

*Torben Thrane\**

## **The Grammars of *Seem***

### **Abstract**

This paper is the ms. of my inaugural lecture at the Aarhus School of Business, 26th September, 1995, with minor modifications. It traces some of the basic assumptions of Chomskyan theories of syntax, as they have determined the requirements of a syntactic description of English sentences containing the verb *seem* at various points in time. It is shown that although the purpose of linguistic *description* must be taken into account when choosing between competing, descriptively adequate grammars, this will not be enough unless the goal of linguistic *explanation* is accessible to independent investigation. It is argued that a reformulation of the issue between functional explanation or not in linguistics may overcome the problem.

### **1. Introduction**

There have been two considerations behind my choice of subject for this lecture. One has been a concern to indicate what ideas I have for the development of the research profile of the department. The second has been my wish to acknowledge the achievement of Bent Jacobsen, my predecessor, and to defer to the research tradition of the department. It seemed to me that both considerations could be honoured if I sketched the evolution of Chomskyan linguistics as I see it. The restriction to Chomskyan linguistics prevents mention of many types of grammar, some of which have been quite influential and well studied. This goes for various types of Dependency, Valency and Categorical grammar, to a lesser degree for Relational and Arc Pair Grammar, and to a still lesser degree for Stratificational and Word grammars of various types. It also precludes mention of the very popular branch of European Functional Grammar initiated by Simon Dik, the Systemic Functional Grammar developed by Halliday, and various versions of Case Grammar and Cognitive Grammar. The omission of all of these should not be taken as

---

\* *Torben Thrane*  
*The Aarhus School of Business*  
*Department of English*  
*Fuglesangs Allé 4*  
*DK-8210 Aarhus V*

an indication that I consider them uninteresting or ‘inferior’ in any way. On the contrary, I think many of them have revealed facts about language that Chomskyan linguistics could not originally handle and therefore simply ignored until it was revised to absorb their results. One of my hopes is to inspire research in all or some of these, in this way perhaps to broaden the theoretical outlook of the department.

Even with the restriction to Chomskyan linguistics, the task I am setting myself would still be an awesome one in just an hour if I had to do it in general terms. Yet it so happens that there is one verb which has played a recurrent and persistent role in this development right from the start. In fact, and putting it rather crudely, the development of Chomskyan linguistics may be seen as the history of the changing views about what it takes to describe English sentences containing the verb *seem*.

So this is part of what I’ll do. I’ll trace the theoretical and analytical requirements thought necessary for writing a grammar of *seem* within Chomsky’s paradigm and one or two theories that have branched off from it, and I’ll pay special attention to the assumptions of linguistic analysis at various times in the evolution. Number [1] in the handout is a statement of what appears to be the most basic assumption of linguistics and grammar of any kind at any time:

- [1] It is the goal of linguistics to provide a *description* of the systematic properties of all languages which is rational, general, consistent, and as simple as possible.

This has been known as the requirement of *descriptive adequacy* since Chomsky’s *Current Issues in Linguistics* (1964). It is taken over from the natural sciences, and it is therefore sometimes taken to imply that there is *one and only one true description* of a language. But it doesn’t. For even though as eminent a scientist as Niels Bohr repeatedly stressed that it is impossible to give an impartial description of nature (‘the very act of observation influences the object observed’), it has nevertheless been a view with some chance of apparent success in the sciences for the simple reason that nature is not there for a *purpose*. To illustrate, when Rutherford - the nuclear physicist - was once asked by the woman sitting next to him at a dinner party what *use* nuclear physics might possibly be, he replied testily: ‘Ma’am, what *use* are babies?’

In contrast, language certainly seems to be here for a purpose - in fact for many different purposes. It has arisen, not just to fill a gap in nature,

but to fulfill a *need*. The prevailing view at the moment is that the prime purpose of language is to convey information, and this may well be so. But the point is not *which* purpose language is supposed to serve. The general point is usually taken to be this: *Whichever* purpose language is supposed to serve, *must* that purpose be borne in mind when you formulate a descriptive theory for it? John Searle, the eminent philosopher of language, says 'yes', Chomsky, the no less eminent linguist, says 'no'.

As the point is formulated, I must agree with Chomsky. But I think the formulation is spurious. As I see it, the point should be formulated like this: What is the purpose of your *description* of language? In other words, what is it you are trying to tell me about *language* by *describing* it the way you do? You may describe the heart as a large muscle of a unique structure with a varying number of cavities, depending on what kind of heart it is. Or you may describe it as a pump. Or you may describe it as the seat of love. Which description is true? I suppose they all are - but in different contexts. They may all tell you something *useful* about the heart.

Someone might object that only the heart-as-muscle description is 'true', for it is the only scientific *description* of it. The heart-as-pump-description is a *functional* one, and the heart-as-loveseat is just romantic dribble. But here, I'm afraid, you'll run into difficulties. What you are advocating is *description for description's* sake. And *that* is not very interesting. What is worse, it is also circular.

To continue the heart story, the biology student who was satisfied with a description, down to the minutest fibre, of the structure of the heart-as-muscle, would no doubt flunk his exam if he had failed to realize that there is a *correlation* between structure-of-muscle and function-as-pump. But now an interesting picture emerges. It turns out that our fictitious student might claim *either* that the function-as-pump *explains* the peculiarity of the structure-of-muscle, *or* that the peculiarity of the structure-of-muscle *explains* the heart's potential as a pump. Now, if he was Searle's student, he would get a good mark for claiming the former, whereas Chomsky would probably fail him. And vice versa.

After this reformulation of the point of debate between functionalism or not, I believe that Chomsky and Searle are both partly right, both partly wrong. Chomsky is right and Searle wrong to the extent that a structural solution to a particular functional problem need not be the only logically possible one - after all, the functional problem of loco-

motion has been solved structurally by the evolution of legs, wings, and fins, depending on context. Chomsky is wrong and Searle right to the extent that it is precisely the *peculiarity* of the structure-as-muscle that is in need of explanation - *why* should the fibre structure of the heart be different from that of other muscles?

However this issue might be settled, the point remains that a scientific description of a phenomenon is only interesting if it tries to *explain* something about that phenomenon. And that is what Chomskyan linguistics tries to do. What *I'll* try to do in this connection is to discuss to what extent the proposed descriptions explain what they set out to do.

With these points in mind, let me now comment on a few presuppositions of the requirement of descriptive adequacy.

- [2a] it presupposes the existence of a language, or notation, in terms of which the description can be stated
- [2b] it presupposes some degree of consensus about what linguistically relevant facts and data are
- [2c] it presupposes the existence of some kind of principle by which we may compare, and choose between, two distinct, but descriptively adequate, grammars

To illustrate [2a]: there is a world of difference between traditional grammatical terminology - as used by Jespersen, for example - and the terminology used by *Government and Binding Theory*. And there is a world of difference between the notational devices used by Lexical Functional Grammar and that used by the Standard Theory of Generative Grammar. Nevertheless, we presume, everyone is concerned to give what they consider a *true* description of the data.

To illustrate [2b]: there is a world of difference as to whether you allow ungrammatical sentences as evidence for conclusions - and hence as linguistic facts - or not. Strictly speaking, only positive evidence can be relevant to *description*. If you admit negative evidence it is because your motives go beyond description proper. This is why I shall admit negative evidence.

Finally, what [2c] says is that there may in fact be two competing grammars that describe the same relevant linguistic data. Therefore, there must be some way to choose between them. Chomsky has always appealed to the notion of *explanatory adequacy* to handle this point. What this means, in essence, is that linguistic theory should help us

decide between competing, descriptively adequate grammars, and at the same time *explain* something about language. Now, there may be all sorts of questions about language that we might want linguistic theory to *explain*. But to Chomsky, what needs explaining is the true nature of our language faculty, and how we get it. If no consensus is reached on this point, or if cognitive faculties are considered out of bounds to the goals of linguistics, then some other goal for linguistic theory must be stated. And here, for reasons already indicated, it is *not* enough just to appeal to descriptive fidelity to the linguistic ‘facts’ - which would be circular.

## 2. The data

Let me now present a list of data that a grammar of *seem* must take into account. Notice that I accept both positive and negative data, for the reason already given.

### [3] **A: Raising and Complementizer Type**

1. John seems a nice chap
2. John seems to be a nice chap
3. It seems (that) John is a nice a chap
4. \*It seems for John to be a nice chap
5. \*For John to be a nice chap seems

### **B: Complementation**

1. John seems (to be) a nice chap
2. John seems (to be) interesting
3. John seems to laugh
4. \*John seems laughing
5. John seems to be laughing
6. John seems (to be) in trouble
7. John seems (to be) far away

### **C: Definiteness Restriction**

1. \*John seems (to be) the nice chap
2. It seems that John is the nice chap
3. \*John seems like the nice chap

### **D: Extraposition and Expletion**

1. It seems (that) there is a discrepancy here
2. There seems to be a discrepancy here
3. \*There seems there is a discrepancy here
4. \*There seems a discrepancy here
5. \*It seems (to be) a discrepancy here
6. \*A discrepancy seems (to be) here

7. It seems (that) John is a nice chap
8. \*There seems that John is a nice chap
9. \*There seems John to be a nice chap
10. \*There seems John a nice chap
11. John seems (to be) a nice chap
12. \*John seems that there is a discrepancy here

**E: Tacit arguments**

1. It seems (to me) that John is a nice chap
2. ?It seems to Harry that John is a nice chap

**F: Comparison**

1. John seems like a nice chap
2. \*John seems like nice
3. \*John seems like to laugh/laughing
4. \*John seems like in trouble
5. \*John seems like far away

I have indicated the grammatical phenomenon that these blocks of examples are supposed to exemplify in the headlines. Thus the A-examples illustrate phenomena usually subsumed under the term *Raising*, the technical term for the view that a structural element has been ‘raised’ by some grammatical process from its ‘original’ position in a complement clause to a position in the matrix clause. The question of *complementizer* type is associated with this. Group B concerns the categorial type of phrasal complement; Group C illustrates some restrictions on the referential form of noun phrases in the complement of *seem*; Group D illustrates the two different kinds of *expletive* element *it* and *there* in *seem* clauses; the small Group E suggests that *seem* from a semantic point of view selects a non-thematic Experiencer role, which is obligatory in the sense that it will always refer to the (direct or indirect) speaker: sentence [E2] must be understood as a report of *Harry’s* utterance of sentence [E1]; and finally Group F displays the fact that *seem* may select the comparative preposition *like*, which only takes NP complements. Perhaps this case is covered already by block [B5].

This, I believe, is an exhaustive list of the grammatically relevant ‘facts’ about the occurrence of *seem* in the main stream linguistics literature. An informal, but descriptively adequate, account of the syntactic potential of *seem* might be:

- [4] *Descriptively adequate 'grammar' for the language fragment [3]:*
- it takes NP [3A1] or expletive *it/there* [3D1, 3D2] as subject
  - it takes a single complement, which may be either a *that*-clause [3D1], an infinitival VP [3B3], a NP [3B1], an AP [3B2], a PP [3B6], or an AVP [3B7]
  - if the subject is a NP, the complement cannot be a *that*-clause [3D12]
  - if the subject is expletive *there*, the complement must be an infinitival VP [3D2-4, 8-10]
  - if the subject is expletive *it*, the complement must be a *that*-clause [3D1,5]
  - if the complement is a NP, it must be indefinite [3C]
  - it selects an Experiencer role which - if expressed - must be in the form of a prepositional phrase headed by *to*; it will always refer to the speaker, directly or indirectly [3E]

Now, if this was all there was to it, we might just stop here and all go home. Indeed, to some people it *is* all there is to it - namely to those who believe that descriptive linguistics is just a matter of collecting a large enough body of data and then sit down and describe what you see, like butterflies. But consider what might happen if a diligent Danish student, who had learnt this grammar by heart, sat down and followed it slavishly when translating from Danish. We might find 'sentences' like those in [3']:

- [3'] B6. \*John seems about her (= John synes om hende)  
 B7. \*She seems beautifully (= Hun ser skøn ud)  
 D2. \*There seems to sit a man (= Der synes at sidde en mand)

None of these would be excluded as ungrammatical - but they are! The problem with the grammar is that it does not try to *generalize*.

### 3. Generality

There are two mutually conflicting concerns in descriptive linguistics. On the one hand, in the interest of generality we should concentrate on properties of sentences which can be subsumed under general principles. Relative to the data in [3], this would make us look for principles to account for the fact that, say, [3A1-3] always have the

same truthvalue. On the other hand, individual lexical items have idiosyncratic properties which would seem to require a non-general, particular description for each lexical item on its own. Linguistic analysis to a large extent is a matter of balancing these two concerns against one another, and always remember there is a trade-off between them.

In traditional descriptive grammar, the emphasis was on the second of these concerns to such an extent that classification of lexical items in terms of their grammatical potential was seen as the main goal of syntactic analysis. To illustrate, the only use that a modern traditional grammar makes of the syntactic properties of *seem* is to establish its classification.

In section 3.8 of Quirk et al. (1982), *seem to* is classified as a semi-auxiliary along with:

- [5a] *be certain to*  
*be (un)likely to*  
*appear to*  
*happen to*  
*turn out to*  
 etc.

There is another subclass of semi-auxiliaries:

- [5b] *be about to*  
*be apt to*  
*be bound to*  
*be going to*  
*be liable to*  
*be sure to*  
*be to*  
*had better/best*  
*have (got) to*  
*come to ('happen to')*  
*fail to*  
*get to*  
*tend to,*  
 etc.

They are all *semi-auxiliaries* because they are like lexical verbs in not admitting inversion and negation without *do*. They are nevertheless *auxiliaries* because they take a naked infinitive and admit passive of it while not allowing passive themselves. The distinction between the two

types of semi-auxiliaries is also based on syntactic criteria. While verbs of type [5a] admit cleft constructions, verbs of type [5b] do not; cf. [5c]:

- [5c] a. he seemed to like it  $\leftrightarrow$  it seemed that he liked it  
 b. he tended to like it  $\leftrightarrow$  \*it tended that he liked it

In section 12.32 of the same work, *seem* is classified as a so-called ‘current’ copula<sup>1</sup> - as opposed to the ‘resultant’ copulas<sup>2</sup> - and both are considered to have more ‘intrinsic’ meaning than the ‘colourless’ copula *be*. Since copulas in general are considered to form an ‘intensive’ relation, we may infer that *seem* is an intensive verb - that is, that it takes a subject predicate complement. And since predicative nominals respect the Definiteness Restriction, [3C1] is ungrammatical. In contrast, [3C2] contains *identitive* ‘be’ - establishing an identity relation between the referents of *John* and *the nice chap*.

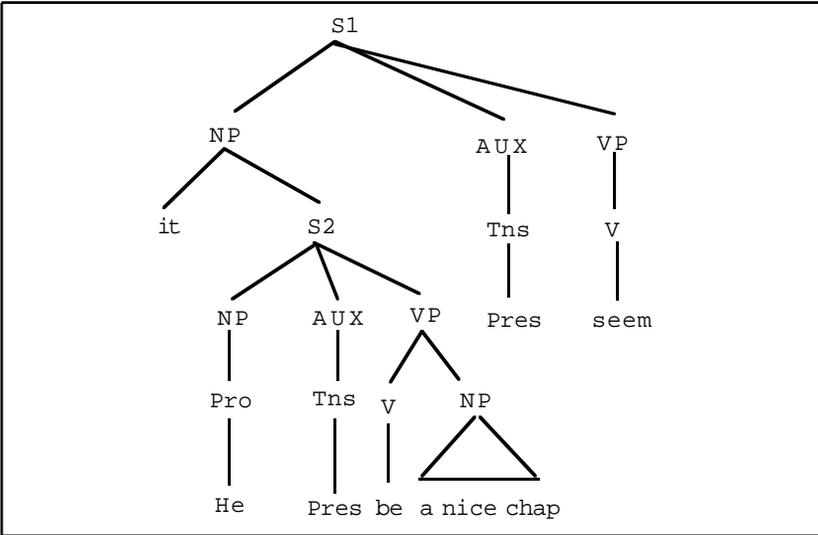
Classification as the primary goal of linguistics was rejected by Generative Grammar, though it was still important to establish syntactically motivated subcategories. Instead, *significant generalization* became the all-important catchphrase. Missing a significant generalization soon became the deadliest sin of a grammar, and transformational rules became the tool for capturing them. Lexical idiosyncracies still had to be taken into account, of course, but they were dismissed as uninteresting. It was ‘just’ a matter of listing them in the lexicon on a form appropriate for the lexical insertion rules to handle. An early lexical entry for *seem* looked like [6]:

- [6] Lexical entry for *seem* (Chomsky 1965:94)  
*seem*, [+V, + \_\_Adjective, + \_\_like-Predicate-Nominal]

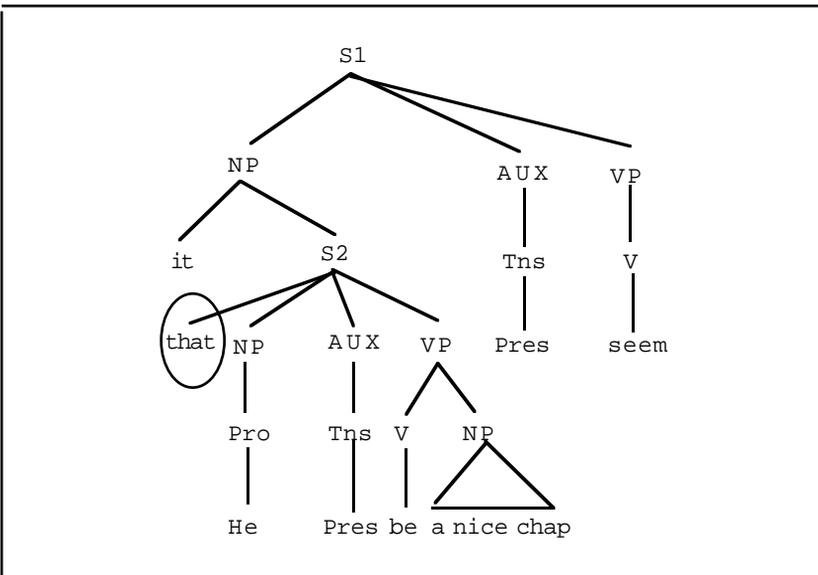
This entry states that *seem* is a verb [+V], and that it subcategorizes for - that is, takes as complements - either an Adjective or the word *like* followed by a nominal with predicative function. Thus it accounts immediately for examples [3B2] and all of [3F], but fails to account for any of the others in [3]. Why? Because the derivation of each of these is seen as a *general* matter of syntax, not a specific matter of lexical properties. The prime target for generalization were Raising and Expletives, and the price was an enormous battery of transformational rules, as you will see from the cartoon in [7a-e], which tells the derivational story of sentences [3A]:

<sup>1</sup> Along with *appear, feel, lie, look, remain, rest, smell, sound, stand, taste*.

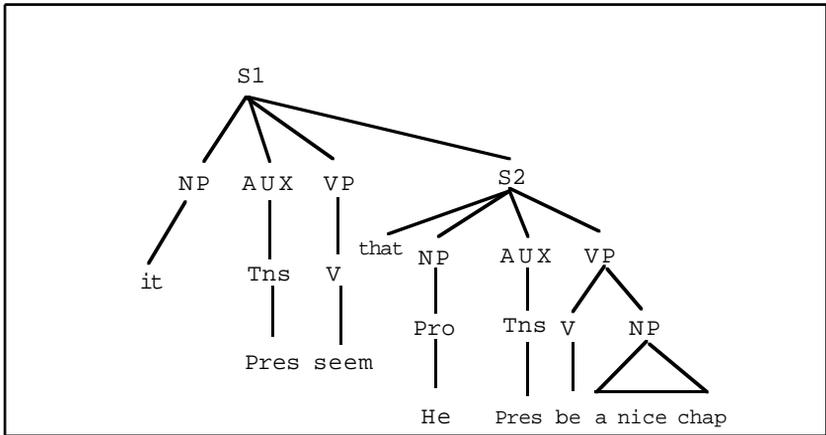
<sup>2</sup> *Become, come, get, go, grow, fall, run, turn*.



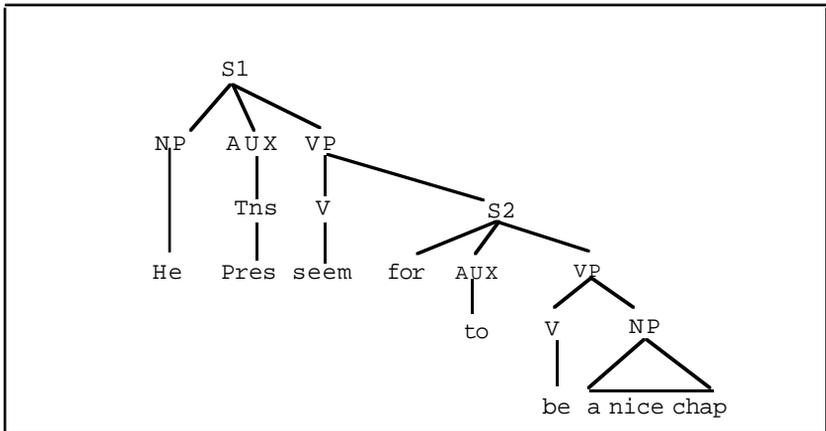
[7a] Deep structure, produced by Phrase Structure Rules and Lexical Rules



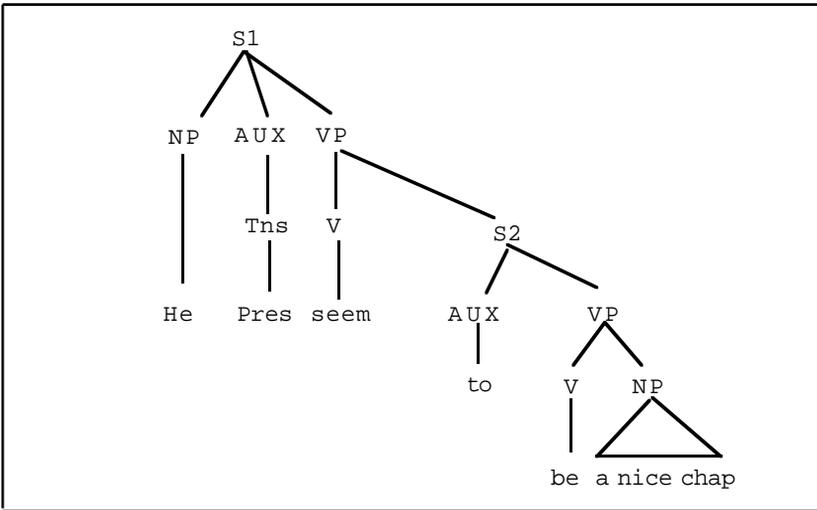
[7b] The tree resulting from application of *Complementizer Placement* (T-Rule) to [7a]



[7c] The tree resulting from application of *Extraposition* (T-rule) to [7b]



[7d] The tree resulting from selection of *for to* as Complementizer, followed by *Raising* (T-rule) to [7b]



[7e] The tree resulting from application of obligatory *Complementizer Deletion* (T- rule) to [7d]

I'll not go into detail over this, but only pinpoint a few salient features. Figure [7a] is a representation of what was called Deep Structure, created by so-called Phrase Structure Rules and Lexical Rules. The rest of the Figures are *derived* step by step from the previous one by a so-called Transformational Rule. So Transformational Rules, which contain very precise instructions, *rearrange* already existing structure by adding or deleting material, or by *moving* bits of structure from one place to another. This reflects one of the very few points of Chomsky's views of technical detail that has remained stable through time: his obsession with geometric arrangement. Indeed, the shaping principle for *every* version of the Standard Theory up until now has been the principle that things *move* in a structure - which therefore must be assumed to have fixed geometric properties in advance.

Now, this may seem farfetched, and highly abstract. But there is no denying that we here have the machinery that will be able to *explain* our intuition that [3A3] in some sense is 'the same sentence' as [3A2], which in its turn is 'felt' to be 'the same sentence' as [3A1]. The difference between [3A2] and [3A3] can be precisely located as a difference in the (random) selection of complementizer. If we select *for* instead of *that*, the T-rule of *Raising* is *not* barred from working as it

was with *that*. When it is brought to work - on a structure like [7b], except that it has *for* instead of *that* as complementizer - we get the result in [7d]. The complementizer *for* demands *Complementizer Deletion*, which will yield the derived structure [7e]. After the application of still further transformational rules not illustrated, we get the final derived tree underlying [3A2]. In order to get [3A1], we just have to apply a transformation - very controversial even at the time - of *Copula Deletion* to [7e]. But as you can see, there is certainly generality in all of this. It will account for the grammaticality judgements in the data sets [3A] to [F], except [C], which are handled by other general rules, and [E], which is simply ignored. But it has a price.

#### 4. The Lexicalist Controversy

That price, as already mentioned, was the enormous descriptive power of transformations. This is the key issue in what is now known as the Lexicalist controversy between the Generative Semanticists, or 'Generativists' as they have come to be called, on the one hand and the Interpretive Semanticists with Chomsky as general on the other. Since the crucial point of disagreement concerned the role of semantics in a grammar, the controversy had farreaching consequences for linguistic theory. And since Chomsky had introduced, and still accepted, the central role of transformational rules, he had in fact given his opponents a powerful tool for writing grammars that *described* the facts correctly. For there is no doubt that a grammar with transformational rules will be descriptively adequate, as we have just seen. But then again, there is no guarantee that there is any *other* motivation for a given T-rule than precisely this: that it *does* describe the facts correctly. And since a lot of different rules might do that, there is no guarantee that this kind of grammar explains *anything* about our mental linguistic faculty, as it was supposed to do.

But let me be a bit more concrete. Relative to our project, the Lexicalist Controversy didn't concern the analysis of *seem* directly. It concerned the question of how sentences with *other* types of verb should be analysed, given the analysis of *seem*.

The basic underlying structure of sentences with *seem*, according to [7], was held to be [8a]:

[8a] [[<sub>S</sub> he AUX be a nice guy] [seem]]

- which has a sentential subject. But suppose instead it was taken to be [8b]:

[8b] [[seem]<sub>S</sub> he AUX be a nice guy]]

- that is a structure with a sentential complement. This might bring sentences with *seem* into alignment with a lot of other sentences, for there are many verbs that take sentential complements. The question was, could there be given a description that would cover all these cases? [9] shows some relevant data:

- [9]
- a1. She persuaded him to become a spy
  - a2. \*She persuaded to become a spy
  - a3. \*She persuaded that he became a spy
  - b1. \*She assumed him to be a spy
  - b2. \*She assumed to be a spy
  - b3. She assumed that he was a spy
  - c1. She expected him to be a spy
  - c2. She expected to be a spy
  - c3. She expected that he was a spy
  - d1. She believed him to be a spy
  - d2. \*She believed to be a spy
  - d3. She believed that he was a spy

In each case we have a sentential complement. But was it the same kind of complement structure in all of [9]?

Chomsky said *no*. In addition, he assumed that *two* transformational rules, Raising and Equi-NP- Deletion, were necessary to cope with [9]. The syntactic differences were attributed to different deep structure properties, plus the sensitivity of individual verbs to reflexivization. Thus we get deep structures for the a-c examples like [10] (details omitted):

- [10]
- a. she persuaded [<sub>NP</sub> him<sub>i</sub>] [<sub>S</sub> he<sub>i</sub> become a spy]
  - b. she assumed [<sub>S</sub> he be a spy]
  - c. she expected [<sub>NP</sub> it [<sub>S</sub> he be a spy]]

But what about [9d]? If we derive it from a structure like [10a], we get in trouble over the different grammaticality status of the examples marked 3 in [9]. If we derive it from a structure like [10b], we get in trouble over the different grammaticality status of the examples marked 1 in [9]. This leaves structure [10c]. The difference between the examples marked 2 in [9], then, is to be accounted for by claiming that

*believe* requires reflexivization of a coreferential complement subject NP, while *expect* requires Equi-NP-Deletion.

In a spirited attack on the Extended Standard Theory, Paul Postal (1974) devoted a book of about 450 pages to the discussion of just one T-rule and its theoretical ramifications - the rule now known as *Raising*. Only this rule is required to account for the data in [8] and [9], he claims, and only one highly abstract kind of deep structure need be postulated, instead of the three already given in [10]. But *only* if we are prepared to give up such dogma as SVO order in deep structure and to introduce a number of abstract layers of lexical transformations. On these assumptions [9a1], for example, has the deep, or semantic, structure in [11].

- [11] CAUSE she [AGREE he [BECOME he a spy]] → she persuaded him to become a spy

The problem for linguistic theory that is highlighted by this controversy is that one cannot argue against either of these views from its failure to *describe* the facts correctly. In fact the Generativists gave some very persuasive and general, but highly abstract descriptions of a wide range of linguistic phenomena, and Postal, I believe, was the first linguist to realize that *seem* must have the obligatory Experiencer Role referred to in [3E].

Therefore, to argue against the Generativists, Chomsky could not argue from their failure to *describe* what they saw as the ‘relevant’ linguistic facts, but had to use other types of argument, concerning falsifiability, simplicity, generality and rationality. The trouble with this is that there was, and is, no generally accepted principle by which he could choose his own model rather than that of the Generativists - except his own notion of explanatory adequacy. And there neither was, nor is, a generally accepted view of what our language faculty might look like, and which might therefore provide him with a sound basis to his argument. This leads us to the next stage in the evolution.

## 5. Strong Lexicalism

The position defended by Chomsky against the Generativists was known as ‘lexicalism’. Its defining characteristic was the insistence that matters of lexical derivation were best regarded as properties of the lexicon and of individual lexical items, rather than of the syntactic



same functional structure, but differ in constituency structure (and lexicon, of course). To mention just one example of the added descriptive power of LFG over the Standard Theory, consider [13]:

- [13] a. In this cave dwells a large, ferocious beast  
 b. In this cave seems to dwell a large, ferocious beast

If subjects are *defined* as NP's, as they are in the Standard Theory, and if Raising is a rule that concerns NP's in subject position, then the Standard Theory may have difficulties with a general account of the relationship between [13a] and [b]. If on the other hand there is no Raising transformation, but only a functional specification to the effect that whatever is Subject of *seems* is also Subject of *to dwell*, there is no problem. And this is quite independent of the potentially controversial question whether the Subject in [13] is *In this cave* (as LFG will have it) or not.

This leads to the fate of the Raising transformation in LFG. Since there *are* no transformations, something else is required to handle the relationship between 'understood' subjects and non-finite complement structures. This is done in the theory of *Functional Control*. I said before that XCOMP was an *open* complement. This means, in effect, that it is a complement which needs something from 'outside', as it were, to be complete. This 'something' is a subject. Consider [9a1] and [9a2] again, repeated as [14a], and compare it to [14b]:

- [14] a1. She persuaded him to become a spy  
 a2. \*She persuaded to become a spy  
 b1. She promised him to become a spy  
 b2. She promised to become a spy

In each case, *to become a spy* is an XCOMP. In LFG jargon, it is functionally *controlled*, in the a-examples by the Object of *persuaded*, in the b-examples by the *Subject* of *promised*. This is a fancy way of saying that the logical subject of *to become* is *him* in a, but *she* in b. The difference emerges from the lexical entries for the two verbs, printed in bold type in [15]:

- [15] a. persuade V ( $\uparrow$  PRED) = 'persuade <( $\uparrow$  SUBJ) ( $\uparrow$  OBJ) ( $\uparrow$  XCOMP)>'  
 ( $\uparrow$  XCOMP SUBJ) = ( $\uparrow$  **OBJ**)  
 b. promise V ( $\uparrow$  PRED) = 'promise <( $\uparrow$  SUBJ) ( $\uparrow$  OBJ) ( $\uparrow$  XCOMP)>'  
 ( $\uparrow$  XCOMP SUBJ) = ( $\uparrow$  **SUBJ**)

The difference between [14a2] and [14b2] is now explained on the basis of a lexical rule that drops the Object and therefore associates the argument with no phonetic realization. Since only material Subjects and Objects are assumed to be controllers, but not empty arguments, this rule has removed the controller from [14a2], leaving the resulting functional structure incoherent.

Once again, we have a grammar which is descriptively correct, on some points even more so than the Standard Theory. What would be the arguments for choosing LFG over another grammar type? Well, according to the founding mother, LFG is meant as a *psychologically more revealing* theory than the Standard Theory. This is an interesting motivation, for it reflects the basis of Chomsky's version of explanatory adequacy, that linguistic theory should explain the cognitive capacity behind linguistic behaviour. But it *still* doesn't work, for we still need an independently motivated account of cognitive capacity, just as we did in Chomsky's case. All we can say now is that there are *two* linguistic theories that each purports to *explain* a cognitive capacity which we don't know much else about, but which both give rise to grammars that *describe* the facts correctly.

But there is another argument. Ron Kaplan's main contribution to the development of LFG was to make sure that it became a *computational* theory of language description, and he in fact wrote a very successful LFG parser. A parser is a computer program that will take natural language sentences as input and output the appropriate structures that describe them in LFG terms. Now, you may agree or disagree that writing a computer program that will parse natural language sentences is the *purpose* of linguistic theory. But you cannot deny the *transparency* of a theory that allows you to do this. In contrast, the Standard Theory was *opaque* in this respect; far too much information was lost in the maze of transformations for anyone to write but the simplest parsers for these grammars. So, to the extent that the so-called Computational Theory of Mind is a good model of cognitive capacity, LFG is a 'better' theory than the Standard Theory.

The strong lexicalist view of LFG has been carried over in such influential new directions as Generalized Phrase Structure Grammar (Gazdar *et al.* 1985) and Head Driven Phrase Structure Grammar (Pollard & Sag, 1992; 1995), probably the most serious rivals of the Standard Theories to date. What is more, LFG is also certain to have

influenced Chomsky in his revision of the Extended Standard Theory which ultimately led to the development of *Government & Binding*.

## 6. Trace Theory

I left Chomsky dangling between the Standard Theory and what is now known as the Extended Standard Theory, the next major development within Chomsky's family of grammars, which saw the introduction of *Trace theory* in *Reflections on Language* from 1976.

The implications of this move for syntactic theory were farreaching. To mention just two, it codified the notion of empty category - that is, syntactically significant elements without material realization in phonetic form, and it solved the problem of 'lost' and irretrievable information that had marred the Standard account. In our context it meant that the D-structure of [3A2] now looks like [16a]:

- [3A2] John seems to be a nice chap  
 [16] a. Y seems [<sub>S</sub> he AUX be a nice chap]

Chomsky explicitly attributes this analysis to Bresnan, the mother of LFG. A T-rule that he calls 'NP-preposing' maps [16a] onto [16b]:

- [16] b. he seems [<sub>S</sub> *e* AUX be a nice chap]

Here *e* is the trace left behind by, and bound by, *he*, which is moved up front by NP-preposing. Chomsky finds support for Trace theory in general, and for the so-called 'Specified Subject Condition' associated with it in particular, from a discussion of examples like [17]:

- [17] a. \*it seems to us that Bill likes each other  
 b. \*Tom seems to us to like each other  
 c. it seems to each of us that Bill likes the other(s)  
 d. Tom seems to each of us to like the other(s)

First of all, he claims, there is no *semantic* reason why the ideas expressed by the grammatical examples c and d should *not* be expressible by a and b; and in particular, why the b sentence should be ungrammatical. There is no *material* subject in the complement sentence *to like each other*, and the 'nearest' NP is *us*, otherwise a perfectly grammatical antecedent for *each other*, as shown in

- [18] they told us to love each other

The reason, he claims, is that although we cannot *see* it or *hear* it, there must nevertheless *be* a subject - and the only candidate is the trace left

after moving *Tom* up front. So *Tom* - or more precisely, its trace - is the Specified Subject, which blocks the association between *us* and *each other*. Since no one has ever explicitly taught us this, and since the relevant negative evidence is never forthcoming for us to react to among the sentences we hear, he concludes that both Trace theory and the Specified Subject Condition are empirically and independently sustained aspects of Universal Grammar.

And this is the true subject of *Government and Binding*, or *Principles and Parameters Grammar*, as it has now come to be called. I shall call it GB.

## 7. GB

Two interdependent features distinguish the development of GB from its predecessors. Firstly, where earlier versions of the Standard Theory had relied heavily on the notion of *rule* of grammar, divided into Phrase Structure Rules and Transformational Rules, GB rejects it, relying instead on the notion of *principle* of grammar - or rather, principles that every grammar of every human language must satisfy. In fact, there is only one 'rule' left - the last gasp of the previously all-powerful Transformational component, called Move- $\alpha$ . The production of a deep structure, previously the job of the Phrase Structure *rules*, is now guided by the *principles* of the so-called X-bar theory.

Secondly, and related to the first point, previous grammars had been organized into components by the two types of rule, the Base, or Phrase Structure Component, and the Transformational Component. Now, the grammar is organized into many more interdependent and cooperative components, each of which represents a particular aspect of syntactic structure. Each component can in fact be seen as a delimited theory of a particular aspect of syntactic structure. In addition to X-bar Theory already mentioned, they are: Trace Theory, which deals with the principles of movement of constituents, Case theory, which deals with the principles for the assignment of syntactic functions to constituents, X-bar-theory, which deals with the principles for assigning semantic functions to arguments, and The Theory of Binding and Control, which deals with the principles for syntactic interpretation of empty categories and anaphoric reference. However, since many of these principles are mutually dependent, the operations of the various components are not

sequentially ordered, but simultaneous. Then all or some of the principles of each component will act as a kind of filter, allowing some strings as grammatical, disallowing others.

Let us see how all this influences the description of sentences with *seem*.

The lexical entry for *seem* in GB is [19]

[19] *seem*, V, <S' ⇒ S>, (Proposition), assigns no case

It says that *seem* is a verb that subcategorizes for a complement S' prime, which may be 'replaced' by (simple) S, that it *s-selects* a single, internal argument that must be a Proposition, and that it does not assign Case to its complement. A radical change from earlier versions of the Standard Theory is that the *important* subcategorization criterion is now what is called *s* (for semantic) - selection of arguments, in this case the selection of *Proposition*.

By X-bar principles, we get a d-structure like [20a] for [3A2-3]:

[20] a. [<sub>NP</sub> ] INFL *seem* [<sub>S</sub> COMP [<sub>S</sub> John INFL [<sub>VP</sub> be a nice chap]]]

Here we have an empty 'landing site' for NP (the subject, which must be filled in English). INFL is the new label for what used to be called AUX, and COMP (for COMPLEMENTIZER) is the sister of S under the maximal projection S'.

Now, since the subject NP node *must* be filled, we have two possibilities, depending on the contents of COMP. If it is *that*, nothing happens except that the upper INFL comes to assign nominative case to an empty NP - which it cannot. Therefore, the expletive *it* is enforced, as the carrier of case. This will give us the s-structure [20b]:

b. *it* seems [<sub>S</sub> *that* [<sub>S</sub> John is a nice chap]]

- which corresponds to the structure of sentence [3A3].

If on the other hand COMP is occupied by *for*, we get the 'pruning' of S' to S, suggested in [19], which opens another possibility. *Seem* is an intransitive verb - and intransitive verbs are assumed not to assign case. However, according to the so-called Case Filter, every phonetically realized NP must be assigned case. *John* in [20a] is phonetically realized, so it must receive case. It cannot get it from *seem*, and it cannot get it from the subordinate INFL, for reasons I won't go into.

Therefore, it must *move*, to a position where it can be assigned case - an instance of Move- $\alpha$  - so we get an s-structure like [20c]:

c. John<sub>i</sub> seems [<sub>S</sub> e<sub>i</sub> to be a nice chap]

- corresponding to [3A2].

I have omitted a lot of detail, but even without that I'm sure you find this fairly tough going. The drive towards Universal Grammar is a drive towards explanatory adequacy at the expense of descriptive perspicuity. Yet the explanatory value of GB is no greater than it was for the rule-based model. Of course it is less of a commitment to a special mental architecture to say that it is described in terms of a set of general *principles* than it would be to say that it is described by a set of specific *rules*. But this does not alter the fact that we still need independent evidence for it.

GB is also a drive towards a much higher degree of abstraction, at least as high as that of the Generativists, which was discarded by Chomsky as 'too abstract'! I suppose this is why it only represents the last step but one in the evolutionary history of the Standard Theory. For the last step - the so-called Minimalist Program, which is currently being developed - is precisely an attempt to minimize abstractness in favour of 'naturalness'. Let us look very briefly at some of the main points of the Minimalist Program.

## 8. The Minimalist Program

The idea that our language faculty is a *computational* system was explicitly introduced in Chomsky's last largescale contribution to GB, *Knowledge of Language*, from 1986; but it is only with the Minimalist Program that the *consequences* of this idea have come to be seen as general constraints on the *directions* in which principles of language structure might be looked for and formulated. The main concern of the Minimalist Program, which began to take form in 1992, and which is further developed in Chomsky (1995) is to *reduce* the machinery established in *Government and Binding*. The guide for this reduction should be considerations of what it might take to *compute* a sentence from a random choice of lexical items, given a number of assumptions of 'naturalness' like 'Least Effort', 'Greed', and 'Procrastinate'. These principles are all explained in relation to computation. Thus, maximal

exploitation of the principle of ‘Least Effort’ is to compute nothing, which would just leave the lexical choices in some random order. Therefore *some* effort is required, and the job is to determine *how* much, and how it is *applied*. However, since *reduction* of descriptive complexity does not in itself entail reorientation of descriptive practice, the Minimalist Program will bring us no closer to an understanding of our linguistic capacity than its predecessors.

But it does bring us back to the beginning. Chomsky explicitly wants the Minimalist Programme to determine how ‘perfect’ language is. Now, this is a totally new concept in Chomskyan linguistics. In fact, it only makes sense to ask for degrees of perfection of anything in relation to some preestablished norm. The preestablished norm, as we might expect in this case, is our current understanding of general cognitive processes, and if it turns out that our language faculty is wholly explainable in terms of these, it is a ‘perfect’ system. As I have tried to stress throughout, we are still a long way away from knowing enough about such processes to conclude one way or the other. But the very fact that Chomsky now seems prepared to entertain the possibility that our language faculty is not ‘perfect’ in this sense, suggests that there may be *other* factors involved. And although he doesn’t actually *say* so, these factors may very well be functional, and we are back with the original debate between Chomsky and Searle, so it may be time to close down the system.

## 9. Conclusion

It has been my concern in this lecture, partly to trace a line of linguistic development, partly to point out how various grammars might seek to describe the same set of linguistic facts. It turned out that all of them could be brought to describe the facts ‘correctly’, even though the assumptions behind them differed significantly. It has therefore been posed as a requirement on linguistics to come up with a principled way to determine which is the ‘better’ grammar, and it has been shown that Chomsky’s principle of explanatory adequacy will not do the trick, essentially because we lack independent knowledge of the capacities he seeks to explain. So I agree with Gerald Gazdar - one of the founders of GPSG and one of the most informed and coherent critics of Chomsky’s - when he says:

If linguistics is truly a branch of psychology (or even biology), it is so far the branch with the greatest pretensions and the fewest reliable results.

[Gazdar et al., 1985:5]

These are all fairly negative conclusions. Isn't there anything *positive* to conclude from this? Yes, there is.

I think that the reformulation of the issue between Searle and Chomsky which I gave at the beginning deserves some consideration. I think, in particular, that further explicitation of the *purpose* of linguistic description will show that there are other interesting standards of explanation than that set by Chomsky. To mention but one example, if the purpose of your description is to reveal the referential properties of language, then you are likely to need different categories and other relationships than if your purpose of description is to make English syntax understandable to second language learners. I am *not* hereby saying that description or theory is unimportant. I am saying that you have to think about what you need them *for*. This might even give some substance to the much-discussed distinction between *English* and *English for special purposes*, a distinction which, I'm afraid, I still find somewhat elusive. And *that*, I think, is not a bad conclusion.

## References

- Chomsky, Noam (1965): *Aspects of the Theory of Syntax*. Cambr. Mass.: MIT 1965.
- Chomsky, Noam (1976): *Reflections on Language*. London, Temple Smith 1976.
- Chomsky, Noam (1986): *Knowledge of Language*. New York: Praeger Books 1986.
- Chomsky, Noam (1995): "Bare Phrase Structure" in: Gert Webelhuth (ed) *Government and Binding Theory and the Minimalist Program. Principles and Parameters in Syntactic Theory*. Oxford: Basil Blackwell 1995, pp. 383-440.
- Gazdar, Gerald et al. (1985): *Generalized Phrase Structure Grammar*. Oxford: Basil Blackwell 1985.
- Pollard, Carl & Ivan Sag (1992): *Information-Based Syntax and Semantics*. Palo Alto: CSLI 1992.
- Pollard, Carl & Ivan Sag (1995): *Head Driven Phrase Structure Grammar*. Chicago: Univ. of Chicago Press 1995.
- Postal, Paul (1974): *On Raising. One Rule of English Grammar and its Theoretical Implications*. Cambr. Mass.: MIT 1974.
- Quirk, Randolph et al. (1982): *A Grammar of Contemporary English*. London: Longman 1982.

