Universals, Occam’s Razor, and the Strong Minimalist Thesis

Norio Suzuki

Abstract

The main purpose of this article is to explore consequences of Chomsky’s (2000, 2001a,b) strong minimalist thesis (SMT) as it pertains to evolution of language, the form and functioning of FL (the faculty of language), the dichotomy between methodological and ontological minimalism in the sense of Martin and Uriagereka (2000) (i.e., Occam’s razor vs. the SMT), substantive and formal universals in the sense of Chomsky (1965), and the choice between Chomsky’s (2000, 2001a,b) ‘phase-wise’ implementation of multiple spell-out and Epstein and Seely’s (1999, 2002) ‘application-wise’ implementation of it. As for narrow syntax, all its descriptive vocabulary is shown to derive from interface conditions (IC) whereas its functioning largely derives from mathematical properties, the mathematical being assumed to be ubiquitous in the natural world (hence, in language as well). This scenario will be seen to lead to the situation required by Chomsky’s (2001b: 3) extremely strong minimalist thesis SMT where the category of “unexplained elements of So” is nonexistent. It will also be argued on the basis of conceptual considerations originating with some bioscience and sociology literature that the concept of ‘thickness’ required by human recognition may play a role in distinguishing between
otherwise almost indistinguishable two ways of implementing multiple
spell-out.

**Keywords:** Strong minimalist thesis; Methodological vs. ontological
minimalism; Form and functioning of narrow syntax; Substantive and
formal universals; Multiple spell-out

1. Introduction

One of the most fascinating foundational concepts of the Minimalist
Program (MP) can be said to be the strong minimalist thesis (SMT) of
Chomsky (2000, 2001a,b).¹ Martin and Uriagereka (2000) make a dis-
tinction between a weak and strong minimalist thesis, calling the for-
mer *methodological minimalism* (Occam's razor) and the latter
*ontological minimalism*. It is quite natural and undeniable that the
study of language has gained much from the former methodology by
seeking simple and nonredundant theories, as in the case of the core
sciences. But Martin and Uriagereka (2000) go on to take into consid-
eration a more subtle take on the MP, what they call ontological mini-
malism, that is, the strong minimalist thesis, stressing the importance
of the fact that Chomsky is not asking, “How good is our theory of FL
(the faculty of language)?” but “How good is FL itself?” when he (2000)
poses the question, “How well is FL designed?” It is clear that this may
immediately lead to the question why language/FL is that way
(Chomsky, 2001b: 2, etc.) We will see in this paper how the adoption of
the SMT may bring about a new view on the concepts of substantive
and formal universals in the sense of Chomsky (1965) and how the part
of FL/language that does not derive from interface conditions must be
accounted for in terms of mathematical properties which underlie the whole natural world, presumably physical and chemical properties also coming into play (Chomsky, 2001b: 2).

2. Evolution and the form and functioning of FL

We can speculate why such a proposition as the SMT has come to be put on the agenda by looking at some important evolutionary considerations in the literature. Discussing the interaction of the three factors involved in the attainment of the properties of a language, Chomsky (2001b: 1) speaks of a genetically-determined initial state: "S₀ itself, a product of evolution." You may need caution in interpreting this phrase when you confirm that S₀ corresponds to FL at the initial state of development, that is, Universal Grammar (UG) in linguistic terms and the language acquisition device (LAD) in acquisitional terms. It may be difficult to imagine that such a thing as FL could evolve with all its abstract mathematical properties, which are rather different from other things amenable to evolution in the biological world. Chomsky notes (2002: 108): "... the system is essentially inserted into already existing external systems: external to the language faculty, internal to the mind". Presumably, it was these already existing external systems (i.e., a sensori-motor system and some kind of system of thought which were there, independently of language) that had been subject to evolution. The two external systems must have reached the level at some point of history where FL could interact with them in such a way that it was usable to them. I therefore interpret Chomsky’s (2001b: 1) phrase “S₀ itself, a product of evolution” as evolution of the whole complex involving at least FL and the two
interfaces when a meaningful relation (i.e., usability of the former by the latter) was established between them.

It would be nice if the specific form of narrow syntactic operations could be predicted or determined by relevant interface conditions alone through the SMT, or it might be the case that the only requirement on the part of the SMT is for narrow syntax to be usable in terms of interface conditions. But these two questions may actually point to the same direction, as can be seen by the following observation of Chomsky (2001b: 2): “If language is to be usable at all, its design must satisfy an ‘interface condition’ IC ...” This observation encompasses almost all the empirical content of the SMT (Chomsky 2000: 96, (2)), bringing about a possibility of determining the form of narrow syntax/FL. Note that the functioning of narrow syntax may be determined by “relations that ... fall out in some natural way from the computational process” (Chomsky 2000: 113)) or/and what Chomsky (2001b: 3) calls “general properties,” Epstein’s (1999) derivational analysis of c-command being an example of relevant ‘relation’ in the sense of Chomsky (2000) and the kinds of computational efficiency that FL selects, which are external to So, falling under Chomsky’s (2001b) “general properties.”

3. Methodological and ontological minimalism

Freidin and Vergnaud (2001: 645) seem to stress the significance of general conditions of conceptual naturalness such as simplicity, economy, or non-redundancy in their discussion of the evolution of linguistic theory, citing in their note Chomsky (1995) as mentioning such conceptual naturalness in motivating the work in generative grammar.
But it does not seem that Chomsky (1995: 1) is as clear as Freidin and Vergnaud (2001) claim about such general considerations of conceptual naturalness in discussing what conditions are imposed on FL, as can be seen by his subtle vagueness when he observes that “Question (B) [general considerations of conceptual naturalness that have some independent plausibility, namely, simplicity, economy, symmetry, nonredundancy, and the like. NS] is not precise, but not without content; attention to these matters can provide guidelines here…” (1995: 1). Note that general considerations of conceptual naturalness in this case may correspond to methodological minimalism. Note that Martin and Uriagereka (2000: 1) admit that a “weak minimalist thesis” (Occam’s razor, ‘methodological minimalism’) is surely a necessary aspect of scientific inquiry, observing at the same time that there is a more subtle take on the Minimalist Program, that is, a “strong minimalist thesis”/‘ontological minimalism.’ And moreover Chomsky (2001b: 2) tries to go beyond explanatory adequacy when he observes that “... asking not only what the properties of language are, but why they are that way”. What Chomsky seeks here is ontological minimalism (i.e., the strong minimalist thesis, which defines language (FL) from the viewpoint of its usability on the part of the interfaces). Is ontological minimalism always compatible and consistent with general considerations of conceptual naturalness, which have generally been fruitfully appealed to and employed in the more developed natural sciences (e.g., theoretical physics)?

4. Substantive and formal universals

Recall that Chomsky (1965: 27-30) classifies linguistic universals as
substantive or formal universals when discussing a theory of linguistic structure that aims for explanatory adequacy. I assume that elements and factors pertaining to ontological minimalism correspond to Chomsky's (1965) substantive universals, which is described as concerning the vocabulary for the description of language. The major difference between the two formulations is that while the notion of substantive universal may only have been instrumental in enumerating (hence, only descriptive) such candidate theoretical constructs as Jakobson's theory of distinctive features, certain fixed syntactic categories (e.g., N, V, etc.), and so forth, ontological minimalism can be said to be explanatory as long as its empirical content can be predicted in terms of the usability of language on the part of the sensorimotor (SM) and conceptual-intentional (C-I) interfaces. That is, the reason for adopting ontological minimalism should be the presence of "interface condition" (IC) that the design of FL must satisfy. Do Chomsky's (1965) formal universals correspond to constructs pertaining to methodological minimalism then? They may not. Chomsky's formal universals may contain such domain-specific constructs as transformation, cyclicity, etc., whereas methodological minimalism consists of such generally conceptual ideas as naturalness, simplicity, economy, non-redundancy, etc. (Occam's razor). What position may the notion of formal universal occupy in the most recent minimalist development then?

Notice that Chomsky (2001b: 3) states that "the goal is to determine just what aspects of the structure and use of language are specific to the language faculty, hence lacking principled explanation at this level," restricting the initial conditions on language acquisition to the following two categories: IC (interface condition; the principled part of S0) and "general properties," and hopefully dispensing with the cate-
category that consists of "unexplained elements of \( S_0 \)," which may lead you to "an extremely strong minimalist thesis SMT." Could it be said that Chomsky's (2001b) category of "general properties" corresponds to Chomsky's (1965) notion of formal universal? According to Chomsky (2001b: 3), "general properties" are external to \( S_0 \) and the kinds of general conditions of computational efficiency that FL selects may belong under the "general properties." But we have seen that formal universals in the sense of Chomsky (1965) are assumed to be domain-specific and include such notions as transformation, cyclicity, etc., which means that they belong under FL/UG/S0. You would be forced to admit that Chomsky's (1965) formal universals belong under the category of "unexplained elements of \( S_0 \)" unless they can be proved to be part of IC. It will, however, be most desirable to dispense with the category of "unexplained elements of \( S_0 \)" if you hope to keep to the SMT. But it would be prima facie difficult to convince yourself that such constructs as transformation and cyclicity derive from SM or C-I interfaces. Recall, however, that "its [the minimalist program NS] task is to examine every device (principle, idea, ...) that is employed in characterizing languages to determine to what extent it can be eliminated in favor of a principled account in terms of general conditions of computational efficiency and the interface condition that the organ [language/FL NS] must satisfy for it to function at all" (Chomsky, 2001b: 3). Notice that 'transformation' may be able to be reformulated in terms of notions such as Merge, Move, etc. according to recent developments in syntactic theory. Merge and Move may be instrumental in providing sufficient diversity of "legible" expressions at the SEM [LF NS] interface (Chomsky, 2001b: 3), it thus being possible to claim that Merge and Move (i.e., transformation) exist due to IC. What about
cyclicity? You would be hard-pressed to derive it from IC. I assume that cyclicity finds its root in mathematical properties and that the latter virtually constitute the empirical content of Chomsky's (2001b: 3) notion of general properties, and thus determine that cyclicity falls under "general properties," hence, outside of S0. The notion of recursiveness may be treated on a par with that of cyclicity. On the other hand, it has been suggested in the literature that the notion of c-command may not function in narrow syntax but in PHON and SEM interfaces (Chomsky, 2000: 116), which may make it possible to derive c-command from IC. The concepts of locality and phase may also be due to IC.

5. Multiple spell-out

You hear in the bioscience literature something like human beings living in the past 0.5 seconds ago (Pollack, 2000/1999). People say in some sociology literature (based on the German sociologist N. Luhmann) something like "the experience of catching a glimpse into a 'crack' being only possible within that 'thickness' of the event" (Yoshizawa, 1995: 167). According to Yoshizawa (1995: 155-156, 164), an event is a radical situation where time and meaning come into existence for the first time, and the event is something that vanishes immediately after it appears. But in order for us humans to understand and live the time and meaning, the present/the radical situation cannot be a point in time but must have a kind of 'thickness.' It may be tempting to interpret the notion of 'thickness' in the latter literature as the time span expressed by '0.5 seconds' in the former. And it may be plausible to take the notion of locality to express something like 'thickness' in
the sense of Yoshizawa (1995) either in terms of time or place, it thus
being possible to derive the notion from IC. The concept of phase has
much to do with that of locality, the former being a concrete
instantiation of the latter in actual terms.

It may, however, be instructive to compare two versions of the notion
of multiple spell-out ('phase-wise' (Chomsky, 2000, 2001a,b) and 'appli-
cation-wise' (Epstein and Seely, 1999, 2002)) in terms of the concept of
'thickness' in the sense of Yoshizawa (1995) because the latter concept
must be deeply involved in the situation concerning the human recogni-
tion of meaning. Roughly (according to Chomsky, 2001b), narrow syn-
tax, the phonological component, and the semantic component
optimally operate cyclically in a single cycle, and the three components
proceed cyclically in parallel with respect to the derivation of the pair
<PHON, SEM>, which are accessed and interpreted by SM and C-I in-
terfaces, respectively. There is no single level of LF in this conception.
The operation of TRANSFER hands the narrow syntactic derivation
over to the phonological component and to the semantic component,
the former operation of which Chomsky (2001b: 4) specifically calls
spell-out, suggesting their integration at the same time. Epstein and
Seely (1999: 59) describe the situation as follows: "One way to think of
this is as an iterated Y-model, i.e. after each transformation applies (be
it Merge or Remerge), the derived representation is fed into both PF
and LF." For the purposes of terminology, I interpret the concept of
spell-out as referring to the transference of the narrow syntactic deri-
vation over either to the phonological component or to the semantic
component, using the term in an extended sense. Now it seems to me
that a transformational application is too small/short to accommodate
any meaningful unit, i.e., not 'thick' enough for any meaning to come
into existence in the terminology of Yoshizawa (1995). On the reasonable assumption that the composition of a meaningful unit must be based at least on some elements that are themselves meaningful, Epstein and Seely’s (1999, 2002) ‘application-wise’ multiple spell-out would be hard-pressed, specifically due to the smallness/shortness of spelled-out elements.

We cannot adopt Epstein and Seely’s multiple spell-out for empirical reasons as well, as can be seen by the following observation of Chomsky (2001b: 20): “We know that S-O [spell-out NS] cannot apply at each stage of cyclic Merge. Relevant information may not yet be available. Suppose, for example, that Merge has constructed (see, OB), where the object OB = that or what. At this stage we do not know whether OB or see is spelled out in situ, or whether they move on overtly to be spelled out in a higher position. If they move on (either sometimes, or always), then S-O plainly cannot apply at this stage.” Since Epstein and Seely’s (1999, 2002) application-wise multiple spell-out would force the implementation of an operation of spell-out in this case as well (with an instance of Merge being a transformational application), it is plainly clear that Epstein and Seely are forced to run into trouble when faced even with such elementary situations as this.

On the other hand, there arise no similar problems with Chomsky’s (2000, 2001a,b) ‘phase-wise’ multiple spell-out, a phase being vP or CP. Notice that vP is where argument/thematic structure of the predicate is expressed and that the force/mood expressed by C is added to the relevant proposition at the level of CP, which leads us to take both phases to be meaningful units which may be ‘thick’ enough in the sense of Yoshizawa (1995) to be understood and lived by us humans. Consider also in this connection Chomsky’s (2001b: 14) observation that
“TRANSFER has a ‘memory’ of phase-length, meaning again that operations at the phase level are in effect simultaneous. It follows that phases should be as small as possible, to minimize memory for S-O ...” It is interesting to note that Yoshizawa (1995) somehow stresses the need for thickness/bigness from the viewpoint of the event which is basically devoid of length by definition, while Chomsky (2001b) calls for the minimal smallness/shortness of phases in terms of memory requirements. It may be the case that we humans can memorize only meaningful things. Then, as long as the issue of meaning is highlighted in such grammatical manipulation as this, we can find another argument in favor of the SMT in that IC requirements are reflected there. And the tension observed between Yoshizawa’s (1995) and Chomsky’s (2001b) requirements may be taken to point to the minimally small/short, if not null or overly instantaneous, time span required for language/human recognition to work, given memory limitations and, presumably, a number of other requirements from which I abstract away due to lack of sufficient understanding.

6. Universals and minimalism

Returning to the substantive vs. formal universals issue (Chomsky, 1965: 27-30), i.e., that of what concept(s) they may correspond to in view of recent developments of minimalist theorizing, it safely can at least be concluded that substantive universals may correspond to Chomsky’s (2001b: 3) principled part of S0 (IC), but you have to be cautious with the treatment of formal universals, which may be prima facie ambivalent between “general properties” and IC, at least in Chomsky’s (1965) version. (I try to eliminate the putatively claimed “unexplained ele-
ments of $S_0$" part of formal universals in conformity with Chomsky's (2001b: 3) "extremely strong minimalist thesis.") I take transformation to put mainly phonological material or UG-provided empty categories into a number of functional positions (both $X^0$ and $XP$; in addition to base-Merge in thematic positions) basically in the sense of Cinque (1999) through Merge or Move. Since the various functional positions are associated with a specific positional meaning, the merged or moved element will gain the specific positional meaning by virtue of being in that position (see Rosengren, 2002; Suzuki, 2002: 22-24). Transformation, therefore, can be said to be a necessary device for individual phonological materials or UG-provided empty categories to obtain the specific meaning associated with a particular functional position.

This supports the view that transformation derives from IC as we have seen above, but there may be a further important aspect to it, i.e., its mode of functioning as is known by the concept of ‘transformational cycle.’ Cyclicity itself has been seen above to fall under “general properties” due to its mathematical properties. Then, in this particular case of transformation, which Chomsky (1965) classifies as a formal universal, we see properties relevant both to IC and to “general properties.” Suffice it to say here, however, that it may not be particularly difficult to tease apart these properties pertaining to two different categories. Much the same may be true of the notion of locality, whose raison d'être has been seen to be due to IC, but whose mode of operation would obey some mathematical principles, which may classify it under “general properties” on the assumption that mathematical principles/properties practically constitute the whole empirical content of “general properties.” Note that Reuland (2001: 441) observes that while
narrow syntax espouses the concept of locality, logical syntax does not. I take this observation to support the SMT on the assumption that narrow syntax is part of natural language, whereas logical syntax has standardly been formed in the tradition of logic (e.g., predicate logic, etc.), i.e., in the tradition of a species of formal language. In spite of some indeterminacy and obscurity arising specifically from historical factors, Chomsky's (1965) case for the dichotomy between substantive and formal universals is sufficiently clear, leading us to safely conclude that substantive universals derive from IC (mainly due to recent developments of minimalist theorizing) and that formal universals must be based on Chomsky's (2001b) general properties, which I specifically interpret as deriving from mathematical principles/properties.

7. Mathematical properties in and out of FL

The major part of Chomsky's (1995: 1) vagueness about general considerations of conceptual naturalness (which we have assumed above to correspond to methodological minimalism) in dealing with FL may presumably derive from the reasonable conjecture that IC factors (the SMT/ontological minimalism/substantive universals) carry much weight with his theorizing of language. But, at the same time, Chomsky is well aware of the presence of something that intervenes primarily in the functioning of the elements that constitute the vocabulary pertaining to IC. It is by now clear that those interveners correspond to Chomsky's (2001b) general properties, which have been assumed to derive from mathematical principles/properties. Then we face two situations in regard to them: one situation where IC factors obey these mathematical principles in their functioning (internal to FL,
which is part of nature), and the other where they have been historically instrumental in developing and refining scientific theories (external to FL; actually, as a way of being of the world). For an illustration of the latter situation, Freidin and Vergnaud (2001: 647) cite P.A.M. Dirac as identifying two main methods within the mathematical procedure: one is to remove inconsistencies, the other, to unite theories that were previously disjoint. We see here the same mathematical principles/properties ruling both over abstract and automatic phenomena such as the mode of operation within FL and scientific developments involving human efforts. When you define Galilean science as the search for mathematical patterns in nature, a significant feature of the generative revolution in linguistics has been the development of a Galilean style in the field (Freidin and Vergnaud, 2001: 647). It looks as if the mathematical, which is ubiquitous in nature (hence, in language as well), somehow guides us through scientific theorizing. Although it has been seen to be instrumental, hence methodological, in this kind of specifically scientific inquiry, the term ‘methodological minimalism’ (for the mathematical) would be misleading since it misses the very important role that it plays in defining the true nature of the natural world.

Footnote

1 Chomsky (2000: 96) presents his “strongest minimalist thesis” as follows:
   (i) Language is an optimal solution to legibility conditions.

References


Yoshizawa, N., 1995. Tasha-no Keiken: Dekigoto-to-shiteno Ai (Experience of others: Love as an event). In: Inoue, S., Ueno, C., Ohsawa, M., Kenda, S.,
Yoshimi, T. (Eds.), Tasha, Kankei, Communication (Others, Relations, and Communication). Iwanami Shoten, Tokyo, pp. 155-174.