How Do We Construct Convergent Numerations?

Joseph Emonds

Abstract

Current versions of Chomskyan syntax take for granted that syntactic derivations depend on prior specification of complete "numerations" of lexical items, which then combine according to the principles of syntax. However, competence models have provided no ways to choose such numerations; they are either chosen randomly or based on the intuition of (ultimately native) speakers. In neither case is there any scientific characterization of these objects, and so syntactic derivations lose their status as a scientific model. This essay claims that numerations in a plausible formal model of language can be conceived as random (or, formally equivalently, pragmatically determined) only if syntactic derivations can supplement them in a highly constrained way: by adding to them items from a special lexical subcomponent of purely grammatical or "closed class" items. Items from this subcomponent, dubbed the Syntacticon in some recent work, seem to have precisely the grammatical properties (insertion into already processed structures, possibly null phonology) needed to make the otherwise randomly selected numerations "converge" to well formed Logical Forms.

In this essay, I will be relating the idea of a **numeration** of lexical items to a corresponding syntactic derivation. An initial definition of a **convergent numeration** (throughout, a "**c-numeration**") is a set of all and only the lexical items that appear in a well-formed syntactic structure. This definition will be modified in the course of discussion.

From the beginning, the fundamental and on-going problem of generative grammar has been to characterise what constitutes a well-formed syntactic structure. As this is the unachieved goal of the enterprise, we cannot take the *results* of generative grammar as a basis for choosing c-numerations; this is circular.[1] Nor are well formed syntactic structures in any mapping relation with some other set of already specified well formed objects, such as "linguistically expressible conceptual structures," since claims that such objects can be formalized are notoriously promissory. So in order to find c-numerations, we need a procedure based on some other formal object, one that is or at least plausibly can be defined independently of generative syntax and independently of logical or semantic representations of syntactic sentences.

That other formal object that provides c-numerations is some formalized model of the lexicon. This essay is devoted to sketching such a model, one that can provide c-numerations, and to outlining the basic relations between this lexical model and c-numerations.

[1] We want to develop models of well-formedness that contain numerations to increase our understanding of syntax, so we cannot simply assume syntax is understood and so can be used to provide numerations.

Theoretical and Applied Linguistics at Kobe Shoin 7, 1–20, 2004.

1. The Concept of Numeration and a Comprehension Model

In terms of a competence model of syntax, there are in fact both trivially adequate and farcically inadequate conceptualizations of c-numerations. Let us consider these two extreme versions in order.

If we consider a c-numeration as a finite set of lexical items from which a candidate for Merge is chosen, at any given stage of a derivation we can always say a numeration has only one member. That is, it consists of a single item with some condition for being lexically inserted that is satisfied by the partially derived syntactic structure. In this sense, successive numerations with just one member can be constructed at each step as a derivation proceeds. But this sense of "numeration" is just a new term for **lexical insertion**, and "selection from a numeration" means simply "successful selection from the lexicon."

Under this conception, for a sentence such as *The guest will sit on my left,* there is no c-numeration of 7 lexical items prior to a derivation, but only 7 instances of lexical selection interspersed in a classical, bottom-up syntactic derivation. If this conclusion is satisfactory, then numeration is simply a more recent and trivially different terminology for the model of lexical insertion introduced in Chomsky (1965).

Most usage of the term c-numeration, however, implies that there exists some way to choose the entire set for a sentence or at least a clause { *guest, left, my, on, sit, the, will* } prior to syntactic computation. There is of course one such way, namely by random selection from the lexicon. Now much informal research (cf. Aitchison (2003)) suggests an individual lexicon contains 20,000 entries; let's be conservative and say 10^4entries. There are then 10^{28} ways to randomly choose 7 word numerations. Perhaps we could use a few common sense computational tricks to cut out 99.99% of the choices, leaving us with only 10^{24} ways. A moment's thought has convinced generations of generative linguists (and Jonathan Swift before them) that random combination of words leads to endless collections of nonsense strings. To avoid such silliness, if we want to retain the term c-numeration for a competence model, we would be led back to some version of a pre-minimalist model of bottom-up, step-by-step lexical insertion.

Of course, we might opt for some intermediate view, such as for example constructing a numeration for each phase, and there might be something like a maximum of three items in each successive numeration, with many, many computational phases, etc. But the only reason to do such things seems to be to try to give the term "numeration" some non-trivial and yet non-absurd meaning. Jockeying back and forth along these lines doesn't seem to advance our understanding of natural language.

It seems to me that syntacticians whose work focuses on natural language and rationally use the term numeration must be thinking along other lines. In my view, the notion of numeration is appealing to generativists because it is unconsciously understood to be a necessary part of some idealized models of performance, that is, of **comprehension** and **production**. Such models are necessarily on the horizon because they might approximate what computational linguists are trying to do, and natural language computation is certainly here to stay. Moreover, with respect to production, it is often repeated that "we must have some idea of the lexical concepts we wish to express or understand before we begin a syntactic derivation," and a numeration seems to embody that kind of intuition.

With respect to a **comprehension model,** a mundane fact about language use, which seems to have escaped comment, is that in an idealized comprehension task, the idea of the language user being furnished with a c-numeration prior to syntactic computation is almost simple common sense. A hearer typically receives the **Phonological Form** ("PF") of a well formed sentence, segments it into words, and then associates these words with an understandable **Logical Form** ("LF").[2] The set of words taken by the hearer to constitute the sentence are a good approximation to a c-numeration, and this indeed precedes syntactic analysis. Thus, no one says they have understood a sentence but haven't understood any of its words (outside of perhaps a few fixed expressions). Of course, hearers may not be able to segment them further into morphemes, or tease apart clitic or contraction sequences. But as far as free morphemes go, not to have understood some word in a sentence is simply not to have (fully) understood the sentence.

Comprehension thus consists of the following sequence of operations on some sentence σ, perhaps carried out in steps on certain subparts σ_i of a longer sentence σ.

(1) **Comprehension Model:**

Step C1. Construct a phonetic sequence π for a heard sentence σ.

Step C2. Using the lexicon and the phonological component, construct a set (or a sequence) of words ν for σ from this π. Call ν a "candidate numeration."

Step C3. Construct an LF with a convergent numeration for σ from the candidate numeration ν at step C2.

It may be that the step C2 includes a possible finer-grained analysis, i.e. that the listener may be constructing word-internal morphological structures for σ as well, using e.g. the right-hand head rule of Lieber (1980), Selkirk (1982) and Williams (1981) for morphological and compound structures in English. But in any case, the step C2 already comes close to yielding a c-numeration of σ. This candidate numeration has been constructed with no circular reference to syntax.[3]

Recall that a c-numeration for σ is supposed to contain all and only the lexical items that occur in σ. While the candidate numeration ν constructed at step C1 contains no extraneous items, it may nonetheless be missing some of the lexical items needed for an LF of σ. In fact, because it is based on phonology, ν is *necessarily* missing any**null morphemes**. Consider for instance the sentence in (2a), which arguably contains at least the six null morphemes shown in (2b), under widely accepted analyses of English.

(2) a. Deer go to that pond less than we English to church.

b. $[_{DP}[_{PLR}$ Ø] deer]] $[_{I}$ Ø] go to that pond less than $[_{DP}$ we English $[_{N}$ Ø]] $[_{I}$ Ø] $[_{v}$ Ø] to [$[_{D}$ Ø] church]

We can give shorthand versions of familiar justifications for the six null morphemes in (2). (i) *Deer* must be formally plural for correct V agreement; as is increasingly common, I locate this plurality in the Determiner. (ii–iii) All English clauses have an I, the second of which here is clearly needed for nominative case assignment to *we*. (iv) *English* is an adjective, which regularly implies the presence of an N in English, which can here only be Ø. (v) The empty V is needed to select the complement P *to*. (vi) *Church* is a count noun requiring a D in English, which can be empty in certain senses and uses. A hearer can thus obviously provide several covert morphemes in the process of constructing a single convergent numeration at step C3.

[2] As in all current models of competence and performance, we are idealizing. The situation of the ideal hearer (without visual clues) is most closely approximated by telephone conversation.

[3] Nor does the model in (1) appeal to what seems to my mind some kind of fantasized formal model of situations, whereby the hearer picks out words on the basis of some presumed representation of "pragmatic context" plus a semantic but a-syntactic grasp of what is being said.

The fact is, much successful communication, which must be considered "normal," proceeds on the basis of a listener typically not hearing or processing any number of overt morphemes at steps C1 and C2. For instance, native speakers/ hearers can easily understand sentence (3a) pronounced as (3b), where the symbol ^{a} stands for a fully reduced schwa:

(3) a. Did you visit some of his relatives on the trip to help him?
    b. J^a viz^{a}t sm ^{a}iz rel^{a}tivz nd^a trip t^a heupim?

Probably no hearers could make sense out of this example even in context if they failed to pick out from the input the words *visit, relative, trip,* and *help*. That is, under typical conditions, none of the words *visit, relatives, trip,* and *help* in (3) can really be "missed" if the sentence is truly to be considered "understood" by the hearer. In some kind of idealized performance model, we probably should say that a successful candidate numeration v *must* contain, roughly speaking, all the "**open class**" items. However, a candidate numeration may or may not contain a number of overt **closed class** items; upon hearing (3) one might easily miss several of them.

For example, with the help of rising intonation, that is [C, +WH], it's plausible to imagine quickly and fully grasping sentence (3) in context, even if only hearing the subsequence *visit—relative—trip—help—him—?* That is, an accurate LF might be constructed from a candidate numeration containing only 2 of the 12 overt grammatical morphemes: *do—past—some—of—pronoun—possessive—plural—on—the—to—pronoun—WH.*[4] Even though many of these are absolutely central to the meaning, they can be somehow "filled in" in passing from the candidate to the convergent numeration in step C3.

Notice that this discussion has amounted to revising Step C3 in (1), as follows:

(4) Step C3′: Construct an LF for σ from the candidate numeration v at step C2, by "supplementing" v with "closed class items" from the lexicon.

There are some problems with formalizing this revision of the algorithm in (1).

(5) Comprehension model problems:
- a. In getting from phonetics π to words v (C2) by using the lexicon, there must be a faster way than checking some 20,000 lexical entries at each step.
- b. We don't yet see how to "supplement" v (C3′), especially since v isn't a unique subset of σ.
- c. We haven't yet characterized the notion of "closed class lexical item" in C3′.

The basis of overcoming the problem (5a) lies I feel in a lexical access principle carefully justified by the psycholinguistic results of Marslen-Wilson (1987, 1990), brought to my attention by K. Dahlgren (personal comm.). The author shows that during processing, words with acoustically similar initial segments are all activated at once. This finding is known as the "cohort theory of lexical access"; the idea is that many, many words from the same "cohort" are accessed upon hearing a first phoneme. Thus, the acoustically similar *beaker* and *beetle* are from the same cohort and activated at once, but *beaker* and *speaker* are not. Emonds (2003) presents further arguments for this view, based on morphological irregularities and oral poetry, and articulates this claim as the following principle:

[4] Alternatively, (3b) might plausibly be heard successfully without rising intonation but by inferring WH at the beginning (signaled by inversion) and by understanding *his* rather than *him*.

(6) **Dictionary Storage**. Dictionary items link constellations of semantic features *f* with phonological forms π. They are organized and hence accessed by virtue of π's initial consonant clusters C*.

As will be reviewed below, the relatively small number of lexical entries that lack purely semantic features are not subject to (6)—these constitute the "closed class items." But since these number probably well under 500, they do not seriously affect the simplification in step C1 that Dictionary Storage allows (Emonds: 2003, 6): "It is clear how the design feature (6) reduces the search space for the hearer—as soon as a candidate for an open class item begins to be pronounced, say the word *steal*, the search space is drastically narrowed to just those Dictionary entries beginning with *st-*."

Next, we can resituate (5b) as a difficult but at least familiar problem by crucially using an informed guess about performance. A lot of research in psychology, with which I have only a layman's superficial acquaintance, has established the existence of a truly "temporary" memory store called **short term memory**. Counter to conversational use of this term, short term memory means "a very few seconds"—it has nothing to do with remembering something only for a couple of hours, until bedtime, or whatever.

I don't think psychologists or linguists have made much of an attempt to relate short term memory and syntactic models. But in fact this memory span can play a crucial role in a comprehension model. It means that a hearer can "keep in mind" as a whole the sentences or clauses of roughly the length that need to be processed as a unit. That is, thanks to the faculty of short term memory, the procedure of "supplementing v with lexical closed class items" need not start at the left or right side; it can *start with any item in* v. The passage C3′ from a candidate to a convergent numeration can begin, for example, with those items whose lexical specifications don't require complements, i.e., those which are most deeply embedded. In fact, this coincides exactly with the way that most competence models since Chomsky (1965) envision bottom up **syntactic derivations**.[5]

This observation, plus the parsimonious assumption that competence and performance models should not needlessly differ, suggests that in an idealized comprehension model step C3', "supplementing v with lexical closed class items," is nothing other than performing a classic syntactic derivation. Reverting now to the term "c-numeration," we can thus revise (1) and (4) as follows:

(7) **Revised Comprehension Model:**

Step C1. Construct a phonetic sequence π for a heard sentence σ.

Step C2. Using Dictionary Storage (6) and the phonological component, construct a set (or a sequence) of words v for σ from this π. Call v a "candidate numeration."

Step C3*. Placing v in short term memory, construct an LF and a c-numeration for σ from v by implementing a classic bottom-up syntactic derivation.

There remains a condition on this model inherited from Step C3′ in (4) which we can state as follows:

[5] Since we are discussing idealized comprehension models, the fact that a hearer might begin with a few wrong items, and hence have to start over a relatively small number of times, is unimportant. In example (3), the only items whose lexical specifications might qualify them as not having potential complements are *deer, pond, English* and *church*. A bottom up derivation starting with *English* would go awry, but the other three choices would be appropriate.

Lexical Proviso on C3*. A candidate numeration becomes a c-numeration at LF by adding (*only*) "closed class lexical items" during a derivation.

Let us now return to the problem noted in (5c): we haven't yet characterized the notion of "closed class lexical item." We have however replaced the undefined processes "construct an LF" or "supplement *v*" with a familiar and partly solved problem, that of specifying well-formed syntactic derivations. But this clarification in itself does not address Problem (5c), which concerns determining the "search space" of closed class lexical items.

The Revised Comprehension Model (7) thus has a claim to being both formal and interesting exactly to the extent that "closed class lexical items" have a non-circular, formal definition, and properties with some empirical content. In fact, my and others' work over the past several years on the structure of a syntactically contentful lexicon has established several clear properties of this set. I will rely on these results to claim that Problems (5a) and (5c), far from being intractable, are currently already essentially solved.

2. A Competence-based Lexical Model: the Dictionary vs. the Syntacticon

Linguists are familiar with a basic dichotomy between open class and closed class lexical items, and are aware that the behavior of items in these two classes differs along a number of dimensions. However, linguists often content themselves with some kind of notion that the differences constitute a sort of "continuum," and do not suggest a sharp split between two distinct linguistic components. Even those who might tend to think of the differences as rather clear and as important for understanding the lexicon and lexical semantics have a tendency to neglect the difference in proceeding with studies of syntax, phonology and language processing.

Nonetheless, I think that work on the differences between open and closed classes shows *not only* that the two types of morphemes indeed differ sharply, but that the differences can best be understood in terms of a generative model using transformational derivations; indeed some properties cannot be formulated (or understood) without situating them within such a model. It then emerges that the many distinguishing properties of each of the two separate lexical components cluster in a clear cut way.

I claim that these differences in the properties of vocabulary impact centrally on syntax, lexical semantics, underlying phonology and processing. Several differences in syntax and lexical semantics have been clarified and justified in my earlier work. I take these syntactic distinctions to be established, and use these results here as needed. My research is just beginning to address the differences between the two vocabularies in phonology and processing; this essay also uses some early results in these areas, e.g., the Dictionary Storage principle (6).

It must be emphasized that the lexical model to be sketched here is entirely justified on the basis of its superior predictions in a competence or well-formedness model. Nothing in the model has been justified on the basis of performance models, idealized or otherwise. Hence if aspects of this theory of the lexicon conform to and confirm the Revised Comprehension Model (7), this constitutes independent support of the model and the lexical theory, not some circular reasoning based on intuited plausibility.

In this work on a competence model, I have followed traditional terminology in calling the entire set of lexical items the "Lexicon," and the proper subset of its open class lexical items the "Dictionary."

(8) A mental **Dictionary** contains open class items: [$_N$ table], [$_V$ disperse], [$_A$ clever], [$_P$ aboard].

All members of the Dictionary differ among themselves by **purely semantic features** f that play no role in syntactic derivations. Open classes are precisely those that have such features, and they are *proper* subsets of *only* the **lexical categories** N, V, A and P (P = pre-/post-positions).[6] I claim that all features of items in other categories such as MODAL, DET, NUMERAL, etc. as well as small subsets of N, V, A and P have no purely semantic features. They have **cognitive syntactic features F** that play a role in *both* semantics and syntax.

(9) The items in closed classes of **grammatical N, V, A, P** and in all other grammatical categories lack f,and are hence *fully characterized* by cognitive syntactic F.

It follows from (8) that the stored representations of such items, while in the Lexicon, are *not* in the Dictionary.

With a distinction between open and closed classes in terms of presence or absence of purely semantic features f, we can now define a lexical component that contains closed class items:

(10) **Split Lexicon.** A **Syntacticon** separate from the Dictionary contains a language's closed class items, i.e., those that lack purely semantic features f.

For concreteness, I exemplify a number of English Syntacticon items: $[_N$ self], $[_V$ get], $[_A$ other], $[_P$ of], $[_C$ if], $[_D$ every], $[_{NUM}$ two], $[_{DEG}$ very], $[_M$ can], $[_{NEG}$ not], $[_{NEG}$ un-], $[_p$ re-], $[_{N,A}$ -ing], $[_V$ -ize], etc.[7] The Syntacticon also turns out to include all the English suffixes that lack inherent stress, as shown I think by the extensive research and argumentation in Emonds (2000, Ch. 3 and 4).

What does it mean for a morpheme to have only cognitive syntactic features F as in (10)? It means that every feature of such a morpheme is used or at least "usable" in some syntactic rule. This entails a strong empirical claim. For any pair of lexical items α and α' characterized with only syntactic features F, defined as those which play a role in syntax (Chomsky, 1965, Ch. 2), α will differ from α' by some feature F′. Hence α and α' will not share whatever syntactic behavior depends on the value of F′. Therefore, we expect every closed class item will be syntactically different from every other, i.e., each will typically exhibit unique syntactic behavior. In fact, two closed class items in a language can have the same syntactic behavior only if effects of any rule(s) using F′ are somehow accidentally unobservable in that language.

(11) **Unique Syntactic Behavior.** Since each closed class item is a unique combination of syntactic features, the hallmark of each such item is unique syntactic behavior.

As examples, every Determiner and every Modal in English can be easily if tediously shown to

[6]Two points might lead to thinking that open and closed classes are not so distinct: There are certainly e.g. grammatical prepositions and "light verbs" that act more like closed class items. On this point, as we will see, the lexical bifurcation cuts across the lexical categories N, V, A and P, rather than including all their members among the open classes. But I claim that any given use of a V or a P either has or does not have a purely semantic feature.

Outside the lexical categories, numerals might seem like an open class. But this confuses language and arithmetic. Syntactically, there are only twenty some separate numeral morphemes. E.g., numbers like English *four-teen* and *forty five are* composed of other morphemes, and as far as grammar is concerned, members of groups like {5, 6, 7, 8, 9} and {*million, billion, trillion*} are in free variation. My speech has no meaningful cardinal above *trillion*; *quadrillion, octillion* and *zillion* are essentially for word play, being semantically imprecise items also in free variation. The first two can be calculated based on knowledge of Latin and arithmetic but such meanings are not part of ordinary usage.

[7]The category assignments here for some items in this list such as *re-* and *-ing* are not obvious, but can be convincingly argued for.

differ in syntactic behavior from every other.[8] In fact, all syntactic features used in derivations such as ANIMATE, MODAL, NEG, PAST, PATH, etc. are also fully cognitive—in fact, I deny the existence of "diacritic" or uniformly uninterpretable syntactic feature types. For this reason, I have chosen the term "*cognitive* syntactic" for these features of *all* categories that lack purely semantic features.

The characteristics of items in the many syntactic categories in the Syntacticon, some termed Specifiers or Functional Categories in various stages of generative development, conjunctions of various sorts, focus particles, pro-forms, etc. are so variable as to lead to a sort of traditional truism that "grammar is so irregular." By this is typically meant is that each such grammatical morpheme seems to be unlike all others. In the framework of this study, this truism is not a hand-waving admission of confusing complexity, but instead a precise consequence of Syntacticon structure: Unique Syntactic Behavior (11). It is akin to saying in chemistry that each element in the Periodic Table has "unique chemical behavior." In the theoretical schema that organizes these elements, each basic element differs from every other in terms of at least one feature (one "electron" in chemistry).

We can note that the "split lexicon" fits in well with an approach to language-particular grammars advocated convincingly in Borer (1984) and Ouhalla (1991).

(12) **Language-particular syntax** is composed wholly of lexical specifications, namely syntactic feature combinations associated with *closed class* items.[9]

Namely, a given Syntacticon *is* a language-particular grammar, and all rules or parameters of a language-particular grammar must take the form of a lexical entry of a closed class item.

The following table of differences between the Dictionary and the Syntacticon is adapted from Emonds (2000, Chapters 3 and 4), with some rearrangements and additions to reflect the present study's concerns with processing and phonology.

The term *open classes* for Dictionary items derives from the fact that there are *many* purely semantic features *f* available for proliferating membership in the categories N, A, V and perhaps P in (b-i), and that adult language users can constantly expand their own Dictionary resources by utilizing these features. This vast store, grouped into (only) four open classes of items, constitutes the Dictionary (Ouhalla's mental lexicon). In contrast, the smaller store of items without *f* is the Syntacticon (Ouhalla's grammatical lexicon). Since the features F on each B are relatively few, they can distinguish relatively few distinct morphemes and hence classes of such items have been called *closed*.

[8]There are 66 pairs of the 12 distinct Modal forms in Modern English: *will, would, can, could, may, might, shall, should, must, ought, need, dare*. Hence 66 (= 11+10+9+...) empirical contrasts are required. 32 contrasts are provided by the simple pattern that only *will, would, can* and *could* appear in tag questions with subjectless imperatives: *Be on time, {could/ *should} you?*. Only *need* and *dare* are negative polarity items, providing 12 additional contrasts. Only *will/ would* contract to their final consonants *'ll* and *'d*. And so on.

[9]I am implementing the conceptualizations in Ouhalla (1991, 7–10): "...Borer's [1984] approach ... associates parameters with individual lexical items, as part of the information included in their lexical entries, rather than with the principles of UG [Universal Grammar]. ... the nature of the lexical information which determines parametric variation [is] nothing other than the usual type of information relating to selection and grammatical features,... it is not information which is available over and above the familiar type of lexical properties; rather, these properties themselves determine parametric variation.

...possibly functional categories and substantives belong to two separate modules of the mind/ brain. ... there should in principle be a distinction between two notions of the lexicon, a grammatical lexicon [the Syntacticon—JE] which contains functional categories and which belongs to the domain of UG, in the sense that its categories are determined by UG, and a mental lexicon [the Dictionary—JE] which contains substantives and which exists independently of UG, that is an autonomous module of the mind/ brain (the conceptual system)."

(13)

	Dictionary	Syntacticon
a. Defining property:		
Items with purely semantic features *f*	YES	NO
b. Syntactic properties:		
i. Grammatical categories in the inventory	N, V, A, P	ALL
ii. **"Late insertion"** possible during syntax and at PF	NO	POSSIBLE
iii. Items with "**alternatively realized**" (AR) features	NO	POSSIBLE
iv. Full suppletion inside paradigms (*go/went*; *bad/worse*)[10]	NO	POSSIBLE
c. Phonological properties:		
i. Items conform phonologically to "primary vocabulary"	POSSIBLE	YES
ii. Bound items have inherent stress and head compounds	YES	NO
iii. Phonetically zero morphemes possible	NO	YES
d. Intermodal and processing properties:		
i. Open classes; adults can coin neologisms	YES	NO
ii. Interface with non-linguistic memory and culture	YES	NO
iii. Processing look-up in terms of initial consonant cluster	YES	NO
iv. Limited to Broca's area of the brain	NO	YES

3. Grammatical Behavior of Syntacticon Items

Let us look now at some differences between the two lexical components in Table (13) that contribute to how the Revised Comprehension Model (7) transforms candidate numerations into c-numerations.

3.1 Late Insertion (b-ii)

Incorporating the Syntacticon into the Revised Comprehension Model is justified only if there are independent reasons for believing that Syntacticon items can quite generally be inserted, as required by the Lexical Proviso C3*, during a syntactic derivation, and that furthermore, *only* these items can be so inserted.

In fact, these are precisely the properties of Split Lexicon model that have been independently justified in syntactic analyses of what is by now quite a number of constructions. These analyses using Late Insertion are seen in for example Emonds (1994, 2000), Caink (1998), Jo (1996), Veselovská (2000) and Whong-Barr (2002). All are based on paradigmatic and other syntactic differences that are predicted by *distinguishing* deep and later insertion. Deep insertion applies only to lexical items that inherently carry purely semantic features or are associated in, say, idiomatic expressions with such features. Most uses of Syntacticon items are

[10]Emonds (2003, section 4) argues extensively that suppletive paradigms for single lexical entries imply that (Syntacticon) items are looked up in terms of syntactic categories, rather than by phonological addressing. Cf. note 14 below.

not of this sort, and hence they are inserted rather in the course of syntactic derivations.

In light of the many analyses cited above, I mention here only one additional example of how different insertion levels for the same morpheme classes can explain familiar but ad hoc traditional grammatical distinctions. Emonds (1995) argues that personal pronouns with c-commanding linguistic antecedents have properties suggesting insertion *during* a syntactic derivation, while in contrast those with pragmatic antecedents appear to be inserted *prior* to derivations (presumably requiring some interfacing with cognition and memory outside of syntax). Finally meaningless, non-referential expletive pronouns are inserted in PF, i.e. they do not contribute to LF. Thus, lexical insertion at three levels explains the otherwise ad hoc three-way division among free, bound, and expletive pronouns.[11]

A theory-internal redefinition of *all* lexical insertion as "late" or "in PF" completely fails to express such differences. Such a move implies that Dictionary items can have the skewed and unique distributions characteristic of Syntacticon items, a massively wrong empirical prediction. What syntax repeatedly exhibits is exactly the split between closed and open class items shown in the Table as (13b-ii). This split is then the crucial step for confirming the possibility of a non-trivial computational procedure for getting from candidate to convergent numerations.[12] Late Insertion thus provides a sufficient basis for the Lexical Proviso C*, and hence warrants incorporating the crucial role of the Syntacticon in the Revised Comprehension Model.

3.2 Alternative Realization (b-iii)

We can better understand the late insertion of Syntacticon items in this Model if we also grasp another central aspect of their grammatical behavior, namely "Alternative Realization," point b-iii in Table (13).

(14) **Canonical position** of features. **Universal Grammar** ("UG") associates a very few cognitive syntactic features F with each grammatical category B. These B are the canonical positions of the F.

(15) **Canonical realization** of features. Syntactic features and categories contribute to Logical Form only in these canonical positions.

(16) **Alternative realization** ("AR"). Syntactic features F of a category B can be alternatively realized in non-canonical positions only by a Syntacticon item in an "adjacent" higher or lower head X^0. They do not contribute to LF in these positions.

"Adjacent" here can be taken as requiring that some projections B^j of B and X^k of X^0 be sisters. There are situations where intervening empty heads between B and X^0 allow AR to reach up or down a tree even further than this, but they need not concern us here.

An important consequence of AR is that it often allows a category whose canonical features contribute to LF to be empty in PF:

[11] While (only) Syntacticon items can be inserted or merged in a domain after its derivation begins, it should be kept in mind that items from both lexical components can satisfy lexical co-occurrence restrictions at the beginning of such derivations and hence can be inserted at the outset of a derivational "cycle" or "phase," more or less along the lines of the deep lexical insertion of Chomsky (1965). That is, Syntacticon items do in certain cases undergo deep insertion, with free pronouns and their pragmatically determined reference instantiating this possibility.

[12] This conclusion has emerged from conversations with L. Veselovská. She further points out, as I hope to demonstrate in more detail in future research, that many analyses going under the rubric of Distributed Morphology have also justified late insertion of closed class morphemes. Although the original paper outlining this approach (Halle & Marantz, 1993) proposes that *all* lexical insertion is "late," the actual empirical (as opposed to theory-internal) justifications for this proposal typically involve bound morphemes, which are all members of the Syntacticon.

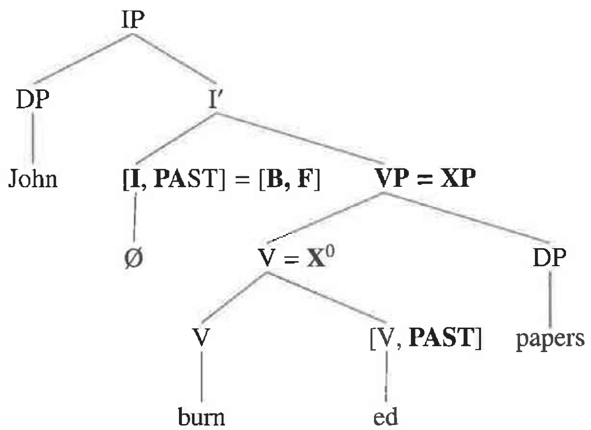
(17) **Invisible Category Principle** ("ICP"). If all of a category B's canonical features F are alternatively realized, the category B can be empty in PF.

As an example of how AR and the ICP work together, consider finite verbal inflections such as English *–ed,* which we can assume is minimally specified as having a host V and a grammatical meaning PAST. As is typical of most (but not all) inflection, it appears in the same positions as does its host category V. The best way to express its verbal distribution is to say that the inflection *is* itself also a V; after all, that is what "V" means—something with the distribution of a verb.[13] The minimal lexical entry (18) expresses these properties, and it is hard to imagine a theory which could boast a simpler characterization of *–ed.*

(18) English Past Tense: ed, V, +PAST, +V__

It is widely agreed that Universal Grammar recognizes that I is the canonical syntactic position of PAST. Consequently, the position of PAST on V licensed by entry (18) in the English Syntacticon exemplifies "lower" Alternative Realization. In the tree below, the AR nodes are in bold: F = PAST, B = I and X^0 = V.

(19)



Because of AR, the ICP (17) allows the I node to remain empty in PF, since its only interpretable feature is PAST.

According to AR (16), a feature such as PAST does not contribute to LF in an alternatively realized position, here on V. Moreover, the "V over V" structure in (19) is plausibly treated the same at LF as a single V. Consequently, a morpheme such as *–ed* has no interpretable features and is hence inserted in a derivation only after Spell Out, in the PF component. Such reasoning in terms of AR and the ICP, it turns out, extends to basically all of what is traditionally known as "inflection" (Emonds: 2000, Ch. 4). That is, *inflections regularly result from Late Insertion from the Syntacticon.*

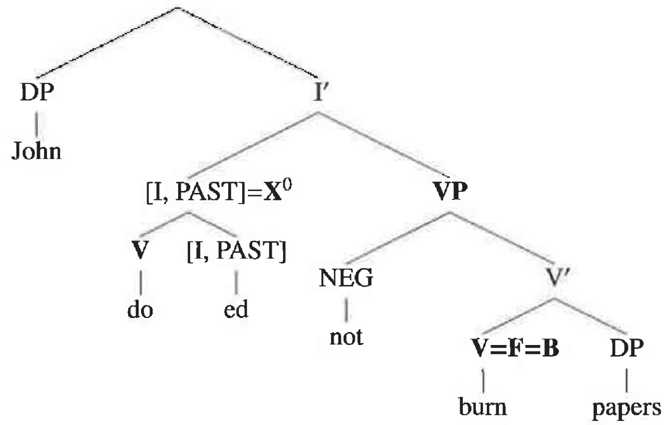
If we transfer this property of a competence model to the Revised Comprehension Model, this seeming conclusion would be that inflections can essentially systematically remain "unheard" during steps C1 and C2, in which Phonology and the Dictionary construct candidate

[13]For a suffix to have the distribution of B and to have a host of category B is not always the same thing. Even among English inflections, the *–ing* of gerundive nominals has the distribution of N (its category is then N) and a host of category V (its word-internal subcategorization is +V__).

numerations. Late Insertion will always be able to rescue such numerations by providing them with appropriate inflections. These in turn come into play by licensing empty categories with canonically realized (i.e., LF-interpreted) features. This line of thought thus helps us understand why inflections notoriously lack phonological salience—as long as the hearer and the speaker share the same Syntacticon, the Revised Comprehension Model allows the hearer to "fill in" the inflections during step C3*.[14]

One further principle interacting with AR plays an important role in a Comprehension model. Note that a category such as V itself can also play the role of F in AR. In (20), an auxiliary *do* alternatively realizes V (=B) under an I^0 (= a higher X^0) that is a sister of VP (=B^{j}). The nodes standing in an AR relation are in bold. In this case, the ICP doesn't come into play, since the V *burn* has other interpretable semantic features *f* besides V.[15]

(20)



Now to illustrate the workings of these principles further, suppose then we eliminate NEG from (20); how do we then choose between (19) and a NEG-less (20): *John burned papers* vs. **John did burn papers*? It appears that, for purposes of a competence model, we can simply count the words and choose the sentence with fewer; this rather simple conception of Economy has widespread and apparently correct consequences (Emonds (1994); Collins (2001)).

(21) **Economy of Derivation.** In the course of a syntactic derivation, realize a given Logical Form by inserting *as few free morphemes (words) as possible.*

The effect of (21) is then correctly that *John burned papers* is chosen over *John did burn the papers.*

The counterpart of Economy for step C3* of the Revised Comprehension Model (7) should be something quite similar to (21). For example: in the course of converting a candidate numeration to a c-numeration, complete a syntactic derivation of LF by the same means: inserting

[14] In fact, the hearer, for example a young child, may have significantly less phonological specification in its Syntacticon entries than the speaker. Emonds (2003, section 4) justifies the following counterpart to Dictionary Storage (6):

Syntacticon Storage. Syntacticon items are organized by virtue of their syntactic categories F.

The upshot of this principle is that the lexical look-ups during step C3* are based on syntactic categories and contexts rather than via phonological addresses. Hence, as long as the hearer's Syntacticon items have the same grammatical specifications as the speaker's, they can be utilized to supplement the candidate numeration and arrive at an LF. Of course, this is not to say that the Comprehension Model excludes the possibility that phonological discrepancies may lead to significant performance errors.

[15] *Do+ed* is of course pronounced *did.*

as few free morphemes as possible. Along these lines, we see that Economy in a competence model and in a comprehension model reduce to essentially the same principle.

A significant part of any "particular grammar," i.e. Syntacticon, consists of how each language uses AR to implement Economy of Derivation in various ways. That is, various inflections and clitics are used to reduce the number of free morphemes required for a given LF expression. Examples of matches of syntactic features with categories in canonical positions are given in columns one and two, while column three shows typical alternative realizations:

(22)

Syntactic features F	Canonical positions B	Typical AR on X^0
Tense and modal features	I	V (lower); C (higher)
STATIVE vs. ACTIVITY	V	I (higher); D (lower)
PERFECTIVE (aspect)	V	I
Quantifier features	D or NUM	D (higher); N (lower)
LOCATION and PATH features	P	V (higher); D (lower)
ANIMATE; COUNT vs. MASS	N	D and NUM
+INHERENT	A	V (Spanish *ser* vs. *estar*)
+DEFINITE; +SPECIFIC	SPEC(DP) or D	V (Romance clitics)

On the other hand, the Syntacticon consists of much more than AR. For example, all the English modals and quantifiers are part of it, and yet plausibly have little to do with AR. As central as its members are to LF, the Comprehension Model espoused here claims that they do not always need to be "heard" to be understood. If they can be inserted during or subsequent to a syntactic derivation (i.e., whenever they express general cognitive rather than the more specific Dictionary meanings), they need not be part of a candidate numeration (i.e., heard) but can be supplied in the course of the syntactic derivation.

3.3 Phonetically Zero Morphemes in the Syntacticon (c-iii)

With respect to zero morphemes, we saw earlier that supplementing a candidate numeration at Step C3* must include finding null grammatical elements in a lexicon. The fact that the Syntacticon must contain all such elements, point c-iii in Table (13), supports the idea of identifying the area of search with the space of the Syntacticon.

It can hardly come as a surprise to a generative audience that languages have syntactically null elements. However, the mere existence of abundant empty categories in syntax doesn't guarantee that the Syntacticon specifies any phonologically null *lexical entries*. To establish the existence of zero morphemes in the lexicon, we must look for contrasts between empty and overt categories in minimally different grammatical systems. I thus give examples of inflections, clitics and free forms that are almost certainly language-particular lexical zero morphemes.

Null inflections

One place where languages seem to vary significantly is in morphological agreements. It is rare for a language to overtly agree only in unmarked forms. Yet English present tense agreement in person and number seems to have just this property. A Syntacticon entry for its unusual null agreement in marked forms is given in (23). The feature –PAST on this empty V alternatively realizes the canonical ±PAST feature of an empty I; μ stands for marked values of agreement features, namely plural and non-third person.

(23) English 3^{rd} singular: V, –PAST, { μNUM/ μPER }, +V__ ; $\pi = \emptyset$

This morpheme constitutes a right hand head of a verbal structure, as shown in (24), and alternates with the overt suffixes *–s* and *-ed* in this position.

(24) The structure induced by the joint effects of (23) and the ICP (17):
$[_{\text{I,-MODAL,}\pm\text{PAST}}\ \emptyset\ ] \ldots [_{\text{V}}$ lexical verb $[_{\text{V,}\pm\text{PAST}}\ \emptyset$/ -s/ -ed]]

Now, must the English Syntacticon stipulate an empty V for both of the two empty items in (24)? The following trade off between the lexicon and the theory of grammar seems to hold:

(25) a. When AR or movement licenses an empty X^i, the Syntacticon need not.
b. When neither AR nor movement licenses an empty X^i, the Syntacticon must.

So the English Syntacticon doesn't stipulate empty Is with tensed verbs, but rather the empty inflection on V as in (23). We can say that a child language learner expects empty categories licensed by movement or AR as in (25a), but has to learn those in (25b).

A NULL CLITIC IN ITALIAN

Romance pronominal clitics can alternatively realize in V all interpreted features of an argument DP or PP, and are thereby able to license such phrasal arguments as empty. But generally, analyses of Romance have not considered the possibility that the licensing clitic itself might be empty.

Rizzi (1986) has effectively established that Italian has empty phrasal direct objects in the position of α in (26), which is interpreted as a null generic pronoun:

(26) a. L'ambizione spesso spinge α a commettere errori.
'Ambition often pushes (one) to make mistakes'
b. Questa musica rende α allegri.
'This music makes (one) happy'

He claims that a null object $[_{\text{DP}}\ \alpha\ ]$ in these examples has the interpreted features [+HUMAN, +PLURAL, –SPECIFIC]. But in Emonds (2000, Ch. 9), I argue that neither UG nor a Syntacticon can directly license such null direct object phrases. Rather, I try to establish that the Italian Syntacticon contains a null verbal clitic that alternatively realizes these features $[_{\text{DP}}\ \alpha\ ]$ on the V itself. By the trade-off in (25), the Italian Syntacticon must then list this clitic, though it does not list the null object DP itself.

FREE FORMS OF X^0 WITH ZERO ALLOMORPHS

If some free morpheme $[_{\text{X}}\ \alpha\ ]$ is obligatorily silent, the Syntacticon must specify its phonology as $\pi = \emptyset$. If α is pronounced optionally, I use the lexical notation (π). English relative pronouns in restrictive relative clauses exhibit such an alternation.

(27) English WH pronouns: a. D, WH, ANIMATE, –__NP; π = (who)
b. D, WH, PLACE, –__NP; π = (where)
c. D, WH, DEF; π = (which)

A restrictive relative pronoun is a specifier of a CP sister of some nominal projection with the same features. The English relative pronouns in (27) can be optionally silent: *The boy (who) I saw; The place (where) John lived; the book (which) I bought there.* In the same configuration, the French counterparts (*qui, où, lequel*) of these pronouns must be overt, though French and English syntax of relative clauses and pronouns is otherwise similar.[16]

Another example of a difference between these two CP systems is the fact the the unmarked English complementizer has a zero allomorph that is allowed by UG when it remains in its basic position directly governed by V. French doesn't allow such a C to be null in any position.

(28) English complementizer: C, +__IP; π = (that)

(29) John persuaded Mary (that) she would easily get the job.
We explained to her (that) her children should stay outside.

Stipulated empty copulas

An empty V is cross-linguistically quite often a copula, either of *be*-type or of *have*-type. (*Be* does not assign accusative case whereas *have* does.) The unpronounced Russian copula may exemplify the *be*-type.[17] Whatever the licensing condition for null allomorphs of *be* and *have* in various languages, they still require language-particular Syntacticon stipulations, since the related systems of Czech/ Russian and Swedish/ Norwegian seem to vary as to when a null allomorph is allowed.

In the Revised Comprehension Model (7), the candidate numerations that emerge from step C2 have been constructed on the basis of the heard phonology of their elements. Hence it is obvious that these numerations cannot contain null morphemes. Yet our brief survey of Syntaction entries demonstrates that language-particular null morphemes must be part of the complete convergent numerations that correspond to Logical Forms. I conclude that these lexical entries must enter comprehension processing during the only remaining step, C3*. That is, the null lexical entries of a Syntacticon are added to a tree during rather than before (or after) its syntactic derivation.

4. A Sketch of a Production Model

We have seen that quite a lot of research on devising a model of syntactic mechanisms in the lexicon has arrived at the following picture: Simultaneous with syntactic derivations that combine a chosen set of lexical items, here termed a "candidate numeration," into a sentence with an LF, more items from a limited set of grammatical items, called the Syntacticon, are added to the original set by point b-ii in Table (13), in order to obtain an LF and a corresponding "convergent numeration." This distinction between an initial set of lexical items and an augmented one is due to the Split Lexicon (10). Moreover, some of these added items can be phonologically null, by point c-iii in Table (13). Previous research has justified this competence model in a range of syntactic analyses, that is, on the basis of accounting for patterns of well-formedness and ill-formedness, with no reference whatsoever to either comprehension or production models.

[16]Chomsky and Lasnik (1977) claim that a necessary condition on "free deletion in COMP" of relative pronouns is identification of the deleted items' features by those of the antecedent sister NPs. This cannot however be a sufficient condition for deletion, as the comparison with French shows. Hence, Syntacticon entries as in (27) are still required.

Non-restrictive relative CPs are not sisters of the NP they modify (Emonds, 1979). As a result, their antecedents fail to identify the features of their relative pronouns, so they can't be omitted.

[17]A contentless and hence possibly null copular V may also include an uninterpreted feature, for example, an alternatively realized PAST feature of I. This seems to occur with the *have*-type copula of Norwegian/ Swedish. The perfect auxiliary *ha* 'have' is "deletable" after past tense modals in both languages (Julien (2002, section 1)). The extension of the *ha/0* alternation into Swedish infinitives seems subject to the same condition of "pastness." Julien's abstract concludes: "In Swedish, . . . *ha* need not be spelled out if *ha* shares its features with some element that is overtly realized . . . ."

But on the performance grounds examined above, this Split Lexicon is essentially what seems to be needed in the Revised Comprehension Model (7). In particular, it accords with the Lexical Proviso on step C3* of the Model. So it appears that we can directly incorporate current versions of syntactic competence models into a final—admittedly very formidable—step C3* of this Comprehension Model: "construct a Logical Form by means of implementing a syntactic derivation"

I now want to make some sketchy remarks about an idealized Production Model. What seems hopeful is that step C3*, if transposed to production, also seems to significantly simplify its schematic conceptualization. For it turns out that by using this "step 3," it becomes plausible to claim that speakers can "say anything they want," at least within the limits of the concepts and categories of the Lexicon.[18] That is, a candidate numeration can be constructed via an a-grammatical selection of words (formally, a random selection; intuitively, a selection determined by "context" and "pragmatics").

For example, suppose a person contemplating current relations between him/ herself and a circle of friends decides, in stream of consciousness fashion, to linguistically express the following string of randomly associated concepts from the lexicon:

(30) Many—friend—recent —nice—me—repay—them—invite—party

Without too gratuitously imputing structures to these thoughts, we might imagine that the concepts could then cluster as in (31a), where X (Y) means that Y depends on or modifies or is relevant to X. Moreover, lexical concepts necessarily come with lexically specified categories, so (31a) actually implies (31b):

(31) a. ((Many (friends)) (recent) nice (me)) (repay (them) (invite (party)))

b. ((Many_D (friends_N)) (recent_A) nice_A (me_D)) (repay_V (them_D) (invite_V (party_N)))

Let's call the idea of associating a random set or sequence of lexical concepts as in (30) step P1 in a Production Model, and an impulse to hierarchically group them as in (31) step P2:

(32) **Production Model:**

Step P1. Construct a sequence of lexical items τ for a new sentence σ. Call τ a "candidate numeration."

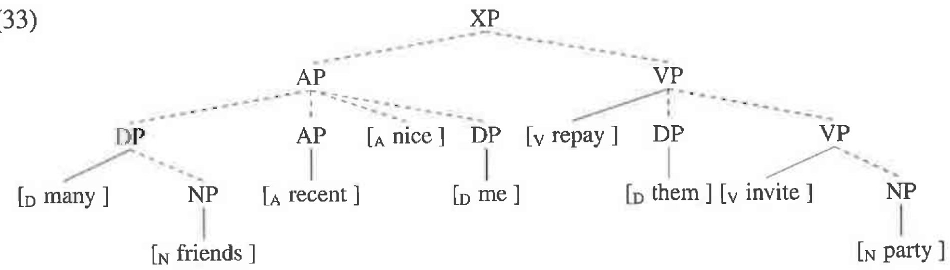
Step P2. Construct a preliminary labeled bracketing (equivalently, a tree) for σ from this candidate numeration, calling it $T(\tau)$.

Although the following step in the production scenario is not entirely crucial to the line of argument here, it may play an important role and in any case it fixes ideas. We can imagine that parentheses and lexical labels as in (31b) algorithmetically translate into endocentric phrases as follows: *a left parenthesis immediately followed by a morpheme indicates a phrasal projection and its left hand head.*[19] The remaining dominance relations in (31b) are represented in (33) with broken lines.

[18]These limits, as set out in detail in Emonds (2000, esp. Ch. 2–4 and 9), are far from trivial. For example, if a speaker wishes to utilize the resources of open class lexical choice, then there are only four grammatical classes available—point b-i in Table (13).

[19]The left-right asymmetry in the algorithm is a function of the example under discussion being English, in which phrases typically have left hand heads.

(33)



For P3, we can now use exactly the step C3* from the Revised Comprehension Model (7).

Step P3. Placing the tree T(τ) short term memory, construct an LF and a convergent numeration for σ from T(τ) by implementing a syntactic derivation.[20]

At this point, just as in the Revised Comprehension Model, we also need the Lexical Proviso for P3, now statable precisely as in (34) in terms of the formally defined grammatical lexicon.

(34) Syntacticon Proviso.
A candidate numeration becomes a c-numeration at LF by adding *(only)* Syntacticon items during a derivation.[21]

We of course need one further phonological step in a Production Model:

Step P4. Use the lexical entries of τ and the phonological component to construct a PF for σ.

By adding appropriate Syntacticon members as needed in the partially specified tree τ in (33), we can generate the following enrichment, with a well-formed LF and PF:

(35) Many of my friends have recently been nice to me, so I'll repay them by inviting them to a party.

It goes without saying that this "first stab" at a Production Model needs to find a set of exact principles by which Syntacticon items are added to a partially structured candidate numeration such as (31) so as to produce final products such as (35). However, I think a not insignificant amount of relevant work in this direction, starting from some skeletal endocentric dominance relations as in (33), is summarized in Emonds (2000, Chs. 7 and 8). These chapters use the principle of Economy of Representation and justify a Logical Form Case Filter, along with Economy of Derivation (21) given earlier.

(36) **Economy of Representation.** Structural requirements such as subcategorization frames are to be satisfied at a level of derivation with as few phrasal nodes as possible.

(37) **Logical Form Case Filter.** At LF, internal arguments YP of X^0 are each specified differently for Abstract Case, where "no case" is one of the Case values.

[20] It may well be the case that many of the choices made in a comprehension model as part of step C3 (insertion from the Syntacticon) will be made earlier in production, as part of choosing a candidate numeration in step P1. However, at the rather programmatic and tentative level of this essay, this difference does not seem to play a central role.

[21] Again, central aspects of this conception, especially the compatibility of Late Insertion from the Syntacticon with a Production Model, are products of conversations with L. Veselovská.

The tension between these principles, namely the lack of structure favored by (36) vs. a necessary minimum of structure imposed by (37), moves in the direction of near unique determination of structures like (35) as a realization of (31). Further restrictions on how syntactic derivations must proceed (including necessary Late Lexical Insertions) are provided by the lexical co-occurrence requirements themselves as well as general principles for well-formed phrase structures, as for example:

(38) a. English VPs must syntactically project to IP, and likewise NPs to DP.
b. APs and VPs require their own subjects (the Extended Projection Principle of Chomsky (1981)).

The importance of these latter principles is examined in some detail in Emonds (2000, Ch. 1).

Perhaps the most striking property of the Production Model (32) is the sharing of the specification of *bottom up syntactic derivations* with both a competence model and a comprehension model. It has often been felt that a production model must necessarily involve top down processing, since the ideas expressed in a sentence first are probably the first to be mentally produced, and leftmost (=first) generally means topmost at least in the trees of head-initial languages. However, the notions of candidate numeration and short term memory suggest that this "necessary conclusion" is flawed. The order of items in a candidate numeration no doubt reflects the order of thought. But the Production Model (32) claims that the *whole* candidate numeration is placed in short term memory, and only then does a syntactic derivation start. There is then no reason for that derivation to have any particular architecture, other than that best served by a competence model. Quite plausibly, a syntactic derivation will respect the left-right order of the candidate numeration as much as possible, but it can presumably do this even though operations begin other than at the left edge.

In conclusion, I summarize by assembling the steps of the envisaged Production Model:

(39) **Production Model:**
Step P1. Construct a sequence of lexical items τ for a new sentence σ. Call τ a "candidate numeration."
Step P2. Construct a preliminary labeled bracketing (equivalently, a tree) for σ from this candidate numeration, calling it T(τ).
Step P3. Placing the tree T(τ) short term memory, construct an LF and a convergent numeration for σ from T(τ) by implementing a syntactic derivation.
Step P4. Use the lexical entries of τ and the phonological component to construct a PF for σ.

(40) Syntacticon Proviso.
A candidate numeration becomes a c-numeration at LF by adding *(only)* Syntacticon items during a derivation.

Because the provisional Production and Comprehension Models proposed in this essay share their crucial and most difficult to formulate step 3 with the generative enterprise itself (the construction of a competence model for syntax), we can concretely hope to develop all three models in tandem. A central feature of these models is that all depend on focusing on and developing our understanding of the Syntacticon, as reflected in (34). And even at this preliminary stage, we have seen that candidate numerations in both comprehension and production tasks can be considered to consist of essentially "random" or "stream of consciousness" lexical choices. The syntactic derivations they trigger will converge to well-formed Logical Forms, provided our models incorporate the possibility that these candidates be augmented by Syntacticon items. The Syntacticon is almost invariably equal to the task of turning the candidates into convergent numerations, in a computationally tractable subset of cases. But if we ignore its properties in constructing our models, the notion "numeration" simply undermines the formal grammar enterprise.

References

Aitchison, Jean (2003). *Words in the Mind.* Blackwell Publishers, London.

Borer, Hagit (1984). *Parametric Syntax.* Foris Publications, Dordrecht.

Caink, Andrew (1998). *The Lexical Interface: Closed Class Items in South Slavic and English.* Ph.D. dissertation, University of Durham.

Chomsky, Noam (1965). *Aspects of the Theory of Syntax.* The MIT Press, Cambridge, Mass.

Chomsky, Noam (1981). *Lectures on Government and Binding.* Foris, Dordrecht.

Chomsky, Noam & Lasnik, Howard (1977). Filters and Control. *Linguistic Inquiry*, **8**, 425–504.

Collins, Chris (2001). Economy Conditions in Syntax. In Baltin, Mark & Collins, Chris (Eds.), *The Handbook of Contemporary Syntactic Theory*, pp. 45–61. Blackwell Publishers, Oxford.

Emonds, Joseph (1979). Appositive Relatives have no Properties. *Linguistic Inquiry*, **10**, 211–243.

Emonds, Joseph (1994). Two Conditions of Economy. In Cinque, G. et al. (Ed.), *Paths toward Universal Grammar: Papers in Honor of Richard Kayne*. Georgetown University Press, Washington.

Emonds, Joseph (1995). Deep Free and Surface Bound Pronouns. In Campos, H. & Kempchinsky, P. (Eds.), *Evolucion y Revolucion en Romance: Festschrift for Carlos Otero*. Georgetown University Press, Washington.

Emonds, Joseph (2000). *Lexicon and Grammar: the English Syntacticon.* Mouton de Gruyter, Berlin.

Emonds, Joseph (2003). Conditions On Empty (Allo-)Morphs: Two Types Of Lexical Access. *Theoretical and Applied Linguistics at Kobe Shoin*, **6**, 1–23.

Halle, Morris & Marantz, Alec (1993). Distributed Morphology and the Pieces of Inflection. In Hale, Ken & Keyser, Jay (Eds.), *The View from Building 20*, pp. 111–176. The MIT Press, Cambridge, Mass.

Jo, Mi-Jeung (1996). Morphosyntactic Roles of the Grammatical Verb *Ha*. *Korean Journal of Linguistics*, **21**, 1179–1204.

Julien, Marit (2002). Optional *ha* in Swedish and Norwegian. *Journal of Comparative Germanic Linguistics*, **5**, 67–95.

Lieber, Rochelle (1980). *On the Organization of the Lexicon.* Ph.D. dissertation, MIT.

Marslen-Wilson, W. D. (1987). Functional Parallelism in spoken word recognition. *Cognition*, **25**, 71–102.

Marslen-Wilson, W. D. (1990). Activation, competition and frequency in lexical access. In Altmann, G.T.M. (Ed.), *Cognitive Models of Speech Processing*. MIT Press., London.

Ouhalla, Jamal (1991). *Functional Categories and Parametric Variation.* Routledge and Kegan Paul, London.

Rizzi, Luigi (1986). Null Operators in Italian and the Theory of pro. *Linguistic Inquiry*, **17**, 501–557.

Selkirk, Elisabeth (1982). *The Syntax of Words.* The MIT Press, Cambridge, Mass.

Veselovská, Lida (2000). Agreement Patterns of Czech Group Nouns and Quantifiers. In Corver, N. & van Riemsdijk, H. (Eds.), *The Function of Content Words and the Content of Function Words*. Mouton de Gruyter, Berlin.

Whong-Barr, Melinda (2002). Small Clauses are Bigger. In *Proceedings of the Linguistics Society of Korea International Summer Conference, Vol. II: Workshops on Complex Predicates, Inversion and OT Phonology.* Thaehaksa, Seoul.

Williams, Edwin (1981). Argument Structure and Morphology. *The Linguistic Review*, **1**, 81–114.

Author's E-mail Address: jeemonds@sils.shoin.ac.jp
Author's web site: http://sils.shoin.ac.jp/~jeemonds/