will be blocked as a matter of principle over a position from which the question word has been moved.*
> (Chomsky 1980b:160)

The proposal was hotly contested in a debate that 'raged within the pages of *Linguistic Inquiry* for nearly a decade' (Getz 2019:119).

> *… the importance of the contraction debate is in our view this: it reveals that [trace theory] is far from the exemplary token of a scientific movement its enthusiasts would have us believe. Rather, this debate indicates that faith in the virtues of [trace theory] is independent of genuine factual accomplishments … and that its most prominent developers do not shy away from citing as instances of explanatory success domains where they have never achieved anything more than descriptive failure.*
> (Postal & Pullum 1982:137)

A different way to approach the contractibility asymmetry is to consider it from a processing perspective.

Let us assume that contraction is subject to a naturalness preference of the following sort, whose articulatory motivation is evident.

> ***Natural Contraction***
> *Contraction of the string XY is most natural when X adjoins to Y without delay.*
> (O'Grady 2005:139ff)

Let us further assume (again uncontroversially) that the resolution of a filler-gap dependency is a real-time operation and that it occurs upon encountering the verb *want* in the pattern below.

> **Who** do you **want _ to** stay?
> ↑
> *The filler-gap dependency is resolved here.*

This suggestion is at least partly confirmed by the presence of two phonological reflexes.

- A prosodic break is more likely after *want* in the filler-gap pattern than in patterns such as *Do you wanna stay?*, where there is no filler-gap dependency.
- The verb is on average 20 milliseconds longer in the pattern containing the filler-gap dependency.

(Warren, Speer & Schafer 2003:41 & 45)

Taken together, these facts suggest that a processing slowdown occurs at the point where the filler-gap dependency is resolved in the uncontractible *want-to* pattern. I propose that this slowdown – not a Case-marked trace – works against the naturalness of contraction.

It is now possible to consider in a fresh light two otherwise puzzling facts about *want to* contraction. A first puzzle is that pre-school children produce the forbidden pattern at a surprisingly high rate in experimental studies involving elicited production.

> *Children use wanna [in the forbidden pattern] nearly half the time (47%) – far more than adults (2%).*
> (Getz 2019:124; see also Zukowski & Larsen 2011)

This may seem surprising, but it actually makes sense. That is because the processing-based impediment to contraction cannot come into effect until learners realize that *wanna* is the contracted version of *want to* – which is far from obvious at the outset.[7]

A piece of evidence for the pedigree of *wanna* comes from the presence of the (barely audible) verbal suffix *-s* in patterns with a third-person singular subject.

> He wan**s**a stay for a week.

However, this clue is counterbalanced by the apparent presence of a past tense marker in patterns such as the following.

> He wanna-**d** a dog
>
> She wanna-**d** Alan to stay.
>
> They wanna-**d** a go right away.

---

[7] The idea that *wanna* is a word in its own right has been frequently proposed over the years; see Getz (2019) for a recent proposal and O'Grady (2008a) for a critical review of the earlier literature.

This is an illusion of course: *wanna-d* is really a contraction of *wanted*, not the past tense of *wanna*. However, the homophony here could easily contribute to a delay in sorting things out.[8] And without an awareness that *wanna* is a contracted form, the principle of Natural Contraction cannot be invoked.

A second puzzle stems from the fact that there is some variation among even adults with respect to the acceptability of contraction in *wh* questions.

> *There exist dialects in which contraction is systematically permitted across the alleged Case-marked traces, so that [sentences like* WHO DO YOU WANNA DRIVE THE CAR*?], ungrammatical for me and many others, are grammatical.*
> (Pullum 1997:96; see also Pullum & Postal 1979:704-05 and Ito 2005)

This is a surprising turn of events, especially if contraction is blocked by an invisible Case-marked trace whose presence is required by principles of Universal Grammar, as Jaeggli's account implies.

On the processing account, however, we need only assume that some speakers of English permit contraction under slightly less than optimal conditions. This proposal comes with an important proviso: no speaker of English should allow contraction only when there is an intervening filler-gap dependency. That is, consistent with Affordability, anyone who accepts the more demanding instance of contraction in the second sentence below should also accept it in the first sentence.

> ***Contraction with no filler-gap dependency***
> Do you wanna drive the car?
>
> ***Contraction that is impeded by a filler-gap dependency***
> Who do you wanna drive the car?

This seems to be correct.[9]

In sum, constraints on filler-gap dependencies – whether they involve impediments to contraction or the presence of island effects – are natural reactions to processing cost. The resolution of a *wh* dependency can block contraction in English. Russian minimizes processing cost by rejecting the Transfer operation needed to extend a *wh* dependency into a finite clause. English has that operation, but can't use it to extend a *wh* dependency into a finite clause that has a second filler-gap dependency in play. And so on.

---

[8] In fact, a search of maternal speech to Adam, Eve and Sarah in the CHILDES corpus reveals an almost identical number of instances of *wants to* and *wanted to* (36 and 37, respectively), in addition to 344 instances of *want to*. However, the corpus gives no clue as to the actual pronunciation of these items.

[9] We should not rule out the possibility that adults who accept the 'forbidden' pattern nonetheless make subtle phonetic adjustments when producing it in their own speech, as an extension of Warren et al.'s study might well show.

Some languages are willing to pay the price of Transfer, and some (e.g., Italian) are even willing to deal with two stored filler-gap dependencies in the same clause. But the number of 'takers' should (and does) decrease as the cost grows.

> *If there is a [processing-based] preference ranking … among structures of a common type …, then there will be … corresponding cut-off points and declining frequencies of languages.*
>
> (Hawkins 2004:256; see also p. 216ff)

Moreover, consistent with Affordability, the adoption of a more costly option comes with the obligation of also allowing less costly options. There are no languages that require all *wh* dependencies to extend into an embedded clause or to always co-occur with a second filler-gap dependency. That wouldn't be natural, and it doesn't happen.

9

Local Coreference

A core tenet of natural syntax is that the algorithms that define the mapping between form and meaning are shaped by processing pressures. The syntax of coreference (also known as 'anaphora'[1]) provides another opportunity to evaluate this point.

> *… anaphora has not only become a central topic of research in linguistics, it has also attracted a growing amount of attention from philosophers, psychologists, cognitive scientists, and artificial intelligence workers… [It] represents one of the most complex phenomena of natural language, which, in itself, is the source of fascinating problems…*
>
> (Huang 2000:1)

> *… [anaphora] figures prominently in a vast amount of works, either as the main research topic, or, perhaps even more frequently, as a diagnostic for constituency, derivational history, and other abstract aspects of grammatical analysis.*
>
> (Büring 2005:ix)

Most work on this subject focuses on reflexive pronouns such as *himself*, *herself* and *themselves*, which introduce 'referential dependencies' that must be resolved with the help of another element, called an 'antecedent.'

> *[Reflexive pronouns] lack the capacity for independent reference, and therefore must depend on another expression for their interpretation.*
>
> (Reuland 2018:82)

Crucially, there are constraints on what and where that 'other expression' can be, as the following contrast helps illustrate.

Mary pinched **herself**.

*__Herself__ pinched **Mary**.

As a first and informal approximation, it appears that a subject can serve as the antecedent for a direct object, but not vice versa.

---

[1] The terms 'coreference,' 'anaphora' and 'binding' overlap in their meaning and are often used interchangeably, as I will also do here.

> *When the subject and object are identical, we use for the latter a so-called reflexive pronoun, formed by means of* SELF*...*
> (Jespersen 1933:111)

> *[Basic] subjects in general can control reflexive pronouns [but not vice versa].*
> (Keenan 1976:315)

This constraint seems to be universal.

> *… there appears to be no language in which the patient argument outranks the agent argument for the purposes of anaphora.*
> (Falk 2006:66)

> *In every ergative language, as in every accusative language, the 'antecedent', i.e. the controller of reflexivity, is the [agent argument].*
> (Dixon 1994:138-39)

This chapter explores the possibility that coreference asymmetries have a natural syntax, in accordance with the Strict Emergentist Protocol. For reasons of space, I will focus my attention on patterns, like the ones above, in which coreference requires a reflexive pronoun rather than a plain pronoun such as *him* or *her*.[2]

9.1 The natural facts

Two long-standing generalizations help define the basic syntax of anaphora.

- The reflexive pronoun requires a co-argument as its antecedent.
- The antecedent must be in some sense more 'prominent' than the pronoun, consistent with the observation above that a subject (agent) can serve as antecedent for a direct object (patient), but not vice versa.

As far as I know, there are really only two ideas about how to go about characterizing the prominence asymmetry.

---

[2] I thus exclude examples such as the following in which coreference can be expressed by either a reflexive pronoun or a plain pronoun. For a detailed discussion of this phenomenon, see O'Grady (2015).

Larry's diary contained [a flattering description of **him/himself**].

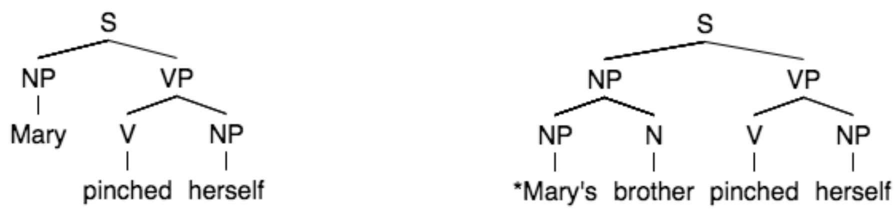
Richard didn't think [that Mary should see those two pictures of **him/himself**].

John was going to get even with Mary. [That picture of **him/himself**] in the paper would really annoy her, as would the other stunts he had planned. (Pollard & Sag 1992:274)

One approach, best exemplified by Principle A of generative grammar, draws on the architecture of syntactic structure. Its key claim is that a reflexive pronoun must look to a higher (so-called ‘c-commanding’) NP for its interpretation.

> ***Principle A*** (paraphrased)
> *A reflexive pronoun must have a higher antecedent in the same clause.*
> (based on Chomsky 1981:188)

The contrasting examples below help illustrate how this principle works.



The first sentence is acceptable since the antecedent (*Mary*) occurs in a higher position than the reflexive pronoun, in contrast to the situation in the unacceptable pattern to its right.

The second approach makes use of argument structure to capture the asymmetries underlying coreference. One way to do this is to organize arguments in terms of their grammatical relations.

> ***Subject < Primary object < Secondary object < Other complements***
> *An anaphor must have a less oblique co-argument as its antecedent, if there is one.*
> (Pollard & Sag 1992:266)

Another way is to posit a hierarchy of thematic roles.

> ***Agent < Location, Source, Goal < Theme***
> *A reflexive pronoun cannot be higher on the thematic hierarchy than its antecedent.*
> (Jackendoff 1972:148; see also Pollard & Sag 1992:297-99)

I too adopt an approach based on argument structure, although without direct reference to either grammatical relations or thematic roles. Instead, I focus entirely on the manner in which the arguments are organized relative to each other within argument structure. For now it suffices to assume that the agent is invariably the first argument of a transitive verb (in any language).[3]

[3] Passivization downgrades the agent by removing it from the first-argument position.

PRED
<ag pat>
1 2

As noted in §5.1, this makes sense from a processing perspective since the activity of an agent typically marks the starting point of the event and is ultimately responsible for the patienthood of the second argument.

> *… the agent is at the head of the causal chain that affects the patient.*
> (Kemmerer 2012:50; see also Bornkessel-Schlesewsky & Schlesewsky 2009:41, Cohn & Paczynski 2013:75, and Sauppe 2016:10)

An algorithm for coreference

We can now formulate the following algorithm for the interpretation of reflexive pronouns.

The Anaphor Algorithm[4]
<α *x*>
↳α

Three points help clarify the purpose and functioning of the Anaphor Algorithm:

- It applies to semantic representations, not strings of words.
- It is triggered by the presence of a referential dependency (represented here as *x*) that is introduced by a reflexive pronoun.
- It resolves the referential dependency by associating it with a 'prior' co-argument, represented in the algorithm by the symbol α. ('prior' = 'prior in argument structure')

A concrete example follows.

---

[4] In its full form, the algorithm should be formulated as follows in order to make room for the possible presence of arguments other than just the reflexive pronoun and its antecedent.

<…α…*x*…>
↳α

Mary told me about **herself**.
I talked to **John** about **himself**.
Larry bought **himself** an ice cream cone.

Mary pinched herself.

Step 0: Projection of the semantic base
PRED
<β>

Step 1: First-Argument Algorithm
Mary… ↦ PRED
<***m***>

Step 2: Predicate Algorithm + Argument Addition
Mary ***pinched***… ↦ **PINCH**
<*m* **_**>

Step 3: Second-Argument Algorithm
Mary pinched ***herself*** ↦ PINCH
<*m* ***x***>

Step 4: Anaphor Algorithm <*m* ***x***>
↳ ***m***

In the final step, the just-encountered reflexive pronoun receives its interpretation thanks to the Anaphor Algorithm, which selects the verbal predicate's first argument (*Mary*) as the only possible antecedent. By acting immediately and locally in this way, the algorithm minimizes the burden on working memory, a major factor in processing cost.

The psycholinguistic evidence suggests that the processor proceeds more or less in this way, quickly selecting the local co-argument as the antecedent over potential competitors.

> *… from the earliest measurable point in time referential interpretations are determined by constraints [on coreference] in tandem with other sources of information.*
> (Clackson, Felser & Clahsen 2011:140; see also O'Grady 2005:167-69)

This finding fits well with the idea underlying the Anaphor Algorithm: an encounter with a reflexive pronoun triggers an immediate search for a co-argument that can resolve the referential dependency at minimal cost.

The problem of unacceptability

On the assumption that the presence of a reflexive pronoun argument automatically triggers the Anaphor Algorithm, there is a natural account for the unacceptability of sentences like the one below.

*Herself pinched Mary.

Here, the usual operations produce the following semantic representation, in which the reflexive pronoun is the first argument, for which there can be no prior co-argument.

PINCH
<***x*** *m*>   *x = herself; m = Mary*
*↵

As a result, the Anaphor Algorithm is unable to do its work and the referential dependency is left unresolved, disrupting the mapping between form and meaning.

The next sentence illustrates another classic contrast.

[Harry's brother] pinched himself.

Here, the co-argument of the reflexive pronoun is *Harry's brother*, not *Harry*. This in turn ensures, correctly, that it will be selected by the Anaphor Algorithm as the antecedent for *himself*.

PINCH
<*b x*>   (*b = Harry's brother*)
↳ ***b***

A third key contrast involves biclausal patterns such as the one below, in which there appear to be two potential antecedents for the reflexive pronoun.

Harry thinks [Fred noticed himself].

On the processing account, the only permissible interpretation is the one in which the referential dependency introduced by the reflexive pronoun is resolved by the referent of *Fred*, its co-argument. This is exactly the result yielded by the Anaphor Algorithm.

Harry thinks [Fred noticed himself].
. . .   NOTICE
<*f x*>
↳*f*

Priority versus linearity

The organization of argument structure is often reflected, at least loosely, in a language's canonical word order: agents precede patients in transitive clauses in over 90% of the world's languages. Crucially, however, the Anaphor Algorithm operates on semantic representations and is therefore not directly sensitive to word order. A striking illustration of this point comes from the VOS language Malagasy, in which we find typologically unusual patterns such as the following, from Keenan (1976:314-15).[5]

[5] As discussed in §5.1, the VOS word order of Malagasy reflects application of the Second-Argument Algorithm before the First-Argument Algorithm.

	PAT	AGT
Manaja	**tena**	Rabe.
respect	self	Rabe

‘Rabe respects himself.’

Although the reflexive pronoun (the second argument) *precedes* its antecedent, the Anaphor Algorithm gives exactly the right result when applied to the corresponding *semantic representation*.

RESPECT
<*r x*> (second arg. = *tena* ‘*self*’)
↳ *r*

Our approach also correctly predicts the *unacceptability* of the following sentence, in which the reflexive pronoun follows its intended antecedent but functions as first argument – making it impossible for the Anaphor Algorithm to resolve the referential dependency.

	PAT	**AGT**
*Manaja	an-dRabe	**tena**.
respect	ACC-Rabe	self

‘Himself respects Rabe.’

RESPECT
<*x r*> (first arg. = *tena* ‘*self*’)
*↵

In sum, the facts from Malagasy are just what we would expect if that language has essentially the *same* system of local coreference as English: despite the differences in word order, the agent is always the first argument, and a reflexive pronoun must look to a prior argument for its reference.

9.2 Ditransitive patterns

English ditransitive patterns offer further evidence in support of the Anaphor Algorithm. As the examples below illustrate, these patterns are characterized by a contrast in the relative ordering of the patient and goal arguments.

Ditransitive patterns

Prepositional ditransitive:	Double object ditransitive:
I threw the ball to Robin.	I threw Robin the ball.

I assume that more than just word order is in play here and that the two patterns have different argument structures.

[This alternation] can be seen as an operation that takes a verb with a semantic structure containing "X causes Y to go to Z" and converts it to a verb containing a structure "X causes Z to have Y."
(Pinker 1989:82)

… the double object construction requires the semantics of caused possession and the TO*-dative construction requires the semantics of caused motion.*
(Yang 2016:191)

In other words, in the prepositional pattern, I act on the ball by making it go to Robin. On this interpretation, the patient (*ball*) is the second argument and the goal (*Robin*) is the third argument).

I threw the ball to Robin.
THROW
<ag pat go>

In the double object ditransitive, in contrast, I act on Robin by having him/her receive the ball. Thus, the goal is the second argument and the patient is the third.

I threw Robin the ball.
THROW
<ag go pat>

If this is right, then there is no fixed thematic-role hierarchy for patient and goal arguments. Rather, they can be ordered in different ways relative to each other, depending on how the event is conceptualized.

Thematic roles are positions in a structured semantic representation.
(Pinker 1989:76)

Arguments are ordered by the lexico-semantics of the verb.
(Weschler & Arka 1998:411-412)

Idioms provide independent evidence for this view. The key insight is that idioms in ditransitive patterns typically consist of the verb and its third argument. Consistent with this observation, we find idioms such as the following.

Prepositional ditransitive – the goal is the third argument:
I **threw** Harry **to the wolves**.
PAT GOAL
'I sacrificed Harry to further my own interests.'

Double object ditransitive – the patient is the third argument:
I **threw** Harry **some crumbs**.
GOAL PAT
'I made a minor concession to Harry to placate him.'
(O'Grady 1998)

As illustrated here, the idiom consists of the verb and its goal argument (*to the wolves*) in the prepositional pattern, but of the verb and its patient argument (*some crumbs*) in the double object pattern – consistent with the two proposed argument structures.

All of this leads to a prediction about coreference: the patient argument should be able to serve as an antecedent for the goal argument in the prepositional pattern, and the reverse should be true in the double object pattern. This is correct, as the next examples show.

Prepositional ditransitive:	Double object ditransitive:
<ag pat go$_{\mathrm{REFL}}$>	<ag go pat$_{\mathrm{REFL}}$>
I described Max to himself.	I showed Max himself (in the picture).
(cf. *I described himself to Max.)	(cf. *I showed himself Max in the picture.)

This is the expected result since the antecedent in each of the acceptable patterns is the second argument and hence prior to the third-argument reflexive pronoun, allowing resolution of the referential dependency by the usual algorithm.

Anaphor Algorithm
$<\alpha\ x>$
$\quad\hookrightarrow\alpha$

The interaction of anaphora with filler-gap dependencies

Further challenges occur in sentences where the syntax of anaphora interacts with the syntax of filler-gap dependencies.

Who did you describe **_** to **himself**?

***Who** did you describe **himself** to **_**?

The second of these sentences exemplifies a 'crossover effect': the *wh* word (the intended antecedent) appears to have moved across the anaphor, which does not happen in the acceptable first pattern. This contrast too follows straightforwardly from the Anaphor Algorithm.

In the acceptable sentence, the *wh* word is the second argument of DESCRIBE.

Resolution of the filler-gap dependency:

Who did you describe **_** . . .?

DESCRIBE
<*y* ***wh*** ...> (*y* = *you*, *wh* = *who*)
↑
The *wh* word is identified as the verb's second argument.

This makes *who* available to assist in resolving the referential dependency when the anaphor (the verb's third argument) is subsequently encountered.

2ND ARG 3RD ARG
Who did you describe **_** to **himself**$_{\mathbf{x}}$?

DESCRIBE
<*y wh* ***x***> (*y* = *you*, *wh* = *who*)
↳***wh***

In the unacceptable sentence, in contrast, the reflexive pronoun is the verb's *second* argument and the *wh* word is the third argument. At the point where the reflexive pronoun in encountered, this leaves only the first argument (*you*) as a potential antecedent.

3RD ARG 2ND ARG
***Who** did you describe **himself**$_x$ …

DESCRIBE
<*y* ***x*** ...> (*y* = *you*)
↳****y***

The result is a person-feature mismatch and an unacceptable sentence. This is just the right outcome.

9.3 Making sense of anaphora

How could a system of coreference built around the Anaphor Algorithm have emerged in the first place? One possibility is that sentence planning is shaped by the structure of the event that is to be expressed, independently of the language's word order conventions.

> *… because the agent drives the event to take place to begin with, [it] becomes the point of attention or anchor for the upcoming information… the agent serves to initiate the construction of an event representation followed by a subsequent linking of the actual event with patients.*
>
> (Cohn & Paczynski 2013:75)

In the case of a transitive event, planning should therefore follow a path that calls for the early projection of an agent. By proceeding from that point, it creates the conditions for a later argument (a patient, for example) to derive its referent from the prior agent argument.

Transitive pattern (*Jerry pinched himself.*)

PINCH	⇒	PINCH	⇒	PINCH
<…>		<*j*…>		<*j* ***x***>
		ag		*ag pat*
				↳*j*

This fits well with MacDonald's (2013) idea that the computational burden of planning and producing utterances promotes choices that reduce processing cost. In the case of anaphora, cost arises from the need to resolve a referential dependency. That cost can be minimized if the referent associated with a position in argument structure that is prior to that of its pronominal co-argument. This essentially gives the syntax of anaphora that I have proposed, with agents prior to patients in argument structure regardless of word order, creating the observed asymmetry in the syntax of coreference and predicting its universality.

The principles that generative grammar uses to regulate coreference are widely acclaimed for their descriptive success and have come to be a showcase example of Universal Grammar – its 'crowning achievement' according to Truswell (2014:215) and a 'window onto the mind' according to others.

> *Within the framework of generative grammar, anaphora has for some time been seen as "the window onto the mind," providing crucial evidence in support of the innateness hypothesis.*
> (Huang 2000:16)

Anaphora does indeed provide a potential glimpse into the language faculty, but what it reveals is not Universal Grammar. On the view outlined in this chapter, neither grammatical principles nor syntactic structure enters into the computation of coreference. Instead, the interpretation of reflexive pronouns is shaped by processing pressures that promote the rapid resolution of referential dependencies – the very requirement embodied in the Anaphor Algorithm. Put simply, coreference has a natural syntax.