

## The Cognitive Basis of Grammar

Peter J. Binkert

For more than 50 years, research on language has been dominated by the thinking of one man, Noam Chomsky. In the early 1950s, Chomsky revolutionized the study of language and redefined the goals of linguistics. He made a fundamental distinction between linguistic competence and linguistic performance, which he defined as (1).

- (1) a. LINGUISTIC COMPETENCE - the unconscious knowledge which native speakers have of their language.
- b. LINGUISTIC PERFORMANCE - the use and acquisition of this knowledge.

What Chomsky meant by this distinction was the following. Native speakers unconsciously know the grammar of the language they speak; for example, they unconsciously know rules such as the rule in (2), a verb must agree with its subject in number. This knowledge underlies the fact that speakers of standard American English consider (2a) and (2b) perfectly grammatical, but reject (2c) and (2d) as ungrammatical.

- (2) A verb must agree with the subject in number (\* means 'ungrammatical').
  - a. He plays soccer.
  - b. They play soccer.
  - c. \*He play soccer.
  - d. \*They plays soccer.

Under any given set of circumstances, during actual performance, a speaker might violate this rule and actually say a sentence like (2c); but, if specifically asked, the speaker will say that (2c) is ungrammatical.

Chomsky also asserted that the linguistic competence which native speakers share is the result of an innate faculty of language that all human beings possess. Children learn their native language with remarkable speed, uniformity, and predictability, regardless of the external circumstances. Chomsky's explanation for this has been that humans possess a special ability to acquire a human language. One of the goals of linguistic theory is to discover the nature of this innate faculty of language.

- (3) "The nature of this faculty is the subject matter of a general theory of linguistic structure that aims to discover the framework of principles and elements common to attainable human languages; this theory is now often called 'universal grammar' (UG)... UG may be regarded as a characterization of the genetically determined language faculty." (Chomsky 1986, p.3)

In other words, linguists analyze the various languages of the world in an attempt to discover general linguistic principles applicable to those languages. Such principles are then considered to be part of UG or the innate language faculty.

As an example of such principles consider (4).

- (4) a. The sentences of all human languages consist of phrases.  
 b. Every phrase has a head, which is its focal element.  
 c. The rules of language refer to phrases.

These principles operate in (5), where the paired brackets separate the phrases.

- (5) a. Ed met the intelligent history teachers.  
 [ [Ed] [met [the intelligent history teachers] ] ]  
 b. Ed met the intelligent history teachers, and Al met them too.  
 (*them* = [the intelligent history teachers])

Notice that the word *them* in (5b) refers to the phrase intelligent history teachers. Thus, pronouns refer to phrases, specifically to phrases with noun heads, or noun phrases. Further, pronouns must agree with the gender and number of the head of the phrase as (6) reveals.

- (6) a. Ed met that woman from Italy, and Al met her too.  
 [NP that woman from Italy] head is *woman*  
 b. Ed met the father of those six boys, and Al met him too.  
 [NP the father of those six boys] head is *father*

Native speakers of English clearly know the principles in (4), after all we all understand how to use pronouns correctly. What I have done in these examples is characterize that knowledge in the form of principles. It turns out that pronominalization works in essentially the same way in other languages. Thus, we might say that (4) is part of UG.

This evening, I would like to reinvestigate some of Chomsky's claims and explore the alternative idea stated in (7).

- (7) Principles like (4) reflect general information processing abilities that humans develop during maturation, rather than an innate language faculty per se.

Specifically, I would like to explore the idea that much of what linguists ascribe to an innate language faculty is reflective of general mental abilities of the type that psychologists have studied for more than a century.

I will explore the idea state in (7) by examining some of the rules of English grammar. Now, I realize that a discussion of grammar is not the sort of thing that the average person looks forward to with wild anticipation. When most people hear the word *grammar*, they groan. They have unfond memories of English class where they had to memorize scores of rules which appeared to be the inventions of the teacher at the front of the room. Frustratingly, in successive years of diagramming sentences, no two teachers ever seemed to have the same set of rules.

The unfortunate result has been that generations of students think that grammar is whimsical, tedious, and completely expendable. In fact it is none of these. Rather, as we will see, the grammar of language is rooted in the cognitive abilities and limitations of human beings. Thus, there is a great deal that we can learn about human mental processes by examining the kinds of grammatical rules that occur in language. For example, consider, again, the rule in (2), a verb must agree with the subject in number.

What this rule says is that the verb must be *plays* if the subject is singular (*he*), and *play* if the subject is plural (*they*).

Why does English have this rule? Surely, it is not something that some teacher invented one day. It is not in the language to make students miserable. Nor is it in the language because some fussy group of individuals thought it would be nice to have. To properly answer the question, we must first note that the rule (2) is stated incorrectly in two ways.

First, agreement between a verb and a subject in English is overtly expressed only in the present tense. Notice (8).

- (8)     a.    He played soccer.  
          b.    They played soccer.

These examples raise an immediate question. If there is no overt agreement between the subject and the verb in the past tense, why do we need it in the present tense? What's wrong with (9a)?

- (9)     a.    \*He play soccer.  
          b.    They play soccer.

On the surface, the ungrammaticality of (9a) seems to be nothing more than a peculiar aspect of English, that kind of thing that is frustrating and bothersome to students.

The second thing wrong with rule (2) is that it is not explicit. It should be stated as (10).

- (10)    The verb must agree with the head of the subject phrase.

Notice the sentences in (11).

- (11)    a.    The boys on that block play soccer.  
          b.    \*The boys on that block plays soccer.

The verb in (11) must agree with *boys* not *block*. No matter how long or complicated a subject phrase is, the verb must agree with its head. Consider (12).

- (12) The six year old boys on that block whose parents think it would be good for them to be involved in a team sport *play* soccer.

Notice in (12) that the verb *play* agrees with the subject *boys*, although there are many other nouns in between. This reveals that a speaker must be able to locate the head of the subject phrase to correctly execute the agreement rule. This is not a trivial matter. Consider (13), where the underlined word is the main verb.

- (13) a. The school shows *play* here.  
 b. The school show *plays* here.  
 c. The school *shows* plays here.  
 d. The schools *show* plays here.

We are now in a position to say something interesting. If we did not have agreement in the present tense, the whole structure of the English sentence would disintegrate. Notice what happens if we remove each instance of final *s* in (13). We get (14).

- (14) a. The school show play here.  
 b. The school show play here.  
 c. The school show play here.  
 d. The school show play here.

Conversely, if we put an *s* on the end of all the nouns and verbs, we get (15).

- (15) a. The schools shows plays here.  
 b. The schools shows plays here.  
 c. The schools shows plays here.  
 d. The schools shows plays here.

Clearly, given a language that allowed sentences like (14) and (15), we would not know what we are talking about. Thus, agreement is not the result of whim. It is not the invention of sadistic teachers. It is in the language by necessity.

The basic problem is that, in English, the singular and plural of nouns is identical to the singular and plural of verbs: *play* is either a singular noun or a plural present as in (16); *plays* is either a plural noun or a singular present as in (17).

- (16) a. The play was good. (*play* is a noun)  
 b. They play soccer. (*play* is a verb)

- (17) a. The plays were good. (*plays* is a noun)  
 b. He plays soccer. (*plays* is a verb)

Notice that the problem disappears if we put the verbs in the past tense.

- (18) a. The school shows *played* here.  
 b. The school show *played* here.  
 c. The school *showed* plays here.  
 d. The schools *showed* plays here.

The reason we no longer have a problem is that there are no nouns which have the same form as a regular past tense, that is, there is no noun like *played* or *showed*. These forms are unambiguously verbs. We now have an explanation for (19).

- (19) During the history of the English language, the past tense has become regular for both singular and plural (*he/they played*), but the present tense has not (*he plays/they play*).

In fact, unless there are dramatic changes elsewhere in the grammar of English, the present tense will never become regular. Again, it is important to realize that the agreement rule in English is not the result of convention or quirkiness. Without the rule, the basic structure of the sentence would fall apart.

The effect of the agreement rule in English is to separate the English sentence into phrases. In fact, many of the constraints on language exist for just this purpose, to organize the words into phrases, which reflects the principles in UG that I described in (4). We can therefore ask why it is so important to organize words into units. I believe that this derives from the severe memory limitations that human's have.

It is a well-known fact that human beings have very limited SHORT TERM MEMORY (STM), which is the immediate memory for new information, such as a telephone number just received from the operator. This contrasts with our virtually limitless LONG TERM MEMORY (LTM), which is the storage and recall of information that is not new, such as one's mother's maiden name.

- (20) a. SHORT TERM MEMORY (STM), which is the immediate memory for new information, such as a telephone number just received from the operator.  
 b. LONG TERM MEMORY (LTM), which is the storage and recall of information that is not new, such as one's mother's maiden name.

- (21) STM is thought to be limited to about seven ( $\pm$  two) bits of information (Miller 1956; see also Graf & Torrey 1966, Mandler 1967, Jarvella 1971, Johnson 1970, 1971). Thus, the lower end of human STM is about five.

Crucially, the amount of material retrievable from STM can be increased if the material is organized in some way, that is, we have (22).

- (22) STM can be facilitated by organizing information into units called CHUNKS (Klatzky 1980, Anderson 1985).

(to use the technical psychological term)

Given the reality of the chunk in psychological studies and the phrase in linguistic studies, it seems plausible to derive (23).

- (23) The sentences of a natural language consist of phrases because human STM limitations require organizing a string of words into chunks to facilitate processing.

Psychologists have demonstrated that a subject's recall of arbitrary material, linguistic or otherwise, can be facilitated substantially by chunking. The exact capacity of STM is a subject of debate, that we need not concern ourselves with tonight. Some researchers place it lower than five. Whatever the exact capacity is, the important point is that STM is severely limited.

Human beings do not do well with unstructured material. (24) is much harder to recall than (25).

- (24) rehearsal during pit orchestra the into directly fell France of south the in town small a from soprano young a

- (25) A young soprano from a small town in the south of France fell directly into the orchestra pit during rehearsal.

In short, there is good reason to believe that some constraints like the rule of agreement are motivated by limitations on man's cognitive capacity. It is worthwhile to explore this in some detail. Consider again the matter of locating the head of the subject phrase in a sentence like (26).

- (26) The six year old *boys* on that block whose parents think it would be good for them to be involved in a team sport *play* soccer.

Somehow, speakers must understand that the head of the subject phrase, which begins with the first word of the sentence and goes until the word *sport*, is *boys*. Thus, out of all the words that precede *play*, the speaker must select *boys* as the head of the phrase and make the verb *play* agree with it. How do speakers locate the head?

I became aware of the magnitude of this problem when I wrote my first computer parsing program. I was trying to teach a computer to understand English, and am still trying to do so. A program which does this is often called a parser. As I began to deal with the problem, I noticed that English had a number of constraints which appeared whimsical, tedious, and completely expendable, not to mention, altogether irritating.

For example, I noticed that phrases with a noun head, or noun phrases, sometimes require a determiner (a word like *the*) and sometimes don't. Consider (27).

- (27) Noun phrases (phrases with a noun as their head or focal element) sometimes require an overt determiner (*the, a, etc.*) and sometimes don't:
- a. He is carrying the/a book.
  - b. He is carrying those books.
  - c. He is carrying books.
  - d. \*He is carrying book.

Why can't we say (27d)? Is this another one of those silly idiosyncrasies that English has? Or is there some specific reason why (27d) is no good? It did not occur to me right away what the reason is, but I did formulate the constraint (28).

- (28) If the head of a noun phrase is a singular count noun, e.g., a noun like *book*, then the noun phrase must have an overt determiner (a word like *the* or *a*); if the head is plural (*books*), an overt determiner need not be present.

This constraint did not do the whole job for the parser. Shortly, I discovered another constraint (29).

- (29) Between the determiner and the head noun of a noun phrase, there can be no other noun that is plural.

This constraint is necessary to account for the data in (30) and (31).

- (30)
- a. She is the teacher of history.
  - b. She is the history teacher.
  - c. She is the teacher of languages.
  - d. \*She is the languages teacher.
- (31)
- a. He submitted an abstract of 500 words. (*words* is plural)
  - b. \*He submitted an abstract of 500 word. (*word* is singular)
  - c. He submitted a 500 word abstract. (*word* is singular)
  - d. \*He submitted a 500 words abstract. (*words* is plural)

Again, why does English have the constraint (29); why aren't the starred sentences in (30) and (31) grammatical? Their ungrammaticality appears completely arbitrary.

After some more work, I discovered the constraint in (32).

- (32) Between the determiner and the head noun of a noun phrase, no other determiner can occur.

This is borne out in (33).

- (33) a. A teacher of this language is hard to find.  
       \*A this language teacher is hard to find.  
       A language teacher is hard to find.
- b. A teacher that clever is hard to find.  
       \*A that clever teacher is hard to find.  
       A clever teacher is hard to find.

What is the reason for all of these constraints?

Well, if you think about it for a minute, you will see that they help speakers to locate the head (they also help frustrated linguists to write parsing programs that work). Remember, that to make the verb agree with the subject, we must find the head of the subject phrase; specifically, we must find out if the head is singular or plural. Given the constraints I have just outlined, that's fairly easy. We have the generalizations in (34).

- (34) a. If the head is singular, then an overt determiner must cue the noun phrase; \**he is carrying book* (27d) is ungrammatical.
- b. If the head is plural, then the determiner need not be overt; however, we still know what the head is because the only allowable plural in a sequence of nouns is the head itself. \**He submitted a 500 words abstract* (31d) is ungrammatical.
- c. There can only be one determiner before the head. \**A this language teacher is hard to find* (33) is ungrammatical.

One can see from the generalizations in (34), that the determiner plays a very important role in English syntax. It acts as a cueing device. In effect, it signals that a particular type of structure occurs, namely, a noun phrase. As soon as a speaker hears a word like *the*, he knows that a noun phrase is coming. Further, since only one determiner is possible before the head, the speaker has a reliable cue that helps to organize or chunk the phrase.

Psychologists have known for many years that a subject's recall of sentences is facilitated by the presence of such cues, not only cues for noun phrases, but also cues for sentences. For example, consider (35) and (36).

- (35) a. I know that he saw me.  
       b. I know he saw me.
- (36) a. The man that I saw left.  
       b. The man I saw left.

Notice that the word *that* can be left out in (35) and (36). In these two examples, the omission of *that* does not interfere with comprehension. In other examples, when *that* is left out, sentences become difficult to understand. Consider (37).

- (37) a. I know the student teachers praise progress.  
b. I know the student teachers praise progresses.

Compare (37) with (38), which includes the cue *that*.

- (38) a. I know that the student teachers praise progress.  
b. I know that the student that teachers praise progresses.

It is well known that subjects have much better access to information when there is an overt cue to structure than when there is no overt cue. The presence of an overt cue facilitates the identification of a particular structure. Speakers can organize and chunk the sentence better.

Again, we find the opposite of this in so-called garden path sentences which omit the cue.

- (39) Garden path sentences are notoriously difficult to process because they result in a miscue of what the units are, that is, having noticed grammatical markers and categorized a construction accordingly, one continues only to find that the categorization was wrong:
- a. The horse raced past the barn fell.  
b. The editor played the tape agreed the story was a big one.

The sentences in (39) can be paraphrased as (40).

- (40) a. The horse that was raced past the barn fell.  
b. The editor that played the tape agreed that the story was a big one.

Thus, the presence of the cue facilitates comprehension. In fact, in certain instances, the cue can never be left out, or comprehension is completely disrupted. Consider (41) and (42).

- (41) a. That he saw me is clear.  
b. \* He saw me is clear.
- (42) a. The man that saw me left.  
b. \*The man saw me left.

Leaving out *that* in (41) or (42) results in a miscue; specifically, speakers think that the main clause is *he saw me* in (41b) and *the man saw me* in (42b), both of which are incorrect. Thus, the noun phrase constraints that I have described seem specifically designed to help speakers organize phrases into units and locate the head. Equally importantly, these constraints tell us why some sentences must be ungrammatical and others grammatical, when there seems no good reason. Consider (43),

where the heads are underlined.

- (43) a. He needs the four foot long *pipe/pipes*.  
 b. \*He needs the four feet long *pipe/pipes*.

Despite the fact that *four* is clearly more than one, we cannot say *four feet* in (43b), we must say *four foot* in (43a). We now know the reason: the only noun that can be plural between the determiner and the head is the head. Remember what happened when we added an *s* to all the nouns before in the examples with *show* and *play*.

Summing up thus far, we have seen that a variety of constraints on phrase structure in English seem to exist to help speakers organize phrases into appropriate units. This in turn appears necessary because of the severe limitations humans have on STM capacity.

Further exploration of the English noun phrase reveals other constraints of the same type. It is possible in English to produce truly gargantuan noun phrases like (44).

- (44) some of those twelve illegal million dollar transatlantic *shipments* of drugs into the NYC of the Mafia's during June

Interestingly, English has two constraints that help us locate the head in such large phrases. The first is (45).

- (45) Any item occurring to the left of the determiner must be separated from it by the word *of*.  
 a. A number of the systems *are* complicated.  
 b. Some of those shipments *were* stopped.

The second constraint is (46).

- (46) If no overt marker is present separating a head noun from its complement, e.g., *against* in *a vote against gun control*, then *of* must be inserted to signal the right boundary of the head.  
 a. The student letters *are* here.  
     The student of letters *is* here.  
 b. His love stories *are* charming.  
     His love of stories *is* charming.

If we represent the above constraints diagrammatically, we have the structure (47).

(47) N3  
 +))))0))))))3))))),  
 ...of.....DET.....N0.....of...  
 .))))0))))-  
 8 8 X 8 8  
 anchor cue head anchor

some of those shipments of drugs

The area marked “X” in (47), the nucleus, is the area in which three of our constraints operate. In the nucleus, there can be only one determiner and only singular nouns. Any violation of these constraints makes it impossible to locate the head unambiguously as we have seen. Further, the word *of* acts like anchor points (Anderson 1985) separating the nucleus from the remaining elements of the noun phrase. We therefore have a variety of cues to help us.

In short, the constraints appear to result from cognitive limitations. Man is not clairvoyant: cognitive limitations require structural principles for locating the heads of phrases in languages like English. The noun phrase constraints I have mentioned are such principles. In short, what appear to be arbitrary constraints on the internal structure of noun phrases are, in fact, well motivated.

If we examine the development of language in children, we see that their familiarity with these constraints develops over time. The data in (48) are sentences taken from a journal I kept many years ago of my son Kevin's development. Notice the progress during the one year span given.

- (48) a. 20-21 months.
- Truck broke.  
Daddy shoe. (=This is daddy's shoe.)
- b. 22-23 months.
- More turkey. (=I want some more turkey.)  
Some coffee. (=Give me some coffee.)  
Many cameras. (=There are many cameras.)
- c. 24-25 months.
- Look at the cough drop.  
Look at this many driver.  
Look at this so many cars.  
Look at this many car.  
Look at this big lollipop.  
Look at this broken door.  
Look that many vitamins.

## d. 26-27 months.

Kevin take big bubble bath.  
 [This is] mommy's nurse's hat.  
 That [is] Kevin's racecar book.  
 Kevin ride that big wheel.  
 Open that [can of] nuts.

## e. 28-29 months.

Put on my other sock, daddy.  
 I have some many. (=I want many vitamins.)  
 I have some eggs. (=I want some eggs.)  
 I go to sit [on the] big chair.  
 This is my hat.  
 I have some more pickles. (=I want some more pickles.)  
 I'm gonna put my sexy pants on.  
 I have to have a tuna fish sandwich.

## f. 30-32 months.

These man are green.  
 There's such a mess in here. I'm such a messer.  
 Them toys fell down.  
 That's a green dump truck.  
 I wanna sing about the bird *that* flies in the sky.  
 (first relative clause)  
 I want to hear the song about the bird.  
 (first posthead preposition phrase modifier)  
 I'm gonna be a steam shovel man when I grow up.  
 I think *the men are* not her. *They* all done. The  
*mens* are done.

One can see from this journal how Kevin's understanding of the constraints on the internal structure of the English noun phrase progressed. This development is in accord with general linguistic growth among all children. It progresses in stages which seem to be motivated by developing cognitive capacity including increases in STM capacity.

Currently, the prevailing opinion in psycholinguistics is that there is little, if any, evidence supporting the idea that STM capacity per se increases over time, that is, the capacity itself may be fixed at birth.

However, studies consistently show that a child's information processing abilities increase with age: "...older children and adults can keep more information in mind and perform mental operations more

rapidly than younger children can" (Shaffer 1989:335). However, the magnitude of the difference varies depending on the task at hand and the ages compared:

(49) Span Studies without regard to "chunkability":

a. Digit span (Dempster 1981, p. 66; 1985, p.220):

Age 2	About 2.25 digits
Age 4	About 3.75 digits
Age 6	About 4.00 digits
Age 8	About 4.25 digits
Age 10	About 4.75 digits
Age 12	About 6.00 digits
Adults	About 7.00 digits

b. Word span (Dempster 1981, p. 67):

Age 2	About 3.00 words
Age 4	About 3.50 words
Age 6	About 4.00 words
Age 8	About 4.25 words
Age 10	About 4.75 words
Age 12	About 5.00 words
Adults	About 5.25 words

c. Improvement in forward digit span (recall in the same order as presentation) between ages 7 and 13 is little more than 1.50 digits (Dempster 1981), whereas improvement in backward digit span (recall in reverse order of presentation) between ages 7 and 13 is about 3.00 digits (Jensen & Osborne 1979).

d. In studies using material that is difficult to chunk, the increases were extremely small (Dempster 1978; Hess & Radtke 1981; Ross 1969).

Thus, more is at stake than capacity per se.

(50) The STM "workspace" may increase because older children can manipulate items better, that is, have better informational processing skills, including better skills for the identification of items, the association of items with each other, the encoding of ordering information, the development of rehearsal strategies, as well as better integration of STM elements with general knowledge from LTM and greater experience with memory tasks. Olson (1973:153) sums up the matter well: "The changes [in STM capacity] are associated with the child's ability to recode or encode, to plan and monitor, to integrate and unitize. Broad limits of information processing capacity, which may be biological in origin, are relatively constant, but how the child operates within these limits undergoes systematic and

profound development."

The breakdown of such cognitive capacity is very much in evidence in agrammatism in adults who have suffered some trauma to the brain.

(51) "Agrammatism generally refers to the tendency...to omit function words... Among the class of items designated as 'function words' are prepositions and determiners. (Kean 1982, p. 241)

As I have indicated in this talk, function words like *the* and *of* are essential to sophisticated adult syntactic competence. Recall (47). Since such words act as cueing devices and anchors, their loss destroys the integrity of the phrasal unit. So-called content words may have more semantic import, but function words are essential to syntactic organization. In many cases of agrammatism, in fact, patients completely omit function words; therefore, sentence structure completely breaks down.

The hypothesis, therefore, is that sophisticated syntax requires full adult STM capacity, a prerequisite to mastering the complete set of syntactic constraints operative in a given language, as well as the dependency and agreement relations that exist between phrasal elements. Full STM capacity includes more than just the ability to deal with seven ( $\pm$  two) chunks of information. It requires, in addition, information processing skills for identifying, associating, ordering, and integrating those units into hierarchical dependencies.

Generally, in linguistics, we justify analyses on the basis of linguistic competence alone. This means, to quote Chomsky again (52).

(52) "Linguistic theory is concerned with an ideal speaker-listener, in a completely homogeneous speech community, who knows its language perfectly and is unaffected by such *grammatically irrelevant conditions as memory limitations...*" (Chomsky 1965, p. 3; italics mine)

Much more is known today about the nature of the human language apparatus than was known in 1957, the year of Chomsky's *Syntactic Structures* and Skinner's *Verbal Behavior*. For example, a number of psychological processes can be directly connected to human linguistic abilities; in particular, man's ability to deal with unstructured expressions is very limited. In a language like English, the structuring processes which are important include chunking or hierarchical ordering, linear ordering, cueing, and anchoring.

(53) Important structuring processes.

- a. Chunking or Hierarchical Ordering (Miller 1956; Graf & Torrey 1966; Mandler 1967; Jarvella 1971; Johnson 1970, 1971; Caplan 1972).

Subjects have better recall of material that can be organized into subunits.

- b. Linear Ordering (Anderson 1985).

Subjects identify strings of elements faster when those strings are ordered closest to a learned string.

- c. Cueing (Tulving & Psotka 1971, Hakes & Foss 1970).

Certain cue words like *that*, in *the girl (that) the man met was tall*, aid in the comprehension of sentences.

- d. Anchoring (Sternberg 1969, Anderson 1985).

Subjects have better access to structure at the beginning and the end of a string of elements.

These processes are reflected in the English constraints on agreement and on the noun phrase structure that I have discussed tonight. The constraints show pointedly that memory limitations are not grammatically irrelevant. Without a connection to such processes, and this is the most important point, the noun phrase constraints appear to be mere peculiarities of English, the kinds of arbitrary rules that have vexed generations of students in English class.

(54) Constraints on Phrase Structure in English.

- a. The division of English sentences into subject and predicate related by agreement is an instance of chunking or hierarchical ordering.
- b. The placement of the subject before the predicate or of the determiner before the head are instances of linear ordering.
- c. The constraints on the nucleus of the noun phrase are instances of cueing.
- d. The placement of the word *of* before and after the nucleus is an instance of anchoring.

Given these relationships between grammatical constraints and psychological processes, more progress might be made in the theory of grammar if grammatical descriptions are related to human cognitive abilities, that is, to abilities that generally are classified as elements of linguistic performance. Doing linguistics in a vacuum with an ideal speaker-listener who is unaffected by memory limitations undermines this progress.

To summarize, we might hypothesize that UG represents general cognitive abilities rather than linguistic abilities per se, that is, what might be innate are a number of generalized information processing abilities applicable to all perceived objects whether they be sounds, words, numbers, pictures, events, etc. Such abilities might include the capacity to sort and integrate data, to identify

and categorize objects in terms of component features, to recognize sequential and hierarchical relationships, to identify boundaries between objects, and so on.

- (55) Information processing abilities which are specified by human biology (innate) and which are linked to maturation.
- a. Sort and integrate data, e.g., words belong to categories like noun and verb.
  - b. Identify and categorize objects in terms of component features, e.g., nouns and verbs have features of agreement (singular or plural).
  - c. Recognize sequential and hierarchical relationships, e.g., sentences are organized into subunits with elements in specific linear orders.
  - d. Identify boundaries between objects, e.g., linguistic units are separated by cues and anchors.

The advantages of recognizing the importance of these general information processing abilities are outlined in (56).

- (56) Advantages of recognizing the importance of these general information processing abilities.
- a. Links the description of language directly to human cognitive and perceptual capacities.
  - b. Characterizes the evolvement of language in children as very much influenced by experience, specifically, by exposure to data and stimuli that are needed to activate, validate, and automatize such abilities.
  - c. Helps to explain the huge dialectal and idiolectal variation that occurs among speakers of the same language, particularly seen in children from differing backgrounds.
  - d. Helps to explain the dissolution of language when information processing abilities are disrupted by disease, aging, or trauma, as well as the failure of language to develop normally when conditions like severe sensory deprivation, retardation, or autism are present.

In short, if human linguistic ability is viewed merely as a particular specialization of general cognitive and perceptual abilities which do vary among humans, then variation in linguistic ability is both possible and expected.

Much more is at stake here than I have indicated thus far, particularly for linguistics. In recent years, Chomsky's program for linguistics has resulted in analyses which are very difficult to reconcile with a reasonable theory of linguistic performance. His current theory, called Government-Binding Theory is based on observations like those given in (57).

(57) Government-Binding Theory (GB Theory).

a. Noun Phrases must be assigned a case.

[1] He saw me. / \*Him saw I.

[2] She knows them. / \*Her knows they.

b. Case is determined by principles of government.

[1] Noun phrases in subject position (governed by a tense) must have nominative case: He/\*him left.

[2] Noun phrases in object position (governed by a verb or preposition) must have accusative case: Look at them/\*they.

In GB-theory, there has been much emphasis on the study of empty categories in languages, that is, categories that are present but phonologically null. One such category is the category PRO, which is postulated to exist to account for the "missing" subject of infinitive constructions like those in (58).

(58) a. John appeared to Bob to be happy. (John is happy)

b. John appeared to Bob to be happy. (Bob is happy)

Notice that determining the "missing" subject is not a trivial matter, any more than determining the head of a noun phrase was. The sentences in (58) look very similar, yet they have very different interpretations. There are many similar pairs; for example, (59).

(59) a. John promised Bob to go. (John will go)

b. John ordered Bob to go. (Bob will go)

Further, we have pairs of sentences like (60).

(60) a. John might want Bob to go. (Bob will go)

b. John might want to go. (John will go)

Native speakers know that the subject of *go* in (60b) is *John*. To account for this, Chomsky and his followers claim that the subject is represented by an abstract empty noun phrase called PRO. In other words, although there is nothing actually there in the acoustic signal of a sentence like *John might want to go*, the category PRO is there abstractly. Chomsky assumes the representation in (61).

- (61) a. John might want Bob to go.  
 b. John might want PRO to go.

This assumption makes the underlying structure of (61b) parallel to (61a). Further, Chomsky can claim that all verbs must have overt subjects in their underlying representations, even if they don't have overt subjects in the actual acoustic signal.

While this symmetry is attractive, it does not help matters. Merely putting PRO into an abstract representation does not tell us what it refers to. If there is only one noun elsewhere in the sentence, the matter seems simple; PRO must refer to that noun, as in (61b). However, if there is more than one noun, as in (58) and (59), how do we know which one it refers to? Further, there are also sentences like (62), where PRO refers to anyone, that is, its interpretation is arbitrary.

- (62) John thinks that it's important PRO to go to college.

To account for these difficulties, GB-theory contains a subtheory, called Control Theory, which specifies how PRO is to be interpreted. I have argued elsewhere that abstract representations with PRO are unnecessary, and so is Control Theory. I have shown that there are other more general principles which predict what the "missing" subject of an infinitive must be. Further, the PRO analysis distorts English syntax.

While it seems attractive to make the underlying structure of all clauses identical, that is, putting in subjects for verbs when the subjects are missing in the acoustic signal, this tactic obscures at least one very important generalization.

Main verbs in English must have an overt subject as we have seen. The reason is that main verbs must agree with that subject in number. If there is no subject there, then how can one make a verb agree with nothing? Infinitive phrases, like *to go to college*, don't have to have specific subjects; their understood subject can be arbitrary. Why?

The reason is interesting. Infinitives lack tense. We do not have sentences like (63).

- (63) a. \*John might have wanted to went.  
 b. \*It is important to will go to college.

Now, the verb in English that carries the tense marker, is also the verb that carries the agreement marker. Consider (64).

- (64) a. He is happy. (*is* is singular and present)  
 b. They were happy. (*were* is plural and past)

Since infinitives lack tense, they also lack agreement; since they lack agreement, it is not necessary for them to have subjects. This explanation is verified elsewhere in the language in imperative sentences which express commands. Consider (65).

- (65) a. Go to college.  
b. (You) go to college.

Notice that the subject of a command, *you*, is generally left out. How come we can leave out the subject of a sentence like this? The answer is the same as in infinitives. Commands lack tense. There is one verb in the language, only one, that proves this. Consider (66).

- (66) a. Be careful.  
b. \*Are careful.  
c. \*Is careful.

Notice that commands are tenseless; since they lack tense, they also lack agreement. Since they lack agreement, they can lack subjects. Thus, making all clauses symmetrical by putting in PRO distorts the syntax. If you put PRO in as the subject of infinitive phrases when a subject is missing, then you leave unexplained that fact that infinitives can be understood as having no specific subject. Again, we see how this simple rule of agreement has such profound consequences for English syntax. It is definitely not a whim.

Thus, even within the theory of linguistic competence, I believe that Chomsky's analysis fails.

Turning to the question of linguistic performance, if Chomsky's PRO analysis is accepted, we are left with a huge problem in terms of language acquisition. The reason is that PRO has none of the characteristics of a noun phrase, though it is supposed to be one.

PRO has the characteristics detailed in (67):

- (67) a. PRO is a phonologically empty noun phrase, that is, it is not pronounced.  
b. PRO can exist in an ungoverned position (unlike ordinary noun phrases).  
c. PRO lacks case (unlike ordinary noun phrases).  
d. PRO can occur where ordinary noun phrases cannot occur:  
    [1] John will persuade Harry PRO to examine Mary.  
    [2] \*John will persuade Harry Bob to examine Mary.  
e. When PRO occurs where an ordinary noun phrase can occur, it is exceptional.  
    [1] John wants him to win.  
    [2] John wants PRO to win.

In [1], *him* must receive case, whereas in [2] PRO must not; yet the two phrases have identical representations.

- f. Since PRO is an empty category, it is invisible to rules of contraction (unlike ordinary noun phrases):

[1] Does John want him to win.  
\*Does John wanna him win?

[2] does John want PRO to go  
Does John wanna go?

- g. However, although PRO is empty and invisible, movement rules can see it, because it must be moved to different positions in the derivation of some sentences:

[1] John wants [ [e] to grow [ Sue to be famous] ]

John wants Sue to grow to be famous.

[2] John wants [ [e] to grow [ PRO to be famous] ]  
John wants PRO to grow to be famous.

Given these data, we must say that PRO is an abstract noun phrase with none of the characteristics of a noun phrase.

In regard to empty categories, Chomsky makes the observation in (68).

- (68) "...there is an intrinsic fascination in the study of the properties of empty elements. These properties can hardly be determined inductively from observed overt phenomena, and therefore presumably reflect inner resources of mind. If our goal is to discover the nature of the human language faculty, abstracting from the effects of experience, then these elements offer particularly valuable insights." (Chomsky 1981, p. 55).

However, given (67), what motivates a child to understand that PRO is a noun phrase since it has none of the characteristics of a noun phrase? It clearly begs the question to assume that PRO must be part of the innate endowment. In short, when syntactic analyses lead to results like (67), it seems appropriate to bring psycholinguistic and neurolinguistic issues into the picture. There is a great deal at stake here: the whole theory of empty categories in language assumes the existence of PRO. What does this theory mean, if, in fact, PRO does not exist?

In conclusion, I feel there is much to be gained from attempting to relate linguistic analyses to cognitive abilities. Not only will it help us to understand what actually happens during the development and dissolution of language, it also will help to keep us on track. Many of the insights of GB Theory are enormously effective. But these insights are embedded in a model that seriously distorts English syntax and is very difficult to relate to a theory of linguistic performance. Notice, however, that it is important to meet Chomsky on his own ground. In all my publications, I first show how Chomsky's model fails as an account of linguistic competence. That is, I don't begin with

the performance issues. Having, I believe, succeeded on autonomous grounds, to show that an alternative model is necessary, I then consider the advantages of my model in terms of psycholinguistic issues. I completely agree with Chomsky on one matter, we must be very clear about what we are doing in so young an enterprise as the formal study of grammar and its relation to mental processes.

## REFERENCES

- Anderson, John R. 1985. *Cognitive psychology and its implications*. 2nd edn. New York: W.H. Freeman and Co.
- Binkert, Peter J. 1990 *The cognitive foundations of syntax: Parts of speech in universal grammar and computer parsing*. ms.
- Brainerd, Charles J., and Michael Pressley. 1985. *Basic processes in memory development*. New York: Springer.
- Brown, Roger. 1973. *A first language*. Cambridge, Mass.: Harvard.
- Camarazza, Alfonso, and Edgar B. Zurif, eds. 1978. *Language acquisition and language breakdown, parallels and divergencies*. Baltimore: Johns Hopkins University Press.
- Caplan, David. 1972. Clause boundaries and recognition of latencies for words in sentences. *Perception and Psychophysics*, 12, 73-76.
- .1987. *Neurolinguistic and linguistic aphasiology*. Cambridge: Cambridge University Press.
- Caplan, David, and Nancy Hildebrandt. 1988. *Disorders of syntactic comprehension*. Cambridge, Mass.: MIT.
- Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge: MIT Press.
- .1981. *Lectures on government and binding*. Dordrecht: Foris.
- .1986. *Knowledge of language*. New York: Praeger.
- Dempster, Frank N. 1978. Memory span and short-term memory capacity: A developmental study. *Journal of Experimental Child Psychology*, 26, 419-431.
- .1981. Memory span: Sources of individual and developmental differences. *Psychological Bulletin*, 89, 63-100.
- .1985. Short-term memory development. In Brainerd and Pressley (1985).
- Fodor, Jerry T. and Merrill F. Garret. 1967. Some syntactic determinants of sentential complexity. *Perception and Psychophysics*, 2, 289-296.
- Gleason, Jean Berko. 1978. The acquisition and dissolution of the English inflectional system. In Camarazza and Zurif (1978).

Graf, Richard, and Jane W. Torrey. 1966. Perception of phrase structure in written language. *American Psychological Association Convention Proceedings*, 83-88.

Hakes, David T. 1972. Effects of reducing complement constructions on sentence comprehension. *Journal of Verbal Learning and Verbal Behavior*, 11, 178-286.

---, and Donald J. Foss. 1970. Decision processes during sentence comprehension: Effects of surface structure considered. *Perception and Psychophysics*, 8, 413-416.

Hess, T.M., and R.C. Radke. 1981. Processing and memory factors in children's reading comprehension skill. *Child Development*, 52, 478-488.

Jarvella, Robert J. 1971. Syntactic processing of connected speech. *Journal of Verbal Learning and Verbal Behavior*, 10, 409-416.

Jensen, A.R., and R.T. Osborne. 1979. Forward and backward digit span interaction with race and IQ: A longitudinal developmental comparison. *Indian Journal of Psychology*, 54, 75-87.

Johnson, Neal F. 1970. The role of chunking and organization in process of recall. *Psychology of language and motivation*, Vol. 4, ed. by Gordon H. Bower.

Johnson, Neal F. 1972. Organization and the concept of a memory code. *Coding processes in human memory*, ed. by Arthur W. Melton and Edwin Martin. New York: Winston.

Kean, Mary-Louise. 1982. Grammatical representation and the description of language processing. In Caplan (1982).

---ed. 1985. *Agrammatism*. Orlando: Academic Press.

Klahr, David, William G. Chase, and Eugene A. Lovelace. 1983. Structure and process in alphabetic retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9, 462-477.

Klatzky, Roberta. 1980. *Human memory*. New York: W.H. Freeman.

Mandler, George. 1967. Organization and memory. *The psychology of learning and motivation*, Vol. 1, ed. by Kenneth W. Spence and Janet T. Spence. New York: Academic Press.

Miller, George. 1956. The magical number seven, plus or minus two: Some limits on your capacity for processing information. *Psychological Review*, 63, 81-97.

Olson, D. 1973. What is worth knowing and what can be taught. *School Review*, 24-43.

Ross, B.M. 1969. Sequential visual memory and the limited magic of the number seven. *Journal*

of Experimental Psychology, 80, 339-347.

Shaffer, David. 1989. Developmental psychology, childhood and adolescence. Pacific Grove, Cal.: Brooks Cole.

Sternberg, Saul. 1969. Memory scanning: Mental processes revealed by reaction time. American Scientist, 57, 421-457.

Tulving, E. and J. Psotka. 1971. Retroactive inhibition in free recall: inaccessibility of information available in the memory store. Journal of Experimental Psychology, 87, 1-8.