ELSEVIER

Lingua 111 (2001) 639–666

Lingua

www.elsevier.com/locate/lingua

Exquisite connections: some remarks on the evolution of linguistic theory☆

Robert Freidin[a,*], Jean-Roger Vergnaud[b]

[a] Program in Linguistics, Clio Hall, Princeton University, Princeton, NJ 08540, USA
[b] University of Southern California, Los Angeles, CA 90048, USA

Received 5 July 2000; revised version 27 March 2001

Abstract

This article discusses the motivation for the shift from GB theory to the Minimalist Program. It explores notions of economy and conceptual naturalness in linguistics and the physical sciences – in particular, symmetries across levels of analysis. As an example of such symmetries, it proposes a new analysis for some Principle C effects. It concludes with some observations on the conduct of science. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Minimalist Program; Economy; Conceptual naturalness; Galilean science; Mathematical method; Binding theory

Lappin, Levine and Johnson's critique of the minimalist program (*Natural Language and Linguistic Theory* 18: 665–671; henceforth LLJ) raises three important issues: the relation between the MP and its predecessors (GB theory in particular), the empirical and conceptual motivation for the MP, and the relation (if any) of theoretical linguistics to the natural sciences. Sadly, LLJ's critique contributes virtually nothing to our understanding of these topics, as the following discussion will demonstrate.

☆ Authors' note: We would like to thank Noam Chomsky, Carlos Otero, Maria Rita Manzini, Martin Prinzhorn, Johan Rooryck, Alain Rouveret, Andrew Simpson, Neil Smith and Maria Luisa Zubizarreta for discussions.

* Corresponding author. Phone: +1 609 258 5403; Fax +1 609 258 4899;
E-mail: freidin@princeton.edu

1. MP vs. GB

As Chomsky and others have pointed out (see Chomsky, 1995b; Freidin, 1997a; Lasnik, 1999), the MP is grounded in the principles and parameters framework, just like the GB theory it has to a large extent superseded. It shares at least three basic assumptions with virtually all of its predecessors within modern generative grammar. First and foremost, the mind/brain contains a language faculty, a component that interacts with other cognitive systems. Next, the cognitive system of language connects with performance systems via levels of linguistic representation, perhaps limited to only two external systems, one involving articulation and perception of the physical linguistic signal and the other involving the encoding or decoding of the meaning of the signal. Under this limitation there are only two interface levels, phonetic form (PF) and logical form (LF). Finally, performance systems do not differ across languages. Even at this fundamental level Chomsky takes nothing for granted, noting that these assumptions are "not at all obvious" (1995: 3).

As both the GB theory and the MP are grounded in the principles and parameters model, they share further specific assumptions [cf. Chomsky, 1995b: 170]:

(i) regarding computational system for human language, C_{HL}, the initial state S_0 contains invariant principles and parameters (i.e. options restricted to functional elements)
(ii) selection Σ of parameters determines a language
(iii) a language determines an infinite set of linguistic expressions (SDs), each pair (π, λ) obtained from the interface levels, PF and LF
(iv) language acquisition involves fixing Σ
(v) the grammar of language states just Σ, lexical arbitrariness and the PF component aside

These shared assumptions have been standard for nearly two decades.

There is one further assumption articulated for the first time in Chomsky (1992), which could easily have been proposed prior to the MP. Chomsky offers as a narrow conjecture the suggestion that there is no variation in the overt syntax or the LF component. Ignoring PF options and lexical arbitrariness, Chomsky suggests that "variation is limited to nonsubstantive parts of the lexicon and general properties of lexical items", in which case "there is only one computational system and one lexicon", apart from the variation mentioned (1995: 170).[1] Thus viewed at the appropriate level of abstraction, there is only one human language. Of course the truth or falsity of this bold conjecture remains to be established.

We come at last to the three assumptions that are unique to the MP: (1) that the interface levels LF and PF are the only relevant linguistic levels, in spite of apparent empirical evidence to the contrary (p. 169), (2) that all conditions are interface

[1] References are to the version of Chomsky (1993) reprinted as chapter 3 of Chomsky (1995b).

conditions (p. 194), and (3) that a linguistic expression is the optimal realization of these conditions (p. 194). These three assumptions constitute what Chomsky (1999) calls the strong minimalist thesis. In contrast to the six basic assumptions that the MP shares with GB theory, the three assumptions unique to the MP do not add up to 'a major paradigm change in the theory of grammar' as will become clear from the following discussion.

For LLJ the contrast between GB theory and the MP is that "the MP adds economy principles" "in addition to local constraints on operations and the structures they produce". While it is correct to highlight the role of economy in the development from GB to the MP, it is inaccurate to claim that economy conditions are an innovation unique to the MP.

The discussion of the role of economy in grammatical analysis begins with the advent of modern generative grammar – i.e. Chomsky's *MMH* (1951). There Chomsky identifies two kinds of criteria of adequacy for grammars – one for the correct description of the structure of the language under analysis, and the other for requirements imposed by its special purposes, "or, in the case of a linguistic grammar having no such special purposes, requirements of simplicity, economy, compactness, etc". (1951: 1). In a footnote, Chomsky supplies the following clarification:

> "Such considerations are in general not trivial or 'merely esthetic'. It has been recognized of philosophical systems, and it is, I think, no less true of grammatical systems, that the motives behind the demand for economy are in many ways the same as those behind the demand that there be a system at all. Cf. Goodman (1943)."

In other words, a grammar is not merely a description of a language; it is moreover an explanatory theory about the structure of a language – i.e., why a language has the properties it does rather than other conceivable properties. It is in this context that considerations of economy, etc. first came into play.[2]

---

[2] It is worth pointing out here that in *MMH* Chomsky's notion of simplicity bears some general similarity to the more current discussions of economy.

> "For the formulation of any relative precise notion of simplicity, it is necessary that the general structure of the grammar be more or less fixed, as well as the notations by means of which it is constructed. We want the notion of simplicity to be broad enough to comprehend all those aspects of simplicity of grammar which enter into consideration when linguistic elements are set up. Thus we want the reduction of the number of elements and statements, any generalizations, and, to generalize the notion of generalization itself, any similarity in the form of non-identical statements, to increase the total simplicity of the grammar. As a first approximation to the notion of simplicity, we will here consider shortness of grammar as a measure of simplicity, and will use such notations as will permit similar statements to be coalesced." (Chomsky, 1951: 5)

To avoid circularity, the notation must be fixed in advance and neutral to any particular grammar.

> "Given the fixed notation, the criteria of simplicity governing the ordering of statements are as follows: that the shorter grammar is the simpler, and that among equally short grammars, the simplest is that in which the average length of derivation of sentences is least". (Chomsky, 1951: 6)

In current work, the 'shortness' of grammars and of derivations is driven by substantive principles of UG.

Applying economy conditions to the selection of derivations within each grammar represents a significant leap from applying them to the selection of grammars. Although economy conditions like Full Interpretation (FI) and Last Resort (LR) were not articulated in the earliest discussions, both were introduced as part of GB, the former in Chomsky (1986) and the latter in Chomsky 1991 (which first appeared in MITWPL #10 in 1989, three years prior to the advent of the MP). What happened in the 1990s, in a nutshell, was that certain economy conditions (e.g. FI) were interpreted in a way that made it natural for them to supplant a significant portion of the older GB principles.

Consider for example the Case Filter. It prohibits any phonetically realized nominal expression that is not marked for Case. Within GB, why this should be is just a matter of stipulation that appears to accord with the facts. From the point of view of economy, Case features are extraneous to interpretation at the LF interface at least and therefore should be eliminated from the derivation before LF. Otherwise, these uninterpretable features at the LF interface will violate FI, causing the derivation to crash. Thus legibility conditions imposed by the cognitive system that interfaces with C_{HL} determines how derivations proceed with respect to Case features. In this way FI and the Case Filter overlap in a way that FI subsumes the empirical effects of the Case Filter, but not conversely.[3] The preference for FI over the Case Filter is just the standard preference for the more general constraint, all things being equal.

In short, the heuristic of eliminating overlapping conditions, which has resulted in much fruitful research over several decades, is one of the central motivations for switching from the analysis of GB to that of the MP.[4]

The same logic provides one of the strongest motivations for eliminating the relation of government from current discussion. Consider the standard case of an ECP violation given in (1), where T stands for the functional category Tense and *t* is the trace of *John*.

(1) *John T is believed [that t T is happy]

Under MP analysis, (1) can be interpreted as a violation of FI in the following way. Suppose that the nominative Case feature of *John* is checked in the embedded clause

A notion of economy is also mentioned in the literature of structuralist linguistics. Hymes and Fought (1981) cite the following from Hockett (1954):

> "A model must be efficient: its application to any given language should achieve the necessary results with a minimum of machinery". (Hockett, 1954: 233)

However it is doubtful that Chomsky's notion of economy was influenced by structuralist linguistics in any significant way. MMH predates Hockett's formulation by at least three years. Moreover, Chomsky's discussion cites the work of philosopher Nelson Goodman as the source of his ideas, not work in structuralist linguistics.

[3] FI prohibits superfluous symbols in general, ruling out vacuous quantification for example, where Case is not an issue.

[4] For a detailed discussion of the rationale that has guided the evolution of generative syntactic theory, see Freidin (1994).

by T, as is the nominative Case feature of T. When *John* moves to the subject position of the matrix clause, Spec-T, only the D-feature of the nominal is available to check the D-feature of the matrix T. The nominative Case feature of *John* has been checked and thereby eliminated, so that the nominative Case feature of the matrix T remains, causing the derivation to crash at LF because of FI.[5] Thus the phenomena that fall separately under the Case Filter and the ECP within GB, are captured under a single principle within the MP analysis – moreover, one that functions as a natural legibility condition that regulates the interface between C_{HL} and other cognitive systems.[6]

Given the greater generality of FI, we would prima facie want to eliminate the ECP in favor of the more general principle.[7] Furthermore, this analysis, which explains deviance on the basis of legibility conditions imposed by cognitive systems that interface with C_{HL}, strikes us as a more promising explanatory account than the postulation of various constraints internal to C_{HL} that basically reflect the complexity of the phenomena in an essentially descriptive fashion.[8] Note incidentally that this approach is highly reminiscent of the one followed in Chomsky and Lasnik (1977). However, even if we discount impressions of what might be a more promising explanatory account, the methodological requirement to eliminate overlapping conditions whenever possible motivates the abandonment of much of the machinery of GB in favor of the MP analysis. Thus the change from GB to the MP is motivated by the same methodology that has always motivated changes in syntactic theory.

---

[5] Under the system of analysis in Chomsky (1999), the ϕ-features of the matrix T in (1) would not be checked because the N *John* would be frozen in place in the complement clause (see p. 5). Therefore ϕ-features of the matrix T would violate FI at LF, rather than a Case feature of that T, which may be an unnecessary postulation. Thus (1) is prohibited because it would involve a single DP entering into two independent Spec-head agreement relations. We assume that such configurations are generally excluded. Whether *John* could actually move to the matrix subject position of (1) is a separate issue. Given that matrix T has a D-feature (EPP) that normally drives movement, then *John* would move to matrix Spec-TP. If it does, then somehow its agreement features must be unavailable for checking the matching features of matrix T, even though they are interpretable at LF and therefore not erased. For evidence that NP-movement may be motivated solely by a D-feature of T as well as further discussion of this issue, see Lavine and Freidin (2001).

[6] The example in (1) is also ruled out by the Case uniqueness requirement of the Chain Condition. Therefore, we might also want to investigate whether the empirical coverage of the Case uniqueness requirement can be subsumed under FI as well. Notice also that certain violations of the θ-uniqueness of chains also fall out from FI with respect to unchecked Case features. For example, in the simplest case (i) the Case feature of T will not be checked.

(i) *Bill T mentioned t.

Because the Case feature of *Bill* is checked in VP, there is no Case feature to check the nominative Case feature of T.

[7] Whether this is feasible depends on a demonstration that other cases of ECP violations can be subsumed in a similar fashion under FI or some other general condition. However, it is a reasonable and promising line of investigation.

[8] The fact that that the level of complexity of the analysis mirrors the complexity of the data constitutes yet another argument against the ECP analysis. This analysis functions more like a technical description of the data than an explanatory account of the phenomena under analysis.

Another motivation for exploring the MP rather than continuing with GB concerns the central topic of phrase structure. The MP introduces bare phrase structure theory, which eliminates the ambivalent top-down and bottom-up view of phrase structure that has been characteristic of the field since the earliest formulations of X-bar theory in the late 1960s. With bare phrase structure there is only bottom up analysis of a specific kind.[9] The theory of bare phrase structure provides a derivational mechanism for syntactic representations, which has been missing from GB since the elimination of phrase structure rules (circa 1980) on the grounds that they are redundant given X-bar theory, the general principles of GB, and the specific properties of lexical items. Furthermore, bare phrase structure theory as incorporated within the MP cannot produce canonical D-structures, hence the elimination of D-structure as a level of representation follows from the nature of C_{HL} rather than a methodological stipulation.

Bare phrase structure eliminates in principle categorial distinctions for levels of phrasal projection, thereby conforming to the Inclusiveness Condition, which restricts computations solely to the elements (features) contained in lexical items. Given this condition, computations cannot introduce new elements such as bar levels, indices, or syntactic categories that are not already part of the lexical items computed.

Chomsky (1995b) is very clear that the Inclusiveness Condition provides one criterion for the 'perfection' of C_{HL}. Although precisely what Chomsky intends by talking about language as a perfect system may not be totally clear in Chomsky (1995b), this discussion is clarified considerably in Chomsky (2000), written in 1998. Here the issue is represented in terms of the language faculty (FL) as a solution to legibility conditions imposed by the cognitive systems that interface with it. An optimal solution would encompass a C_{HL} restricted to just the properties of lexical items involved in computations plus just those few operations required for derivations that connect LF with PF (perhaps just the elementary operations of adjunction/concatenation (for Merge and Move) and deletion (for feature checking)). If the behavior of derivations is controlled solely by legibility conditions imposed by other cognitive systems at the interfaces, then C_{HL} can be reduced to these bare necessities, excluding additional internal machinery like the Case Filter and the ECP.[10] Thus in terms of simplicity, economy, and non-redundancy, the MP is clearly preferable to GB.

---

9 The top-down analysis instantiated via phrase structure rules actually became suspect within GB when it was realized that phrase structure rule function as language specific stipulations of properties that were already accounted for by general principles in conjunction with the specific properties of lexical items. Therefore, given the redundant character of phrase structure rules, it was assumed that they existed neither in individual grammars nor in UG. However, without phrase structure rules in GB, there appears to be no explicit way to derive phrase structure representations, though there were explicit conditions on the form of such representations (i.e., X-bar theory). Thus bare phrase structure answers the crucial question: if not via phrase structure rules, then via what? For this reason alone, the MP constitutes an important alternative to GB.

10 At present this bare necessities view does not account for locality conditions on movement, which appear to be conditions internal to C_{HL} itself rather than the effects of legibility conditions imposed on C_{HL} by other systems that interface with it.

2. On conceptual naturalness

Appeal to general considerations of conceptual naturalness such as simplicity, economy, or non-redundancy is not unique to generative grammar.[11] It has been employed fruitfully in the more developed natural sciences – in particular, theoretical physics. The discussion of physics that follows attempts to elucidate this notion in a way that, ultimately, should illuminate its role in contemporary theoretical linguistics.

Consider, for example, Einstein's principle that all physical laws must be Lorentz-invariant. As Putnam (1962) notes: "This is a rather vague principle, since it involves the general notion of a physical law. Yet in spite of its vagueness, or perhaps because of its vagueness, scientists have found it an extremely useful leading principle". This is because they have "no difficulty in recognizing laws": a law of nature will be an equation relating "real magnitudes" that has "certain characteristics of simplicity and plausibility". In other words, determining whether Einstein's principle may be applied to any particular case will involve 'general considerations of conceptual naturalness'.

In a different area, Bohr's quantum mechanical 'Correspondence Principle' (circa 1913) is arguably rooted in such considerations. It states that, in the classical limit, the results obtained from quantum mechanics should converge with those obtained from classical mechanics.[12] According to some physicists, the research work carried out during the years 1919–1925 that finally led to quantum mechanics may be described as systematic guessing guided by the Correspondence Principle. This is then a case where considerations of conceptual naturalness appear to have played a direct role in the progress of science.

The appeal to conceptual naturalness manifests itself also in the quest for mathematical beauty, which motivates many a theoretical physicist as Dirac[13] notes:

"Theoretical physicists accept the need for mathematical beauty as an act of faith. There is no compelling reason for it, but it has proved a very profitable objective in the past. For example, the main reason why the theory of relativity is so universally accepted is its mathematical beauty". (Dirac, 1968)

---

[11] Chomsky (1995b: 1) observes that one question that has motivated the work in generative grammar is that of the conditions that are imposed on the language faculty by virtue of "general considerations of conceptual naturalness that have some independent plausibility, namely, simplicity, economy, symmetry, non-redundancy, and the like".

[12] More precisely, the principle states that the classical theory is in agreement with the experiments for processes which depend on the statistical behaviour of a large number of atoms and which involve states where the difference between neighbouring states is comparatively little.

A more specific formulation is found in Bohr (1913) for the emission and absorption of spectral lines. There, the principle is taken to postulate a general conjugation of each of the various possible transitions between stationary states with one of the harmonic oscillation components in which the electrical moment of the atom can be resolved.

The Correspondence Principle has been a staple of textbooks on quantum mechanics. However, recent experiments have shown it to be incorrect.

[13] P.A.M. Dirac shared the 1933 Nobel Prize for physics with E. Schrödinger.

In the natural sciences, while hypothesis formation may be guided by appeals to conceptual naturalness, any given hypothesis will carry weight only to the extent that it can be subjected to the inexorable test of experiment. This is the essence of the scientific method, which governs physics and linguistics alike. But there is no chosen method for elaborating the scientific hypotheses themselves. The scientific method is not concerned with that, nor could it be, for it is not possible to set up explicit rules or criteria in this area. This does not mean that 'anything goes'. But it does mean that there is a lot of diversity in the ways scientists deal with problems and arrive at solutions.

Dirac discusses this diversity of methods in theoretical physics:

"One can distinguish between two main procedures for a theoretical physicist. One of them is to work from the experimental basis. For this, one must keep in close touch with the experimental physicists. One reads about all the results they obtain and tries to fit them into a comprehensive and satisfying scheme.

The other procedure is to work from the mathematical basis. One examines and criticizes the existing theory. One tries to pinpoint the faults in it and then tries to remove them. The difficulty here is to remove the faults without destroying the very great successes of the existing theory.

There are these two general procedures, but of course the distinction between them is not hard-and-fast. There are all grades of procedures between the extremes". (Dirac, 1968)

Dirac designates the two types of procedures as "experimental" and "mathematical", respectively. He then proceeds to give several examples of the mathematical procedure:

"Maxwell's investigation of an inconsistency in the electromagnetic equations of his time led to his introducing the displacement current, which led to the theory of electromagnetic waves. ... Einstein noticed a difficulty in the theory of an atom in equilibrium in blackbody radiation and was led to introduce stimulated emission, which has led to the modern lasers. [this is Einstein, 1917; RF&JRV] But the supreme example is Einstein's discovery of his law of gravitation, which came from the need to reconcile Newtonian gravitation with special relativity". (Dirac, 1968)[14]

Dirac's notions also apply to the founding work in quantum mechanics between 1913 and 1925. The following description is striking:

"Whether one follows the experimental or the mathematical procedure depends largely on the subject of study, but not entirely so. It also depends on the man. This is illustrated by the discovery of quantum mechanics.

Two men are involved, Heisenberg and Schrödinger. Heisenberg was working from the experimental basis, using the results of spectroscopy, which by 1925 had accumulated an enormous amount of data. Much of this was not useful, but some was, for example the relative intensities of the lines of a multiplet. It was Heisenberg's genius that he was able to pick out the important things from the great wealth of information and arrange them in a natural scheme. He was thus led to matrices.

Schrödinger's approach was quite different. He worked from the mathematical basis. He was not well informed about the latest spectroscopic results, like Heisenberg was, but had the idea at the back of his mind that spectral frequencies should be fixed by eigenvalue equations, something like those that fix the

[14] One could add to this short list Dirac's own discovery of the correct laws of relativity quantum mechanics, which was arrived at simply by guessing the equation (see Feynman, 1965: 57).

frequencies of systems of vibrating springs. He had this idea for a long time, and was eventually able to find the right equation, in an indirect way". (Dirac, 1968)

The 'mathematical procedure' typically arises in what Husserl has called the 'Galilean style of science', in recognition of its origins in the work of Galileo. Weinberg (1976) characterizes this style as follows:

"... we have all been making abstract mathematical models of the universe to which at least the physicists give a higher degree of reality than they accord the ordinary world of sensation."

More generally, one can define Galilean science as the search for mathematical patterns in nature.[15] As Chomsky notes, implementing the Galilean style entails a "readiness to tolerate unexplained phenomena or even as yet unexplained counterevidence to theoretical constructions that have achieved a certain degree of explanatory depth in some limited domain, much as Galileo did not abandon his enterprise because he was unable to give a coherent explanation for the fact that objects do not fly off the earth's surface" (1980, 9–10).

A significant feature of the Generative Revolution in linguistics has been the development of a Galilean style in that field. And, to a great extent, the recent developments within MP must be viewed in this light – specifically, as Dirac's mathematical procedure (method) at work within linguistics. Dirac has identified two main methods within the mathematical procedure itself: one is to remove inconsistencies, the other, to unite theories that were previously disjoint (see Dirac, 1968). In linguistics, the inconsistencies primarily concern overlapping grammatical conditions, as discussed earlier, which conflict with the basic assumption that C_{HL} has an optimal design. Note further that this assumption itself relates directly to the quest for mathematical beauty, which informs the Galilean style.

One aspect of Dirac's mathematical procedure as applied in linguistics involves the effort to extend and deepen the mathematical formalism used to express syntactic concepts and syntactic principles. We will refer to this facet of the Minimalist endeavor as the 'Generative Program' for the study of language (GP) because it originates in Chomsky's foundational work in the fifties and sixties and has been essential to the development of the Galilean style in linguistics.[16] However, it should be obvious that linguistics and physics are at very different stages of mathematical maturation. From this perspective, it is useful to distinguish the 'Galilean character' of an area, i.e., how much of the subject matter can be analyzed mathematically, from what one could call its 'Pythagorean character', how much of mathematics is put to use in the Galilean treatment. Linguistics and physics have the same Galilean character, although they obviously differ in Pythagorean character.[17]

The difference in mathematical status between physics and linguistics partly reflects the more general difference between physics and biology – especially from the perspective that generative grammar is ultimately a branch of theoretical biology,

[15] Of course, there are many different types of mathematical patterns: algebraic, geometrical, analytical, topological, ... etc.

more specifically, of theoretical developmental biology. In biology, the genetic code rather than mathematics has been the tool of choice for explaining life.

This, however, appears to be a historical accident, not the result of some principled difference between biology and the physical sciences. Mathematics has a central explanatory role to play in biology, as discussed in Stewart (1998), whose title, *Life's other secret*, is intended as contrapuntal to 'life's first secret', which is the genetic code:

"The mathematical control of the growing organism is the *other* secret – the second secret, if you will – of life. Without it, we will never solve the deeper mysteries of the living world – for life is a partnership between genes and mathematics, and we must take proper account of the role of *both* partners. This cognizance of both secrets has run like a shining thread through the history of the biological sciences – but it has attracted the mavericks, not the mainstream scientist. Instead of thinking the way most biologists think, these mavericks have been taking a much different approach to biology by thinking the way most physical scientists and mathematicians think. This difference in working philosophy is the main reason why understanding of the deeper aspects of life has been left to the mavericks." (Stewart, 1998: xi)

The main message of d'Arcy Thompson, one of the great mavericks in biology, is that "life is founded on mathematical patterns of the physical world".[18] Thus one role of theoretical biology is to identify such mathematical patterns and elucidate the way they function in organisms:

"The role of mathematics [in biology] is to analyze the implications of models – not 'nature red in truth and complexity', as Tennyson did not quite say, but nature stripped down to its essence. Mathematics pursues the necessary consequences of certain structural features. *If* a planet can be considered a uniform sphere, what would its gravitational attraction be like? ... *If* the movement of cells in some circumstances is controlled by physical forces and does not greatly depend on complicated internal features such as mitochondria, what will the cells do? From this point of view, the role of mathematics is not to explain biology in detail, but to help us separate out which properties of life are consequences of the deep mathematical patterns of the inorganic universe, and which are the result of more or less arbitrary initial conditions programmed into lengthy sequences of DNA code." (Stewart, 1998: 243–244)

It is worth noting at this point that Chomsky was aware that both approaches, separately or jointly, might account for the human language faculty. In criticizing the empiricist view of language acquisition in the first chapter of Chomsky 1965 (written in 1958–1959, as mentioned in Huybregts and van Riemsdijk, 1982), he notes:

"...there is surely no reason today for taking seriously the position that attributes a complex human

[16] See Chomsky, 1955, 1957, 1965. Following standard practice, we view language as one component of the human mind. Thus the study of language via GP concerns human cognition, and human biology more broadly.

[17] Considering that modern mathematics with its dazzling complexity evolved in great part from the study of numbers, it is to be expected that a science that is concerned with quantitative relations like physics will tend to make maximal use of the structures found in mathematics. This is all the more so since there exist many connections between the different parts of mathematics.

[18] See Stewart, 1998: 243. Stewart's book, which features a quotation from Thompson (1942) at the beginning of each chapter, constitutes a contemporary commentary on Thompson's seminal ideas.

achievement entirely to months (or at most years) of experience, rather than to millions of years of evolution or to principles of neural organization that may be even more deeply grounded in physical law ..." (p. 59)

However, twenty years later, Chomsky is openly skeptical of a purely genetic approach to evolution.

"It does seem very hard to believe that the specific character of organisms can be accounted for purely in terms of random mutation and selectional controls. I would imagine that biology of 100 years from now is going to deal with evolution of organisms the way it now deals with evolution of amino acids, assuming that there is just a fairly small space of physically possible systems that can realize complicated structures." (Huybregts and van Riemsdijk, 1982: 23)

From this point of view, the more promising approach is "d'Arcy Thompson's attempt to show that many properties of organisms, like symmetry, for example, do not really have anything to do with a specific selection but just with the ways in which things can exist in the physical world" (Huybregts and van Riemsdijk, 1982: 23).[19]

The mathematical perspective informs the Generative Program (GP), in effect, 'the study of language's other secret'. Thus Chomsky's mathematical work defines a central facet of GP, beginning with his construction of the foundations of modern generative grammar in Chomsky (1951) and (1955–1956).[20]

Because the MP is a particular implementation of GP, the notion of 'perfection' often invoked within MP is *ultimately* a mathematical notion, calling for a higher level of mathematical formalization in syntax.[21] The Minimalist conjecture that C_{HL} is a 'perfect system' is a tentative claim about the form and the complexity of each computation. The claim is (i) that each computation can be represented as an abstract mathematical structure completely defined by interface (output) conditions and ii) that this structure is an extremum in some mathematical space. A natural metric for the comparison of computations is their complexity as measured by their length. Note that, if the only constraints on C_{HL} are those that follow from legibility conditions at the interfaces, then it is unavoidable that some notion of computational cost should be part of the definition of 'effective' computations, since, within such a sys-

[19] See Jenkins (2000) for further discussion.

[20] In the early 1950s Chomsky had developed a mathematical understanding of natural language, which he then brought to bear on current issues in automata theory – in particular, demonstrating the inadequacy of finite state automata as a model of natural language (Chomsky, 1956, 1959) and investigating more broadly the relation between automata and grammars. The famous "Chomsky Hierarchy" of formal grammars (and corresponding formal languages) is due to him (Chomsky, 1959), and so are the proofs of associated central theorems about regular grammars (Chomsky and Miller, 1958) and context-free grammars (Chomsky, 1962), all results that have ever since been a staple of textbooks in computer science. Chomsky (1962), for example, establishes the equivalence of context-free languages and push-down automata (which was proved independently by M.P. Schützenberger and by J. Evey). For additional clarification concerning the context of Chomsky's contributions to computer science, see Otero (1994).

tem, it is always possible to combine a computation with a 'vacuous one' (i.e., one that has a null effect). The unidirectionality of movement (if it is a fact) would then be a particular design feature aimed at reducing the likelihood of vacuous steps.

Considerations of economy have a long standing legitimacy in the physical sciences. It was in physics that an economy principle of any depth was first advanced.[22] This was the *principle of least time*, discovered by Fermat circa 1650.[23] That principle states that, out of all possible paths that it might take to get from one point to another, light takes the path which requires the *shortest time*.[24] Fermat's principle is a particular instance of the general physical principle of 'least action'. Another important economy principle of physics is "the idea that the inorganic world is fundamentally lazy: it generally behaves in whatever manner requires the least energy" (Stewart, 1998: 16). That idea was for Thompson (1942) a central principle underpinning the mathematics of growth and form found in living organisms.

Comparing Fermat's principle with Snell's theory of light,[25] Feynman notes that such economy principles have a special philosophical character distinct from causal explanations of phenomena.

"With Snell's theory we can 'understand' light. Light goes along, it sees a surface, it bends because it does something at the surface. The idea of causality, that it goes from one point to another, and another, and so on, is easy to understand. But the principle of least time is a completely different philosophical principle about the way nature works. Instead of saying it is a causal thing, that when we do one thing,

---

21 We stress 'ultimately' because the MP is a research program based on specific conjectures, not a theory or even a framework. As Chomsky has noted (1998), "there are minimalist questions, but no specific minimalist answers". It should go without saying that whatever minimalist answers we might discover will only be found by actively pursuing the questions posed by the MP. Furthermore, it should be noted that Chomsky has been quite clear about the provisional nature of the MP, saying explicitly that it could turn out to be wrong, or equally problematic, premature (i.e. in much the same way that Einstein's search for a unified field theory was premature, though not wrong if developments in string theory succeed (see Greene, 1999)).

22 The first 'economy principle' acknowledged within the Western intellectual tradition actually is the maxim known as Ockham's razor: *Entia non sunt multiplicanda praeter necessitatem.* The prominent 14th century philosopher and logician William of Ockham (c. 1295–1349) has traditionally been credited with this principle (hence the name). However, the historical record suggests otherwise. We quote from Kneale and Kneale (1962):

> "No doubt this [the maxim RF-JRV] represents correctly the general tendency of his philosophy, but it has not so far been found in any of his writings. His nearest pronouncement seems to be *Numquam ponenda est pluralitas sine necessitate*, which occurs in his theological work on the *Sentences* of Peter Lombard". (Kneale and Kneale, 1962: 243)

See also Boehner, 1958.

23 Fermat had been preceded by Hero of Alexandria, who had stated that the light travels in such a way that it goes from a point to a mirror and then to another point in the shortest possible *distance*.

24 Actually, as Feynman points out, the principle as stated is incorrect, since it would predict that light emanating from a point in front of a mirror should avoid the mirror! There is a more exact formulation that avoids this problem and coincides with Fermat's original formulation in the case of refraction of light. See Feynman R.P., R.B. Leighton and M. Sands, 1963, Chapter 26.

25 Willebrord Snell, a Dutch mathematician, found the formula describing the change of angle of a ray of light that goes from one medium into another.

something else happens, and so on, it says this: we set up the situation, and *light* decides which is the shortest time, or the extreme one, and chooses that path". (Feynman et al., 1963: 26–7)

Feynman's observation extends to all economy considerations developed in the natural sciences. Economy principles fall under what 17th and 18th philosophers called 'final causes', as opposed to 'efficient causes'.[26] Efficient causes are essentially mechanistic in nature like those invoked in a Newtonian account of the dynamics of a point particle, for example, or Snell's account of refraction as described by Feynman above. Final causes involve a deeper level of understanding, as Feynman notes:

"Now in the further development of science, we want more than just a formula. First we have an observation, then we have numbers that we measure, then we have a law which summarizes all the numbers. But the real *glory* of science is that *we can find a way of thinking* such that the law is *evident*". (Feynman et al., 1963: 26–3)

Thus, the distinction between efficient and final causes is locally one of levels of analysis and globally one of levels of explanation.

The notion 'level' (of analysis, of explanation) is evidently crucial. The natural sciences provide instances where successful explanatory theories that had been developed at a certain level were later unified with theories at some other level. This is the case for classical thermodynamics, which is deducible from statistical mechanics (hence a reduction). Also the unification of structural chemistry with physics was made possible by the development of quantum mechanics, which provided a common foundation (see Chomsky, 1995a, and Smith, 1999 for discussion). However, the explanatory import of a theoretical principle at some given level L is in general relatively independent of the possibility of unifying L with other levels. A case in point is that of the principle of 'least action' mentioned above (the general principle subsuming Fermat's principle of least time), which is reducible to other principles in every area where it applies (see Jourdain, 1913 and Lanczos, 1970 for discussion). Thus, it applies in classical mechanics, where it is known as 'Hamilton's principle'. And, indeed, Hamilton's principle is an alternative formulation of classical mechanics, equivalent to the Newtonian formulation. As it turns out, though, the Hamiltonian formulation has desirable features not found within the Newtonian formulation. For example, the Hamiltonian formalism can be generalized to all types of coordinates and, furthermore, is more convenient than Newton's equations when the system is complex. But the real importance of the Hamiltonian formalism arises from the fact, both, that it can be generalized to classical electricity and magnetism (with an appropriate Lagrangian) and that it constitutes the point of departure for the quantization of physical systems (see the discussion in Cohen-Tannoudji et al., 1996: 1476–1491, for example).

There may be deep reasons for this remarkable generality. The following excerpt from Toffoli (1999) is intriguing in that respect:

---

[26] See the discussion in Thompson (1942): Chapter 1.

"We are taught to regard with awe the variational principles of mechanics [such as Hamilton's principle RF-JRV]. There is something miraculous about them, and something timeless too: the storms of relativity and quantum mechanics have come and gone, but Hamilton's principle of least action still shines among our most precious jewels.

But perhaps the reason that these principles have survived such physical upheavals is that after all they are not strictly *physical* principles! To me, they appear to be the expression, in a physical context, of general facts about *computation,* much as the second law of thermodynamics is the expression, in the same context, of general facts about *information.* More specifically, just as entropy measures, on a log scale, the number of possible microscopic *states* consistent with a given macroscopic description, so I argue that action measures, again on a log scale, the number of possible microscopic *laws* consistent with a given macroscopic behavior. If entropy measures *in how many different states you could be in detail* and still be substantially the same, then action measures *how many different recipes you could follow in detail* and still behave substantially the same." (Toffoli, 1999: 349–350)

If this is on the right track, the computational significance of the Hamiltonian formalism supersedes any deduction of it in any particular subdomain.[27]

The computational nature of economy considerations provides a link between physics and linguistics, at least metaphorically. Whether it is stronger than that will have to be determined by a future neuroscience that can validate the physical approach to complex mental structures as suggested by Chomsky, extending the views of d'Arcy Thompson. In any event, economy considerations contribute substantially to what constitutes the 'perfection' of the computational system in both domains. Whether these considerations for each domain turn out to be related or the same remains an empirical question for the future.

In linguistics, there are several ways the 'perfection' of C_{HL} could be manifested in terms of economy conditions. Shortness of derivation is only one symptom of perfection. Another manifestation, possibly equivalent in some cases, would be the existence of symmetries across levels of analysis, given that such symmetries enhance the economy of computations.

---

[27] In the light of this discussion, the following statement (LLJ: 666) appears to be profoundly in error: "Finally, one may suggest that the notion of perfection that Chomsky has in mind is based upon an analogy with the minima and maxima principles of physics. So, for example, air pressure in a soap bubble produces a spherical shape as the optimal geometric design for distributing this pressure. Similarly, light reflecting of a mirror takes the path of least time between two points. If this is, in fact, the sort of optimality that Chomsky has in mind, then it has no place in the theory of grammar. Minimization/maximization principles are derived from deeper physical properties of the particles (waves, vectors, etc.) which satisfy them. They follow from the subatomic structure and attributes of these particles, and are not themselves basic elements of the theory. Hence they have no independent explanatory status within physics, but are reducible to other principles. By contrast, the MP takes economy conditions to be essential elements of the grammar and the optimality which they encode to be one of its defining properties."

LLJ claim that because the empirical content of a principle X is deducible from other more elementary considerations, X has "no independent explanatory status". They suggest that this applies to linguistics as well as physics and therefore that the economy principles discussed in linguistics cannot legitimately be considered part of the theory of language. In the case of linguistics, this suggestion is moot because as yet no deductive relation with more elementary considerations has been established. Therefore it is both natural and rational to consider economy conditions as fundamental. In the case of physics, the point appears to be mistaken as the text above indicates.

To illustrate, consider the following well-known contrast in anaphoric interpretation for the paradigm in (2):[28]

(2) a. Mary thinks she solved the problem.
b. She thinks Mary solved the problem.

While *Mary* in (2a) may be construed as anaphoric with *she*, this is not a possible construal for (2b). Exactly how we account for this depends crucially on what representations are available. Prior to the Minimalist Program these anaphoric representations would be given in terms of co-indexing generated by a rule of Index NP (see Freidin and Lasnik, 1981 for discussion). Thus the construals under discussion would be given as (3a) and (3b) respectively, where the viable construal of (2b) is given as (3c).

(3) a. $Mary_i$ thinks she_i solved the problem.
b. *She_i thinks $Mary_i$ solved the problem.
c. She_j thinks $Mary_i$ solved the problem.

However, given the Inclusiveness Condition (4), which we take to be central to the Minimalist Program, indices are not legitimate elements of representations.

(4) *Inclusiveness Condition* "Outputs consist of nothing beyond properties of the lexicon (lexical features)" (Chomsky, 1995b: 225).

Therefore the construals of (2) indicated in (3) will have to be represented another way. We shall assume that a definite pronoun is a definite description with a silent NP component (cf. Postal, 1966 and Brody, 1982). Specifically, we posit the following underlying representation for a pronoun:

(5) *[$_{DP}$ [+def] φ ~~NP~~]*, with φ the agreement features of the nominal expression and ~~NP~~ the silent NP component.[29]

For example, the pronoun *she* has the representation in (6):

(6) [$_{DP}$ [+def] [3rd person, singular, feminine] ~~NP~~]

The form *she* is taken to be the PF realization of the morphosyntactic representation in (7):

(7) [$_{DP}$ [+def] [3rd person, singular, feminine]]

---

[28] The discussion of Principle C follows from Vergnaud, 1998, which is a hand-out for a lecture given at UCLA, explores various notions of 'multiplicity of representations'.

[29] The DP constitutes a definite description where the head D is indicated by the feature [+def]. At this point whether the agreement features φ are associated with D or N is left open.

The NP component of the pronoun determines its interpretation: two different interpretations of a pronoun reflect two distinct underlying representations of that pronoun. For example, the sentence in (2a) is represented as in (8) when *she* is construed as anaphoric with *Mary*, but as in (9) when *she* is construed as referring to *Clea*[30]:

(8) Mary thinks *[[+def] [3rd pers., sg., fem.] ~~Mary~~]* solved the problem.
(9) Mary thinks *[[+def] [3rd pers., sg., fem.] ~~Clea~~]* solved the problem.

We propose to relate the interpretive contrast in (2) to symmetries in the representations of the structures involved.

The defining property of a pronominal element like *she* in (2) is that its PF representation is invariant under substitution of its NP component. Call this the *pronominal symmetry*:

(10) *Pronominal Symmetry*
Let *pro* be some singular pronoun analyzed as *[$_{DP}$ [+def] ϕ (~~NP~~]*. The PF representation of *pro* is invariant under the substitution in (i):
(i) NP → NP′

---

30 Questions arise in the case of such structures as that in (i) (discussed in Jacobson, 1977):
(i) [the man who loved her$_2$]$_1$ kissed [his$_1$ wife]$_2$
We have the following descriptions for the constituents *his wife* and *her* in (i):
(ii) a. [his$_1$ wife]$_2$ = [[[+def] ϕ_1 ~~man~~]'s ϕ_2 wife], with ϕ_2 and ϕ_2 the agreement features for *man* and *wife*, respectively
b. her$_2$ = [[[+def] ϕ_1 ~~man~~]~~'s~~ ϕ_2 ~~wife~~]
We assume that;
(iii) The DP "the man's wife" is [+def] for the same reason that "a man's wife" is [–def] (technical details aside).
(iv) The structure *[[[+def] ϕ_1 ~~man~~]'s* ϕ_2 ~~wife~~] in (iib) is ambiguously realized as *his* or as *her*
(v) In (iv) above, *his* is the PF realization of *[+def] ϕ_1 's* and *her*, of *[+def] ϕ_2*
Independent principles (having to do with contrast at surface structure) determine which one of the alternative forms in (v) is realized. An important fact discovered by P. Jacobson is that *her* in (i) must be analyzed as a copy of the whole antecedent *his wife*, and not merely as a copy of *[+def] wife*. In other words, *her* in (i) must be described as in (iv). Call this the *Principle of Anaphora Interpretation*:
(vi) Definite anaphora only holds between complete DPs.
An analogous principle was postulated in Vergnaud 1974. The principle in (vi) entails the ungrammaticality of (vii) on the indicated reading (see Brody, 1982):
(vii) [his$_2$ employer]$_1$ respects [her$_1$ secretary]$_2$
The constituents *his employer* and *her secretary* in (vii) are described as in (viii):
(viii) a. [his$_2$ employer]$_1$ = [[[+def] ϕ_2 ~~secretary~~]'s ϕ_1 employer]
b. her$_1$ secretary]$_2$ = [[[+def] ϕ_1 ~~employer~~]'s ϕ_2 secretary]
Note that the above assumptions require that the head of the relative clause construction in (i) (*the man*) be analyzed as a DP. The impossibility of (ix) is presumably related to that of (x):
(ix) *[The President]$_i$ said that [he]$_i$ that had been elected could not resign
(x) *[A man]$_i$ came in. [The man]$_i$ that looked tired sat down.
The example (x) contrasts with (xi):
(xi) [A man]$_i$ came in. [The man]$_i$ sat down.

No matter what representation is assigned to NP, the PF representation of *pro* remains constant. We formalize this as in (11):

(11) Let *pro* be some occurrence of a pronominal item in a structure Σ:
(i) pro = [$_{DP}$ [+def] [nth person, α number, β gender] ~~NP~~].
Define *pro(NP′)* to be the pronominal element obtained by substituting *NP′* for *NP* in (i). Define the *range of pro*, denoted by *<pro>*, to be the set of all pronouns *pro(NP′)* for which *NP′* has the same agreement feature specifications as *NP*. Note that *<pro>* includes *pro* itself. It also includes such descriptions as *pro(scientist that discovered radioactivity)*, *pro(Clea)*, *pro(trigger-happy officer)*, etc.[31]

Thus all elements in the range of *pro* share the same PF representation.

Now, there is a general principle in grammar that items in a structure are not interpreted in isolation, but always with respect to some larger domain. Technically, grammar constructs an interpretation of the head and of the specifier of *x* only at the level of some constituent properly containing *x*. Call this the *Generalized Phase Conjecture*, in reference to the analysis proposed in Chomsky (1999):

(12) *Generalized Phase Conjecture (GPC)*
Given some constituent *C*, the head of *C* and its specifier are interpreted at the level of a constituent *[$_P$ ... C ...]*, where P properly contains C. *P* is called the *phase for C*.

Chomsky (1999) considers a notion of phase that seems appropriate for the interpretation of expressions involving displacements. We conjecture that a different notion of phase exists for the assessment of anaphoric relations. Specifically, the phase for a pronominal expression *pro* is its c-command domain.[32]

Considering now the paradigm in (2), let us call P_{she} the phase for the pronoun *she*. For the form in (2a), P_{she} is the embedded TP *[she solved the problem]*. The pronominal symmetry associated with *she* carries over to P_{she}: the PF representation of the phase of *she* is invariant under substitution of NP in the representation of *she*, quite obviously. We assume this to be a general requirement for phases, stated as (13):

(13) *The Principle of Phasal Coherence (PPC)*
Given some structure Σ containing constituent *x*, let P_x be the phase for *x* (i.e., the minimal phase containing *x*). Then, every interpretive symmetry of *x* must also be a symmetry of P_x.

---

[31] The range of an expression is more than a mere list. To wit, it is a semi-lattice structure (see Vergnaud and Zubizarreta, 2000). Note that the range of an expression may also include descriptions outside the linguistic modality, e.g., visual descriptions (see Vergnaud and Zubizarreta, 2000).

[32] For the purpose of that discussion, we assume that c-command is reflexive, so that the c-command domain of *x* contains *x*.

Given the PPC, the PF invariance of a pronoun *pro*, which constitutes the pronominal symmetry, must in general carry over to the phase of *pro* P_{pro}. We evaluate satisfaction of the PPC by extending the notion of 'range' to P_{pro}:

(14) In a structure Σ, let *pro(NP)* be some occurrence of a pronominal item and let P_{pro} be the phase of *pro(NP)* in Σ. Denote by $P_{pro}(NP')$ the constituent obtained from P_{pro} by substituting *NP′* for *NP* in *pro(NP)*. Define the *range of* P_{pro} *relative to pro*, denoted by $<P_{pro}, pro>$, to be the set of all constituents $P_{pro}(NP')$ for which *NP′* has the same phi-feature specifications as *NP*. Note that $<P_{pro}, pro>$ includes P_{pro} itself.

Accordingly, the range of *pro* in (15a) establishes the set of parallel structures in (15b):

(15) a. $<pro> = \cup_i$ {[$_{DP}$ [+def] ϕ ~~NP~~$_i$]}
b. $<P_{pro'}, pro> = \cup_i$ { P_{pro}([$_{DP}$ [+def] ϕ ~~NP~~$_i$])}

Then:

(16) The pair (*pro*, P_{pro}) satisfies the PPC only if all structures in $<P_{pro}\ pro>$ share the same PF.

In this way, (2a) satisfies the PPC.[33]

Consider next whether ii) satisfies the PPC. In that structure, the phase P_{she} is the matrix TP containing in addition to the pronoun a second DP which could, but need not, relate to the interpretation of the pronoun. In the case where *she* in (2b) is interpreted as *Mary*, the corresponding representation is that in (17) (with the phi-features *[3*rd *person, singular, feminine]*):

(17) [[[+def] ϕ ~~*Mary*~~] thinks *Mary* solved the problem]

The structure in (17) contains the accidental chain in (18):

(18) (~~Mary~~, Mary)

The set of parallel structures established by the range of *pro* in this case includes one structure in which a pair of expressions are anaphorically linked, to wit the structure

[33] Observe that, in a different domain, the rule of Quantifier Raising (QR) can be taken as a manifestation of the PPC. QR ensures that the inherent variable interpretation associated with a quantified expression *QNP* is carried over to the phase of *QNP*. The existence of such symmetries among linguistic representations might suggest an approach to the study of the underlying neural system very much in the spirit of that pioneered by Stewart (1998, 1999) for the study of animal locomotion (see Stewart, 1998: Chapter 9, for example).

in (17). This conformation is subject to the Parallelism Principle[34] as formulated in (19):

(19) *Parallelism Principle for anaphora*
Let $\{\Sigma_i\}_{i \in I}$ be a set of parallel structures and let (*N,pro*) be a pair of nominal expressions with *pro* a pronoun such that *N* and *pro* are anaphorically linked in Σ_p, $p \in I$. Then, either (i) the anaphoric link remains constant across all structures in $\{\Sigma_i\}_{i \in I}$ (the case of 'sloppy identity') or (ii) the value of *pro* remains constant across all structures in $\{\Sigma_i\}_{i \in I}$.

Given the definition of the range of *pro*, case (ii) doesn't apply to the set of parallel structures $<P_{pro}, pro>$. The application of case (i) amounts to revising the definition of the range $<P_{pro}, pro>$ as follows:

(20) In a structure Σ, let *pro(NP)* be some occurrence of a pronominal item and let P_{pro} be the phase of *pro(NP)* in Σ. Denote by $P_{pro}(NP'[i])$ the constituent obtained from P_{pro} by substituting *NP′* for some occurrence *NP[i]* of *NP* in P_{pro}. The range of P_{pro} relative to *pro*, denoted by $<P_{pro}, pro>$, is defined as the maximal set of structures $P_{pro}(NP'[i])$ that verifies (i)–(iii):
(i) *NP′* has the same phi-feature specifications as *NP*
(ii) $<P_{pro}, pro>$ includes P_{pro}
(iii) $<P_{pro}, pro>$ obeys the Parallelism Principle.

If we apply this definition to the structure in (17), then, the range of P_{she} relative to *she* includes such structures as those in (21):

(21) $<P_{she}, she>$:
(i) [[[+def] ϕ ~~Clea~~] thinks *Clea* solved the problem]
(ii) [[[+def] ϕ ~~Susan~~] thinks *Susan* solved the problem]
(iii) Etc.

Because the set $<P_{she}, she>$ in (21) is not PF invariant, the PPC is violated.[35] In this way, the construal (3b) of (2b) is excluded. Note that, in the case of the construal in (3c), the range of P_{she} may not include the structure in (17) – the structure where *she* is anaphoric with *Mary* – by (ii) and (iii) of (20). It is easy to check that (2b), under construal (3c), satisfies the PPC: no matter the value of *pro* within the admissible range, the pronoun and its phase will both remain PF invariant. In essence, Principle C reflects a conflict between Parallelism and Phasal Coherence: in the

[34] Introduced by Chomsky in class lectures during the mid-seventies.

[35] By contrast, no violation of the PPC occurs in the case of (2a) because P_{she} is the embedded TP and therefore does not contain an accidental chain, even though the matrix TP does. The notion of 'range' introduced in the text analysis could be developed so as to provide an account of the anaphoric interpretation of pronouns in the spirit of that of Lasnik (1976), where there is no specific rule of grammar that establishes an anaphoric reading between a pronoun and an antecedent.

case of a structure such as (17), there is no coherent definition of 'range of a phase' that can satisfy both principles.[36]

To summarize, Principle C follows from the interaction of the Principle of Phasal Coherence, related to QR, with the Parallelism Principle.[37] This account immediately extends to the contrast in (21) if the chunk *him* in *himself* is treated as a pronoun falling under the analysis above:

(22) a. Clea believes Luc$_i$ to have introduced himself$_i$ to Mary.
b. Clea believes himself$_i$ to have introduced Luc$_i$ to Mary.

The above account also extends to the following paradigm from French:

(23) a. Le juriste$_1$ sait très bien qu'il$_i$ est en difficulté.
'The jurist$_i$ knows very well that he$_i$ has a problem.'
b. Il$_i$ sait très bien que le juriste$_i$ est en difficulté.
'He$_i$ knows very well that the jurist has a problem.'
c. *Il$_i$ sait très bien que le juriste qu'il$_i$ est est en difficulté.
'He$_i$ knows very well that the jurist that he$_i$ is has a problem.'
d. Pierre$_i$ sait très bien que le juriste qu'il$_i$ est est en difficulté.
'Peter$_i$ knows very well that the jurist that he$_i$ is has a problem.'
e. *Il$_i$ sait très bien que le juriste que Pierre$_i$ est est en difficulté.
'He$_i$ knows very well that the jurist that Peter$_i$ is has a problem.'
f. *Le juriste qu'il$_i$ est sait très bien que Pierre$_i$ est en difficulté.
'The jurist that he$_i$ is knows very well that Peter$_i$ has a problem.'

[36] A fundamental aspect of the account of Principle C proposed in the text is that it centrally relies on the PF distinction between the full-fledged nominal expression *[[+def] φ NP]* and its pronominal counterpart *[[+def] φ ~~NP~~]*. The prediction is then that Principle C is inoperative in elided constituents such as deleted VPs. The prediction appears to be correct, as shown by the grammaticality of the construal in which *he* is anaphoric with *John* for the form in (i) (see Fiengo and May, 1994: 220):

(i) Mary loves John, and he thinks that Sally does, too.

Indeed, the representation of pronouns proposed in the text would obviate the need for any 'vehicle change' in the sense of Fiengo and May (1994) (see also Lappin, 1991 and Lappin and McCord, 1990). Note however the contrast between (ii) and (iii):

(ii) Mary believes that John is eligible and Sally claims he does, too.
(iii) Mary believes John to be eligible and Sally claims he does, too.

The pronoun *he* may be construed as anaphoric with *John* in (ii), but not in (iii). This suggests that an independent LF constraint is at work in the case of (iii), presumably the Principle of Disjoint Reference (see Chomsky, 1976). If this is on the right track, then the behavior of Principle C with respect to reconstruction phenomena needs to be reconsidered.

[37] Given the representation of pronouns postulated, the text account reduces Principle C to the law that states that only the most c-commanding element in a chain may be realized at PF. Conversely, one can take that law to be a subcase of Principle C.

Note that the Parallelism Principle itself can be described as a principle of 'symmetry preservation'. The relevant symmetry in that case is the invariance of LF interpretation under the permutation of anaphorically linked expressions.

g. Le juriste que Pierre_i est sait très bien qu'il_i est en difficulté.
'The jurist that Peter_i is knows very well that he_i has a problem.'

h. Le juriste qu'il_i a nommé sait très bien que Pierre_i est en difficulté.
'The jurist that he_i appointed knows very well that Peter_i has a problem.'

(23a–b) show the standard contrast for disjoint reference under c-command as in (2). In surprising contrast, (23c) allows the coreferential interpretation between the pronominal matrix subject and the complement subject via the pronoun in the relative clause. The same anaphoric behavior obtains when pronominal matrix subject is replaced by an R-expression, as in (23d). However, disjoint reference obtains again if the pronoun in the relative clause in (23c) is replaced by an R-expression, as illustrated in (23e). Note that (23f) results from transposing the matrix and complement subjects in (23d) and thus this pair is exactly parallel to (23a–b). This analysis extends to the pair (23e,g). The example in (23h) is grammatical as expected, in contrast to (23f).

The paradigm in (23) shows that the constituent *[le juriste que DP est]* 'the jurist that DP is' has the same anaphoric behavior as the *DP* in it. For the purpose of applying Principle C, it behaves as a pronoun when *DP* is a pronoun and as a name otherwise.[38] *DP* in turn behaves as if it occupied the position of the head modified by the relative clause. Noting that the predicate *juriste* and its subject *DP* within the restrictive relative clause construction *[le juriste que DP est]* share the same set of phi-features ϕ, we shall assume that the notion of symmetry extends to such pairs of constituents:

(24) Let (C, C') be a pair of constituents in Σ that share the same phi-features. Let S be some interpretive symmetry that holds of the pair (C, C'). The PPC is extended to such a case, requiring that S also hold of the minimal phase for (C, C').

Consider in this light the form *[le juriste que pro(NP) est]*, with *pro(NP)* a pronoun. The PF of the pair *(juriste, pro(NP))* remains invariant under the substitution of *NP'* for *NP* in the pair *(juriste, pro(NP))*. By the above extension, *pro(NP)* establishes a range not only for its own phase, but also for the phase of the raised predicate *le*

---

[38] Technically, the generalization encompasses other cases – e.g., (i).

(i) *Pierre_i sait très bien que le juriste que Pierre_i est est en difficulté.
'Peter_i knows very well that the jursit that Peter_1 is has a problem'

(i) has the same anaphoric behavior as (ii).

(ii) *Pierre_i sait très bien que Pierre_i est en difficulté.
'Peter knows very well that Peter has a problem'

Surprisingly, (iii) is not as deviant as (ii) or (i).

(iii) ?Le juriste que le Pierre_i est sait très bien que le Pierre_i est en difficulté.
'The jurist that the Peter_i is knows very well that the Peter_i has a problem'

(iii) may constitute an exception to this generalization. However, the coreferential interpretation of a pair of R-expressions may not be within the purview of Principle C under the appropriate formulation. See Lasnik (1991) and Freidin (1997b) for some discussion.

juriste, since *pro(NP)* and *juriste* share the same phi-features. The PPC gives rise to the contrast between (23g) and (23h). Note that (24) entails that, in a similar fashion, the notion of symmetry may be extended to the pair (*Mary*, *she*) in the structure in (3a), since the DP *Mary* and the pronoun *she* share the same phi-features. However, in that case, if an *NP* different from *Mary* is substituted for *Mary* within the pair (*Mary*, ~~*Mary*~~), the PF of the pair is altered (we assume that the substitution takes place across the board). No PF invariance obtains and the PPC is then not relevant to the pair (*Mary*, *she*).

To the extent that the kind of analysis proposed above is viable, it provides a modest illustration of what is being referred to as the 'perfection' of the grammatical system. The possibility then arises that the abstract analytical principles involved in the formal definition of a computation turn out to have exactly the right empirical consequences.[39] This is an exciting prospect, which, combined with that of potentially rich mathematical developments, is stirring imaginations. The authors of this note understand the excitement, and share in it. Uriagereka's *Rhyme and reason* is a particular expression of that entirely natural and legitimate reaction. In essential respects, linguistics is no different from other fields in natural sciences at comparable stages of development.[40]

3. Methodological issues

But apart from prospects for any line of research, there is the more concrete methodological question of how to proceed. LLJ propose as a model Arthur Holly Compton's Nobel Prize winning work on the quantum theory of the scattering of X-rays and γ-rays by light elements. Compton discovered that, when X-rays of a given frequency are scattered from (essentially) free electrons at rest, the frequency of the scattered X-rays is not unaltered, as the classical theory would predict, but decreases with increasing scattering angle. He described this effect by treating the rays as relativistic particles of energy $h\nu$ and momentum $h\nu/c$, and by applying the usual energy and momentum conservation laws to the collision. One could characterize Compton's contribution as an instance of Dirac's 'experimental procedure', working from the empirical data to arrive at the theoretical conclusions.[41] This is in contrast

[39] Some intriguing proposals in this area are put forth in Jenkins, 2000 (see pp. 151–170, for example). See also Fukui (1996) for another line of investigation into these topics.

[40] Perhaps linguistics is, in this regard, roughly comparable in character to structural chemistry in the years preceding its unification with physics (see Chomsky (1995a) for important remarks on structural chemistry in connection with the issue of 'reductionism'), or with the initial development of quantum physics prior to the Heisenberg/Schrödinger formulations.

[41] In that sense, both Heisenberg's research and Compton's Nobel prize winning work belong in the same 'Diracian category', i.e., the 'experimental procedure'. However, one cannot stress enough the difference in scope between the contributions of the two scientists. Compton was concerned with the interpretation of a particular experiment. Heisenberg was a theoretician trying to construct a general account of all the spectroscopic data available at the time in order to get a better understanding of the laws of nature. Heisenberg ended up constructing a new physics.

with the general style of the MP, which tends to operate in the reverse direction using the mathematical procedure.

The actual history of the quantum theory of radiation-matter interaction provides a more complicated picture, though. It is not just a story about the triumph of the experimental procedure, but also one that involves the triumph of the mathematical procedure as well. In fact, Compton (1923) is preceded by an article by Einstein (Einstein (1917), cited in Dirac (1968) as an example of the 'mathematical procedure'; see the quote in section 2 above). This article indeed "addresses questions of principle without offering any new experimental conclusion or prediction" (Pais, 1982: Ch. 21). By general admission, this article was a fundamental contribution, which has served as a basis for subsequent research on absorption, emission and dispersion of radiation (with the notable exception of Compton).[42] One of its central results concerned the exchange of momentum in radiation-matter interactions. This result can be stated as follows:

(25) In conditions of thermal equilibrium, an exchange of energy $h\nu$ between radiation and matter that occurs by a transition process (between two stationary states of an atomic system) is accompanied by an exchange of momentum of the amount $h\nu/c$, just as would be the case if the transition were accompanied

[42] It seems that Compton was unaware of Einstein's results at the time he was developing his theory. His only reference to Einstein in connection to the question of the interaction between radiation and matter is to Einstein (1905). Compton presented his theory at the 1 December 1922 meeting of the American Physical Society held at the University of Chicago. It is somewhat surprising that he would not have known of Einstein's (1917) article at the time, since there was really a free flow of information and ideas in physics between Europe and the US at the beginning of the 1920s. Thus, Debye could learn quickly about Compton's experimental results, as did Einstein. There were also common channels of publication (e.g., *The Philosophical Magazine*, in which A. Compton had previously published, as had many major contributors to quantum mechanics). More surprisingly, nowhere in his whole work does he refer to Einstein's 1917 paper.

One may surmise that Compton arrived at his result by a different route, namely, from classical electrodynamics (as it turns out, in conjunction with O.W. Richardson, Compton's first research advisor at Princeton University). Within classical electrodynamics, an electromagnetic wave carries momentum, giving rise to 'radiation pressure'. This is how it works. Suppose an electromagnetic wave is acting on an electric charge. The electric component in the wave makes the charge oscillate. This oscillation in turn interacts with the magnetic component of the wave, creating a force in the direction of the propagation of the wave. The value of the induced momentum is equal to the energy absorbed by the charge divided by the speed of light c. The division by c merely reflects the fact that the strength of the magnetic field associated with the wave is that of the electric field divided by c (see, for example, Feynman et al., 1963: 34–10&11). A particular instance of radiation pressure is that of an atom emitting an energy W in some direction. Then, according to classical electrodynamics, there is a recoil momentum $p = W/c$.

There is a big leap between the classical theory and the quantum hypothesis put forth in Einstein (1917), in Compton (1923), and in Debye (1923), though. The classical relation is a statistical relation, defined over averages of fluctuating quantities. More seriously, it only applies to directional waves. In case of a spherical wave, there should be no recoil. Einstein's fundamental insight was to consider that every exchange of energy was accompanied by an exchange of momentum. Correlatively, he was led to assume that the radiation was always directional, even in, e.g., the case of spontaneous emission, which was classically described as a spherical wave. Einstein assumed that that was a case where direction was only determined by 'chance'.

by the starting or stopping of a small entity moving with the velocity of light c and containing the energy $h\nu$. (adapted from Bohr et al., 1924)

This conclusion constituted a fundamental innovation, a conceptual 'mutation', since, by associating momentum quanta with energy quanta, it amounted to defining light-quanta as particles, on a par with, e.g., electrons: light-quanta, like electrons, were taken to be entities endowed both with energy and momentum. Previous discussions of the interaction between radiation and matter had solely been concerned with the exchange of energy (cf., for example, Einstein, 1905).[43]

This central result in Einstein (1917) anticipates Compton's account. Conversely, Compton's discovery helped clinch the argument in Einstein's article for ascribing a certain physical reality to the theory of light-quanta.[44] The mathematical and experimental procedures are, in the best circumstances, mutually reinforcing.

Thus there are several complementary facets to the story of the quantum theory of radiation-matter interaction. (To tell the story properly, one would need the combined talents of the author of *The Alexandria Quartet* and of the author of *Blackbody Theory and the Quantum Discontinuity, 1894–1912*.) Theoretical strands and contributions are intertwined in a complex pattern of interactions that are hard to disentangle. What is sure is that a conceptual 'mutation' happened, by which the light quantum postulated by Planck in 1900 and in Einstein (1905) was granted the fundamental attributes of particles, namely energy and momentum. Several strands of work contributed to that mutation. It would be nonsensical to single out any of them as the most representative. What we really have here is a kind of 'ecological system'.

The development of scientific hypotheses can actually be advantageously compared to the evolution of species. New hypotheses are put forward and concepts created, which are then subjected to selection. Normally only the fittest survives. Selectional pressure, e.g., in the form of such an experiment as Compton's, is then crucial to the development of successful hypotheses and theories. At the same time, selection is not the 'exclusive means of modification'.[45] Quite obviously, there must exist a source of new hypotheses (i.e. scientific 'mutations'). So we are led to distinguish two types of scientific events: 'selectional events' (often, but not exclusively, experiments) and 'mutational events'.[46] Analogously, we can distinguish

[43] Einstein's article included several other far-reaching assumptions, from which Einstein was able to derive both Planck's radiation law and Bohr's frequency condition, among other results. Indeed, one of the article's central contributions was to establish a bridge between blackbody radiation and Bohr's theory of spectra.

[44] Debye, who knew about Compton's evidence against the classical theory, independently derived the equations describing the scattering of a photon off an electron at rest (Debye, 1923). We note that, in his article, Debye expressed his indebtedness to Einstein's 1917 theory.

[45] Cf. the introduction to Darwin (1859), where the last sentence reads as follows:

i) "Furthermore, I am convinced that Natural Selection has been the main but not exclusive means of modification".

[46] One issue of importance is how one recognizes a 'mutational event'. For example, was the atomism of Leucippus and Democritus a 'mutational' scientific hypothesis? Of course, a preliminary question is whether it was a scientific hypothesis at all. Given that one could not really conceive of any serious

between 'selectional contributions' and 'mutational contributions'. If we follow Dirac (1968), 'mutational contributions' in turn would be of two types, 'experimental' (i.e. like Heisenberg's, which was based on the experimental data) or 'mathematical'. Note that the distinction between 'selectional contributions' and 'mutational contributions' is one of roles, not essences, relative to scientific situations. The same contribution might count as selectional or mutational depending on the set of hypotheses considered.[47]

Compton's Nobel Prize winning contribution was both a selectional and a mutational one.[48] However, it was obviously quite different in character and in scope (see note 41) from that of the fundamental 'mutational contributors' to quantum mechanics, including Bohr, Born, Dirac, Einstein, Heisenberg, Jordan, Planck, Pauli, and Schrödinger. While Compton was initially busy defending classical physics, they were building an alternative framework, often with few certainties at hand. From the evolutionary point of view above, both approaches were necessary to ensure the required level of 'ecological diversity'. One cannot emphasize this point too much. Science needs a diversity of styles for its continued progress. This applies as much to linguistics as to physics, perhaps more so given the relative immaturity of the field.

Chomsky's work on the MP has been from the outset grounded in the mathematical procedure he employed so successfully to launch modern generative grammar in the early 1950s. In essence it constitutes a distillation of the mathematical procedure applied to linguistic theory that Chomsky envisioned in the opening paragraph of *Syntactic structures*: "The ultimate outcome of these investigations should be a theory of linguistic structure in which the descriptive devices utilized in particular grammars are presented and studied abstractly, with no specific reference to particular languages". From one perspective, the MP offers the most promising approach to this goal, far off though it still remains. Nonetheless, based on the considerable empirical successes of its predecessor, which to a large extent are still

---

experimental testing of the (rather vague) philosophy of the Greek atomistic school, it would not qualify as a scientific hypothesis. In general, there are various criteria by which one may evaluate and compare mutational hypotheses. One factor is the 'likelihood' of an hypothesis, i.e., the degree of expectation of the hypothesis given the state of the knowledge at the time. We could call this the 'originality' of the hypothesis. Thus, when the Danish astronomer Roemer proposed that the apparent discrepancies between the measured movement of the moons of Jupiter and Newton's Law were merely an observational illusion due to the noninstantaneous character of light propagation, this was an original hypothesis at the time. Another criterion in evaluating an hypothesis is its complexity, i.e., how much mathematical or conceptual elaboration it involves, and how much revision of the existing system of knowledge it requires.

[47] It should be clear that such notions as 'scientific mutation' or 'scientific selection' are intended to apply to all sciences, not only to Galilean sciences. Thus, Darwin's hypothesis concerning natural selection and evolution qualifies as a 'mutational' hypothesis.

[48] It was a selectional contribution in two distinct ways. It established the reality of the light quantum as a true particle. But also, because it could not account for the angular dependence of the scattered X-ray intensities, the total scattering cross-section as a function of energy, or the exact state of polarization of the Compton scattered X-rays, it emphasized the need for a more basic quantum mechanics, to be soon developed by Born, de Broglie, Dirac, Heisenberg and Schrödinger.

incorporated under current proposals, there is good reason to continue exploring the MP to discover whatever insights it can provide as well as whatever its actual limitations may be.

References

Boehner, P., 1958. Collected articles on Ockham. E.M. Buytaert (ed.). St. Bonaventure, NY: Franciscan Institute.
Bohr, N., 1913. On the constitution of atoms and molecules. Philosophical Magazine 26, 1–25, 476–502, 857–875.
Bohr, N., H.A. Kramers and J.C. Slater, 1924. The quantum theory of radiation. Philosophical Magazine 47, 785–802. [A German version appeared in Zeitschrift für Physik 24, 69–87 (1924).]
Brody, M., 1982. On circular readings. In: N.V. Smith (ed.) Mutual knowledge, 133–146, New York: Academic Press.
Chomsky, N., 1951. Morphophonemics of modern Hebrew. [Published New York: Garland, 1979.]
Chomsky, N., 1955–1956. The logical structure of linguistic theory. New York: Plenum.
Chomsky, N., 1956. Three models for the description of languages. IRE Transactions on Information Theory 2, 113–124.
Chomsky, N., 1957. Syntactic structures. The Hague: Mouton.
Chomsky, N., 1959. On certain formal properties of grammars. Information and Control 2, 137–167.
Chomsky, N., 1962. Context-free grammar and pushdown storage. Quarterly Progress Report 65, 187–194. Research Laboratory in Electronics, Cambridge, MA.
Chomsky, N., 1965. Aspects of the theory of syntax. Cambridge, MA: MIT Press.
Chomsky, N., 1976. Conditions on rules of grammar. Linguistic Analysis 2, 303–351.
Chomsky, N., 1980. On binding. Linguistic Inquiry, 11, 1–46.
Chomsky, N., 1986. Knowledge of language. New York: Praeger.
Chomsky, N., 1991. Some notes on the economy of derivation and representation. In: R. Freidin (ed.), Principles and parameters in comparative grammar, 417–454. Cambridge, MA: MIT Press. [an early version appeared in MIT Working Papers in Linguistics 10, 1989; reprinted in Chomsky, 1995b.]
Chomsky, N., 1993. A minimalist program for linguistic theory. In: K. Hale and S.J. Keyser (eds.), The view from Building 20, 1–52, Cambridge, MA: MIT Press. [Reprinted in Chomsky, 1995b.]
Chomsky, N., 1995a. Language and nature. Mind 104, 1–61.
Chomsky, N., 1995b. The minimalist program. Cambridge, MA: MIT Press.
Chomsky, N., 1998. Minimalist inquiries: The framework. In: R. Martin, D. Michaels and J. Uriagereka (eds.), Step by step: Essays on minimalist syntax in honor of Howard Lasnik, 89–155. Cambridge, MA: MIT Press, 2000. [An early version appeared as MIT Occasional Papers in Linguistics 15, 1998.]
Chomsky, N., 1999. Derivation by phase. MIT Occasional Papers in Linguistics 18.
Chomsky, N., 2000 New horizons in the study of language and mind. New York: Cambridge University Press.
Chomsky, N. and G.A. Miller, 1958. Finite-state languages. Information and Control 1, 91–112.
Chomsky, N. and H. Lasnik, 1977. Filters and control. Linguistic Inquiry 8, 425–504.
Chomsky, N. and H. Lasnik, 1993. The theory of principles and parameters. In: J. Jacobs, A. von Stechow, W. Sternefeld and T. Vennemann (eds.) Syntex: An international handbook of contemporary research, 506–569, Berlin: Walter de Gruyter. [reprinted in Chomsky 1995b.]
Cohen-Tannoudji, C., B. Diu and F. Laloë, 1996. Mécanique Quantique, Paris: Hermann.
Compton, A.H., 1923. A quantum theory of the scattering of x-rays by light elements. The Physical Review 21, 483–502.
Darwin, C., 1859. On the origin of species by means of natural selection. London: John Murray.
Debye P., 1923. Zerstreuung von Röntgenstrahlen und Quantentheorie. Physikalische Zeitschrift 24, 161–166.

Dirac, P.A.M., 1968. Methods in theoretical physics. In: From a life of physics; Evening lectures at the International Center for Theoretical Physics, Trieste, Italy. A special supplement of the International Atomic Energy Agency Bulletin, Austria.

Einstein, A., 1905. Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt. Annalen der Physik 17, 132–148.

Einstein A., 1917. Zur Quantentheorie der Strahlung. Physikalische Zeitschrift 18, 121–128. [First printed in 1916 in Mitteilungen der Physikalische Gesellschaft Zürich 16, 47–62.]

Feynman, R., 1971. The character of physical law. Cambridge, MA: MIT Press.

Feynman R.P., R.B. Leighton and M. Sands, 1963. The Feynman lectures on physics, Vol. 1. Reading, MA: Addison-Wesley.

Freidin, R., 1994. Conceptual shifts in the science of grammar: 1951–1992. In: C.P. Otero (ed.), Noam Chomsky: Critical assessments, vol. 1, 653–690. London: Routledge.

Freidin, R., 1997a. Review article: N. Chomsky. The minimalist program. Cambridge, MA: MIT Press. 1995. Language 73(3), 571–582.

Freidin, R., 1997b. Binding theory on minimalist assumptions. In: H. Bennis, P. Pica and J. Rooryck (eds.), Atomism and binding, 141–153. Dordrecht: Foris.

Freidin, R. and H. Lasnik, 1981. Disjoint reference and wh-trace. Linguistic Inquiry 12, 39–53.

Fukui, N., 1996. On the nature of economy in language. Cognitive Studies 3, 51–71.

Goodman, N., 1943. On the simplicity of ideas. Journal of Symbolic Logic 8(4): 107–121.

Greene, B., 1999. The elegant universe: Superstrings, hidden dimensions, and the quest for the ultimate theory. New York: Norton.

Hockett, C., 1954. Two models of grammatical description. Word 10, 210–234.

Huybregts, R. and H. van Riemsdijk, 1982. Noam Chomsky on the generative enterprise. Dordrecht: Foris.

Hymes, D. and J. Fought, 1981. American structuralism. The Hague: Mouton.

Jacobson, P., 1977. The syntax of crossing coreference sentences Doctoral dissertation, University of California, Berkeley.

Jenkins, L., 2000. Biolinguistics. Cambridge, England: Cambridge University Press.

Jourdain, P.E.B., 1913. The principle of least action. Open Court Publishing Company.

Kneale, W. and M. Kneale, 1962. The development of logic. Oxford: Oxford University Press.

Lanczos, C., 1970. The variational principles of mechanics, 4th edition. Toronto: University of Toronto Press.

Lappin, S., 1991. Concepts of logical form in linguistics and philosophy. In: A. Kasher (ed.) The Chomskyan turn, 300–333, Oxford: Blackwell.

Lappin, S. and M. McCord, 1990. Anaphora resolution in slot grammar. Computational Linguistics 16: 197–212.

Lasnik, H. 1976. Remarks on coreference. Linguistic Analysis 2, 1–22.

Lasnik, H., 1991. On the necessity of binding conditions. In: R. Freidin (ed.), Principles and parameters in comparative grammmar. 7–28, Cambridge, MA: MIT Press.

Lasnik, H., 1999. Minimalist analysis. Malden, MA: Blackwell.

Lappin, S., R. Levine and D. Johnson, 2000. The structure of unscientific revolutions. Natural Language and Linguistic Theory 18, 665–671.

Lavine, J. and R. Freidin, 2001. The subject of defective T(ense) in Slavic. To appear in Journal of Slavic Linguistics.

Otero, C., 1994. Chomsky and the cognitive revolution of the 1950s: the emergence of transformational generative grammar. In: C.P. Otero (ed.), Noam Chomsky: Critical assessments, vol. 1, 1–36, London: Routledge.

Pais, A., 1982. 'Subtle is the Lord ...': The science and life of Albert Einstein. Oxford: Oxford University Press.

Postal, P., 1966. On so-called 'pronouns' in English. In: F. Dinneen (ed.) Report of the seventeenth annual round table meeting on linguistics and language studies, 177–206, Washington, D.C.: Georgetown University Press.

Putnam, H., 1962. The analytic and the synthetic. In: H. Feigl and G. Maxwell (eds.), Minnesota Studies in the Philosophy of Science, III, 358–397. Minneapolis, MN: University of Minnesota Press.
Smith, N., 1999. Chomsky: ideas and ideals. Cambridge, England: Cambridge University Press.
Stewart, I., 1998. Life's other secret. New York: Wiley.
Stewart, I., 1999. Designer differential equations for animal locomotion. Complexity 5, 12–22.
Thompson, D.W., 1942. On growth and form: A new edition. Cambridge, England: Cambridge University Press.
Toffoli, T., 1999. Action, or the fungibility of computation. In: A.J.G. Hey (ed.), Feynman and computation, 349–392. Reading, MA: Perseus Books.
Vergnaud, J.-R., 1974. French relative clauses. Doctoral dissertation, MIT, Cambridge, MA.
Vergnaud, J.-R., 1998. On two notions of representations in syntax. Manuscript.
Vergnaud, J.-R. and M.L. Zubizarreta, 2000. Intervention effects in French wh-in-situ construction: Syntax or interpretation? Manuscript.
Weinberg S., 1976. The forces of nature. Bulletin of the American Academy of Arts and Sciences, 29(4), 13–29.