ucla

Occasional Papers in Linguistics

Number 11

edited by Filippo Beghelli and Murat Kural

JANUARY 1993

UCLA Occasional Papers in Linguistics
Number 11
January 1993

Recent Papers in Syntax, Semantics, and Computational Linguistics

edited by

Filippo Beghelli and Murat Kural

Table of Contents

Edward L. Keenan, Semantic Order & Semantic Answers to Syntactic Questions . . 1

Murat Kural, V-to (-I-to) -C in Turkish 17

Anoop K. Mahajan, On Gamma Marking Adjunct Traces in Hindi 55

Jeannette Schaeffer, Dutch Nominal Compounds and the Innateness of Level-Ordering 65

Edward P. Stabler, The Finite Connectivity of Linguistic Structure 87

Andi Wu, A Minimalist Universal Parser 135

SEMANTIC ORDER
&
SEMANTIC ANSWERS TO SYNTACTIC QUESTIONS

Edward L. Keenan

iyw8elk@oac.ucla.edu

Let C be a grammatical category {S, NP, VP, ...} of natural language, and consider the collection of things we may reasonably think of expressions of category C as denoting (= being semantically interpreted as). Experience shows that in general this set is not just some random collection -- rather its elements are *ordered* in a certain, usually quite specific, way.

We shall be concerned in this article with several semantic generalizations about English (and hopefully natural language in general) which build on the notion of a semantic order. The generalizations all concern, at least as special cases, the interpretation of NPs or quantifiers. Most of them are available in the literature, though some more accessibly than others.

Our purpose here is expository: to make these generalizations accessible to the non-specialist and to exhibit the sense in which semantic analysis may contribute to the solution of problems which arise in a syntactic setting. In addition we unify the generalizations by presenting them all in an order theoretic perspective.

1. On the notion of order

As an illustrative example consider (tensed) VPs, such as those italicized in (1):

(1) a. John *laughed*
b. John *laughed loudly*
c. John *both laughed and cried*
d. John *either laughed or cried*

The VPs of these Ss are semantically related in some obvious ways. Extending standard usage, the VP of (1b) "entails" that of (1a), meaning that whenever *laughed loudly* is true of an individual then so is *laughed*. One standard way to say this is as follows:

In a given a situation σ, a VP of the sort in (1) is true of some (possibly all) of the objects we might be talking about in σ. For p a VP expression write $\mathbf{p}_\sigma$ for the set of things p is true of in σ. Write E_σ, called the *universe* of σ, for the set of things we might be talking about in situation σ. (We usually omit the subscript σ since we are not comparing different situations).

And we define: a VP p *entails* a VP q iff for all situations σ, $\mathbf{p}_\sigma \subseteq \mathbf{q}_\sigma$. So to say that *laughed loudly* entails *laughed* is to say (omitting subscripts) that in all situations σ, **laughed loudly** $\subseteq$ **laughed**. That is, any object that *laughed loudly* is true of is an object that *laughed* is true of.

exercise Verify informally that *both laughed and cried* entails *laughed* and that *laughed* entails *either laughed or cried* #

Note that *laughed* does not entail *slipped,* since there are situations in which **laughed** $\not\subseteq$ **slipped**, that is, situations in which someone laughed who didn't slip. But there are also situations in which **laughed** $\subseteq$ **slipped**, that is, ones in which everyone who laughed did slip. Thus in a given situation some VP denotations stand in the subset relation and some do not.

Now the subset relation is a basic *order* relation: *transitive* (if $A \subseteq B$ and $B \subseteq C$ then $A \subseteq C$), and *antisymmetric* ($A \subseteq B$ and $B \subseteq A \Rightarrow A = B$). Standardly,

Def 1 A binary relation R defined on a set E is called an *order* relation iff

i. R is transitive
(viz. for all a,b,c ε E, aRb & bRc $\Rightarrow$ aRc) and

ii. R is antisymmetric
(viz. for all a,b ε E, aRb & bRa $\Rightarrow$ a = b)

Note that antisymmetry rules out that distinct objects each bear the relation to the other, but it allows that a given object bear the relation to itself. Indeed the subset relation is *reflexive*, meaning that for each set A, $A \subseteq A$.

And in general the order relations we are linguistically motivated to consider below are reflexive. We often use $\leq$ as a symbol denoting a reflexive order relation. (Note e.g. that the natural $\leq$ relation among numbers in arithmetic is a reflexive order). Anticipating a more general use, we write $\leq_{VP}$ for the reflexive order relation on VP denotations defined above. (So $\leq_{VP}$ is just the subset relation as we presented it.)

A second basic order relation is the "implication" order $\leq_S$ defined on possible Sentence denotations:

Let x and y be possible S denotations (in a situation σ). Then x $\leq_S$ y iff an arbitrary sentence of the form *if p then q* is true when p denotes x and q denotes y. For example, thinking of Ss as denoting either **T** ("True") or **F** ("False"), we see that the $\leq_S$ relation is completely given by:

(2) $\mathbf{T} \leq_S \mathbf{T}$, $\mathbf{F} \leq_S \mathbf{T}$, and $\mathbf{F} \leq_S \mathbf{F}$.

The only case where a truth value x fails to bear the "implication" relation to a truth value y is when x = **T** and y = **F**. So to show that x $\leq_S$ y we must merely show that if x = **T** then y = **T** (since if x = **F** then x $\leq_S$ y no matter what truth value y is).

And as with VPs, we define: a sentence p *entails* a sentence q iff for all situations σ, $\mathbf{p} \leq_S \mathbf{q}$, where **p** and **q** are the respective denotations of p and q in σ.

One sees by inspection of (2) that $\leq_S$ as defined is reflexive (= for all x ε {**T**,**F**}, x $\leq_S$ x). Equally no two different truth values each stand in the $\leq_S$ relation to the other, so $\leq_S$ is antisymmetric. And transitivity is checked by cases in (3). To show that if x $\leq_S$ y and y $\leq_S$ z then x $\leq_S$ z, for x, y, z ε {**T**,**F**}, it is sufficient to consider the choices of values for x, y, z which make the "and" clause true.

(3)

x $\leq_S$	y	and y $\leq_S$	z	x $\leq_S$	z	??
T	**T**	**T**	**T**	**T**	**T**	yes
F	**T**	**T**	**T**	**F**	**T**	yes
F	**F**	**F**	**T**	**F**	**T**	yes
F	**F**	**F**	**F**	**F**	**F**	yes

Our concern now is with denotations of NPs, such as those italicized in (4). Such NPs combine with VPs to form Ss and may naturally be interpreted by functions which map VP denotations to S denotations (the truth value of such a sentence then being the value that the NP function assigns to the VP denotation).

(4) a. *John* is asleep

b. *Most students* can read

c. *More students than teachers* read the Times

Def 2 Now, if F and G are possible NP denotations (in a situation σ) we say that F $\leq_{NP}$ G iff for all possible VP denotations **p**, $F(\mathbf{p}) \leq_S G(\mathbf{p})$. (That is, if $F(\mathbf{p}) = \mathbf{T}$ then $G(\mathbf{p}) = \mathbf{T}$).

fact The $\leq_{NP}$ relation defined above is (for each σ) a reflexive order relation (see below).

And the entailment relation on NPs is defined as before: An NP A entails an NP B iff for all situations σ, $\mathbf{A} \leq_{NP} \mathbf{B}$, where **A** and **B** are the respective denotations of A and of B in σ.

For example, *every student and some teacher* entails *every student* since for any VP q, if *every student and some teacher q* is true then, obviously, *every student q* is true. And more generally,

(5) For C ε {NP, VP, S} and for x,y expressions of category C,

i. *both x and y* entails *x* and

ii. *x* entails *either x or y* #

The $\leq_{NP}$ order builds on the $\leq_S$ order on the set, {**T**,**F**} in which NP functions take their values. This way of inheriting orders is fully general:

Def 3 Let A be any set and let $\leq_B$ be a reflexive order on a set B. Then we define a relation $\leq$ on $[A \rightarrow B]$, the set of functions from A into B, as follows:

For all $f,g \in [A \rightarrow B]$,

$$f \leq g \text{ iff for all } b \in B,\ f(b) \leq_B g(b)$$

fact: $\leq$ as defined is a reflexive order.

2. Some semantic generalizations

2.1 Constraints on interpreting lexical items

We will establish here a very non-trivial semantic constraint on the interpretation of lexical NPs -- one that extends with some success to lexical expressions of other categories.

To set up the generalization let us consider first the following entailment paradigm (cf. Aristotle).

(6) a. All socialists are vegetarians
 b. Some doctors are socialists
 ∴ Some doctors are vegetarians

We understand (6) to mean that the first two Ss jointly entail the third. That is, in any situation in which the first two are interpreted as true the third is also interpreted as true.

Query Which NPs X can replace *some doctors* everywhere in (6) preserving the entailment (changing plurals to singulars if necessary)?

NPs which satisfy the **Query** are called *increasing*. For example *Mary* is increasing: if all socialists are vegetarians and Mary is a socialist then, obviously, Mary is a vegetarian. Some other increasing NPs are given in (7), as the reader is invited to check:

(7) a. she, this cat, more than two cats, at least one cat, some cat, every cat, the (ten) cats, John's (ten) cats, most cats, several cats, more than half the cats, his cat, every student's bicycle
 b. at least two of the ten cats, most of John's cats, at least two thirds of the students, more than five of John's cats
 c. John and some student, at least two teachers and more than ten students, either a student or a teacher, most liberal and all conservative senators

To give a properly general account of these entailment facts consider the semantic representation of (6) below (given a situation σ):

(8) a. **socialist** $\subseteq$ **vegetarian**
 b. (**all doctor**)(**socialist**) = **T**
 ∴ (**all doctor**)(**vegetarian**) = **T**

Generalizing from (8) we see that the function **all doctor** *preserves* the order relation on its arguments in the sense that if **p** $\leq_{VP}$ **q** then (**all doctor**)(**p**) $\leq_S$ (**all doctor**)(**q**). That is, **all doctor** is *increasing* as defined below:

(9) Let A and B be ordered sets and F a function from A into B.

a. F is *increasing* (= *order preserving*) iff for all $a,a' \in A$,
   if $a \le a'$ then $F(a) \le F(a')$

b. F is *decreasing* (= *order reversing*) iff for all $a,a' \in A$,
   if $a \le a'$ then $F(a') \le F(a)$

c. F is *monotonic* iff F is increasing or F is decreasing

And the NPs in (7) are increasing in the sense that in all situations σ they denote increasing functions. And we now state:

Gen 1: Lexical NPs are monotonic -- in fact monotonic increasing with at most a few exceptions.

Here is a snapshot of the lexical NPs of English: they include one productive subclass, the *proper nouns*: *John, Mary, ..., Siddartha, Chou en Lai, ...* ('productive' here means that new members may be freely added without changing the language significantly). They also include listable sprinklings of (i) personal pronouns -- *he/him,..* and their plurals *they/them*; (ii) demonstratives -- *this/that* and *these/those*; (iii) possessive pronouns -- *his/hers .../theirs*; and (iv) a few "indefinite pronouns" as *all* in *A good time was had by all, some* in *Some like it hot,* and *many* and *few* in *Many are called but few are chosen.* Some linguists would include here *everyone, everybody, everywhere*; *someone, somebody, somewhere*; and *none, noone, nobody, nowhere,* though these expressions appear to have meaningful parts.

Excluding *none, noone, nobody,* and *nowhere,* which are properly decreasing, the lexical NPs noted above are increasing.

We shall discuss decreasing NPs in **Gen 2** below. Here let us just note that the NPs in (10) below are not monotonic.

(10) a. every student but not every teacher, every student but John, exactly five students, between five and ten cats, no student but John, John but neither Bill nor Sam, most of the students but less than half the teachers

b. either fewer than five students or else more than a hundred students, approximately a hundred students, more students than teachers, exactly as many students as teachers

Note that in any given situation the NPs in (10) will denote perfectly reasonable functions from possible VP denotations to possible S denotations, but those functions are not monotonic. Thus **Gen 1** is a strong semantic claim about natural language -- many functions that are denotable by NPs in English are not denotable by lexical NPs.

If we think of **Gen 1** as a constraint on the interpretation of human languages then it helps to explain how children learn languages quickly with limited exposure to limited data. They must learn the meanings of the expressions they use. And while learning the meanings of syntactically complex expressions is facilitated by knowning the meanings of their parts, learning the meanings of lexical items is not facilitated in this way. But it is facilitated if the child need consider only monotonic (increasing) denotations for his lexical NPs.

For further generalizations concerning constraints on denotations of lexical items see Keenan (1987). We turn now to our second generalization.

2.2 Negative polarity items

To characterize the set of expressions judged grammatical by native speakers of English, we must distinguish the grammatical expressions (11a) and (12a) from the ungrammatical (11b) and (12b).

(11) a. John hasn't ever been to Moscow
b.*John has ever been to Moscow

(12) a. John didn't see any birds on the walk
b.*John saw any birds on the walk

Npi's (negative polarity items) such as *ever* and *any* above, do not occur freely; classically [Klima 1964] they must be licensed by a "negative" expression, such as *n't* (= *not*). But observe:

(13) a. No student here has ever been to Moscow
b.*Some student here has ever been to Moscow

(14) a. Neither John nor Mary saw any birds on the walk
b.*Either John or Mary saw any birds on the walk

(15) a. None of John's students has ever been to Moscow
b.*One of John's students has ever been to Moscow

The *a*-expressions here are grammatical, the *b*-ones are not. But the pairs differ with respect to their initial NPs, not the presence vs. absence of *n't*.

The linguistic problem: define the class of NPs which license the npi's, and state what, if anything, those NPs have in common with *n't/not*.

A syntactic attempt to kill both birds with one stone is to say that just as *n't* is a "reduced" form of *not* so *neither...nor...* is a reduced form of [*not* (*either...or...*)], *none* a reduction of *not one*, and *no* a reduction of *not a*. The presence of *n-* in the reduced forms is thus explained as a remnant of the original *not*. So on this view the licensing NPs above "really" have a *not* in their representation, and that is what such NPs have in common with *n't*. Moreover NPs built from *not* do license npi's:

(16) Not a single student here has ever been to Moscow
Not more than five students here have ever been to Moscow

However, as Ladusaw [1983] has taught us, this solution is insufficiently general: The initial NPs in the *a-* sentences below license npi's; those in the *b-* sentences do not. But neither present reduced forms of *not*.

(17) a. Fewer than five students here have ever been to Moscow
b. *More than five students here have ever been to Moscow
a. At most four students here have ever been to Moscow
b. *At least four students here have ever been to Moscow
a. Less than half the students here have ever been to Moscow
b. *More than half the students here have ever been to Moscow

An hypothesis which does yield correct results is a semantic one discovered by Ladusaw (1983), building on the earlier work of Fauconnier (1979). (See also Zwarts (1981)).

Gen 2 *The Ladusaw-Fauconnier Generalization (LFG)*

Negative polarity items occur within an argument of a monotonic decreasing function

To check that an NP is decreasing verify that (18) is valid when substituted for X.

(18) All linguists can dance
X can dance
∴ X is a linguist (are linguists)

This test shows that the NPs in (13) - (17) which license npi's are decreasing whereas those that do not are not.

Further the LFG yields correct results on structures like (19) and (20) below, not considered by Ladusaw or Fauconnier.

(19) No player's agent should ever act without his consent
*Every player's agent should ever act without his consent
Neither John's nor Mary's doctor has ever been to Moscow

(20) None of the teachers and not more than three of the students have ever been to Moscow

(19) draws on the fact that possessive NPs, ones of the form [X's N] such as *John's doctor*, inherit their monotonicity from that of the possessor X. Viz., X'*s doctor* is increasing (decreasing) if X is. (20) testifies that conjunctions (and disjunctions) of decreasing NPs are decreasing.

Finally we may observe from a linguist's perspective that the LFG is quite general. Denotation sets for most categories of expression in English are ordered (Keenan & Faltz, 1985) and thus most expressions of functional types are classifiable as increasing, decreasing or non-monotonic. We may expect then to find npi licensers in many categories, and we do.

A crucial case of course is that of ordinary negation *not* (*n't*). In general it denotes a complement operation in the set in which its argument denotes. E.g. at the VP level *didn't laugh* denotes E - **laugh**, the set of objects under discussion that are not in the **laugh** set. So **not (n't)** maps each subset **p** of E to E - **p**. And one shows easily that if $\mathbf{p} \subseteq \mathbf{q}$ then $E - \mathbf{q} \subseteq E - \mathbf{p}$. Which is just to say that the denotation of *not* (*n't*) is decreasing.

Thus the LFG finds a non-trivial and independently verifiable property which NPs like *no student* have in common with simple negation.

For further refinement see Nam (1992) and Zwarts (1990).

2.3 Partitives and definite NPs

We consider partitive NPs like *at least two of the students, all but one of John's children* and *most of those questions*. They appear to be of the form [DET_1 of NP], and more generally [DET_K (of NP)^K], like *more of the students than of the teachers.*

The linguistic issue: For which choices of DET_1 and NP is the sequence ⟨DET_1 of NP⟩ a grammatical NP? Some partial answers:

(21) a. *[at least two of X]* is a grammatical NP when X = the boys; the ten or more boys; these boys; these ten boys; John's cats; John's ten or more cats; my cats; the child's toys; that child's best friend's toys

b. *at least two of X* is ungrammatical when X = each boy; all boys; no boys; the boy; some boys; most boys; exactly ten boys; ten boys; no children's toys; most of the houses; at least nine students, more students than teachers, five of the students

Whether an NP of the form DET_1 + N occurs grammatically in the partitive context [*two of* __] depends significantly on *its* choice of DET_1. DET_1s acceptable here were first characterized semantically in Barwise and Cooper [1981]. We build on their analysis below.

Note that we might naively refer to NPs which occur naturally in these partitive contexts as "definite plural". So what is at issue is how to characterize that notion. We propose a semantic characterization. One problem that must be correctly handled here is the following: by various criteria NPs like those in (22) are definite plural, but, as indicated, they are at best problematic in (+count) partitive contexts, (23).

(22) a. the student and the teacher
b. this student and that student
c. John and Bill

(23) a. *?all/both of the student and the teacher
b. *most of this student and that student
c. *?none of John and Bill

We respond to this problem below by characterizing the NPs which occur in (+ count) partitive contexts in terms of the Det_1s used to build them, rather than directly in terms of denotational properties of the NP itself.

Now observe that Det_1s like *most, every, more than ten, at least and not more than ten, ...* combine with common nouns to form NPs. Semantically then we may interpret them by functions mapping common noun denotations to NP denotations. We shall take common noun denotations to be sets of objects (e.g. in a situation σ with universe E, the students in E are the objects in the set denoted by *student*. Here are some illustrative examples in an obvious notation:

(24) **some**(p)(q) = **T** iff $p \cap q \neq \varnothing$
every(p)(q) = **T** iff $p \subseteq q$
(**the ten**)(p)(q) = **T** iff $|p| = 10$ and $p \subseteq q$
most(p)(q) = **T** iff $|p \cap q| > |p - q|$

So e.g. *most p's are q's* is true iff the number of p's who are q's is greater than the number of p's who are not q's.

To avoid certain trivial cases in our characterization of "definite plurals" we note the (largely obvious) definitions of trivial functions: An NP function F is *non-trivial* in σ iff there are subsets **q**,**q'** of E such that F(**q**) = **T** and F(**q'**) = **F**. A Det_1 denotation **g** is *non-trivial* in σ iff for some $\mathbf{p} \subseteq E$, **g**(**p**) is non-trivial. And a Det_1 expression g is non-trivial iff for some situation σ, the denotation **g** of g is non-trivial in σ.

Lastly, an NP function F is said to be a *principal filter* iff for some $\mathbf{s} \subseteq E$, F(**q**) = **T** iff $\mathbf{s} \subseteq \mathbf{q}$. In such a case F is said to be *generated* by **s**.

For example in a situation with many cats the NP *the cats* denotes the filter generated by **cat**. So does the NP *the two or more cats*. If John has exactly two cats then *John's two cats* is the filter generated by **cat which John has**. We now propose an answer to our query:

Def 4 A Det_1 expression g is *semantically definite* iff g is non-trivial and for each situation σ and each $\mathbf{p} \subseteq E$ such that **g**(**p**) is non-trivial, **g**(**p**) is the filter generated by some non-empty $\mathbf{s} \subseteq \mathbf{p}$. If **s** always has at least two elements **g** is called *definite plural.*

(25) *Some semantically definite plural Det_1s*

the ten, ten two or more, the_{pl}, John's ten, John's two or more, $John's_{pl}$, these, these ten, these ten or more, John and Bill's ten, ...

We might note here that *every* is not semantically definite, and that the Det_1's *the one, John's one* are semantically definite but not definite plural.

Gen 3 An NP is grammatical in plural partitive contexts iff it is of the form [d N] where d is semantically definite plural or it is a conjunction or disjunction of such NPs.

We note that NPs in (22) such as *this student and that teacher* are excluded by this definition.

2.4. Existential NPs

Consider *Existential There* (ET) Ss like those in (26):

(26) There wasn't more than one student at the party
There are more dogs than cats in the garden
There was noone but John in the building at the time

Such Ss are typically used to affirm, deny or query the existence of objects (e.g. students) with a specified property (e.g. being at the party). NPs like *more than one student* which naturally occur in such Ss will be called *existential NPs*. So NPs italicized in (27) are not existential, as the Ss are either ungrammatical or assigned an unusual interpretation.

(27) *There wasn't *John* at the party
*There were *most students* on the lawn
*?There wasn't *every student* in the garden

The linguistic problem: define the set of existential NPs in English. Barwise & Cooper (op cit) again were the first to propose a semantic solution to this problem, and as in the previous case, located the solution in the nature of the DET_1s rather than the NPs themselves. The solution presented below is original here but draws on theirs and on Keenan [1987b]. See Reuland and ter Meulen [1987] for extensive discussion of the empirically problematic issues here.

We construct the existential NPs from ones built from *intersective* Determiners. To say that **more than ten** is intersective is to say that we can decide whether **more than ten** p's are q's just by checking $p \cap q$, the p's who are q's. We need not for example concern ourselves with p's which fail to be q's, as we must when checking whether **all** p's are q's.

Equally to say that a two place determiner such as *more...than...* is intersective is to say that we can decide whether more students than teachers are vegetarians is true just by checking the students who are vegetarians and the teachers who are vegetarians. We need know nothing about students or teachers who fail to be vegetarians. Formally,

Def 5 A function **g** mapping k-tuples of sets to possible NP denotations is *intersective* iff for all k-tuples $(p_1,\ldots,p_k)$ and all sets q,q' if $p_i \cap q = p_i \cap q'$, all

$1 \leq i \leq k$, then $F(p_1,...,p_k)(q) = F(p_1,...,p_k)(q')$.

Gen 4 NPs which occur naturally in ET contexts are ones built from intersective Dets or they are boolean compounds (in *and, or, not, neither...nor...*) of such NPs.

Since (**more than ten**)(p)(q) = 1 iff $|p \cap q| > 10$ we see that it is intersective and thus *more than ten cats* is, correctly predicted to occur naturally in ET contexts. Some further examples of intersective DET_1s:

(28) some, a, exactly ten, fewer than ten, not more than ten, no, between five and ten, at most ten, at least two but not more than ten, just finitely many, uncountably many, no...but John, more male than female

One checks that NPs built from these Dets occur naturally in existential Ss, as do their boolean compounds. Equally one proves easily that boolean compounds (in *and, or,* and *not*) of intersective Dets are intersective. So we predict that NPs such as those in (29) occur naturally in ET contexts, a correct prediction.

(29) at least two and not more than ten cats, either exactly two or exactly four cats

By contrast the Det_1s displayed below are not intersective and do not naturally occur in existential contexts:

(30) most, all, all but two, every...but John, two out of three, less than half the, at most twenty per cent of the, the ten, John's ten, these, my

Equally one checks that cardinal comparative DETs like *more...than..., fewer...than..., exactly as many...as..., more than twice as many...as...,* as in (1c) are intersective functions of type $\langle\langle 1,1\rangle,1\rangle$. E.g. whether *fewer students than teachers are vegetarians* is true is determined by the sets **student** $\cap$ **vegetarian** and **teacher** $\cap$ **vegetarian.** Thus we correctly predict that *There are fewer students than teachers in the garden* is natural.

A closing remark: Our purpose here has been to present generalizations which rely on the underlying order in denotation sets. It may not be obvious here however just how the notion of an intersective Det is built specifically on the underlying order. That is because we took the denotations of common nouns and VPs as sets because of the familiarity of this notion. But we could essentially without change have taken Ns and VPs to denote functions from E into {**T**,**F**}. But then would could not have literally referred to "intersections" of such functions.

So the crucial point here is to recognize that intersection in set theory is characterizable purely order theoretic terms. Specifically $\mathbf{p} \cap \mathbf{q}$ is the *greatest lower bound* of $\{\mathbf{p},\mathbf{q}\}$, where that notion is defined for ordered sets in general as follows:

Def 6 Where $\leq$ is an order relation on a set A, and $K \subseteq A$,

a. an element $\alpha \in A$ is a *lower bound* (lb) for K iff for all $k \in K$, $\alpha \leq k$.

b. α is a *greatest lower bound* (glb) for K iff

(1) α is a lb for K, and

(2) for all lbs ß for K, $ß \leq \alpha$.

c. If a subset {x,y} of B has a glb, it is noted $(x \wedge y)$.

fact The orders we have considered, e.g. $\leq_C$ for $C \in \{S, NP, VP\}$ are ones in which for all x,y in the set, {x,y} has a glb.

Then the properly general definition of (one place for simplicity) intersective Det_1 functions would be:

(31) **g** is intersective iff for all **p**, **p'**, **q**, **q'**

if $\mathbf{p} \wedge \mathbf{q} = \mathbf{p'} \wedge \mathbf{q'}$ then $\mathbf{g(p)(q)} = \mathbf{g(p')(q')}$

In this way we see that the notion of intersectivity is built on the underlying order.

Indeed the **fact** above enables us to appreciate a final linguistic generalization, much elaborated in Keenan & Faltz (1985):

Gen 5 For x,y expressions of category C, the expression *x and y* is interpreted as the greatest lower bound of the denotations of *x* and of *y*. In other words, the common meaning that *and* in all its uses is as a greatest lower bound former. (Similarly *or* is a least upper bound former).

Using **Gen 5** then we account for the common meaning of *and* in (many of) its diverse occurrences without having to say that the non-sentence level occurrences are derived from coordinate Ss by some kind of reduction rules, ones that are of necessity non-paraphrastic given that e.g. *Exactly two students both came early and left late* is not a paraphrase of *Exactly two students came early and exactly two students left late.*

In conclusion: We have exhibited several semantic generalizations about English which are defined in terms of the underlying semantic order on the denotations of expressions of a fixed category. Several of these generalizations provide a reasonable (but never perfect) answer to queries that were first raised in a purely syntactic context.

Bibliography

Barwise, J. and R. Cooper (1981) 'Generalized quantifiers and natural language' in *Linguistics and Philosophy* 4. 1981; 159 - 219

Fauconnier, G. (1979) 'Implication Reversal in Natural Language' in F.Guenthner and S. Schmidt (eds) *Formal Semantics for Natural Language* 1979, Reidel

Keenan, E.L. (1987) 'Lexical Freedom and Large Categories' in *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers* J. Groenendijk et al (eds)

____. (1987b) 'A semantic definition of 'indefinite NP' ' in Reuland and ter Meulen (1987), pp. 286 -317

Keenan, E.L. and L. Faltz *Boolean Semantics for Natural Language* Reidel 1985

Klima, E. (1964) 'Negation in English' in J.A. Fodor and J.J. Katz (eds) *The Structure of Language* Prentice-Hall

Ladusaw, W. (1983) 'Logical Form and Conditions on Grammaticality' in *Linguistics and Philosophy* 6, pp.389-422.

Nam, S. (1992) *Another type of Negative Polarity Item* ms. Dept. of Linguistics, UCLA

Reuland, E.J. and A. ter Meulen (1987) *The Representation of (In)definiteness* MIT Press

Zwarts, F. (1981) 'Negatief polaire Uitdrukkingen I.' *Glot* vol 4. pp. 35 - 132

____. (1990) 'The Syntax and Semantics of Negative Polarity' in *Views on the Syntax-Semantics Interface II*, S. Busemann (ed) to appear.

V-TO(-I-TO)-C IN TURKISH*

Murat Kural

izzyfk6@mvs.oac.ucla.edu

1. Introduction.

There are two major claims in this paper on Turkish syntax that have implications for the theory of agreement in Universal Grammar. The first claim is the alternative classification of the subordinate Infls of Turkish in section 2, where *-Iš -* [1] is analyzed as the gerundive; *-mE-* and *-mEK* as the infinitive; *-DIK-* as the past; and, *-EcEK-* as the future tense morphemes. This is followed by the second claim in section 3 that the final *-K-* found in these morphemes belongs to the C^0 category. Because the subject-verb agreement marker (Agr) follows *-K-*, this latter claim raises the issue of accounting for the morpheme order in the verbal complex. The empirical evidence discussed in section 4 leads to the conclusion that contra Pollock (1989) and Chomsky (1991), Agr is not an independent head in syntax. Section 5 takes up some of the related issues and remaining problems.

2. Subordinate Infl Categories.

The following shows the conventional analysis of the subordinate Infl categories in Turkish.[2] To my knowledge, the earliest version of this classification can be found in Underhill (1976):

---

* I would like to thank Joseph Aoun, Hilda Koopman, Dominique Sportiche, and Tim Stowell for their valuable input and discussion during the preparation of this paper.

[1] The capital letters indicate phonological variants. Stops and affricates are devoiced when they follow an devoiced consonant, or when they are in the syllable-final position (Final Devoicing). Vowels harmonize with the quality of the first preceding vowel. An intervocalic velar stop /k/ becomes the velar semivowel /ɰ/ (orthographically, 'ğ') after the first binary foot from the left. More on this particular rule in section 2.

[2] The term 'subordinate context' refers to argumental clauses throughout this paper. Adjunct clauses behave somewhat more differently with respect to agreement and subject Case.

(1) Category — Conventional analysis

Category	Conventional analysis
a. -DIK-	gerundive
b. -EcEK-	gerundive
c. -mE-	gerundive
d. -mEK	infinitive
e. -Iš-	deverbal nominal

For Kornfilt (1984), *-DIK* and *-mE-* are both participial forms; a factive nominal and an action nominal marker respectively. She analyzes *-EcEK-* as the future tense morpheme. For Kennelly (1990) *-DIK-* and *-EcEK-* are aspectual markers under Infl that are distinct from main clause tenses.

In this section, I will argue for the following classification:

(2)

Category	Alternative analysis[3]
a. -DIK-	past; cf. main clause past -DI-
b. -EcEK-	future; cf. main clause past -EcEK-
c. -mE-	infinitive
d. -mEk	infinitive
e. -Iš-	true gerundive; equivalent to English *-ing*

The evidence for (2) involves the external distribution of clauses headed by these morphemes, their temporal interpretation, and internal properties.

The reason that the morphemes *-DIK-*, *-EcEK-*, and *-mE-* in (1) are considered to be gerundives has to do with the following properties of complement clauses in Turkish:

(3) a. Subject bears the genitive Case in this context.
b. Subject-verb agreement is in the nominal paradigm.
c. All subordinate clauses are and must be Case-marked.

These properties are illustrated in (4) below (relevant parts bold-faced):

(4) a. Ahmet-∅ ev-e git-ti-∅
A.-**nom** home-dat go-past-**agr**
'Ahmet went home'

[3] If there is an analysis of Turkish subordinate Infls in the literature that has reached to the same classification as (2) that I am not yet aware of, my apologies for not citing the work.

b. *pro* [Ahmet-**in** ev-e git-tiğ-**i**]ni duy-du-m
1.sg A.-**en** home-dat go-DIk-**agr-acc** hear-past-agr
'I heard that Ahmet went home'

c. *pro* [Ahmet-**in** ses-**i**]ni duy-du-m
1.sg A.-**gen** voice-**agr-acc** hear-pst-agr
'I heard Ahmet's voice'

The subject *Ahmet* bears the null morpheme -*∅* nominative Case in (4a), but the genitive *-In* in (4b), parallel to the NP in (4c). Similarly, the 3.singular agreement is the null morpheme -*∅* in the main clause (4a), but *-I-* in (4b) and (4c). Finally, the subordinate clause itself is accusative Case-marked in (4b), same as the corresponding NP in (4c) (/n/ is there only to separate the two vowels). The properties in (3) will be refered as the 'gerundive bahavior' of subordinate clauses when discussed in section 5.

2.1. The gerundive status of -Iš:

I will establish in this section that verbs with *-Iš-* are not derived nominals by contrasting them with truely nominalized verbs. This will not be sufficient to show that *-Iš-* is actually a gerundive marker, which will be done later in this section, in comparison with the other Infl elements.

2.1.1. Structural Case.

As Kennelly (1987) observes, verbs with *-Iš-* can assign structural Case, but those with true nominalizers cannot:

(5) a. [Ahmet-in [bu araba]yɨ al-ɨš-ɨ]
A.-gen this car-acc buy-Iš-agr
'Ahmet's buying this car'

b. *[Ahmet-in [bu araba]yɨ al-ɨm-ɨ]
A.-gen this car-acc buy-Im-agr
'Ahmet's purchase of this car'

(6) a. [[Bu firma]nɨn [Encyclopaedia Britannica]yɨ bas-ɨš-ɨ]
this firm-gen E.B.-acc publish-Iš-agr
'This firm's publishing the Encyclopedia Britannica'

b. *[[Bu firma]nɨn [Encyclopaedia Britannica]yɨ bas-ɨm-ɨ]
this firm-gen E.B.-acc publish-Im-agr
'This firm's publication of the Encyclopaedia Britannica'

Nominals cannot assign structural Case to their complements in Turkish, and there is no Case-marking dummy preposition that would correspond to the English *of*. Thus, all nominals inside another nominal must move to the [Spec, DP] (or NP) position to receive the genitive Case, which is why complex nominals like (7) are not allowed in this language:

(7) John's picture of Mary

Instead, one finds either *John's picture* where John is the possessor or the agent, (8a), or *Mary's picture* where Mary is the theme, (8b):

(8) a. [John-un resm-i]
J.-gen picture-agr
'John's picture'

b. [Mary-nin resm-i]
M.-gen picture-agr
'Mary's picture'

(9) *[John-un Mary (C/P) resm-i]
J.-gen M. picture-agr
'John's picture of Mary'

The complex nominal in (9) cannot be saved by any possible Case and/or postposition combination.[4] This means that verbs with *-Im-* in (5b) and (6b) pattern with true nominals,[5] while those with *-Iš-* maintain their verbal character, which enables

---

[4] Actually, the nominals in (5b), (6b), (9) would be fine with non-specific complements:

(i) [Ahmet-in araba alɨ-ɨm-ɨ] cf. (5b)
A.-gen car buy-Im-agr
'Ahmet's purchase of a car'

(ii) [[Bu firma]nɨn ansiklopedi bas-ɨm-ɨ] cf. (6b)
this firm-gen encyclopedia publish-Im-agr
'This firm's publication of an encyclopedia'

(iii) [Ahmet-in manzara resm-i] cf. (9)
A.-gen scenery picture-agr
'Ahmet's picture of (some) scenery'

This quite possibly indicates that nouns are capable of assigning the partitive Case of Belletti (1987), or that non-specifics incorporate to the head noun in the analysis proposed by Enç (talk at USC, 1991). However, this is somewhat oversimplified, as will be apparent below.

[5] Perhaps unsurprisingly, there are various nominalizer morphemes in Turkish, whose choice is lexically specified and has to be learned. Among these, there is also an *-Iš-* that ought not to be confused with the gerundive *-Iš-*. The difference can be seen in *görüš* from the verb *gör* 'see', which can mean either 'opinion' (nominal), or 'seeing' (gerundive). Thus, forms like *giriš* 'entrance' from *gir* 'enter', *cɨkɨš* 'exit' from *cɨk* 'exit', and *yürüyüš* 'walk' from *yürü* 'walk' must be considered ambiguous between

them to assign structural Case.

2.1.2. Interaction with causatives, passives and negatives.

The morpheme *-Iš-* can appear with verbs in the causative form, which is something true nominalizers cannot do:

(10) a. öl 'die'
b. öl-dür 'kill'

(11) a. öl-üm 'death'
b. ?öl-üš 'dying'
c. *öl-dür-üm 'kill (nominal)'
d. öl-dür-üš 'killing'

The nominal *ölüm* 'death' is derived by adding *-Im-* to the verb *öl* 'die' in (11a), which is impossible with the causative form *öldür* 'kill' in (11c). The morpheme *-Iš-* however, is possible with the causative *öldür* 'kill' in (11d). The marginality of *ölüš* 'dying' in (11b) seems to be the blocking effect of the corresponding *ölüm* 'death' in (11a) (Aronoff 1976).

Irrespective of whether category changing derivational morphology is to be handled in the lexicon or syntax, the fact remains that the selectional properties of derivational morphemes (the nominalizer *-Im-*) are far more restrictive than inflectional morphemes (the gerundive *-Iš-*). This means that *-Im-* can be specified by (or select) only the base form *öl* 'die', but not the more complex version *öldür* 'kill'.[6]

Similarly, verbs with *-Iš-* are allowed in passives, whereas verbs with true nominalizers are not:

(12) a. çöz 'solve'
b. çöz-ül 'be solved'

---

nominal and gerundive versions.

[6] There are actually some exceptional cases like *yatırım* 'investment' from *yat-ır* 'invest' (where *yat* means 'be deposited'), and *yaptırım* 'sanction' from *yap-tır* 'cause to do/make'. Strictly speaking, these are not counter-examples since they are still impossible with multiple causatives, otherwise possible in Turkish: **yat-ır-t-ım* from *yat-ır-t* 'cause someone to invest', or **yap-tır-t-ım* from *yap-tır-t* 'cause someone cause someone else to do/make'. The real issue raised by forms like *yatırım* 'investment' and *yaptırım* 'sanction' is how they are lexicalized by the language learner.

(13) a. çöz-üm 'solution'
b. ?çöz-üš 'solving'
c. *çöz-ül-üm 'passive-solution' (intended)
d. çöz-ül-üš 'being solved'

Again, the selectional requirements of the true nominalizer *-Im-* (whether lexical or syntactic) are too narrowly specified to co-occur with the syntactic passive of *çöz* 'solve'. The Infl-level *-Iš-* however, can take such forms and it is not sensitive to the presence of the passive morpheme on the verb.

Also, unlike true nominalizers, *-Iš-* is possible with negatives:

(14) a. yɨk 'demolish'
b. yɨk-ma '(do) not demolish'

(15) a. yɨk-ɨm 'demolishment'
b. ?yɨk-ɨš 'demolishing'
c. *yɨk-ma-yɨm 'not demolishment' (intended)
d. yɨk-ma-yɨš 'not demolishing'

If category changing derivational morphology is lexical, *-Im-* would not be able to appear with the unambiguously syntactic negatives. If it turns out that nominalization is a syntactic phenomenon derived by head-movement, (15c) would still be ruled by the fact that negation is at the Infl-level, whereas *-Im-* must be a much lower head that directly selects the verb.[7]

2.1.3. Modification by frequency adverbs.

Verbs with *-Iš-* can be modified by frequency adverbs, but those with true nominals cannot:

(16) a. [[Bu ulke]de bebek-ler-in sɨk sɨk öl-üš-ü]
this country-loc baby-plr-gen frequently die-Iš-agr
'Babies' frequently dying in this country'

b. *[[Bu ülke]de bebek-ler-in sɨk sɨk öl-üm-ü]
this country-loc baby-plr-gen frequently die-Im-agr
'Babies' frequent death in this country' (intended)

[7] The selection here must be direct since nominalizers are highly restrictive in terms of the verbs that they can convert to nouns, i.e., they are irregular and must be lexically specified.

(17) a. [Ahmet-in arada bir koš-uš-u]
A.-gen occasionally run-Iš-agr
'Ahmet's occasionally running'

b. *[Ahmet-in arada bir koš-u-su]
A.-gen occasionally run-I-agr
'Ahmet's occasional run' (intended)

This means that verbs lose their verbal properties with nominalizers *-Im-* in (16b) and *-I-* in (17b) (/s/ separates the two vowels when the second one is an agreement marker). In contrast, verbs remain verbal enough to be modified by frequency adverbs when they appear with *-Iš-*.

The evidence so far shows that *-Iš-* cannot be a true nominalizer, and that it must be an Infl-level element that preserves the verbal properties of verbs. To see that it is actually the gerundive morpheme, we must compare it with the other Infl elements.

***2.2. The infinitival status of* -mE- *and* -mEK.**

Verbs with *-mE-* and *-mEK* also maintain their verbal behavior: they assign structural Case; they co-occur with causatives, passives, and negatives; and they can be modified by frequency adverbials. An obvious difference between *-mEK* and all other Infl elements in Turkish is that only *-mEK* gives the citation forms of verbs. Actually, this is not a particularly strong argument that *-mEK* is the infinitive morpheme, since languages can differ on how they morphologically encode the citation forms.

In this section, I will be mainly concerned about forming a natural link between *-mE-* and *-mEK*, and distinguishing them from the gerundive *-Iš-*. The evidence will be based on the internal and external distribution and the temporal interpretation of *-mE-*, *-mEK*, and *-Iš-*.

2.2.1. The external distribution of *-mE-* and *-mEK* inside VPs.

The distribution of *-mE-* and *-mEK* with verbs that select both is predictable: They select *-mE-* in contexts of subject-verb agreement, where the subject needs Case (lexical NPs, pro, or the operator-traces in relative clauses),[8] and *-mEK* in non-

---

[8] It is very likely that operator-traces in relative clauses are actually resumptive pronouns.

agreement contexts (PRO-control):[9]

(18) Ahmet-∅ [PRO araba kullan-may]ɨ ist-iyor-∅
A.-nom car use-mEK-acc want-pres-agr
'Ahmet wants to drive cars'

(19) Ahmet-∅ [Berna-nin araba kullan-ma-sɨ]nɨ ist-iyor-∅
A.-nom B.-gen car use-mE-agr-acc want-pres-agr
'Ahmet likes Berna to drive cars'

Also note that both sentences have the interpretation of an unrealized 'car-driving' event relative to the 'wanting' time (more on this in section 2.2.3). The semantics of both clause types and their complementary distribution motivate the conclusion that the core infinitival is *-mE-*, and that *-mEK* is composed of *-mE-* plus a final *-K-* (also see Kennelly 1990), whose Case and agreement correlation will be discussed in section 3. Under this view, verbs that select only *-mE-* or *-mEK* (but not both) can be reinterpreted as verbs specifying their complements as PRO-control or non-control contexts.

2.2.2. The external distribution of *-mE-* and *-mEK* inside NPs.

Unlike their distribution inside VPs, PRO-control clauses can only be headed by *-mE-* inside NPs (or DPs), where *-mEK* remains very marginal, and *-Iš-* is excluded. A lexical NP or pro is not allowed in these clauses:

(20) a. [Ahmet-in [PRO koš-ma] çaba-sɨ]
A.-gen run-mE attempt-agr
'Ahmet's attempt to run'

b. [Ahmet-in [PRO uyu-ma] dileğ-i]
A.-gen sleep-mE wish-agr
'Ahmet's wish to sleep'

(21) a. ?*[Ahmet-in [PRO koš-mak] çaba-sɨ]
A.-gen run-mEK attempt-agr
'Ahmet's attempt to run'

[9] The /y/ corresponding to /k/ in (18) is due to a regular phonological process that will be discussed later on. Note also, that *iste* 'want' is unique in that it allows its subordinate PRO-control clause to appear without the accusative marking, which is actually the more preferred version. This process must somehow be related to the non-specificity effects in Enç (1991b).

b. ?*[Ahmet-in [PRO uyu-mak] dileğ-i]
A.-gen sleep-mEK wish-agr
'Ahmet's wish to sleep'

(22) a. *[Ahmet-in [PRO koš-uš] çaba-sɨ]
A.-gen run-Iš attempt-agr
'Ahmet's attempt in running' (intended)

b. *[Ahmet-in [PRO uyu-yuš] dileğ-i]
A.-gen sleep-Iš wish-agr
'Ahmet's wish of sleeping' (intended)

Recall that Turkish has no strategy to provide Case to the complement of a noun.[10] The nominal behavior of *-Iš-* clauses in (22) with respect to Case suggests that they are gerunds ruled out by the Case Filter, assuming along the lines of Abney (1987) that gerunds are IPs under DPs.[11] Since the *-mE-* clauses in (20) need not be Case-marked, they cannot be gerunds, contra Underhill (1976) and much subsequent work. The *-mEK* clauses in (21) have a status that is intermediate betwen the *-mE-* clauses in (20) and the *-Iš-* clauses in (22). This will have an explanation in section 5.

<u>2.2.3. The temporal interpretation of *-mE(K)* and *-Iš-*.</u>

The temporal location of events denoted by verbs in *-mEK* clauses is determined by the higher tense. Verbs in *-Iš-* clauses on the other hand, typically denote events that have occured by the matrix event time. This is parallel to the case of infinitives and gerundives in English, where the contrast becomes more apparent with verbs like *remember* and *forget:*[12]

---

[10] Although non-specific complement nouns are licensed inside nominals in Turkish, the facts are a little more complex since the type of noun seems to make a difference in cases like:

(i) [Ahmet-in uyku arzu-su]
A.-gen sleep desire-agr
'Ahmet's desire of sleep'

(ii) *[Ahmet-in košu arzu-su]
A.-gen run desire-agr
'Ahmet's desire of run'

(iii) *[Ahmet'in uyku dileğ-i]
A.-gen sleep wish-agr
'Ahmet's wish of sleep'

(iv) *[Ahmet'in košu dileğ-i]
A.-gen run wish-agr
'Ahmet's wish of run'

Thus, whichever way non-specifics are licensed inside NPs (partitive Case or incorporation), the mechanism must be sensitive to the type of nouns involved in the structure.

[11] Note that this account works only if gerundive DPs are specific in the relevant sense.

[12] The verbs *hatɨrla* 'remember' and *unut* 'forget' select only PRO-control infinitives, and agreeing gerunds, hence the absence of a complete contrast in terms of PRO and pro. They also select tensed clauses, i.e. *-DIK-*and *-EcEK-*, whose interpretation is much closer to *-Iš-* clauses, although not identical.

(23) a. Ahmet-∅ [PRO ev-e git-mey]i hatɨrla-yacak-∅
A.-nom home-dat go-mEK-acc remember-fut-agr
'Ahmet will remember to go home'

b. Ahmet-∅ [*pro* ev-e gid-iš-i]ni hatɨrla-yacak-∅
A.-nom home-dat go-Iš-agr-acc remember-fut-agr
'Ahmet will remember going home'

(24) a. Ahmet-∅ [PRO Berna-yɨ öp-mey]i hep unut-uyor-∅
A.-nom B.-acc kiss-mEK-acc always forget-prs-agr
'Ahmet always forgets to kiss Berna'

b. Ahmet-∅ [*pro* Berna-yɨ öp-üš-ü]nü hep unut-uyor-∅
A.-nom B.-acc kiss-Iš-agr-acc always forget-pres-agr
'Ahmet always forgets kissing Berna'

Ahmet will remember performing the act of going home in (23a), and he always forgets performing the act of kissing Berna in (24a). On the other hand, he will remember a certain instance of going home in (23b), and he always forgets a certain instance of kissing Berna in (24b). It is likely that these readings come from the semantics of control infinitives and gerunds; e.g., the unrealized event interpretation of *-mEK* squares well with the future modality analysis of infinitives by Bresnan (1972) and Stowell (1982).

Another distinction is that *-mE(K)* clauses are not definite, whereas events denoted in *-Iš-* clauses are existentially presupposed, similar to the Poss-ing gerunds of English (Portner, 1992).

(25) a. [Ahmet-in yarɨn evlen-me-si]∅ herkes-i šašɨrt-acak-∅
A.-gen tomorrow marry-mE-agr-nom everyone-acc surprise-fut-agr
'For Ahmet to marry tomorrow will surprise everyone'

b. [Ahmet-in yarɨn evlen-iš-i]∅ herkes-i šašɨrt-acak-∅
A.-gen tomorrow marry-Iš-agr-nom everyone-acc surprise-fut-agr
'Ahmet's marrying tomorrow will surprise everyone'

It is not presupposed in (25a) that Ahmet will actually have a wedding tomorrow. In (25b) by contrast, it is necessarily understood as an event that is known to have been scheduled for tomorrow. These interpretations are consistent with the analysis of *-Iš-* as the gerundive and *-mE(K)* as the infinitive morpheme.

The differences will be pointed out later on in section 2.3.3.

2.3. The tense status of -DIK- *and* -EcEK-.

Verbs with *-DIK-* and *-EcEK-* are also Case assigners; they co-occur with causatives, passives, and negatives; and they can be modified by frequency adverbials. In this section, I will analyze *-DIK-* and *-EcEK-* as past and future tense morphemes on the basis of their temporal reference, and compare them with the infinitive *-mE(K)* and the gerundive *-Iš-*.

2.3.1. Temporal reference and morphological similarities.

First, note that the subordinate Infl *-EcEK-* locates its event in the future relative to the matrix tense, without any sequence of tense effects:

(26) *pro* [Ahmet-in uyu-yacağ-ɨ]nɨ anla-dɨ-m
1.sg A.-gen sleep-EcEK-agr-acc realize-past-agr
'I realized Ahmet would be sleeping'

The subordinate sleeping event is temporally fixed at a time that is in the future of the matrix realization event. The subordinate *-EcEK-* has always been analyzed in the literature as an Infl with some future value.[13] This is not so surprising since the matrix future tense morpheme is also *-EcEK-*:

(27) Ahmet-∅ uyu-yacak-∅
A.-nom sleep-fut-agr
'Ahmet will sleep'

It is reasonable to assume then, that the subordinate *-EcEK-* is the same as the matrix *-EcEK-*, and that they are both future tense morphemes.

Second, note that the subordinate Infl *-DIK-* provides past reference as does the matrix past tense *-DI-*. If *-DIK-* is composed of the matrix past *-DI-* and an additional *-K-*, *-EcEK-* and *-DIK-* would both be [+ tense] Infls, which explains why verbs that select for one also select for the other.[14]

A complicating factor here is that *-DIK-* also has a covert present reading. That is, it places the subordinate event either in the past relative to the matrix event, or simultaneous with it:

---

[13] For George and Kornfilt (1981), *-EcEK-* is a future gerund; for Kornfilt (1984), it is a future factive nominal; and for Kennelly (1990), it is [+ future] aspect marker.

[14] There are some exceptions, such as *um* 'hope', but it is plausible that the semantics of these verbs require a future complement event, i.e., one can only hopes about future events.

(28) Ahmet-∅ [Berna-nɨn uyu-duğ-u]nu san-ɨyor-∅
A.-nom B.-gen sleep-DIK-agr-acc think-pres-agr
a. 'Ahmet thinks Berna slept'
b. 'Ahmet thinks Berna is sleeping'

(29) Ahmet-∅ [Berna-nɨn uyu-duğ-u]nu anla-dɨ-∅[15]
A.-nom B.-gen sleep-DIK-agr-acc realize-past-agr
a. 'Ahmet realized Berna had slept'
b. 'Ahmet realized Berna was sleeping'

(30) Ahmet-∅ [Berna-nɨn uyu-duğ-u]nu anla-yacak-∅
A.-nom B.-gen sleep-DIK-agr-acc realize-fut-agr
a. 'Ahmet will realize that Berna was sleeping'
b. 'Ahmet will realize that Berna is sleeping'

The (a) interpretations are the true past readings of *-DIK-*, where the subordinate sleeping time is in the past relative to the matrix thinking and realizing time. So in (30a) for instance, the sleeping event is temporally vague with respect to the utterance time: it may take place during, before, or after the utterance time as long as it preceeds the matrix realizing time.

In the (b) interpretations on the other hand, the subordinate sleeping event is simultaneous with the matrix thinking and realizing events. This is the covert present tense reading of *-DIK-* that motivates Kennelly (1990) to claim that *-DIK-* is a [- future] aspectual marker. Setting aside a full analysis of the temporal reference of *-DIK-*, two quick comments are in order: a) Unlike ordinary aspect, *-DIK-* does not refer to the internal logic of an event, e.g., its type or stage, its completion, duration, recurrence, so forth. Rather, it simply temporally orders a subordinate event with respect to its matrix event. b) The covert present interpretation of *-DIK-* can be understood in terms of 'variable-tense binding' most recently discussed by Ogihara (1989), Enç (1991a), and Stowell (1992), as long as it is ambiguous between past and variable-tense readings. Thus, it can be assumed that the subordinate *-DIK-* is the matrix past *-DI-* with an additional *-K-*.

2.3.2. The interpretation of *-DIK-* and *-EcEK-* vs. *-mE(K)*.

With a verb like *söyle* 'tell', clauses containing *-DIK-* and *-EcEK-* are interpreted as reporting a past, present, or future fact, see (31); while those with

[15] *Anla* 'realize' would be incompatible with the present tense in (28) for aspectual reasons.

-mE(K) are interpreted as indirect imperatives, ordering the lower subject to perform the action denoted by the verb (32):[16]

(31) a. Ahmet Berna-ya [*pro* okul-a git-tiğ-i]ni söyle-di-∅
(nom) B.-dat 3.sg school-dat go-DIK-agr-acc tell-past-agr
'Ahmet told Berna that he/she went to the school'

b. Ahmet Berna-ya [*pro* okul-a gid-eceğ-i]ni söyle-di-∅
(nom) B.-dat 3.sg school-dat go-EcEK-agr-acc tell-past-agr
'Ahmet told Berna that he/she will go to the school'

(32) Ahmet Berna-ya [*pro* okul-a git-me-si]ni söyle-di-∅
(nom) B.-dat 3.sg school-dat go-mE-agr-acc tell-past-agr
'Ahmet told Berna to go to the school'

This is exactly the situation in English, indicating that there is a semantic correlate between verbs of saying and the imperative reading of infinitives.

Also, *-DIK-* and *-EcEK-* clauses that are complements to verbs like *remember* and *forget* are interpreted as reporting temporally ordered events; while *-mE(K)* clauses refer to the performance of these acts:

(33) a. Ahmet [*pro* okul-a git-tiğ-i]ni unut-tu-∅
(nom) 3.sg school-dat go-DIK-agr-acc forget-past-agr
'Ahmet forgot that he went to school'

b. Ahmet [*pro* okul-a gid-eceğ-i]ni unut-tu-∅
(nom) 3.sg school-dat go-EcEK-agr-acc forget-past-agr
'Ahmet forgot that he would go to school'

(34) Ahmet [PRO okul-a git-me]yi unut-tu-∅
(nom) school-dat go-DIK-agr-acc forget-past-agr
'Ahmet forgot to go to school'

In (33a), Ahmet went to school, but later forgot about that fact. In (33b), he forgot the fact that he would later go to school, and in (34), he forgot to perform the act of going to school. The latter two are different in that (33b) is compatible with a world in which Ahmet ultimately went to school at the time he was supposed to be there, as is (33a), but not (34). The readings here are more or less identical to the readings one gets with their equivalents in English in tensed and infinitival clauses.

[16] The indirect object (*Berna'ya* in (31)) is optional with *söyle* in *-DIK-* and *-EcEK-* clauses. It is glossed as 'tell' when the indirect object is present, and as 'say' when it is absent.

2.3.3. The interpretation of *-DIK-* vs. *-Iš-*.

Recall from section 2.2.3, that events denoted in *-Iš-* clauses are necessarily presupposed, meaning that they must have occurred before the utterance time (though we have seen in (25b) that this is not always true):

(23) b. Ahmet-∅ [*pro* ev-e gid-iš-i]ni hatırla-yacak-∅
A.-nom 3.sg home-dat go-Iš-agr-acc remember-fut-agr
'Ahmet will remember going home'

(24) b. Ahmet-∅ [*pro* Berna-yı öp-üš-ü]nü hep unut-uyor-∅
A.-nom 3.sg B.-acc kiss-Iš-agr-acc always forget-pres-agr
'Ahmet always forgets kissing Berna'

Sometimes this holds for *-DIK-* clauses as well, but even in such cases, they are not interpreted differently. The act of going home and kissing Berna for instance, must have occurred respectively in (35a-b) also:

(35) a. Ahmet-∅ [*pro* ev-e git-tiğ-i]ni hatırla-yacak-∅
A.-nom 3.sg home-dat go-DIK-agr-acc remember-fut-agr
'Ahmet will remember he went home'

b. Ahmet-∅ [*pro* Berna-yı öp-tüğ-ü]nü hep unut-uyor-∅
A.-nom 3.sg B.-acc kiss-DIK-agr-acc always forget-prs-agr
'Ahmet always forgets that he kissed Berna'

The difference is the following: In (35a), he will remember only a fact, so it is possible that all he will remember is that he arrived home that day but have no idea how he got there. In (23b), he will remember the incident, so he will have to remember every subpart of the going home event. Similarly, in (35b), he forgets that the kissing event ever happened, so he must have forgotten everything about the kiss. In (24b), he may remember that the kissing event occurred, but have no memory how it happened or how it felt.

The difference can also be observed with verbs of perception, where the *-Iš* - clause must denote the type of thing that is directly perceivable:

(36) *pro* [Ahmet-in [bu kitab]ı oku-yuš-u]nu duy-du-m
1.sg A.-gen this book-acc read-Iš-agr-acc hear-past-agr
'I heard Ahmet's reading this book'

In (36), I physically heard Ahmet's voice reading this book aloud,[17] while in (37), all I heard was the fact that Ahmet had read this book at one point in his life, which he may have done quietly and on his own.

In the following, I will assume that the evidence presented so far is conclusive in making the case for the classification (2) in the introduction.

3. The morpheme *-K-*.

The analysis that the subordinate Infls *-mEK* and *-mE-* are the same element, and that the subordinate *-DIK-* contains the matrix *-DI-*, leaves the curious segment *-K-* unidentified. It occurs with the infinitival *-mE-* in PRO-control, and the past tense *-DI-* in subordinate contexts. I will argue in this section that *-K-* is the C^0 (COMP) category. Before that though, I will discusss the phonological evidence that the future *-EcEK-* must also have an additional *-K-* in subordinate contexts, but not in main clauses.

In Turkish, intervocalic /k/ regularly becomes the velar semivowel /ɰ/ (orthographically 'ğ') after the first binary foot from the left. This is a very unstable segment that may become /y/ or /w/, or it may lengthen the preceding vowel and then delete, as in the context of *-EcEK-*. The matrix *-EcEK-* further undergoes an optional rule that shortens the long vowel created by the ɰ-deletion (in addition to an /a/ to /u/ raising below):

(38) *pro* koš-acağ-ɨm
1.sg run-fut-agr
'I will run'

pronounced: a. ?/koʃudʒaam/
b. /koʃudʒam/

In fact, (38b) is the preferred version. Notice now that this optional vowel shortening does not apply in the subordinate *-EcEK-*:

(39) [*pro* koš-acağ-ɨm]∅ belli
1.sg run-EcEK-agr-nom obvious
'That I will run is obvious'

[17] It is possible that the semantics of gerunds is derived by an operator that semantically nominalizes an event to create an expression that denotes an (abstract) object in universe.

(39) [*pro* koš-acağ-ɨm]∅ belli
1.sg run-EcEK-agr-nom obvious
'That I will run is obvious'

pronounced: a. /koʃudʒaam belli/
b. */koʃudʒam belli/

Since the vowel shortening in (38b) is blocked in the subordinate context, the underlying form of the verb must be phonologically distinct in (39). The vowel shortening rule targets the environment of /k/ in *-EcEK-*, so something in that area must hold the two forms distinct. An additional *-K-* after the subordinate *-EcEK-* in (39) would set the two forms apart.[18]

In the following, I will assume that there is an additional *-K-* with the the subordinate *-EcEK-* but not with the matrix *-EcEK-*, i.e., *-EcEKK-* vs. *-EcEK-*, and proceed to explain why this *-K-* must be the C^0 category. Note that this correctly removes the gerundive *-Iš-* from the picture under Abney's (1987) hypothesis that they are IPs subordinated under DPs.

3.1. The ECM configuration.

A limited number of verbs in Turkish display the same type of ECM behavior observed with the *believe* class of verbs in English, where the subject of the lower clause receives accusative Case from the higher verb. The difference is that the ECM verbs in Turkish select only tensed clauses:

(40) Ahmet-∅ [Berna-nɨn uyu-du-ğ-u]nu san-ɨyor-∅[19]
A.-nom B.-gen sleep-past-comp-agr-acc think-pres-agr
'Ahmet thinks Berna slept'

(41) Ahmet-∅ [Berna-**yɨ** uyu-du] san-ɨyor-∅
A.-nom B.-**acc** sleep-past think-pres-agr
'Ahmet thinks Berna slept'

---

[18] Although Turkish allows geminates, there is no word-internal /kk/ sequence in precisely the context of intervocalic /k/ → /ɰ/ (barring noun compounds). There are some instances of /k/ + /k/ concatenation in the phonology that do not form a single word, such as:

(i) *pro* o kadar acɨk-tɨ-k ki
1.pl that much get.hungry-past-agr ?
'We got so hungry'

The sequence /adʒɨktɨk#ki/ does not define a word: a) *ki* does not bear stress (Turkish is a final syllable stress language); and, b) the vowel in *ki* does not harmonize.

[19] From here on, I will use the proposed labels for the subordinate Infls.

There are two dialects in Turkish with respect to the ECM in (41). The so-called 'Dialect A', which is what Kornfilt (1977) investigates, does not allow agreement in the lower clause, while the 'Dialect B' does (see Brendemoen and Csato 1984). As the basis for discussion, I will take the judgements in Dialect B, where the lower agreement is optional:

(42) Ahmet-∅ [ben-**i** uyu-du-**m**/∅] san-ɨyor-∅
A.-nom I-**acc** sleep-past-**agr** think-pres-agr
'Ahmet thinks I slept'

In both dialects, the lower subject may get the nominative Case instead of the accusative as long as it agrees with the lower verb:[20]

(43) Ahmet-∅ [ben-∅ uyu-du-**m**/*∅] san-ɨyor-∅
A.-nom I-**nom** sleep-past-**agr** think-pres-agr
'Ahmet thinks I slept'

The relevant point here is that in the more liberal Dialect B, ECM applies only when the *-K-* is missing in the structure, as in (42), regardless of the status of the lower agreement. Assuming ECM is possible just in case of CP-deletion, such a direct correlation between the ECM and the absence of *-K-* indicates that it must be the C^0 head of the subordinate CP.

3.2. Case-marking in infinitives.

If we take the missing *-K-* as an indication for the missing CP, the distribution of *-mE-* and *-mEK* would have a natural interpretation under Raposo's (1987) theory of Case assignment. This system was designed for the obligatory I-to-C movement with the agreeing infinitives of European Portuguese. He motivates this movement by assuming that an infinitival Infl (T^0 or Agr) cannot assign Case to [Spec, TP] unless it is Case-marked by the higher verb. The I-to-C movement of the verb enables Infl to be Case-marked at C^0.

Raposo's idea can be worked out for the situation at hand, provided that an I-to-C movement somehow fails to produce the sufficient conditions for assigning Case to the lower subject in Turkish, and a missing CP solves this problem.[21] This means that CP should remain intact in PRO-control structures since PRO does not need to

[20] The structure in (43) is termed as the Direct Complement Clauses in George and Kornfilt (1981), and Kornfilt (1984), and as Structure C by Kennelly (1992).
[21] In section 5.2, I will discuss the possibility that the C^0 is empty in these cases.

be Case-marked in the first place:

(44) Ahmet-∅ [PRO düš-me-k]ten kork-uyor-∅
A.-nom fall-inf-comp-abl fear-pres-agr
'Ahmet is afraid to fall'

When the subject needs to be Case-marked on the other hand, the CP must be absent because Case is unavailable TP-internally, and will remain so unless the subordinate Infl is Case-marked by the higher verb:

(45) Ahmet-∅ [Berna-nɨn düš-me-si]nden kork-uyor-∅
A.-nom B.-gen fall-inf-agr-abl fear-pres-agr
'Ahmet is afraid for Berna to fall (approximately)'

A possible reason why this cannot be accomplished under an I-to-C movement will be given in section 5.2. What matters here is that the data above are consistent with the proposal that *-K-* is the C^0 in Turkish.

3.3. Disjoint reference and -K-.

The absence of *-K-* has the effect of extending the disjoint reference domain of the lower subject pronoun one clause up:

(46) a. Ahmet-$\emptyset_i$ [*pro$_i$* Ankara-ya git-ti-ğ-i]ni san-ɨyor-∅
A.-nom 3.sg A.-dat go-past-comp-agr-acc think-prs-agr
'Ahmet thinks he went to Ankara'

b. *Ahmet-$\emptyset_i$ [*pro$_i$* Ankara-ya git-ti] san-ɨyor-∅
A.-nom 3.sg A.-dat go-past think-prs-agr
'Ahmet thinks he went to Ankara'

(47) *Ahmet-$\emptyset_i$ [*pro$_i$* Ankara-ya git-me-si]ni ist-iyor-∅
A.-nom 3.sg A.-dat go-inf-agr-acc want-pres-agr
'Ahmet wants him to go to Ankara'

It is unclear how the binding domain of pronouns is computed in Turkish.[22] Whichever way this is achieved though, the above shows that *-K-* must be a part of the definition. Assuming it is the C^0, the CP its presence indicates in (46a) can be plausibly determined as the maximal projection that closes off the binding domain for pronouns. Its absence in (46b) and (47) seems to make the first maximal projection dominating the clause (the matrix VP) the binding domain, which would derive the disjoint reference effects under the VP-internal subject hypothesis (e.g.,

[22] See George and Kornfilt (1981), Kornfilt (1984), Enç (1985), Kennelly (1991) among others.

Koopman and Sportiche, 1990).

3.4. V-to-C movement and its consequences.

Taking *-K-* as the C^0 category that follows T^0 in the verbal complex entails the movement of V^0 to C^0 in Turkish, which explains the following.

3.4.1. Case-marking on the clause.

The verbal complex that consists of V-T-C must be in the highest head position in order to receive the Case-marking:

(48) Ahmet-∅ [*pro* ev-e koš-tu-ğ-um]u bil-iyor-∅
A.-nom 1.sg home-dat run-pst-comp-agr-acc know-prs-agr
'Ahmet knows that I ran home'

(49) Ahmet-∅ [*pro* ev-e koš-tu-ğ-um]a inan-ɨyor-∅
A.-nom 1.sg home-dat run-pst-comp-agr-dat believe-prs-agr
'Ahmet believes that I ran home'

Leaving the precise mechanics of it aside, what is relevant here is that Case is assigned locally within the higher clause, and the verbal complex V-T-C must be high enough to be accessible to the Case morphology. This would not have been possible if the verbal complex were not at the highest head position of the clause, and there were a head between the final position of the verb and the head of the XP that is Case-marked by the higher verb.

3.4.2. Postverbal constituents.

Under neutral intonation, material inside a clause cannot scramble to the postverbal position in subordinate contexts in Turkish:

(50) Ahmet-∅ t_i git-ti-∅ okul-a_i
A.-nom go-past-agr school-dat
'Ahmet went to school'

(51) *Berna [[Ahmet-in t_i git-ti-ğ-i]ni okul-a] duy-du-∅
B.-nom A.-gen go-pst-comp-agr-acc school-dat hear-pst-agr
'Berna heard that Ahmet went to school'

There is extensive evidence that postverbal constituents are CP-adjoined in Turkish (Kural 1992). Moving the verb to the highest head at S-structure, C^0, would force postposed elements to unambiguously adjoin the highest projection, CP, instead of TP. CP-adjunction can be ruled out in (51) by the general prohibition against adjunction to arguments (Chomsky 1986).

3.4.3. The absence of CED effects with subjects.

A verb that has moved to C^0 would govern the subject position, [Spec, TP], and nullify the CED effects for subjects (Huang 1982). The following shows that this is what happens in Turkish:

(52) [Op_i [Ahmet-in t_i git-me-si]nin ben-i üz-dü-ğ-ü] ev
A.-gen go-inf-agr-gen I-acc sadden-pst-cmp-agr house
'The house [which [that Ahmet went to t] saddened me]'

(53) [Op_i [*pro* [[t_i anne-si]nin herkes-le konuš-tu-ğ-u]nu
1.sg mother-agr-gen everyone-with talk-pst-cmp-agr-acc
duy-du-ğ-um] adam
hear-past-comp-agr man
'The man [who I heard [that [t mother] talked to everyone]]'

If the V^0s *üzdüğü* 'saddened' in (52) and *konuštuğu* 'talked' in (53) are at their respective C^0s, they can properly head-govern [Spec, TP] and allow the relative clause operator to extract from both subjects. In terms of Chomsky (1986), the verbs would be able to L-mark their subjects from C^0 since they are also thematically linked to the verb. The subject CP in (52) and NP (or DP) in (53) would then no longer be barriers to the movement of the operator and/or the proper head-government of the traces it creates.

3.4.4. Subject Negative Polarity Item (NPI) licensing.

Subject NPIs are licensed by negation on the verb in both main and subordinate contexts in Turkish:

(54) a. Kimse-∅ koš-ma-dɨ-∅
noone-nom run-neg-past-agr
'Noone ran'

b. *Kimse-∅ koš-tu-∅
noone-nom run-past-agr
'Noone ran'

(55) a. *Ahmet-∅ [kimse-nin koš-tu-ğ-u]nu san-ɨyor-∅
A.-nom noone-gen run-past-comp-agr-acc think-pres-agr
'Ahmet thinks noone ran'

b. Ahmet-∅ [kimse-nin koš-ma-dɨ-ğ-ɨ]nɨ san-ɨyor-∅
A.-nom noone-gen run-neg-pst-cmp-agr-acc think-pres-agr
'Ahmet thinks noone ran'

Without specifics, I will base the discussion here on the observation that NPI-licensing requires S-structure c-command. The NPI *kimse* 'noone' in (54b) and (55a) needs to be licensed by negation through c-command, and the raising data below show that NPIs are not licensed once they move outside the scope of negation to a Case position at S-structure:

(56) a. Kimse-$\varnothing_i$ [t_i vur-ul-du] san-ɨl-m-ɨyor-∅
noone-nom shoot-pas-past think-pas-neg-pres-agr
'Noone is thought to have been shot'

b. *Kimse-$\varnothing_i$ [t_i vur-ul-ma-dɨ] san-ɨl-ɨyor-∅
noone-nom shoot-pas-neg-past think-pas-pres-agr
'Noone is thought to have been shot'

This implies that the negation picked up by the verb as it moves ends up at a position where it can c-command the subject in [Spec, TP] in both main and subordinate clauses in Turkish, which is predicted under V-to-C movement.

Kornfilt (1984) observes that subject NPIs can only be licensed by the higher negation in ECM clauses, but not by the lower negation:

(57) a. *Ahmet-∅ [kimse-yi koš-ma-dɨ] san-ɨyor-∅
A.-nom noone-nom run-neg-past think-pres-agr
'Ahmet thinks noone ran'

b. Ahmet-∅ [kimse-yi koš-tu] san-m-ɨyor-∅
A.-nom noone-nom run-past think-neg-pres-agr
'Ahmet does not think anyone ran'

This has two possible explanations: a) after CP-deletion, the verb cannot go as high and c-command the subject NPI in [Spec, TP]; or, b) accusative Case in ECM is assigned in the higher clause, and the lower NPI subject ends up higher than negation when it moves out of [Spec, TP] to the matrix clause. Either way, negation would fail to c-command the subject NPI in (57a).

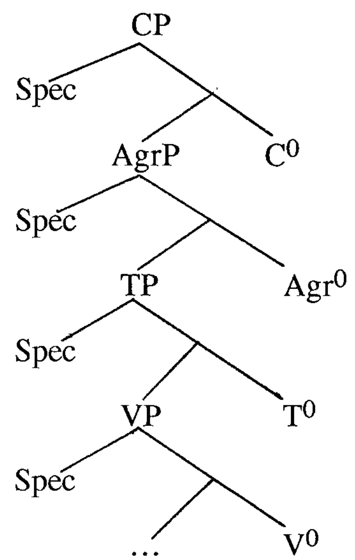
4. The status of subordinate agreement.

Consider the order of morphemes on the subordinate verb below:

(58) *pro* [Ahmet-in koš-**tu**-**ğ**-**u**-]nu bil-iyor-um
1.sg A.-gen run-**past**-**comp**-**agr**-acc know-pres-agr
'I know that Ahmet ran'

The Agr *-u-* appears as the outmost element in the clause, followed only by the Case morpheme *-nu-*. The issue here is how Agr can be outside C^0 if (59) is the basic clausal configuration (Belletti 1990, Chomsky 1991):

(59)



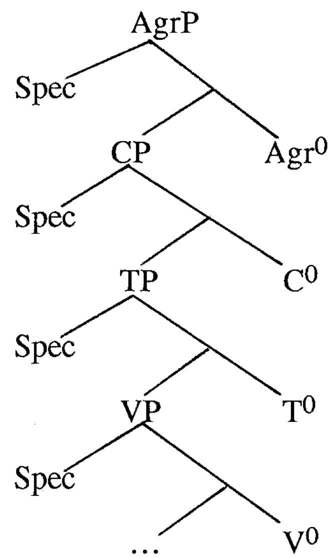
This structure assumes with Pollock (1988) that Agr is a syntactic head that projects an AgrP. The V-T-C-Agr order in (58) could be derived from (59) only if Agr^0 moves to C^0 independent of the V-T complex. However, this would violate not only the Head Movement Constraint (HMC) by having V-T skip Agr^0 to get to C^0, but also the Mirror Principle (MP; Baker 1985) since the order of V-T-C-Agr in the morphology of (58) would not reflect the syntactic order of V-T-Agr-C in (59).

I will consider various possible solutions to this paradox in the following section, and show that they are not viable. This will lead to the conclusion that the V-T-C-Agr order in Turkish is empirical evidence that Agr is not an independent head that projects an AgrP.

4.1. Reversing AgrP and CP.

One possible strategy would be to maintain the HMC and the MP and change the structure in (59) by reordering AgrP and CP:

(60)



This representation faces four main problems.

First, suppose that the current thinking is correct, and subject-verb agreement is an instance of Spec-head relationship established at AgrP. This means that subjects must be in [Spec, AgrP], which is higher than the [Spec, CP] in (60). However, object Wh-phrases take unambiguous scope over subject QPs in Turkish, i.e., no family of questions reading below:

(61) Herkes-∅ kim-i gör-dü-∅?
everyone-nom who-acc see-past-agr
'Who did everyone see?'

a. For which x, x a human, everyone saw x?
b. *For every y, y a human, who did y see?

(62) *pro* [herkes-in kim-i gör-dü-ğ-ü]nü sor-du-m
1.sg everyone-gen who-acc see-past-comp-agr-acc ask-pres-agr
'I asked who everyone saw'

a. I asked for which x, x a human, everyone saw x.
b. *I asked for every y, y a human, who y saw.

If the subject QPs were in the [Spec, AgrP] of (60), they would take scope over the object Wh-phrases in the [Spec, CP] at LF. This would incorrectly derive the (b) readings in (61) and (62), even though subject QPs cannot take scope over object Wh-phrases in Turkish, ambiguously or otherwise. This suggests that subjects must be lower than [Spec, CP] at LF.

Second, note that under neutral intonation, material inside a clause can scramble to the presubject position in subordinate contexts, but not to the postverbal position (Kural 1992):

(63) Berna-∅ dün [okul-a$_i$ Ahmet-in t$_i$ gid-eceğ-∅-i]ni
B.-nom yesterday school-dat A.-gen go-fut-comp-agr-acc
duy-du-∅
hear-past-agr
'Berna heard yesterday that Ahmet would go to school'

(51) *Berna [[Ahmet-in t$_i$ git-ti-ğ-i]ni okul-a] duy-du-∅
B.-nom A.-gen go-pst-comp-agr-acc school-dat hear-pst-agr
'Berna heard that Ahmet went to school'

Recall from section 3.4.2 that (51) is ruled out because verbs are in the highest head position at S-structure, forcing postverbal elements to adjoin the highest clausal projection, i.e., AgrP in (60), CP in (59). Preposed constituents would have been ruled out for the same reason if subjects were at the highest specifier position, i.e., [Spec, AgrP] in (60), [Spec, CP] in (59). However, this cannot be the case as (63) is well-formed.

Third, consider the binding facts briefly mentioned in connection with (46) and (47) in section 3.4, repeated below.

(46) a. Ahmet-∅$_i$ [*pro$_i$* Ankara-ya git-ti-ğ-i]ni san-ıyor-∅
A.-nom 3.sg A.-dat go-past-comp-agr-acc think-prs-agr
'Ahmet thinks he went to Ankara'

b. *Ahmet-∅$_i$ [*pro$_i$* Ankara-ya git-ti] san-ıyor-∅
A.-nom 3.sg A.-dat go-past think-prs-agr
'Ahmet thinks he went to Ankara'

(47) *Ahmet-∅$_i$ [*pro$_i$* Ankara-ya git-me-si]ni ist-iyor-∅
A.-nom 3.sg A.-dat go-inf-agr-acc want-pres-agr
'Ahmet wants him to go to Ankara'

The absence of *-K-* extends the binding domain of a lower subject pronoun. Under (60), the subject would be in [Spec, AgrP] which is higher than C^0, i.e. *-K-*. This would imply that the binding domain of the subject pronoun is computed on the basis of a projection lower in the structure, which is an unlikely situation. This indicates that subjects must be lower than C^0 (*-K-*).

Finally, recall from section 3.1 that Dialect B allows ECM even when agreement morphology is present in the structure:

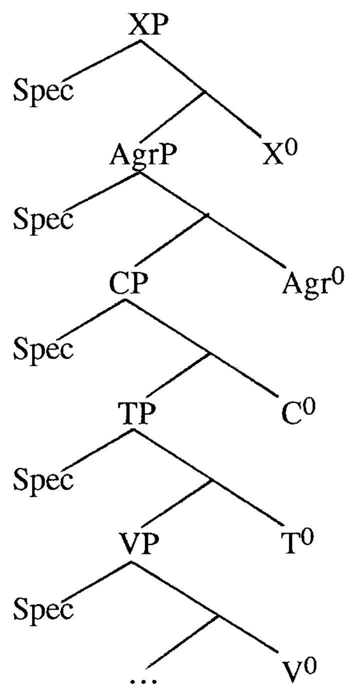
(42) Ahmet-∅ [ben-**i** uyu-du-**m**/∅] san-ıyor-∅
A.-nom I-**acc** sleep-past-**agr** think-pres-agr
'Ahmet thinks I slept'

CP cannot be deleted in (60) without an AgrP-deletion. Yet the subject may optionally agree with the verb in (42). If AgrP is where agreement is established, this means that it must be higher than the CP in the structure.[23]

4.2. An empty projection dominating AgrP.

Consider now (64), which is almost identical to (60), but with an extra projection XP dominating AgrP. It is crucial here that the head and specifier positions of XP both be empty, and that the V-T-C complex further move from Agr^0 to X^0 before it picks up the Case morphology:

(64)



This would handle the scrambling facts above: The Verb would be in the highest head position, the empty X^0, forcing postposed constituents to adjoin the maximal projection of the clause, XP. The subject on the other hand, would remain lower than the highest specifier position, the empty [Spec, XP], allowing constituents to be preposed in the subordinate context.

Note that this solution still requires the subject in [Spec, AgrP] to be higher

23 This may be a weak argument if ECM indicates CP-transparency instead of CP-deletion.

than the [Spec, CP], and we have seen above that this gives the wrong scope readings for subject QPs and object Wh-phrases. To fix this problem, we might assume that C^0 is the empty X^0 in (64) instead of the lower *-K-*. This way, object Wh-phrases would move to the [Spec, XP] for the correct scope readings. However, this would not solve all the problems.[24] Recall for instance, that agreement is possible in the ECM clauses of Dialect B:

(42) Ahmet-∅ [ben-**i** uyu-du-**m**/∅] san-ɨyor-∅
A.-nom I-**acc** sleep-past-**agr** think-pres-agr
'Ahmet thinks I slept'

Under the CP-deletion approach to ECM, the process would be removing both the null X^0 above the Agr^0 and the *-K-* (C^0) below the Agr^0 in (42) simultaneously, while leaving the Agr^0 itself behind.[25]

Any theory based on a configuration like (64) must provide a principled reason for the strict co-occurrence of the XP and the lower projection headed by *-K-*. The co-occurrence of X^0 and *-K-* is the one in which X^0 is absent from the structure in exactly the cases where *-K-* is absent, i.e., ECM clauses and infinitivals (sections 3.1 and 3.2). This is needed to explain why the binding domain of a pronoun in [Spec, AgrP] of (64) is extended only in these contexts, characteristically defined by the absence of *-K-* (section 3.3), and how the lower subject optionally gets the accusative Case when *-K-* is missing in tensed clauses (section 3.1).

It may be possible to devise a clever selectional procedure that would derive (42), where X^0 selects *-K-* across an AgrP. The question would still remain though, why X^0 occurs only in structures where *-K-* occurs, when we know that *-K-* is not a very selective C^0 in the first place, and that it occurs with both tensed and infinitival (PRO-control) clauses. In a way, the gist of this solution is to place CP <u>both</u> higher <u>and</u> lower than the AgrP, which is a roundabout way of saying that CP is lower than AgrP in its actual position, but it is selected higher than the AgrP. As such, this approach provides no real insight to the problem and remains to be an ad hoc solution.

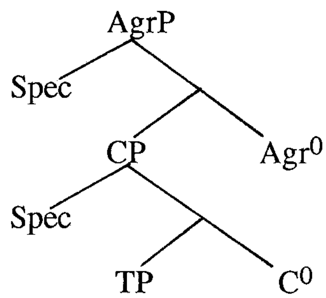
---

[24] For the purposes of discussion, I put aside the question what the XP or *-K-* stand for in either solution related to (64), as well as how the presence of the null XP is posited by the language learner. These questions are not trivial at all.
[25] It is not clear what CP-transparency would mean when X^0 is empty to begin with.

4.3. Agreement under government?

The final strategy that I will consider here is to maintain the order in (60), but to assume that agreement obtains under government:

(65)



In this scheme, the subject will have to be in [Spec, CP] so that Agr^0 can govern the subject to trigger agreement.

This proposal fails to derive the correct scope relations between subject QPs and object Wh-phrases since it places them both in [Spec, CP]. Apart from the issue of having two constituents at the same position, this system would also incorrectly assign equal scope to such QPs and Wh-phrases for an unavailable family of questions reading in (61) and (62).

Further, note that although the verb in the highest head position (Agr^0 in this case) properly head-governs a preposed constituent in terms of CED, it agrees only with the subject, and never with a preposed object:

(66) [Op$_i$ [*pro* [[t$_i$ Kitab-ɨ]nɨ$_j$ Ahmet-in t$_j$ oku-du-ğ-u]nu
1.sg book-agr-acc A.-gen read-past-comp-agr-acc
bil-di-ğ-im] adam
know-past-comp-agr man
'The man whose book I know Ahmet read'

(67) ?[Op$_i$ [*pro* [[Bu kitab]ɨ$_j$ [t$_i$ anne-si]nin t$_j$ oku-du-ğ-u]nu
1.sg this book-acc mother-agr-gen read-pst-cmp-agr-acc
bil-di-ğ-im] adam
know-past-comp-agr man
'The man whose mother I know read this book'

(68) Ahmet-∅ [[bu kitab-ɨ] ben-im oku-du-ğ-**um**]u
A.-nom this book-acc I-gen read-past-comp-**1.sg.agr**-acc
(*oku-du-ğ-**u**]nu) bil-iyor-∅
(read-past-comp-**3.sg.agr**-acc) know-prs-agr
'Ahmet knows that I read this book'

In (66) the preposed object is properly head-governed and can be extracted from. Extraction from subjects is slightly deviant under object preposing, but it is still acceptable, cf. (67). Verbs never agree with preposed objects in Turkish, cf. 3.sg agreement is unacceptable in (68), as opposed to the well-formed 1.sg agreement. This indicates that subject-verbs agreement must be established as a Spec-head relationship in a position reserved uniquely for subjects, such as [Spec, TP] or [Spec, VP].

4.4. The head status of Agr.

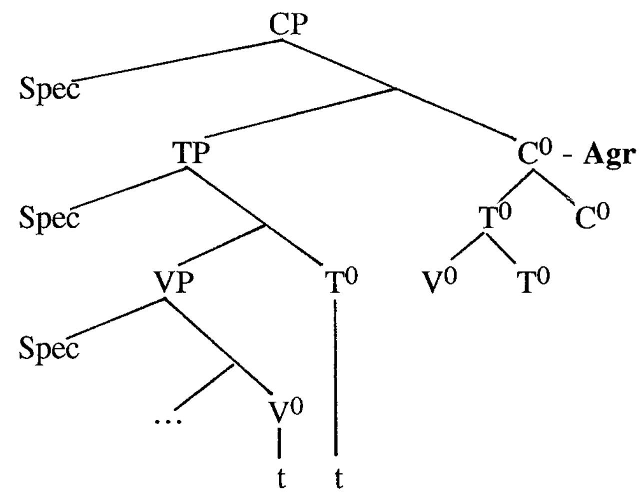
The essence of the paradox is that the verb appears to be located at the highest head position with Agr as its outmost morpheme, but the subject is at least one specifier position lower than that. Agreement needs to be established under a Spec-head relationship (section 4.3), but what triggers it must be lower than CP. The agreement morpheme however, appears outside the higher C^0, which violates the HMC and the MP.

It seems at this point that we are forced to loosen the HMC and MP to allow the lower Agr to appear higher than the C^0 in Turkish. The issue is then why the principles of UG should make exception to Agr. Claiming that Agr is an exception to the HMC and MP is equivalent to saying that it does not behave like syntactic heads. This implies that contra Pollock (1989) and Chomsky (1991), Agr is not an independent head in syntax, and as such, the HMC and MP do not apply to Agr. In a way, this paradox leads us back to the pre-Pollockian conception of Agr as a bundle of syntactic features.

Suppose now that the agreement morpheme is a spell-out of the Agr-features (or the index) picked up under Spec-head relationship at VP or TP in the syntax, and which in this case, are carried to C^0 by the verb via the V-to-C movement.[26] After successive head-movement, the structure will be as (69) below, where the bold-faced Agr results from the spell-out function:

[26] Alternatively, we can say that the C^0 is inserted as a fully-inflected morpheme at D-structure, and its features are checked at S-structure after the V-to-C movement.

(69)



The first question this raises is how the Agr-features of the subject picked up by the verb at a lower projection are accessible to the higher head C^0. This could be achieved by a feature percolation mechanism, whereby the adjoined head passes its features up to its host head. Such a mechanism is already needed for the licensing of subject NPIs, e.g., *kimse* 'noone', by the Neg^0 buried deep inside the verbal complex V-Neg-T-C-Agr:

(55) b.	Ahmet-∅	[kimse-nin	koš-ma-dɨ-ğ-ɨ]nɨ	san-ɨyor-∅
	A.-nom	noone-gen	run-neg-pst-cmp-agr-acc	think-pres-agr

'Ahmet thinks noone ran'

The second question is why Agr appears immediately after C^0, but not between T^0 and C^0. Quite possibly, this is how agreement operates in UG. That is, what agrees in natural languages is not the V^0 per se, but the verbal complex formed at S-structure by V-movement.[27] This means that the Agr-features picked up by the verb must be spelled out morphologically at the point where the verb (or Infl/Tense) reaches the final head within its clause at S-structure to form the language particular verbal complex (V-T-C in Turkish, V-T in English). In this view, Case morphology is outside Agr because it is not a clause-internal component participating in the formation of the verbal complex, but rather, it is a property of the entire clause.

This predicts that Agr must be the outmost morpheme on the verbal complex in languages with V-movement. A language in which Agr appears between the V^0

[27] The fact that agreement appears on *do* in English do-insertion suggests that the more precise formulation must involve the inflectional system rather than the verb itself.

and T^0, Neg^0, or C^0 would constitute a counter-example.

5. Related Issues.

Some of the issues related to the analyses above are addressed below.

5.1. The gerundive behavior of subordinate clauses.

I mentioned in the beginning that subordinate clauses display the gerundive characteristics in (3), repeated below:

(3) a. Subject bears the genitive Case in this context.
  b. Subject-verb agreement is in the nominal paradigm.
  c. All subordinate clauses are and must be Case-marked.

There is a longstanding tradition in the literature on Turkish equating these subordinate clauses with gerunds. This is made explicit by Kennelly (1990, 1992), who analyzes them as IPs subordinated under DPs (her Structure D) after Abney (1986). The counter-arguments in section 2 were mostly based on the clause-internal, temporal interpretations of *-Iš-*, *-mE(K)*, *-DIK-*, and *-EcEK-* clauses, i.e., properties related to their Infl nodes (T^0s). The fact remains though, that they behave like gerunds in the sense of (3).

I will follow Kennelly (1990) in taking the head of the projection that dominates the subordinate TPs as the locus of these properties, except for a minor difference. I will assume this head is a C^0 with a nominal character in Turkish unlike its Germanic or Romance counterparts. The difference may be merely terminological as it can be plausibly argued that what I call a nominal C^0 is the same thing as Kennelly's D^0. As long as the category dominating *-Iš-* clauses, which have the semantics of English Poss-ing gerunds, is distinct from the one dominating *-mE(K)*, *-DIK-*, and *-EcEK-* clauses, both systems would produce equivalent results in terms of (3).[28]

Like Kennelly's D^0, the nominal C^0 also needs to be Case-marked since it is a [+N] category. This would explain not only the Case-marking on the clause, i.e., the property (3c), but also why *-mEK* and *-Iš-* clauses cannot be complements to nouns,

[28] The raising of the subordinate subject from [Spec, IP] to [Spec, DP] in Kennelly (1990) is a separate issue. It is basically required by the specific conception of Case and agreement as an indication for an S-structure Spec-head relationship that is being contested here.

as shown in section 2.2.2.[29]

Next, consider the Case and agreement facts. The abstract Agr-features picked up by the verb could not discriminate between nominal and verbal agreement paradigms. However, the [+N] content of C^0 can define the domain of agreement as a nominal domain when the verb reaches C^0 and the Agr-features are spelled out (I will tentatively assume that this is a direct relationship, and revise it below).[30] This would derive the nominal agreement, (3b), and the genitive Case on the lower subject, (3a).

5.2. CP-deletion vs. empty C.

There are at least three reasons why one may claim that the CP is not missing in agreeing infinitives (*-mE-*), but rather the verb moves to some empty C^0 dominating the TP: a) The verbal complex in these clauses bears Case morphology. b) Material can be extracted from their subjects, (70). c) Negation on the infinitival verb licenses subject NPIs, as in (71):

(70) [Op_i [[[Ahmet-in t_i git-me-si]nin ben-i üz-me-si-]nin
A.-gen go-inf-agr-gen I-acc sadden-inf-agr-gen
herkes-i kɨzdɨr-dɨ-ğ-ɨ] ev
everyone-acc anger-past-comp-agr house
'The house [which [that [that Ahmet went to t] saddened me] angered everyone]'

(71) Ahmet-∅ [kimse-nin uyu-ma-ma-sɨ]nɨ ist-iyor-∅
A.-nom noone-gen sleep-neg-inf-agr-acc want-pres-agr
'Ahmet wants noone to sleep'

However, this proposal would fail to provide a natural explanation for the extension of binding domain for subject pronouns in (47) from section 3.3:

(47) *Ahmet-$\varnothing_i$ [*pro$_i$* Ankara-ya git-me-si]ni ist-iyor-∅
A.-nom 3.sg A.-dat go-inf-agr-acc want-pres-agr
'Ahmet wants him to go to Ankara'

For a uniform account of disjoint reference effects in Turkish, an empty C^0 would

[29] The fact that *-mEK* clauses are slightly better than *-I* - clauses inside NPs (DPs) might indicate that *-K-* is not a fully-nominal category.

[30] Note that for Case to be assigned from C^0 to [Spec, TP] without involving the mechanism of government (section 4.3), agreement must entail coindexation and should not rely entirely on feature sharing, so that the Case assigner (whichever it may be, T^0 or Agr) can assign the Case to the index, and have it received by the subject.

have to behave like the CP-deletion/ECM configuration in pulling the binding domain of subject pronouns one clause up. It is not obvious how and why the lexical content of C^0 should have such an effect in Turkish, but not in English (cf. *that*-deletion clauses).

On the other hand, the CP-deletion analysis can handle these facts as follows. a) The verbal complex V-T bears Case morphology because an infinitival Infl cannot assign Case to subject on its own (Raposo 1987), and it needs to be Case-marked by the higher verb. b) The subject position is properly head-governed by the higher verb or its trace (Stowell 1981, Chomsky 1986), which voids the CED effects in (70). c) We assumed above that NPIs are licensed by a c-commanding negation at S-Structure. It has been suggested (Haegeman and Zanuttini 1990) however, that they are licensed by entering into a Spec-head relationship with negation at LF. In this view, the subject NPI would be licensed in (71) as long as it preserves its S-structure Spec-head relationship with the negation on the verb at LF.[31]

The subject NPIs like *kimse* 'nobody' in (72) would be licensed in the tensed CP-deletion clauses in a similar fashion, i.e., by remaining in the Spec-head relationship with the verb that carries the Neg^0 at the TP level:

(72) Ahmet-∅ [kimse-∅ uyu-ma-dɨ-∅] san-ɨyor-∅[32]
A.-nom noone-nom sleep-neg-past-agr think-pres-agr
'Ahmet thinks noone slept'

One issue that remains is that agreeing infinitives also display the gerundive behavior in (3). These properties were tied in with the nominal character of the C^0 in the case of tensed clauses. However, this would not be available for agreeing infinitives if they must all undergo CP-deletion. Suppose now, that what defines the nominal agreement domain is not the nominal C^0 *-K-* itself, but the Case

[31] NPI-licensing via Spec-head relationship is generally assumed to be established at NegP (Haegeman and Zanuttini 1990). But since NegP is lower than TP in Turkish, this relationship must hold at some other projection. Otherwise, subject NPIs would never be licensed because they have to move past [Spec, NegP] to [Spec, TP] for Case reasons.

[32] Note that if this is how the subject NPIs are licensed in agreeing infinitives and tensed CP-deletion clauses, we need to assume that the accusative lower subject *kimse* 'noone' in ECM structures, cf. (56a) above, is actually Case-marked inside the higher clause.

(i) *Ahmet-∅ kimse-yi [t koš-ma-dɨ] san-ɨyor-∅
A.-nom noone-acc run-neg-past think-pres-agr
'Ahmet thinks noone ran'

This would support the current return to the Subject-to-Object Raising analysis of ECM.

morphology on the clause. That is, nominal agreement may be the only available agreement in Case-marked domains in Turkish. With tensed clauses, the Case morphology is required by the nominal C^0; while with agreeing infinitives, it is needed for having the subject Case-marked. On the other hand, genitive Case appears to be the only Case available in such Case-marked nominal agreement domains.[33]

5.3. Tensed CP-deletion.

The tensed CP-deletion construction (Structure C of Kennelly 1992) is characterized by nominative Case on the subject, agreement in the verbal paradigm, and the absence of Case-marking on the clause, all of which are illustrated in (73) below:

(73) Ahmet-∅ [Berna-∅ uyu-du-∅] san-ɨyor-∅
A.-nom B.-**nom** sleep-past-**agr** think-pres-agr
'Ahmet thinks Berna slept'

In the absence of the nominal C^0 *-K-* and any other reason for the clause or its Infl to be Case-marked. The clause being a non-Case-marked domain, agreement appears in the verbal paradigm, and the tensed Infl (T^0 or Agr) is capable of assigning nominative Case. These facts are consistent with the system of Case and agreement proposed in the prior subsection.

These clauses have unique properties at the lower Infl level. Unlike other subordinate clauses, they allow all five matrix tenses, e.g., the present in (74);[34] and they license multiple tense, as in (75) (Kennelly 1992):

(74) Ahmet-∅ [*pro* [bu kitab]ɨ oku-yor-um] san-ɨyor-∅
A.-nom 1.sg this book-acc read-pres-agr think-pres-agr
'Ahmet thinks that I am reading this book'

---

[33] The subject receives nominative Case in tensed clauses inside the time adjuncts with *sonra* 'after', even though the whole clause is Case-marked for the ablative. However, this is the only instance in Turkish where the verb does not agree with the subject:
(i) [[Ahmet-∅ ev-e git-ti-k-**ten**] sonra]
A.-**nom** home-dat go-past-comp-**abl** after
'After Ahmet went home'
Thus it seems that a tensed Infl can assign nominative Case in the absence of agreement despite the Case-marked nominal C^0 *-K-* (which may be a language-particular property).
[34] *-DI-*: past; *-mIš-*: evidential past; *-Iyor-*: present; *-EcEK-*: future; and, *-Ir-*: generic.

(75) *pro* [Ahmet-∅ uyu-yor-du-∅] san-ɨyor-um[35]
1.sg A.-nom sleep-pres-past-agr think-pres-agr
'I thought Ahmet was sleeping'

Under the CP-deletion analysis, this indicates that the nominal C^0 perhaps because of its selectional properties, restricts the kinds of tenses allowed in regular subordinate clauses, and rules out the occurence of multiple tenses. These restrictions do not apply in the absence of CP. An alternative to this view is proposed by Kennelly (1992), who analyzes the tensed CP-deletion construction as IPs dominated by CPs (her Structure C), as opposed to the DPs in regular subordination. In her system, all the properties mentioned in this subsection are licensed by the true CP and disallowed by the DP.[36]

It is hard to find empirical evidence that would decide between these proposals. The distribution of relative clauses comes close though. They allow only regular clauses (76a), and not CP-deletion clauses (76b):

(76) a. [Op_i [Ahmet-in t_i oku-du-ğ-u]] kitap
A.-gen read-past-comp-agr book
'The book that Ahmet read'

---

35 The multiple tense construction involves the inflected copular forms *idi-*, *imiš-*, and *ise-* (past, evidential past, and conditional). The initial *i-* is categorized as the defunct copula by traditional grammarians. These forms can constitute a single word with the preceding verbs or non-verbal predicates where the *i-* is dropped and the vowels are harmonized.

36 Kennelly (1992) further notes that the lower subject NPIs are not licensed by higher negation, (i), and that the reciprocal *birbirleri* 'each other' in the subject position of the lower clause cannot take an argument in the higher clause as its antecedent, (ii):

(i) *Ahmet-∅ [kimse-∅ uyu-du-∅] san-m-ɨyor-∅
A.-nom noone-nom sleep-past-agr think-neg-pres-agr
'Ahmet thinks noone slept'

(ii) *Adam-lar-∅ [birbirleri-∅ uyu-du-∅] san-ɨyor-∅
man-plr-nom eachother-nom sleep-past-agr think-pres-agr
'The men think each other slept'

These data seem to indicate that material inside these clauses cannot escape at LF: *kimse* 'noone' in (i) cannot A'-move to the higher clause to satisfy the Neg-criterion; and in (ii), the distributor in *birbirleri* 'each other' cannot move out to pied-pipe its antecedent (Heim, Lasnik, and May 1991; Kennelly 1991). Wh-phrases however, can escape these clauses at LF:

(iii) Ahmet-∅ [kim-∅ uyu-du-∅] san-ɨyor-∅?
A.-nom who-nom sleep-past-agr think-pres-agr
'Who does Ahmet think slept?'

Therefore, A'-movement outside these clauses itself cannot be barred. This suggets that something else must be responsible for (i) and (ii). I will leave this matter unresolved.

b. *[Op$_i$ [Ahmet-∅ t$_i$ oku-du-∅]] kitap
A.-nom read-past-agr book
'The book that Ahmet read'

This would suggest that the clauses of the type in (76b) do not have their CPs intact, and they can only be licensed by verbs of the *san* 'think' class.

The issues relating to the tensed CP-deletion clauses as opposed to regular subordinate clauses are far from being resolved. There is no doubt that a further investigation of these structures will prove fruitful.[37]

6. Conclusion.

In this paper, I motivated a certain classification of subordinate Infls in Turkish, which singled out the segment *-K-* attached to some of those Infl elements. I proposed an analysis where *-K-* is the C^0 in Turkish, one that has a more nominal character, compared to the C^0 in English, French, so forth. The fact that Agr appears higher than C^0, while the subject is lower than [Spec, CP] led to the conclusion that Agr is not an independent head in syntax, but rather, it is the S-structure spell out form of the abstract Agr-features picked up by the verb during its movement to C^0. This was posited as a property of UG, whereby Agr is spelled out at S-structure after the language particular verbal complex is formed via head-movement.

[37] For example, Kennelly (p.c.) has suggested that tensed CP-deletion clauses (her Structure C) may ultimately be displaying the behavior of non-specific arguments (Enç 1991). Among others, this would also explain why they cannot scramble, unlike regular subordinate clauses.

Bibliography

Abney, S. 1987. *The English Noun Phrase in Its Sentential Aspect.* PhD Dissertation. MIT.

Aronoff, M. 1976. *Word Formation in Generative Grammar.* LI Monographs. MIT Press. Cambridge, Massachusetts.

Baker, M. 1985. The Mirror Principle and Morphosyntactic Explanation. *Linguistic Inquiry,* 16, 373-415. MIT Press. Cambridge, Massachusetts.

Belletti, A. 1987. The Case of Unaccusatives. *Linguistic Inquiry:* 19, 1-34. MIT Press. Cambridge, Massachusetts.

______. 1990. *Generalized Verb Movement.* Rosenberg and Sellier, Torino.

Brendemoen, B. and E. A. Csato. 1984. The Head of S in Turkish: A Comparative Approach to Turkish Syntax. In A. A. Koç and E. E. Taylan eds., *Proceedings of the Turkish Linguistics Conference.* Boğaziçi University Publications. Istanbul, Turkey.

Bresnan, J. 1972. *Theory of Complementation in English Syntax.* PhD Dissertation. MIT.

Chomsky, N. 1986b. *Barriers.* LI Monographs. MIT Press. Cambridge, Massachusetts

______. 1991. Some Notes on Economy of Derivation and Representation. In R. Freiden, ed., *Principles and Parameters in Comparative Grammar.* MIT Press. Cambridge, Massachusetts.

Enç, M. 1987. *Agreement and Governing Categories.* ms, University of Southern California.

______. 1991a. *On the Absence of the Present Tense Morpheme in English.* ms. University of Wisconsin, Madison.

______. 1991b. The Semantics of Specificity. *Linguistic Inquiry,* 22, 1-25. MIT Press. Cambridge, Massachusetts.

George, L. and J. Kornfilt. 1981. Finiteness and Boundedness in Turkish. In F. Heny, ed., *Binding and Filtering.* MIT Press. Cambridge, Massachusetts.

Haegeman, L. and R. Zanuttini. 1990. *Negative Concord in West Flemish.* Ms. Universite de Geneve.

Heim, I., H. Lasnik, and R. May. 1991. Reciprocity and Plurality. *Linguistic Inquiry.* 22, 63-101. MIT Press. Cambridge, Massachusetts.

Huang, C.-T. J. 1982. *Logical Relations in Chinese and the Theory of Grammar.* PhD Dissertation. MIT.

Kennelly, S. 1987. Turkish Gerunds. In H. E. Boeschoten and L. Th. Verhoeven, eds., *Studies on Modern Turkish: Proceedings of the Third Conference on Turkish Linguistics.* Tilburg University Press.

______. 1990. *Theta Government in Turkish.* Talk presented at GLOW workshop, London.

______. 1991. *Reciprocity in Turkish: Evidence against Agr as a Definition of Finiteness.* Talk presented at WCCFLX, Arizona State University.

______. 1992. *Turkish Subordination: [-CP, -Tense, +Case].* Talk presented at 6th Conference on Turkish Linguistics, Eskisehir.

Koopman, H. and D. Sportiche, 1990. *The Position of Subjects.* ms. UCLA.

Kornfilt, J. 1977. A Note on Subject Raising in Turkish. *Linguistic Inquiry;* 8, 736-742. MIT Press. Cambridge, Massachusetts.

______. 1984. *Case Marking, Agreement and Empty Categories in Turkish.* Doctoral Dissertation, Harvard University.

Kural, M. 1992. *Properties of Scrambling in Turkish.* ms, UCLA.

Ogihara, T. 1989. *Temporal Reference in English and Japanese.* PhD Dissertation. University of Texas at Austin.

Pollock, J.-Y. 1989. Verb Movement, Universal Grammar, and the Structure of IP. *Linguistic Inquiry:* 20, 365-424. MIT Press. Cambridge, Massachusetts.

Portner, P. 1992, *Situation Theory and the Semantics of Propositonal Expressions.* PhD Dissertation. University of Massachusetts at Amherst.

Raposo. E. 1987. Case Theory and Infl-to-Comp: The Inflected Infinitives in European Portuguese. *Linguistic Inquiry.* 18, 85-109. MIT Press. Cambridge, Massachusetts.

Stowell, T. 1981, *Origins of Phrase Structure.* PhD Dissertation. MIT.

______. 1982. The Tense in Infinitives. *Linguistic Inquiry.* 13, 561-570. MIT Press. Cambridge, Massachusetts.

______. 1992. *Aspects of Tense Theory.* Talk presented at the GLOW Colloquia, Lisbon.

Underhill, R. 1976. *Turkish Grammar.* MIT Press. Cambridge, Massachusetts.

ON GAMMA MARKING ADJUNCT TRACES IN HINDI

Anoop Mahajan

mahajan@cognet.ucla.edu

1.0. Lasnik and Saito (1984) distinguish between arguments and adjuncts along the following lines: argument traces created at SS must be γ-marked at SS while γ-marking for adjunct traces must be delayed till LF.[1] In this paper, I examine some aspects of adjunct extraction in Hindi that seem to indicate that such a distinction would yield undesirable results. The evidence that I present suggests that adjunct traces should be γ-marked in the same fashion as argument traces. This implies that Lasnik and Saito's analysis must be modified. I will suggest a minor revision of Lasnik and Saito's theory that will have the desired result. This revision will not appeal to SS/LF asymmetry in γ-marking of arguments and adjuncts and is therefore compatible with theories such as Chomsky (1992) which dispense with the level of SS.

2.0. The following sentences from Hindi show that both argument and adjunct wh-phrases can be freely extracted from finite clauses.

(1) kis-ko	raam-ne	socaa	ki	mohan-ne	__	dekhaa
who	Ram-erg	thought		Mohan-erg		saw

'Who did Ram think Mohan saw?'

---

[1] In this paper 'γ–marking' is to be understood to mean assigning '+γ feature' unless specifically qualified to convey a different meaning.

(2) kɛse raam-ne kahaa ki mohan-ne ___ gaaṛii ṭhiik kii
how Ram-erg said Mohan-erg car fixed
'How did Ram say that Mohan fixed the car?'

Non wh-phrases can also be fronted out of finite clauses as shown below:

(3) siitaa-ko raam-ne socaa ki mohan-ne ___ dekhaa
Sita Ram-erg thought Mohan-erg saw
'Sita, Ram thought that Mohan saw'

(4) bahut dhiire se raam soctaa hɛ ki mohan-ne ___ gaaṛii ṭhiik kii
very slowly Ram thinks be-pres Mohan-erg car fixed
'Very slowly/in a very slow manner, Ram thinks that Mohan fixed the car'

2.1. Any analysis of the sentences given above must deal with the fact that the extraction domain, i.e., the embedded finite clause, is in a non-canonical position. That is, while Hindi is a head final language in which normal arguments precede a head, finite embedded clauses always appear following the matrix sentence. Most studies in Hindi syntax therefore recognize the fact that finite embedded clauses are obligatorily extraposed (this is true of recent works like Mahajan (1990) and Srivastav (1991); for a somewhat different view, see Davison (1987)). However, if the embedded clauses in (1)-(4) above have been extraposed, they are not L-marked in their SS position.[2] They should therefore be barriers for extraction and (1)-(4) should be ungrammatical. However, this is not the case. There are at least two possible solutions to this problem.

2.1.1. Under the first solution, one could argue that the post verbal clause is in fact in its base generated position. That is, finite clauses in Hindi are base generated to the right of the subcategorizing head. This solution would treat Hindi as underlyingly SOV or SVO depending on whether the object is a simple NP or a clause. This solution thereby accounts for the grammaticality of (1)-(4). The embedded clause is L-marked by its sister V and is therefore not a barrier.

2 Mahajan (1990) and Srivastav (1991) discuss facts like these in their treatment of wh-extraction. The solutions suggested to handle argument extraction is similar in both of these studies.

Extractions of arguments as well as adjuncts can therefore take place as they do in English. This solution is however undesirable for several reasons. First, it provides no reason why clausal arguments must be generated to the right of a head while the non-clausal arguments are generated to the left. Secondly, the finite clauses do not always appear as sisters of the subcategorized verbs. If there are any auxiliaries in the matrix sentence, they must intervene between the verb and its clausal complement. The auxiliaries cannot follow the embedded clause as would be expected if the verb and its clausal complement formed a constituent. This is shown in (5) and (6) below.[3]

(5) raam soc rahaa thaa [ki mohan-ne siitaa-ko dekhaa hogaa]
Ram think prog. be-pst Mohan-erg Sita seen be-fut
'Ram was thinking that Mohan must have seen Sita'

(6) *raam soc [ki mohan-ne siitaa-ko dekhaa hogaa] rahaa thaa
Ram think Mohan-erg Sita seen be-fut prog. be-pst.
'Ram was thinking that Mohan must have seen Sita'

The third reason for not treating the post verbal embedded clauses as sisters of their subcategorizing heads is that it is possible in Hindi to have post verbal clauses accompanied by an expletive like NP in the preverbal position. This is exemplified in (7) below:

(7) raam yah soc rahaa thaa [ki mohan-ne siitaa-ko dekhaa hogaa]
Ram it think prog. be-pst Mohan-erg Sita seen be-fut
'Ram was thinking that Mohan must have seen Sita'

If one assumes, following Mahajan (1990) and Srivastav (1991), that the expletive holds the real argument position, then the post verbal clause must obviously be in a non-argument position. This assumption would immediately account for the fact that extraction out of the embedded clause in (7) is bad.

---

3 The only way around this problem would be to treat auxiliaries and verbs as forming a complex verbal projection which heads a phrase that takes the clause to its right as its argument. As noted in Mahajan (1990), the verb and the auxiliaries that follow do not form a constituent. For instance, it is possible, in some cases, to separate the verb and the auxiliaries. The intervening material could be negation (which is not an affix in Hindi) or even full NPs.

(8) ??? kis-ko raam-ne yah socaa ki mohan-ne maaraa
who Ram-erg this thought Mohan-erg hit
'Who did Ram think this that Mohan hit?'

Thus, there is some evidence for treating post verbal embedded clauses as non L-marked constituents. There is no clear evidence, apart from the extraction facts themselves, that the post verbal clause in (8) and (1) occupy structurally distinct positions. If the post verbal clause in (8) is not L-marked then there is no reason why the post verbal clause in (1)-(4) should be L-marked. This implies that we have to find a reason for the grammaticality of (1)-(4) and this brings us to our second solution for the problem noted above.

2.1.2. For the extraction of wh-phrases in (1)-(2) as well as non wh-phrases in (3)-(4), we require the embedded clause to be L-marked. Suppose that the embedded clause does originate as a left sister of the verb. This is what one would expect in this head final language. At this stage of the derivation, the embedded clause is L-marked, therefore it is not a barrier. I suggest that extractions of the wh as well as non wh-phrases takes place from the embedded clause at this stage of the derivation. The embedded clause is extraposed following this extraction. Note that this solution correctly distinguishes (1) from (8). In (1), the extraposed clause is base generated in an argument position and is therefore in an L-marked position at some stage of derivation. In (8) however, the extraposed clause could not have been base generated since the expletive 'yah' holds the argument position. This means that the extraposed clause in (8) was never in an L-marked position and therefore extractions from such clauses would always be ill-formed.

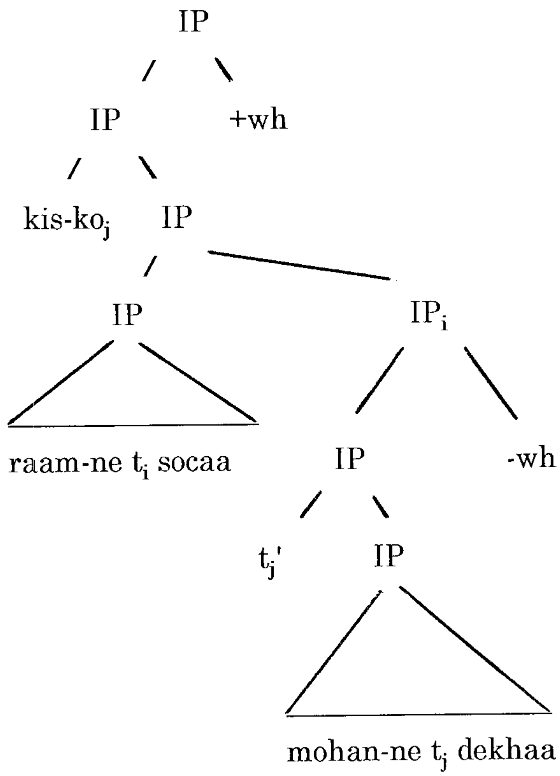
The final landing site of the wh-phrase in sentences like (1)-(4) has not been identified so far in this paper. In Mahajan (1990), it is suggested that wh-phrases adjoin to the matrix clause (see also Bains,1987). In that analysis, there is no difference between the landing site of wh-phrases (as in (1)-(2)) or topics (as in (3)-(4)).[4] Adopting that analysis, the s-structure of a sentence like (1) would have a

---

[4] Srivastav (1991) suggests that the final landing site of wh-phrases in Hindi is SPEC CP. It should be noted that if S is a cyclic node, then wh-movement to SPEC CP followed by extraposition (adjunction to S) would violate strict cyclicity. On the other hand adjunction to S of a wh-phrases followed by adjunction to the same S of the extraposed clause does not

representation like (9). (I will assume here that Hindi finite clauses are IPs rather than CPs.[5]:

(9)



I am assuming, following Mahajan (1990), that adjunction to IP is possible both as a landing site as well as a mechanism to void the barrierhood of an IP. In the derivation of (1), the wh-phrase would first adjoin to the embedded IP and then adjoin to the matrix IP. The embedded clause (IP$_i$) would then extrapose. Since we are dealing with an argument in (1) (and in (3)), γ-marking takes place at SS. Both t_j and t_j ' are γ-marked at SS prior to extraposition (though in this case t_j can be γ-marked by V and γ-marking t_j' is not critical since this trace can be deleted).

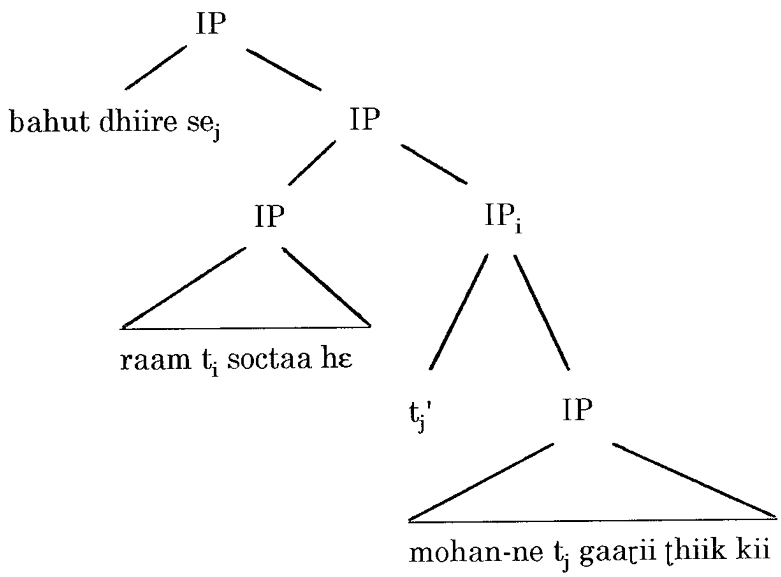
This brings us to the main point of this paper- the derivation of sentences like (2) and (4). Let us consider the derivation of (4) ((2) would have a similar derivation). The SS would look like (10):

---

violate strict cyclicity.

[5] I will assume here that Hindi finite clauses are IPs rather than CPs. This is consistent with Mahajan (1990) though nothing crucial in this paper hinges on this assumption.

(10)



Recall that in Lasnik and Saito's theory, γ-marking for adjuncts must wait till LF. Therefore, under that assumption, t_j as well as t_j' are not γ-marked at SS. At LF however, the embedded clause has been extraposed. It is therefore no longer L-marked. This implies that while t_j can be γ-marked, t_j' which is adjoined to the extraposed IP will receive a -γ feature since IP_i is a barrier. This analysis then predicts that (4) (and (2)) should be ungrammatical. However, there is no difference between the grammaticality of (1)/(3) and (2)/(4). In particular, Lasnik and Saito's approach leads us to expect an argument/adjunct asymmetry in these cases. However, no such asymmetry exists for such cases in Hindi.

It should be noted that it is not the case that there are no argument-adjunct asymmetries in Hindi. Mahajan (1990) notes that when one examines extractions out of clauses which are base generated in an adjoined position, i.e., cases like (8), argument/adjunct asymmetries manifest as would be expected. Thus (11) and (12) below are considerably worse than (8) if the wh-phrases are construed as originating in the embedded clause:

(11) * kɛse raam-ne yah socaa ki mohan-ne gaaṛii ṭhiik kii
how Ram-erg this think Mohan-erg car fixed
'How did Ram think that Mohan fixed the car?'

(12) *kab raam-ne yah socaa ki mohan-ne gaaṛii ṭhiik kii
when Ram-erg this think Mohan-erg car fixed
'When did Ram think that Mohan fixed the car?'

This is to be expected since (8) is merely a subjacency violation as an object is being extracted from a non L-marked clause. (11) and (12), on the other hand, are cases of adjunct extraction which violate ECP as well as subjacency.

4.0. The problem that sentences like (2) and (4) pose for Lasnik and Saito's analysis can be resolved only if we assume that adjunct extractions are identical to argument extractions in simple cases in Hindi. Both take place prior to extraposition and in both the cases γ-marking takes place at a level prior to extraposition. This analysis then supports a derivational view of syntax in which the level of γ assignment for arguments and adjuncts is not distinguished.

5.0. Lasnik and Saito's theory requires that adjuncts be assigned γ features at LF primarily to account for sentences like (13):[6]

(13) *Why$_1$ do you wonder who$_2$[$_{IP}$t$_{1'}$ [$_{IP}$ t$_2$ left t$_1$]]?

If γ-marking for adjuncts was to take place at SS then there is a derivation which allows (13) to escape ECP. Under this derivation, t_1 will be γ-marked by the IP adjoined t_1' at SS and at LF t_1' (which will receive minus γ-marking at SS) can be deleted.[7] However, this derivation should not be allowed because (13) is clearly ungrammatical. If we delay γ-marking the adjuncts till LF then t_1' cannot be deleted since it is required at LF to γ-mark t_1. At LF then, t_1' becomes the offending trace making (13) ungrammatical.

However, the Hindi evidence that we have seen indicates that delaying γ-

---

[6] I will not explore the details of the lack of that-trace effects with adjuncts in this paper, a point that may be of some importance in this discussion. Let me simply note here that the solution to the lack of that-trace effects with adjuncts under my approach will require recasting Lasnik and Saito's discussion in terms of a "Barriers" type approach with a further stipulation that adjuncts in English be allowed to adjoin to IP while subjects may not do so.

[7] If IP is no longer a barrier after adjunction as we have assumed then one may need to invoke a relativized minimality type account for sentences like (13). I will, however, follow Lasnik and Saito's assumptions in the analysis of (13).

marking of the adjunct traces will have an undesirable effect. Both, the Hindi as well the English data that we have seen so far is compatible with a minor modification of Lasnik and Saito's approach. What we need for sentences like (13) is that t_1' should be present at LF even if γ assignment applies at SS. t_1' would then receive a -γ feature at SS but will not be able to delete at LF. This result could be achieved by some principle that ensures that intermediate adjunct traces can never be deleted (unlike intermediate argument traces which must be deleted at LF). Chomsky (1989;63) makes a suggestion based on "the least effort principle" that has the desired result. According to Chomsky, the presence of intermediate traces of arguments at LF will produce illegitimate "heterogeneous chains" that consist of an adjunct chain and an A-bar-A (operator-variable) chain; the A-bar position being occupied by the intermediate trace. This illegitimate chain can become legitimate only by deletion of the intermediate trace. This approach then requires intermediate argument traces to be deleted. On the other hand, since an adjunct chain containing an intermediate trace is already a legitimate object, the intermediate trace (t_1'in our example (13)) need not be deleted and therefore according to the "least effort" approach, it must not be deleted.

6.0. To conclude, I have presented some evidence from Hindi that indicates that adjuncts must be assigned a γ feature at a level of derivation prior to LF and in that respect their behavior is indistinct from that of arguments. This indicates that Lasnik and Saito's (1984) proposal regarding γ-marking of adjuncts needs to be modified. I have outlined a proposal that circumvents the problem that Lasnik and Saito's approach faces.

It should be noted that Lasnik and Saito's proposal regarding the SS/LF distinction for γ-marking arguments and adjuncts is embedded within the theory that requires a level of SS. The proposal that I have outlined is rather different. It requires that γ-marking for arguments as well as adjuncts take place derivationally and that γ-marking can be followed by other movement/deletion processes. Thus, in this model there is no need for a distinct level of SS, a conclusion that may be of significance in light of Chomsky's (1992) proposal regarding the elimination of SS as a distinct level of representation.

BIBLIOGRAPHY

Bains, G.S. (1987) Complex Structures in Hindi-Urdu: Explorations in Government and Binding Theory, Ph.D. dissertation, N.Y.U.

Chomsky, N. (1986) Barriers, MIT Press, Cambridge.

____. (1989) "Some Notes on Economy of Derivation and Representation," in I. Laka and A.K. Mahajan, eds., MIT Working Papers in Linguistics, Vol 10, 43-74.

____. (1992) A Minimalist Program For Linguistic Theory, MIT Occasional Papers in Linguistic, Number 1, MIT.

Davison, A. (1987) "Syntactic Constraints on Wh-in-situ: Wh-questions in Hindi-Urdu," ms.

Lasnik, H. and M. Saito (1984) "On the Nature of Proper Government," Linguistic Inquiry 15, 235-289.

Mahajan, A.K. (1990) The A/A-bar Distinction and Movement Theory, Ph.D dissertation, MIT, Cambridge.

Srivastav,V. (1991) Wh Dependencies in Hindi and the Theory of Grammar, Ph.D. dissertation, Cornell University, Ithaca.

DUTCH NOMINAL COMPOUNDS AND THE INNATENESS OF LEVEL-ORDERING

Jeannette Schaeffer

schaeffe@cognet.ucla.edu

0. Introduction

Many Generative Linguists assume that language acquisition is determined by a biological innate capacity for language, the Language Acquisition Device (henceforth LAD). The LAD contains abstract principles that can be filled in language specifically. These abstract principles are reflected in universal properties of the grammars of natural languages: the Universal Grammar (henceforth UG). In order to gain insight into the LAD (and thus in UG) it is necessary to study the grammars of natural languages. If it is assumed that the number of possible grammars is determined by the LAD, it is clear that investigating not only adult states of linguistic knowledge but also child grammars are of direct interest for linguistic research (for child grammars are grammars of natural languages as well). In principle, child language is a source for (dis)confirmation of theories about the LAD.

In this paper I would like to consider the plausibility of the phenomenon of "level-ordering" (Kiparsky (1982)) as a UG principle. If level ordering is an innate principle, then we expect all natural languages to be submitted to it. I will call this idea the Innate Level Ordering Hypothesis (henceforth ILOH). In particular, I will discuss the morphological properties of compound formation. Level-ordering predicts that irregular plurals may be formed at level 1, prior to compounding at level 2. However, regular plurals formed at level 3 may not precede compounding. Both English adult and child language behave perfectly according to the ILOH: regular plurals do not occur within compounds (*rats-eater), but irregular plurals do (mice-eater).

However, (adult) Dutch seems to provide a lot of counterexamples against the ILOH, since regular plurals within compounds frequently appear, for example:

(1) bij-en korf
'bee-pl(?) hive'

(2) boek-en kast
'book-pl(?) case'

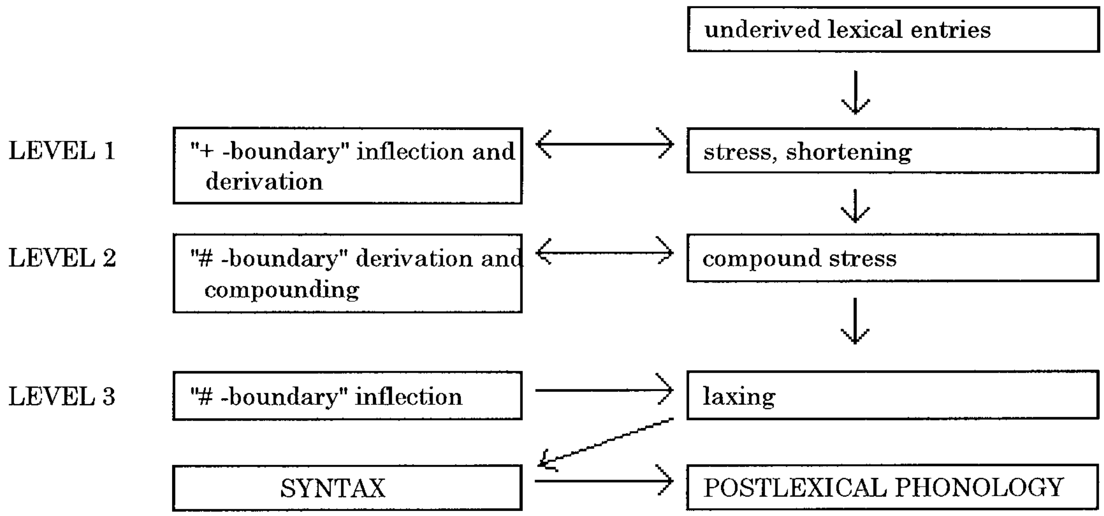
I will show that this counterevidence is only apparent by arguing that most "linking morphemes" in Dutch compounds are not real plural morphemes. Dutch acquisition data confirms this idea.

The paper is organized as follows: in section 1 I will give a more elaborate explanation of Kiparsky's notion of level-ordering, focused on English compounds. Section 2 gives an outline of Peter Gordon's (1985) study of compounds in English child language, which confirms the ILOH. The third section will deal with the problematic Dutch data, followed by the conclusion in section 4.

1. Level ordering

The notion of ordered levels appears in the work of Allen (1978) and Siegel (1977) and has been extended by Kiparsky (1982). Basically, it says that properties of word formation can be coherently accounted for by positing ordered "levels" of rule application. Each level is associated with a set of phonological rules for which it defines the domain of application. The output of each word-formation process is submitted within the lexicon itself to the phonological rules of its level. For present purposes, I shall assume the three-level version of Kiparsky (1982) and I will focus on the morphological part. Level 1 is said to include primary (+) affixes (e.g. +ian, +ous, +ion) that characteristically deform their hosts phonologically by stress shifting, vowel reduction, alternation and so on, and are often semantically idiosyncratic in being non-compositional (e.g., the meaning of populat+ion appears to go beyond a simple semantic composite of populate and the nominalizing +ion affix). Also included are irregular inflections (e.g., tooth → teeth, ox → oxen), pluralia tantum (e.g. clothes, scissors, alms) and possibly others. Level 2 contains secondary (#) affixes of derivational morphology (e.g., #er, #ism, #ness) and is the site of compounding. The third level shows neither semantic idiosyncrasy nor stem deformation (e.g., car → cars). So, for example, a word like Darwin-ian-ism is fine, but words like Darwin-ism-ian are unattested, since a level 1 affix (such as +ian) can precede a level 2 affix (such as #ism), but not vice versa. The three levels are schematized below.

(3)



Rule application proceeds through the three levels such that rules at a later level may not be applied prior to those at a previous level.

One very interesting prediction from this model, noted by Kiparsky (1982) and Gordon (1985) among others, is that one should not find regular plurals "inside" compounds. That is, once a compound is formed at level 2, its constituents cannot be inflected at level 3. In other words, the left branch of a compound may not be inflected, but the compound itself may be inflected to the right. However, since irregular inflections are at level 1, they should be allowed inside compounds. This prediction is supported by the difference in acceptability of mice-infested versus *rats-infested, since the former includes a level 1 plural and the latter, a level 3 plural. Pluralia tantum (level 1) also find their way inside compounds in some cases (e.g., clothes-basket), although reduction is possible in other cases (e.g. scissor-legs).

3. The acquisition of English

Assuming that level ordering is the correct way of characterizing lexical structure, we should consider how a child could ever learn such an organization. What evidence in the linguistic input would lead inductively to setting up this system? It would seem that of all the hypotheses available, there would be little to persuade an open-minded learner to choose this, rather than some other path. In other words, the idea of level ordering raises a learnability problem. For example, most compounds that the child hears involve singular forms inside compounds. This forms positive evidence for the child with respect to the constraint of not having regular plurals inside compounds. However, on the other hand, it does not provide positive evidence regarding the possibility of placing irregular plurals inside compounds. In fact, the child almost never hears compounds containing irregular plurals.

Gordon (1985) carried out an examination of certain high frequency items with irregular plurals (mouse, man, tooth, foot and goose), using the Kucera and Francis (1967) word count of about one million words. This revealed that while these forms were listed in a total of 28 compound types in non-head (left) position (token frequency: 153), in only two cases was the noun listed in its irregular plural form (token frequence: 3). This compares with a plural-to-singular ratio of 1181:1436 for the irregular nouns not occurring inside compounds. Yet, despite this "poverty of stimulus", the child acquires the knowledge that, for example, mice is a level 1 plural and that claws is a level 3 plural. She appears to be aware of the fact that unproductive/ irregular forms are level 1 and therefore separately lexicalized, and that the more productive forms are assigned to level 2 and 3. This is furthermore confirmed by some more recent experiments, carried out by Peter Gordon (1989), which found general support for a systematic relation between productivity and level assignment in child language.

As a solution to this problem Gordon (1985) suggests that level ordering is an innate structural property of the lexicon. Given this hypothesis, there are a number of developmental predictions that arise with respect to the appearance of plurals within compounds (Gordon (1985)):

I. If rules of compounding and regular inflection are correctly assigned to levels 2 and 3 respectively, then as soon as the child acquires the regular plural morphology and shows evidence of regularization (e.g., by overgeneralization to irregular forms), the regular forms should be reduced to singulars inside compounds. For example, one should find rat-infested but not rats-infested being produced by the child.
II. As soon as the child stops overregularizing an irregular form (e.g., *mouses) and uses the appropriate plural (mice), then such forms should be (optionally) allowed inside compounds (e.g., mice-infested).
III. As soon as the child learns that pluralia tantum are irregular in the sense of having no singular form, then they too should optionally occur inside compounds (e.g., clothes-dryer).

In order to test these predictions, Gordon executed an experiment in which noun-agentive compounds (e.g., rat-eater) were elicited from 33 children (individually) of age 3;2 to 5;10. He examined the following three noun-types:

(i) regular plural;

(ii) regular plural;

(iii) pluralia tantum.

Initially, the child was introduced to a Cookie Monster puppet and was told:

> *Do you know who this is? ... It's the Cookie Monster. Do you know what he likes to eat? Answer: Cookies.) Yes___and do you know what else he likes to eat?___He likes to eat all sorts of things...*

Objects were then brought out and the child was asked if the Cookie Monster would like to eat X (where X was the name of the stimulus). They were then asked: "What do you call someone who eats X?" (Answer: An X-eater.) With this prodecure, it was possible to elicit compounds of the form teeth-eater/rat-eater and so on. All the subjects were properly trained in producing compounds and for each child it was ascertained that the irregular plurals she used were true plurals. This was done because irregular plural forms are often used by children as if they were singulars, thus producing one mice, two mices etc.

For regularly pluralized nouns, subjects overwhelmingly showed the correct pattern of reduction inside compounds (e.g., rat-eater) at all ages with 161/164 such patterns. Subjects were categorized as supporting the predicted pattern if all regular plurals were reduced inside compounds. When children overregularized an irregular noun (mouse → mouses) they similarly reduced to the singular form in compounding (mouse-eater) on 86/88 items. As soon as the subjects produced the correct irregular pattern (mouse → mice) they immediately showed evidence that these irregulars were allowable inside compounds. 36/40 responses in this category were of the form mice-eater. For the pluralia tantum, it was predicted that these should be optionally allowed inside compounds in their plural form while their regular counterparts should be reduced. This was confirmed by the test results, too. Basically, there were two patterns found among the pluralia tantum; one in which reduction to a singular form occurred (scissor-eater, glass-eater), and the other in which reduction was not prevalent (clothes-eater, pants-eater). An overall analysis showed that there was a significantly greater tendency to produce pluralia tantum inside compounds than regular plurals.

In sum, the results of this experiment provided strong confirmation of the predictions being made on the basis of the idea that level ordering is innate, or part of UG.

3. Dutch - a problem?

3.1 Dutch plural

As I mentioned in the Introduction, it looks like Dutch data provides counter-evidence against the ILOH, since regular plurals within compounds seem to appear frequently. For example,

> tand-en#borstel ('tooth-pl#brush'), boek-en#kast ('book-pl#case), bloem-etje-s#behang ('flower-dim-pl#wallpaper'), varken-s#hok ('pig-pl#sty), steden#bouw ('town-pl#building').

The phenomenon of the linking morpheme between the two parts in Dutch compounds is a very controversial issue. Some linguists say there is no regularity at all, which, in terms of lexical phonology/morphology, implies that every single compound must have its own lexical entry; others provide classifications (with a lot of exceptions) (van den Toorn (1981; 1982); Mattens (1984)). To my knowledge, nobody has come up with a satisfactory explanation so far.

In his article about the innateness of level ordering, Gordon (1985) mentions the fact that Dutch could provide "a considerable embarassment to the present proposal for the innateness of level-ordering". The explanation he gives boils down to the idea that Dutch plurals are comparable to the English irregular plurals in their level assignment. In other words, he claims that Dutch plural is not productive and therefore the form of the plural would have to be listed with each lexical item rather than being applied productively in the strong sense (i.e., as an independently stated rule). This idea is based on the following reasoning: Unlike English where there is basically one form of the plural (-s), Dutch has two basic forms: -en, as in boek-en ('books') and -s as in varken-s ('pigs'). There is also a rarer form, -eren as in kind-eren ('children') plus other even less productive forms. While -en is the most common form of the plural, it appears that -s is not exactly rare. If -en is not sufficiently dominating in frequency, it could not in any sense be a "default" value for realization of the plural (as appears to be the case for -s in English).

As Collins (p.c.) points out, there are a number of problems with the above explanation. First, it is not clear what frequency has to do with either the default or productive status of a rule. Second, it is generally agreed by Dutch linguists that -en is the default plural marker in Dutch, mainly because its application is far wider than that of the -s plural (cf. Booij (1977); Donaldson (1987); van Marle (1985)).[1] Van Marle (1987) proposes the following formulation of the Pluralization Rule in Dutch:

```
(4) a. PL  →  eren  /  X] ___            for X =  rund, ei, kind, etc.

    b. PL  →  s     /  ə  C] ___
                          |
                       [+son]

                    /  V] ___

                    /  [-native] ___

    c. PL  →  en
```

These rules are ordered disjunctively from (a) to (c). (a) represents a small class of about 13 nouns that pluralize in -eren. (b) says that there are three main stem classes which take -s plural; stems ending in a sonorant consonant, stems ending in a vowel and borrowings. (c) represents the elsewhere case, which applies in cases where (a) and (b) have not applied.

There appear to be exceptions in both directions: there are stems that we would expect to have -en endings that take -s endings and vice versa. For example, artikel ('article') ends in a sonorant and yet we have the plural artikel-en. Similarly, kok ('cook') ends in a non-sonorant obstruent, yet it has the plural kok-s. These exceptions must simply be memorized.

The above story describes the Dutch pluralization phenomenon in quite a simplified

---

1 Although there is a tendency to use more and more -s plurals, e.g. appel-s ('apples') instead of appel-en. Eastern Dutch dialects even seem to **prefer** -s plural, e.g. arm-s ('arms') instead of Standard Dutch arm-en; artikel-s instead of Standard Dutch artikel-en.

way. There are a few other things which play a role in plural formation, such as stress and stem modification. These issues are rather controversial and for the purposes of this paper I will stick to the account given above.

In order to find out whether pluralization in Dutch was a productive phenomenon Collins executed two different experiments with a Dutch native speaker. The first one concerned the plural of proper names. Plurals of nonce proper names were made up from common nouns from the various plural classes. The informant was asked to pluralize both the common noun and the nonce proper name which was made up from that common noun. For example:

(5)	NOUN	NORMAL PLURAL	PLURAL AS NAME
a.	broer 'brother'	broers 'brothers'	Broeren 'Brothers'
b.	kok 'cook'	koks 'cooks'	Kokken 'Cooks'
c.	engel 'angle'	engelen 'angles'	Engels 'Angles'

It turned out that both -en and -s plurals are used to pluralize proper names. Since the made up proper names could not have been memorized with their plural forms, this indicates that -s and -en pluralization are both productive. Furthermore, the results of this experiment showed that if the form of the plural is memorized (e.g. engel-en, kok-s), then the plural of the corresponding proper name conforms to the general rules of -s and -en affixation and not the memorized plural.

A similar effect was found in the second experiment, in which Collins elicited plural forms of phrases. In English one can pluralize syntactic phrases in certain situations. For example, one could say: "Too many 'I love you's, will ruin the relationship". In Dutch you can do the same thing. The subject was asked to form a question of the following frame: "How many more 'X's, before the story ends?" where X ranges over syntactic phrases. For example:

(6) How many more 'I repaired the roof's, before the story ends?

In all of the phrases there was a noun in the last position of the phrase. Again, the nouns were chosen from the various plural classes. The results reflect the data obtained in pluralizing proper names. Both -s and -en were used as plurals, indicating that both are productive forms, since phrases cannot have memorized plural forms associated with them. Furthermore, both -s and -en were used as a plural in environments you would expect according to the general Pluralization Rule: in stems where the -s plural is memorized (kok-s), the -en plural was used for the phrase. Similarly, in cases where the -en plural is memorized (engel-en), the -s plural was used on the phrase.

Concluding, Dutch pluralization is a productive process, and thus Gordon's reasoning with respect to the plural morpheme on the left branch of Dutch compounds is proved to be invalid. This means that we are still left with the problem that Dutch seems to allow plurals inside compounds, something that should not be permitted according to the idea of level ordering.

Annie Senghas, in her talk at the Boston University Conference on Language Development (1991), tried to explain the Dutch counterexamples in a different way. She proposed that Dutch and German inflections may occur at level 1 or 2 and may therefore end up inside compounds. Also, there would be particular circumstances where the product of syntax can be fed back into the level-ordered morphological process. Thus, after compounds are formed at level 2, they enter the syntax (via level 3) where they form a new input for level 2. Back at level two, regular plural inflection of the compound's left branch can take place. As Kiparsky points out, some limited recursion from syntax back into morphology must be assumed anyway. Cases such as a hands-off policy, the save-the-whales campaign, and computer systems analyst involve phrases embedded in compounds. Syntactic phrases are fed back into level 3 where regular plurals are formed. Senghas' account contains the implicit assumption that idiosyncratic inflections and derivations are assigned to level 1, whereas regular inflection and derivation takes place at level 2, instead of level 3, as Kiparsky claims.

A few objections against Senghas' proposal can be made. First, it seems a contradiction in terms to say that level ordering is innate, but that certain morphological processes are assigned to one level in one language, but to another level in another language (e.g. pluralization in German/Dutch: level 1; pluralization in English: level 2 or 3). According to me this contradicts the idea of innateness and universality of level ordering. Second, the idea of having a loop back from syntax into morphology seems to weaken the (strong) concept of level ordering. If it is possible for a certain morphological form to be fed back into an earlier level, then it seems to me that the whole effect of ordering levels is defeated.

3.2 The nominal case-system in Dutch

In the preceding section we saw that neither Gordon's nor Senghas' ideas seem to provide an adequate explanation for the occurrence of plurals inside Dutch compounds. Notice that both accounts start out with the assumption that the linking morphemes inside Dutch compounds are plural markers. If this assumption is wrong, it is no wonder that neither analysis is adequate.

In fact, I would like to propose that the linking morphemes (-s and -en) do not represent plural in most cases, but that they are (the remainder of) genitive case affixes. This is exemplified by (7 a-d):

(7) a. dorp-s-gek
village-GEN-fool

b. varken-s-hok
pig-GEN-sty

c. bij-en-korf
bee-GEN-hive

d. man-e-schijn
moon-GEN-shine

Since genitive case inflection in Middle Dutch and plural affixes are identical in phonological shape, it is not surprising that there is confusion with respect to the function of the linking-morpheme in Dutch compounds. The assumption that these linking morphemes represent case rather than plural becomes more plausible if we look at the case-inflection system in Middle Dutch. Van Loey (1960) gives the following inflectional paradigms for common nouns in Middle Dutch:

(8) INFLECTION I

masculine/neuter sg.			**masculine/neuter pl.**
NOM	worm	('worm')	worm-e
GEN	worm-s		worm-e
DAT	worm-e		worm-en
ACC	worm		worm-e

feminine sg.			**feminine sg.**
NOM	daet	('action')	dad-e
GEN	daet		dad-e
DAT	daet		dad-en
ACC	daet		dad-e

INFLECTION II

masculine/neuter sg.			**masculine/neuter pl.**
NOM	cnape	('boy')	cnape-n
GEN	cnape-n		cnape-n
DAT	cnape		cnape-n
ACC	cnape		cnape-n

feminine sg.			**feminine pl.**
NOM	siele	('soul')	siele-n
GEN	siele(-n)		siele-n
DAT	siele(-n)		siele-n
ACC	siele		siele-n

Note that there are two types of nouns, which are inflected differently: inflection II type nouns end in an -e, inflection I type nouns do not.

In modern Dutch, we still find some archaic forms which reflect this case-marking system:[2]

---

[2] The only Modern Dutch instances of a pre-posed genitive/possessive left are proper names, such as Freds gitaar ('Fred's guitar'); Annes boek ('Anne's book'). Unlike English, a post-nominal possessive is possible, too in these cases: de gitaar van Fred ('the guitar of Fred'); het boek van Anne ('the book of Anne'). This could be an indication of the fact that the genitive/possessive -s is disappearing, at least in the syntax of

(9) a. de heer de-s huiz-es
the lord/master the-GEN house-GEN
"the master of the house"

b. 's Hertog-enbosch
(the)-GEN Duke-GEN woods/forest
"the woods of the Duke" (= name of town in the South of The Netherlands)

c. 's Heer-enbroek
(the)-GEN lord-GEN brook
"the brook of the lord" (= name of town in the East of The Netherlands)

d. 's land-s wijs, 's land-s eer
(the)-GEN country-GEN way, (the)-GEN country-GEN honor
"the country's way of life is the country's honor" (= proverb)

e. Ledigheid is de-s duivel-s oorkussen
idleness is the-GEN devil-GEN pillow
"idleness is the parent of vice" (= proverb)

As Collins points out, in a language where the plural is phonologically identical to a linking morpheme (in this case, a case morpheme) it is predicted that one of the two should disappear. This indeed seems to happen in Dutch. As I said before, case inflection is only found in archaic expressions and is no longer a productive morphological process in common nouns. Furthermore, and even more important for my proposal, most new Dutch compounds are formed without any linking morpheme (van den Toorn (1982); Mattens (1984)). This is predicted by the loss of case-inflection in Dutch. Since there is no disappearance of plural-inflection in Dutch, the tendency that new Dutch compounds do not contain a linking morpheme cannot be explained by any theory which takes the "linking morpheme in Dutch compounds = plural marker" idea as a starting point. Some examples of relatively new "linking morpheme-less" Dutch compounds are given below:

Modern Dutch.

(10) a. bom-brief
'bomb-letter'

b. Dam-slaper
'Dam-sleeper' (Dam = name of square in Amsterdam)

c. adem-test
'breath-test'

d. jeugd-sentiment
'youth-sentiment'

e. kantoor-tuin
'office-garden'

f. softenon-kind
'softenon-child' (softenon = type of medicine)

g. wereld-winkel
'world-shop'

h. toren-flat
'tower-flat'

i. hitte-schild
'heat-shield'

j. wapen-wedloop
'arm-race'

The idea that linking morphemes in Dutch compounds represent the relic of case rather than plural is furthermore conformed by the fact that Dutch children hardly use any linking morphemes in their compounds, neither in existing, nor in compounds made up by the children themselves. In other words, Dutch children seem to avoid linking morphemes in compounds. The compounds made up by the children themselves are especially interesting, because these show that the children really see the word as a compound, whereas other compounds could be considered just as one lexical item, without any morphological segmentation. It is not surprising that Dutch children do not use linking morphemes in compounds. Linking morphemes represent case, but case is disappearing in Dutch, so there is hardly any positive evidence for Dutch children regarding case. From their input, they cannot figure out that case is overtly expressed and therefore there is no reason for them to use case-inflection in compounds. A search for self-made up compounds through the Dutch data in the CHILDES data system yielded the following results:[3]

(11) a. CHILD: allemaal werkje-vuur (T 2;9)
'all work-DIM-fire'
MOTHER: allemaal vuur-werk
'all fire-work'
b. CHILD: dat is rozijntje-brood (T 2;8)
'that is currant-DIM-bread'
MOTHER: nee, saucijz-e-brood he?
no, sausage-bread huh?
c. CHILD: das koffie-broek (T 2;9)
'that 's coffee-trousers'
MOTHER: koffie-broek???
'coffee-trousers???'
CHILD: ja
'yes'
MOTHER: koffie tubruk
'coffee "tubruk"' (= black coffee)
CHILD: nee! koffie-broek!!
'no! coffee-trousers!'

If these compounds were present in adult language, at least the first and the second would have a linking morpheme: werkje-vuur would be werkje-s-vuur; rozijntje-brood would be rozijntje-s-brood.

Nevertheless, in spite of the above story, there exist compounds in Dutch whose left branch appears to express real plural and not only case. These compounds seem to provide counterevidence against the theory proposed above. However, a closer look at them shows that all left branches of this type of compounds are pluralia tantum. Some examples are given in (12):

(12) a. paperass-en-berg

[3] This corpus contains longitudinal data from a Dutch boy, Thomas, who was recorded at least three times a week, from age 1;11 - 3;0.

'paper-PL-mountain'

b. war-en-huis
'ware-PL-house' (= department store)

c. chemicali-ën-fabriek
'chemical-PL-factory'

d. goeder-en-wagon
'good-PL-wagon'

e. kler-en-kast
'cloth-PL-closet'

Following Kiparsky (1982) I claim that pluralia tantum are inherently marked with [+PLURAL] at level 1 (or maybe even before level 1) and can therefore occur inside compounds, just as in English (clothes-brush, arms-race, odds-maker, etc.). Thus, the only way to have a real plural morpheme inside Dutch compounds is a plurale tantum. The rest of the linking morphemes always represent case. [4]

3.3 Back to Level Ordering

One could wonder now, whether the level ordering problem in Dutch is solved. Replacing the idea that linking morphemes in Dutch compounds are plural markers by the idea that they represent case inflection does not seem to make any difference with respect to level ordering. Supposedly, case inflection is regular and therefore assigned to level 3, just like regular plural formation, and thus, after compounding.

I would like to propose that the linking morphemes in Modern Dutch compounds are a remainder of case-inflection in Middle Dutch, but that they no longer represent the actual syntactic case. In other words, both linking morphemes ((-e)s and -en) in Modern Dutch still denote the semantic interpretation that was related to the Middle Dutch genitive case, but the actual syntactic case-inflection process is lost.[5]

---

[4] The spelling of linking morphemes in Dutch compounds is a different discussion. Since -en in Standard Dutch is optionally pronounced as a single schwa, I claim that both -e and -en as linking morphemes in compounds represent the same thing, namely genitive (or dative) case. Because of the pronunciation and the fact that the remainders of case inflection are no longer recognized by speakers of Modern Dutch, the case affix -en is sometimes spelled as the linking morpheme -e and sometimes as -en.

[5] Van der Horst (1981) reports that genitive case in Middle Dutch can represent various meanings. Besides denoting possession ("possessive genitive"), it can denote the agent of an action ("subjective genitive"), or the patient of an action ("objective genitive"). Furthermore, it can refer to a part of something

Let us assume that "X (left branch of compound) + linking morpheme + Y (right branch of compound)" should be interpreted roughly as: "Y, characteristically owned/produced by X". This becomes more clear if we compare the Dutch compounds with their English equivalents, which contain a possessive 's. Both the Dutch compounds and their English equivalents have a generic interpretation:

(13)

	DUTCH	ENGLISH
a.	(een) tovenaar-s-hoed	(a) magician's hat
b.	(een) bakker-s-diploma	(a) baker's certificate
c.	(een) boer-en-zoon	(a) farmer's son
d.	(een) ber-e(n)-vel	(a) bear's skin
e.	(een) mann-en-wereld	(a) man's world

The above examples show that the English equivalents of the Dutch compounds are not compounds (and thus not formed in the lexicon), but nouns with a preposed possessive (and thus formed in the syntax), which is also a relic of genitive case in Old/Middle English. However, semantically, they are equal. They also show that both the -(e)s and the -en linking morpheme can express a "genitive" (in the semantic sense) relation.

The idea that the linking morphemes in Modern Dutch represent a semantic notion rather than the actual syntactic case is further confirmed by the fact that there no longer seems to be a rule which predicts the occurrence of -(e)s or -en inflection, like there was in Middle Dutch. As we saw earlier on, Middle Dutch nouns that ended in an -e received -n inflection in the genitive, whereas nouns that did not have a final -e got

---

bigger ("partitive genitive"). Examples of each of them are given below:

(i) possessive genitive:
des coninc-s sale
'the king's room'

(ii) subjective genitive:
de grote daden God-s
'the great deeds/actions by God'

(iii) objective genitive:
de vreze des her-en
'the fear for the lord'

(iv) partitive genitive:
een pot bier-s
'a pot of beer'

It seems to me that at least the first three meanings are still found in Modern Dutch compounds, either represented by a linking morpheme or not.

-(e)s genitive inflection. Since case and therefore this rule is no longer productive in Modern Dutch, it is predicted that in principle, nominal compounds can freely choose between the various linking morphemes (because all of them equally represent the semantic notion related to genitive case). Therefore, it is not surprising that we find compounds which can have either -(e)s or -en as their linking morpheme, meaning exactly the same. This is illustrated in (14):

(14) a.	lerar-en-vergadering 'teacher's meeting'	leraar-s-vergadering 'teacher's meeting'
b.	olifant-en-huid 'elephant's skin'	olifant-s-huid 'elephant's skin'

Comparing English and Dutch, it seems that in English, the relic of genitive case, the possessive 's, still occurs as a syntactic process, but that in Dutch the remainder of genitive case appears on a lexical level. This leads us to the claim that in Dutch, an intermediate generation between Middle and Modern Dutch has changed the genitive case process from a syntactic into a lexical process. I propose this Modern Dutch lexical process to take place on level 2 as a word-formation rule.[6] This word-formation rule can be incorporated into the Compounding Rule, so that Dutch compounds, including the linking morphemes, are accounted for by the Compounding Rule at level 2, just like in English. This new Compounding Rule can be stated as follows:

(15) *Compounding Rule*:

$$X + \text{-(e)s/-en/}\varnothing + Y \rightarrow Y\text{, in possessive, subjective or objective relation to } X$$

Thus, in principle, Modern Dutch compounds are free to choose an -(e)s, -en or an empty linking morpheme. This is confirmed by a little experiment I executed, in which I asked Dutch native speakers to create new compounds. One of the results was that some of them used -(e)s, where others used -en and vice versa. For example, a hat, characteristically owned by a teacher was a leraar-s-hoed ('teacher's hat') according to one native Dutch speaker, but a lerar-e(n)-hoed according to another one (the ∅-linking

[6] Recall (from footnote 2) that there is still one syntactic instance of genitive case in Modern Dutch, namely proper names as preposed possessives, such as Freds gitaar ('Fred's guitar'). I must thus stipulate that Dutch still has some sort of a "Preposed Possessive Rule" (although a very restricted one) in order to account for these cases.

morpheme variant was not attested in this case).

The example given above exemplifies the possessive relation. The subjective relation is exemplified by god-s-wonder ('miracle caused by (a) God'), and an objective relation is found in mens-en-angst ('fear of people') for example. Note that also the "∅-option" for the linking morpheme is incorporated in the Compounding Rule. This is necessary, since compounds without a linking morpheme (or, with an empty linking morpheme) can express the same kind of relations between their left and their right branch (e.g. boek-∅-verkoper ('seller of books'), which expresses an objective relation).

To sum up, the Compounding Rule correctly expresses the fact that in Modern Dutch a possessive, subjective or objective relation between Y and X in a compound can be represented by either an -(e)s morpheme, an -en morpheme or a ∅ morpheme. As I mentioned before, this implies that the Modern Dutch speaker has different options to create new compounds. Therefore it can happen that one compound occurs in two or maybe even three different variations (cf. example (14)), or that one and the same left branch can be followed by either a -(e)s, a -en or a ∅ morpheme, as is illustrated in (16):

(16) a. maan-s-verduistering
'moon-s-darkening' (= moon eclipse)

b. man-e-schijn
'moon-e-shine'

c. maan-licht
'moon-light'

As Mattens (1984) points out, a language can offer different linguistic possibilities to express a compound, but it is the choice out of these possibilities made by the native speakers that determines which form is judged to be correct and which one is not. This does not mean that the other options are "incorrect" because they violate a linguistic principle. On the contrary, they are grammatically fine, but they are just not chosen by the community speaking that language. In other words, the acceptability of a compound with one of the available linking morphemes (-(e)s, -en or ∅) is determined by social factors, rather than by linguistic factors.

3.4 Some apparent problems

There are a few phenomena in Dutch which seem to form exceptions to the claim that

Dutch compounds can in principle choose freely between the three linking morphemes (namely, -(e)s, -en or ∅).

One of them is the case in which the left branch of the compound is a diminutive. (Modern) Dutch seems to have generated a "sub-regularity", which requires that the linking morpheme be -s if the left branch of the compound is a diminutive. For example:

(17) a. bloem-etje-s-behang
'flower-DIM-s-wallpaper'

b. huis-je-s-melker
'house-DIM-s-milker' (= slum lord)

In other words, no free choice out of the three linking morphemes is allowed in these cases. This sub-regularity could be captured by assuming that Dutch diminutives are inherently marked with a certain feature, let us call it [+S], which means that it "subcategorizes" for an -s in case it is further suffixed. The examples below show that this idea holds not only for the linking morpheme in compounds ((18 a)), but also for any other suffixation to a diminutive ((18 b and c)):

(18) a. bloemetje-s-behang (linking morpheme)

b. bloemetje-s (plural morpheme)
'little flowers'

c. bloemetje-s-achtig
'flower-DIM-ish'

Secondly, another apparent exception needs to be mentioned. Dutch has one group of common nouns which can get the suffix -er (but not necessarily) if it occurs inside compounds, and which receives the suffix -eren in case of plural. The following shows that it is extremely difficult, if not impossible, to discover any regularity in this group:

(19) a.	*singular*		*plural*
	rund		rund-eren
	'cow'		'cows'
	kalf		kalv-eren
	'calf'		'calves'
	lam		lamm-eren
	'lamb'		'lambs'
	ei		eier-en
	'egg'		'eggs'
	kind		kind-eren
	'child'		'children'
b.	rund-er-gehakt	*versus*	rund-vlees *but*
	'cow-?-ground beef'		'cow-meat'
c.	kalf-s-gehakt and		kalf-s-vlees
	'calf-s-ground beef'		'calf-s-meat'
d.	lam-s-vlees,		lam-s-bout
	'lamb-s-meat'		'lamb-s-meat'
e.	ei-er-dopje	*versus*	ei-geel
	egg-?-cup-DIM'		'egg-yellow' (= egg yolk)
f.	kind-er-stoel,		kind-er-meisje
	'child-?-chair'		'child-?-girl'

With respect to the compounds containing -er I would like to propose that the linking morphemes are empty and that the affix -er is a modification of the stem (along the lines of Collins (p.c.). Assuming this, a compound like kinderstoel has a ∅, rather than an -er linking morpheme. Again, I claim that social factors decide which linking morpheme (-(e)s, -en or ∅) should be used, also in this type of compounds. Nevertheless, it is not clear to me how the -er stem-modification should be incorporated in the system of lexical phonology/morphology or perhaps even in the Compounding Rule. We might have to stipulate that these type of compounds are all lexicalized. Furthermore, it seems to be strictly prohibited for a stem, modified by -er, to take an -s linking morpheme. Native Dutch speakers uniformly reject newly formed compounds in which a left branch in -er is followed by the -s linking morpheme, e.g. *rund-er-s-oog ('cow's eye'), *ei-er-s-eter ('egg's eater'). I leave this problem open to future research.

4. Conclusion

In this paper, I showed that Dutch compounds do not necessarily form counter-evidence against the hypothesis that level-ordering is innate. I claimed that the linking morphemes inside Dutch compounds are not plural markers, but the lexical residue of Middle Dutch genitive case. Furthermore, I proposed that Dutch has undergone a language change in which the syntactic process of prenominal genitive formation was converted into a lexical process, namely a word-formation rule on level 2. In this word-formation rule (the "new" Compounding Rule in (15)), the semantic information of the Middle Dutch genitive case is maintained, but the actual syntactic case properties are lost.

The "new" Compounding Rule in (15) expresses the fact that Modern Dutch can choose whether it uses the -(e)s, the -en or the ∅ linking morpheme in order to form compounds. This choice is in principle free; however, non-linguistic factors seem to determine which linking morpheme is accepted inside a certain compound and which one is not.

Since the linking morphemes -(e)s and -en are a relic of Middle Dutch genitive case, which is no longer productive in Modern Dutch, the theory described above correctly predicts that new compounds in Dutch are often formed without a linking morpheme (or, with an empty linking morpheme) and that Dutch children tend to avoid using linking morphemes in their novel compounds. Both phenomena follow from the fact that Dutch speakers (adults as well as children) have no access to the Middle Dutch case system, and therefore fail to see the original function of the linking morpheme so that they can omit it without violating any rule in their grammar.

Thus, an analysis of Dutch compounds along the proposals in this paper, is fully compatible with Kiparsky's level-ordering model and therefore supports the idea that level-ordering is innate, or universal.

Bibliography

Allen, M. (1978) *Morphological investigations.* Unpublished doctoral dissertation, Department of Linguistics, University of Connecticut, Storrs, CT.

Booij, G.E. (1977) *Dutch morphology: a study of word formation in generative grammar.* Peter de Ridder Press, Lisse, The Netherlands.

Gordon, P. (1985) "Level-ordering in lexical development". *Cognition,* 21, 73-93.

____. (1989) "Levels of Affixation in the Acquisition of English Morphology". *Journal of memory and language,* 28, 519-530.

Horst, J.M. van der (1984) *Kleine Middelnederlandse Syntaxis.* Huis Aan De Drie Grachten, The Netherlands.

Kiparsky, P. (1982) "Lexical Morphology and Phonology". In" The Linguistic Society of Korea (ed.) *Linguistics in the Morning Calm.* Hanshin Publishing Co., Seoul.

Kucera, H., and W.N. Francis (1967) *Computational analysis of present day American English.* Brown University Press, Providence, RI.

Loey, A. van (1960) *Middelnederlandse Spraakkunst.* Wolters, Groningen, The Netherlands.

Marle, J. van (1985) *On the paradigmatic Dimension of Morphological Creativity.* Foris, Dordrecht.

Mattens, W.H.M. (1984) "De voorspelbaarheid van tussenklanken in nominale samenstellingen". *De Nieuwe Taalgids,* 77, 333-343.

Senghas, A., J. Kim, S. Pinker and C. Collins (1991) "Plurals-Inside-Compounds: Morphological Constraints and their Implications for Acquisition". Paper presented at the 16th Annual Boston University Conference on Language Development, Boston, MA.

Siegel, D. (1977) Topics in English morphology. Unpublished doctoral dissertation, Department of Linguistics, MIT, Cambridge, MA.

Toorn, M.C. van den (1982) "Tendenzen bij de beregeling van de verbindingsklank in nominale samenstellingen, I en II". *De Nieuwe Taalgids,* 75, 24-33 and 153-160.

THE FINITE CONNECTIVITY OF LINGUISTIC STRUCTURE

Edward P. Stabler, Jr.
stabler@cognet.ucla.edu

1. Overview

While there is no interesting bound on the number of nouns that can occur in acceptable English noun compounds, there is a very low bound on the number of times that causative morphemes can be iterated in the verbal compounds of agglutinative languages. In languages that allow iteration at all, double causatives are slightly awkward, triple causatives are very awkward, and further iteration is impossible, even though we may be able to affix many tense, aspect and emphatic elements. Turning to the crossing dependencies in the clause-final verb clusters of Dutch, we find a similar bound. Clusters of verbs from two or three embedded clauses are acceptable, but clusters formed from deeper embedding are much more awkward. In the productive V-V compounding in Chinese resultative constructions, we again find that some V-V-V compounds, though awkward, can be accepted, while longer compounds become impossible. In languages that allow multiple wh-extractions at all, two extractions are fine, but more become progressively more unacceptable. We propose a very simple, unified account of these diverse phenomena: there is a finite bound on the extent to which acceptable constituents in any language can be related or "connected" to other positions in acceptable linguistic structures. In particular, we make the following rather programmatic proposal:

(1) There is a natural, finite typology of linguistic dependencies such that acceptability degrades quickly when more than two relations of the same type connect any constituent α (or any part of α) to positions external to α.
For example, when more than two or three $\bar{A}$-chains, A-chains, or head movement chains cross a constituent boundary, structures become unacceptable. Case-marking, constituency and θ-marking relations of the same type (agent, patient,...) across any boundary are also limited to two or three.

Miller and Chomsky (1963) made famous another similar bound: when we center embed more than two clauses, acceptability decreases rapidly. And verb-final clusters in some West Germanic languages other than Dutch exhibit nested rather than crossing dependencies, still showing very similar limits. The connectivity bound (1) does not predict this; it is perfectly compatible with arbitrarily deep center embedding. But the following hypothesis subsumes (1) and predicts these limitations on center embedding as well:

(2) Parsing involves the explicit recognition of all significant grammatical relations, and partial structures constructed by the human parser respect the same connectivity bounds as completed constituents.

This proposal also solves an old puzzle: How can we place a natural, finite bound on our parser that rules out deeply center-embedded structures but does not at the same time rule out many perfectly acceptable structures? This puzzle is solved by stating our connectivity bound in terms of a appropriate typology of linguistic dependencies. Most familiar parsers (top-down, bottom-up, left-corner, head-driven) do not respect (2) even in processing acceptable sentences, but we quickly sketch the outlines of one that does, a parser that is "frugal" in the sense that it never allows too many outstanding dependencies, never more than two or three of any given type. In fact, we argue that the bounds (1) and (2) can be regarded as providing a binary connectedness hypothesis about processing, relative to an appropriate finite typology of relations, where the bound has a certain squishiness for which we offer a certain speculative account.

The perspective developed here has a number of interesting consequences which are briefly noted. It fits well with theories that assume multiclausal sources for verbal complexes containing valency-changing elements, and suggests that the parsing of complex words may be rather more like syntactic parsing than has previously been assumed. More importantly, though, the remarkably modest resource demands of frugal parsers suggest that we should abandon the predominant perspective on human parsing: it is no longer helpful to think of the human parser as an infinite machine with some arbitrary and linguistically uninteresting finite resource bounds imposed on it by biological necessity. On the contrary, it appears that our linguistic resources have a definite finite structure which imposes particular limitations on processing, not like a machine with a potentially infinite, homogenous tape or stack, but like a machine with finitely many registers whose roles are very tightly constrained. This view lends itself to connectionist implementations while still allowing fully recursive and compositional treatments of language. Finally, we observe that recent arguments for the intractability of human language processing all crucially involve unbounded violations of (1) and (2), as one should expect.

2. A puzzle about valency-increasing affixes

In Swahili, there are a couple of causative suffixes, one of which is seen in (4):[1]

(3) Msichana a-li-u-fungu-a mlango.
girl SUBJ-PAST-OBJT-open-IND door[2]
'The girl opened the door'

(4) Mwalimu a-li-m-fungu-lish-a msichana mlango.
teacher SUBJ-Past-Obj-open-make-IND girl door
'The teacher made the girl open the door'

It is impossible, though, to causativize the verb twice:

(5) * Mwalimu a-li-m-fungu-lish-ish-a msichana mlango.

Considering the fact that it makes perfect sense to have someone make the girl open the door, this inability to iterate the causative morpheme may be surprising, but morphological restrictions of this sort are common. However, there are other languages which allow limited morphological iteration of causatives, and even languages that allow a particular causative affix to be iterated. This has been documented in Hungarian (Hetzron, 1976), Turkish (Zimmer, 1976), Kashmiri (Syeed, 1985), Kannada (Schiffman, 1979, p88), Kuki (Mahajan, 1982), Amharic (Hetzron, 1976), Awngi (Hetzron, 1969), Chichewa (Alsina, pc), West Greenlandic (Fortescue, 1984, §2.1.3.1.3) and other languages. These are of particular interest, because they may indicate what limits there are on iteration when there is no reason to think that general morphological, syntactic and semantic principles disallow iteration in general.

Consider the following constructions from Bolivian Quechua, an SOV language with the causative suffix *-chi*:[3]

---

[1] These first 2 examples are essentially those of Comrie (1976, p287), with minor changes for the dialect of my consultant, Deogratias Ngonyani. The inability to iterate the Swahili causative which I observe in (3) is also noted in other Bantu dialects by Givón (1976, pp337-339) and Abasheikh (1978, p133).

[2] In this paper, SUBJ stands for subject marker, OBJT stands for object marker, TOP stands for topic marker, IND for indicative, PROG for progressive, FUT for future, NEG for negative, CMP for complementizer, 1s for first person singular, 3s for third person singular, 3p for third person plural, OBJ for objective case, DAT for dative, ACC for accusative, BEN for benefactive, GEN for genitive, LOC for locative, DEL for delimitative, DUR for durative, EMP for emphatic, COM for comitative, and EUPH for euphonic elements.

[3] The Quechua judgements in this paper are those of Jaimé Daza, from Cochabamba. (7) is from a popular folk song.

(6) Tata-y-pis Mama-y-pis wañu-sa-nku
father-POSS1s-EMP mother-POSS1s-EMP die-PROG-3p
'My father and mother are dying'

(7) Tata-y-ta-pis mama-y-ta-pis yarqay-manta wañu-chi-sa-nku
father-POSS1s-OBJ-EMP mother-POSS1s-OBJ-EMP hunger-from die-make-PROG-3p
'They are starving my father and mother'

(8) Řiku-ni
see-1s
'I see it'

(9) Řiku-chi-ni
see-make-1s
'I show it'

Double causatives are also found in Bolivian Quechua, and they are typically semantically regular, though they are often slightly awkward:

(10) Řiku-chi-chi-ni
see-make-make-1s
'I have it shown'

(11) Susanitapaj leche-ta t'impu-chi-chi-ni
Susanita-BEN milk-OBJ boil-make-make-1s
'I boil milk for Susanita'

(12) ?? Tata-s-ni-y-ta wañu-chi-chi-sa-nku
father-PLURAL-EUPH-POSS1s-OBJ die-make-make-PROG-3p
'They are having my parents killed'

Interestingly, we seem to hit some sort of complexity boundary here. There is some variability among speakers, but in general verbs with more than two occurrences of *-chi* are extremely awkward or impossible:[4]

---

[4] A construction with three occurrences of *-chi* is listed in Herrero and Sánchez de Lozada's (1978, p216) descriptive grammar of Cochabamba Quechua:

(a) Susanitapaj t'impuchichichiy lecheta
'Have someone make boiling milk for Susanita'

However, the translation given by Herrero and Sánchez de Lozada for this triple causative is the one they give, and the one expected, for the simpler form *t'impuchichiy*. It is interesting that Mohanan (1982, p570) also lists a triple causative in Malayalam, but gives it the same translation as the double causative. And in Turkish as well, verbs like *göster–* (show) can take 2 causatives (yielding a verb that means "make someone have something shown"), and when further causative affixes are added they do not introduce additional intermediate causers, but add only an emphatic or humorous effect (Murat Kural, p.c.). In short, iteration beyond two causatives ceases to be valency-increasing. In

(13) * Řikuchichichini

(14) * Tatasniyta wañuchichichisanku

The possibility of two causatives as in (10-12) shows that the problem here is not simply due to an inability to repeat an affix, and it is certainly not due to some absolute upper limit on the number of affixes that can appear on a verb stem. Many more complex forms are perfectly acceptable:

(15) Suldadu-s wañu-chi-chi-lla-sa-nku-ña-puni.
soldier-PLURAL die-make-make-DEL-PROG-3p-DUR-EMP
'soldiers are still just having people killed as always'

Nor does the lack of productivity have any apparent semantic explanation. Notice that the causal morpheme makes a regular semantic contribution in all of the acceptable examples shown here, and so it is puzzling that we do not accept and interpret (13) in the regular way.[5] This lack of productivity in morphological causatives illustrated by (13-14) has been noted before in various dialects of Quechua[6] and in every other language in which morphological causatives have been studied (e.g., Comrie, 1981, p160). The collection of languages known to respect the bound includes both verb-final languages like Quechua, and also verb-initial languages like Amharic (Hetzron, 1976) and Arabic (Comrie, 1976, p286). Our first puzzle is to explain why this should be so:

(16) Why does the acceptability of complex verbs degrade quickly when they incorporate more than two causative affixes?

One response to (16) is to propose that there is an absolute upper bound on the number of arguments any verb can take, a bound on "semantic valency." Since causatives, unlike affixes of tense, aspect, and emphasis, increase valency, a bound is predicted. This idea does not explain, though, why transitives and intransitive verbs seem to have the same limits on iteration of the causative affix, as illustrated in the examples above with the intransitive *wañu-* and the transitive

Quechua such forms are certainly very awkward and quite rare. My consultant from Cochabamba, Jaimé Daza, finds (a) just as bad as (13) and (14). And I have been unable to find any triple causatives at all in any other Quechua literature. Hetzron (1969, §2.2.1; 1976, p383) describes two Awngi constructions as triple causatives, but in each of them the first "causative" forms a transitive from an intransitive form, and it is not clear that this step has the same semantic import as the later iterations. It would be interesting to study these exceptional constructions more carefully with speakers who find them acceptable.

[5] Like other languages with morphological causatives, some Quechua causatives have irregular, idiomatic meanings. No surprise here. All languages have idiomatic phrases.

[6] See Muysken (1977, pp125f), Weber (1989, p164), Cole (1985, p183).

řiku- in Quechua.[7] And in the second place, three causatives are not allowed even in the presence of an apparently valency-decreasing affix like a reflexive. In Quechua, reflexives seem to make both double and triple reflexives unacceptably complex:

(17) Řiku-chi-ni waway-ta dujtur-man
see-make-REFL-PROG-1s child-OBJ doctor-GOAL
'I had the doctor see my child', 'I had my child seen by the doctor'

(18) Řiku-chi-ku-ni
see-make-REFL-PROG-1s
'I have myself seen' (e.g. by a doctor), 'I give myself away'

(19) * Řiku-chi-chi-ku-ni
* Řiku-chi-ku-chi-ni
* Řiku-ku-chi-chi-ni

(20) * Řiku-chi-chi-chi-ku-ni
* Řiku-chi-chi-ku-chi-ni
* Řiku-chi-ku-chi-chi-ni
* Řiku-ku-chi-chi-chi-ni

So while we can agree that there is a limit on the number of valency changing affixes that any stem can host, the reason is not simply that there is an upper bound on the number of arguments that any verb (complex or not) can take.

A second idea about (16) is that the limit on causativization is due to the fact that there is a bound on the number of positions in any clause where the arguments of the causatives could go, perhaps for case reasons.[8] This idea does not seem quite right either. Even in languages where all the intermediate agents need not and even cannot be mentioned with overt NPs, the restriction mentioned in (16) holds. For example, in Quechua it is perfectly acceptable to have causatives with

---

[7] Comrie (1976, p286) suggests that, when all the causatives allowed by all languages are considered, causative forms of transitive verbs are less common than causative forms of intransitives. It is not quite clear how to assess this idea. In languages with quite productive causativization like Quechua or Turkish, how can we count the numbers of forms allowed? And in languages where all causative verbs are just lexical items, would we expect an interesting theoretical basis for a trend toward causative intransitives? In any case, since the facts about causatives vary so significantly across languages, it is hard to know what would follow from the conjectured trend.

[8] Something like this is suggested by Givòn's (1976, p337) speculation that the reason iteration of causatives is blocked in Bantu languages is due to "the lack of sufficient case markings to differentiate the semantic function of the various object nominals following the verb, since every application of lexical causativization increases the transitivity of the verb by one nominal object." Although we do not take exactly this view, we agree with Givón's idea that the limitation is "probably motivated by speech processing considerations."

no overt NPs at all, as we saw in (9-10). Furthermore, this idea fails to explain why languages with very rich case marking and pre- or post-positional systems like Quechua do not allow more iterations than those with impoverished systems. In Quechua, for example, besides the objective case marker, other postpositions can be used to mark intermediate agents, as in (17) or,

(21) Eulalya-wan uj punchitu-ta awa-chi-ko-rqa-ni
Eulalia-COM a poncho-OBJ weave-make-REFL-PAST-1s
'I had Eulalia weave a poncho'

And finally it is worth noting that this proposal shares a defect with the previous one: it does not explain why transitives and intransitives have the same limits on causativization. So, in general, we could not use a lack of role-markers in a language to account for the severe restriction on causative iteration unless we had a language where every argument had to be overt, and we had an independent reason to suppose that only some very small subset of these role-markers could be used in any language.

Two basic types of theoretical approaches to morphological causatives and other valency-changing affixes can be distinguished in the literature. One approach with a long history maintains that these constructions are derived from biclausal syntactic structures by incorporation of the verb from the lower clause into the higher causative. Double causatives then would call for underlying structures having three clauses. Recent prominent views of this sort are provided by Marantz (1984) and Baker (1988), for example. Baker (1988) argues that causatives are formed by verb raising in the syntax, an instance of head movement. We can schematize the basic idea of such approaches with a picture like the following:

(22) *Syntactic Causativization*
$[_{VP} \ldots [_{V}$ make-V$_i] \ldots [_{VP} \ldots$ t$_i \ldots]]$

Baker treats other reflexive and reciprocal markers similarly, as independent syntactic units that are incorporated into the verb:

(23) *Syntactic Reflexivization*
$[_{VP} \ldots [_{V}$ V-self$_i] \ldots [_{NP} \ldots$ t$_i \ldots]]$

The order of constituents in these schemata is irrelevant, of course, and for present purposes it does not matter whether we suppose that the parts of a causative complex come together by movement. The important point is that causatives are treated like verbs with their own syntactic phrasal projections and argument positions, somehow incorporating or merging with verbs from an embedded phrase which also has its own argument positions. At some level of representation, there is a connection between the valency-changing elements of the complex verb and other syntactic positions.

An alternative approach treats the combination of the causative morpheme with a verb as a lexical operation. Causatives induce a certain lexical mapping between argument structure and syntactic expression, a certain "morpholexical operation on argument structure." We could use the following sort of picture to indicate that causativization adds an argument position:[9]

(24) *Morpholexical Causativization*

$$\begin{array}{c} \emptyset \\ \Downarrow \\ \langle \theta \ldots \theta_{causer} \ldots \rangle \end{array}$$

Reciprocalization and reflexivization are similarly treated as morpholexical operations that "suppress" one role of the verb, as indicated in the following schema:

(25) *Morpholexical Reflexivization*

$$\begin{array}{c} \langle \theta_i \ldots \theta_i \ldots \rangle \\ | \\ \emptyset \end{array}$$

This approach thus fits with the general program of resisting the "syntacticization of grammatical phenomena" by providing a purely lexical account of causativization.

In the present context, it is clear that nothing in either of these basic approaches to causatives immediately solves the puzzle (16). On the syntactic approach, morphological causatives are derived from complex syntactic structures, and yet periphrastic causatives do not seem to be subject to the same sort of restriction:

(26) The private killed the reporter.
The sergeant always made the private kill the reporter.
The general made the sergeant make the private kill the reporter.
The president made the general make the sergeant make the private kill the reporter.

9 This schema for causativization is modeled on Bresnan and Moshi's (1990) schema for applicative, and the following schema for reflexivization is exactly the one they suggest for reciprocalization. Clearly, these representations suppress details about what the operations involve. In his study of causatives in Chicheŵa, Alsina (1992) proposes that the Chicheŵa causative denotes a three place relation between a causer, a patient, and the caused event, and that when this morpheme combines with another "embedded" predicate, the patient argument of the causative is fused with some argument of the embedded predicate. For present purposes, this account fits the scheme (24), since the net increase in arguments is one. Mohanan (1982) also proposes a morpholexical analysis of causatives in Malayalam.

The corporate executives made the president make the general make the sergeant make the private kill the reporter.

No one makes the corporate executives make the president make the general make the sergeant make the private kill the reporter.

...

So the syntactic approach could formulate its own version of the puzzle:

(27) **Syntactic Source Puzzle:** Since complex periphrastic causatives like (26) are acceptable, why does the restriction on morphological causatives mentioned in (16) hold?[10]

Similarly, the puzzle remains unsolved in the morpholexical approach:

(28) **Morpholexical Source Puzzle:** Since it is plausible that at least some morpholexical operations can iterate, as we see in double and triple causatives, why does the restriction mentioned in (16) hold?

This paper will argue that the restriction mentioned in (16) follows from a very simple and intuitive complexity bound on syntactic structures. Since this bound can be independently motivated, (16) actually supports a syntactic analysis like Baker's rather than threatening it. That is, this independently motivated bound immediately solves the Syntactic Source Puzzle (27), whereas the morpholexical approach would have to propose another, separate but very similar bound on lexical complexity to answer (28).

3. A hypothesis about syntactic connectivity

Syntactic representations are usually depicted as trees, but this is very misleading. Actually, many relations are assumed in the syntax, relations which do not correspond to any of the arcs shown in constituent structure trees at any level of representation. We get a better image of syntactic structure when we imagine enriching our tree representations with arcs corresponding to every relation among constituents: relations of theta-marking, agreement, case-assignment, antecedence in chains, control, and perhaps others. Phrasal structures in human languages then look not like

[10] Baker (1988, p71) points out, we can always assume that there are special morphological filters which simply rule out unacceptable forms that we would have predicted, on syntactic grounds, to be well-formed. Two points about this idea. First, this move is unappealing unless it is really necessary. There are certainly regularities here that we just do not understand yet, and it is not clear exactly what morphological filtering will be needed. Second, notice that even in the extreme case where every morphological causative is assumed to be learned separately, we would still face a version of our basic puzzle in trying to explain why languages tend not to call on the child to learn double or triple causatives, and never quadruple causatives. Pinker (1984, §8) provides a nice discussion of some of the peculiarities of causatives and some ideas about how lexically-based variations on productivity might be learned.

bare trees but like trees tangled with vines.[11] Let's consider the nature of these structures more carefully.

We assume that there are, in addition to basic constituency relations, a finite number of dependencies or "licensing relations" among constituents. One goal of linguistic theory is to identify these relations and show how they determine significant properties of linguistic structures. How many such relations are there in any structure? An n-place relation can be regarded as a set of n-tuples, so considering any linguistic structure S with constituent α, we define the following measure of complexity:

(29) The **connectivity of α in S** is k if and only if k n-tuples must be removed from S to completely disconnect α (and all constituents of α) from the rest of the structure.
In a complete structure S, we will say that **the connectivity of S** is k just in case k is the maximum connectivity of any constituent α in S.[12]

In these terms, we will argue that although there is no finite bound on the connectivity of grammatical structures, there is a connectivity bound in acceptable structures:

(30) There is a finite bound k such that, in any language, no acceptable structure has a connectivity greater than k.[13]

This immediately implies that if we enrich an acceptable n-node tree structure with arcs representing all the other linguistic relations to make it completely explicit, we will need no more than kn arcs altogether. That is, there is a linear upper bound on the number of arcs in an explicit structure as a function of the number of constituents. In fact, we will move to a number of much stronger and more specific claims about the connectivity bounds in the course of a brief survey of

[11] It is of course possible to require that all relations be represented "in the architecture of the tree," so that, in effect, an arc relating any two nodes passes through the least common ancestor of those two nodes. This is a particular hypothesis about how dependencies are represented. Roughly speaking, the idea that all dependencies should be handled by feature-passing mechanisms, as in systems like GPSG (Gazdar et al., 1985) and most unification-based models of grammar (Shieber, 1992), is an instance of hypothesis. In these frameworks, if the typology of features was appropriate, the the proposal stated below would imply a bound on the complexity of the feature structures required at any node in acceptable sentences.

[12] Connectivity is a standard measure of structural complexity in mathematics (Gould, 1988), which we here relativize to the linguist's notion of phrase structure constituency. Notice that this definition applies to n-ary relations for any n.

[13] Notice that a structure can be unacceptable for reasons other than connectivity! Given a particular k, this condition provides a sufficient, but not a necessary condition for unacceptability.

some constructions that might appear to threaten this claim. First let's see how such bounds could account for the limitations on morphological causatives discussed in the previous section.

We assume that complex causative verbs are verbs, X0 constituents in syntactic structure. We do **not** assume that there is any interesting bound on the complexity of X0 constituents in general. For example, English noun compounds can contain 2 or 4 or 8 or more nouns and still be perfectly acceptable:

(31) The customer called to ask about [$_N$ ticket validation].
The customer called to ask about [$_N$ parking ticket validation].
The customer called to ask about [$_N$ parking lot ticket validation].
The customer called to ask about [$_N$ grocery parking lot ticket validation].
The customer called to ask about [$_N$ grocery store parking lot ticket validation].
...

Noun compounds are not subject to the restriction on the iteration of morphological causatives, we conjecture, because additional nouns in a noun compound do not generally increase the connectivity of the complex.[14] When we incorporate a verb into a causative, on the other hand, we increase the connectivity of the causative complex.

Our claim is that we see a connectivity bound in the limitations on causative constructions in languages like Bolivian Quechua or Turkish. Notice, to begin with, that each causative is related to a causer which may or may not be explicitly expressed in the sentence. Recent theories say that the causatives assign agent θ-roles to NP positions. So if V has an agent θ-role to assign, then V+*make* assigns two agent θ-roles. Some recent theories assume that this assignment is done indirectly, with each verb assigning its θ-roles from a different position, before it is moved into the complex, perhaps by head movement. In either case, though, causativization increases connective complexity, either by increasing the number of agent θ-role assignments directly, or by increasing the number of head-movement relations that involve the complex. So we conjecture that the following (repeated from (1)) holds in all languages:

[14] Ward et al. (1991) point out that internal elements of noun compounds and other X0 elements may sometimes increase connectivity because they may be involved in anaphoric relations, as in:

(b) Although [cocaine$_i$ use] is down, the number of people using it$_i$ routinely has increased.
[McCarthy$_i$ites] are now puzzled by him$_i$.

However, the anaphoric relations here are pronoun-antecedent relations, which are not local in the way most grammatical relations are, and we will suggest in §5.3 below that they are not be subject to the same connectivity bounds as other relations.

(32) There are a finite number of different types of θ-roles and the acceptability of any constituent degrades quickly when it, or any parts of it, assign more than two or three roles of the same type.

And the acceptability of any constituent degrades quickly when it, or any parts of it, are involved in more than two or three head-movement chains.

The former claim is less than precise until we provide a clear typology of the possible θ-roles, but we have focused on a clear case.[15] Causative affixes introduce argument positions, and plausibly two occurrences of the same affix will be related to two argument positions by the same type of θ-role assignment. So given an analysis of causatives like Baker's, or any other in which these argument positions are external to the verbal complex, the noted restrictions on causative iteration are entailed by (32), and we have solved the puzzle (16). That is, assuming that each causative is related to an agent position outside the X0 complex, either by movement or by θ-assignment, we predict acceptability to be substantially degraded in any complex that has two or more causative elements. Notice that we have also solved the puzzle (27), since the non-morphological causative constructions like (26) do not have any constituent that is connected by assignment of more than one agent θ-role to the rest of the structure. In other words, this approach realizes the common sense idea that the difference between morphological causatives and the non-morphological constructions is that the morphological complexes can pile up unacceptably many unsatisfied requirements on the rest of the structure.

Before looking at other constructions with regard to whether the very general claim (32) is really independently supported, or even tenable, notice that the morpholexical approach is not rescued by this proposal. Causativization adds an argument, increasing connectivity, but if that argument is suppressed by a reflexive or reciprocal, connectivity is decreased again. So the morpholexical approach has no account of the contrast between the acceptable double causative (10) and the triple causative with a reflexive (20). In fact, adding a reflexive morpheme even to the double causative as in (19) seems to make it slightly more, not less awkward. This is predicted by the syntactic approach, since reflexivization, like causativization, increases the connective complexity of the verbal complex. The morpholexical approach, on the other hand, has the same difficulty with this fact as does the simple valency-based idea discussed above. More generally, if (1) and (2) apply uniformly here and in the syntax, as this paper will argue, this supports syntactic and mixed

[15] The latter claim is also imprecise until we provide a clear characterization of head-movement chain. We discuss this in §4.1, just below.

approaches to agglutinative and polysynthetic word structure, leaving a puzzle for approaches that assume valency-changing morphological operations are purely lexical.

4. Other X0 connectivity bounds

The hypothesis (32) is quite strong and very general, so we attempt only a very preliminary defense of it in this paper. We begin by considering other X0 complexes. There are languages with even richer verb incorporation structures than we see in Quechua, but in these languages we find similar bounds on the number of valency-increasing elements in V-V complexes. For example, Fortescue (1984, §2.1.3.1.3) describes such bounds in West Greenlandic. Going the other direction, let's consider briefly whether there are similar bounds on X0 complexes in languages that do not have multiple verbs forming a single phonological word.

4.1. Verb clusters in West Germanic

It is natural to consider at this point the famous causative constructions in West Germanic languages, the clause-final "verb clusters." According to some analyses, these clusters are formed by verb raising, in which a verb in an embedded clause moves to get its inflection and then up into the higher clause and the resulting complex can then move to combine with the verb of the higher clause, and so on.[16] For example, in Dutch we have:

(33) dat [Jan [Piet [Marie t_3] t_2] t_1] [zag [laten zwimmen$_3$]$_2$]$_1$
that Jan Piet Marie saw make swim
'that Jan saw Piet make Marie swim'

Furthermore, we find similar clause-final clusters in German, but in the reverse order:

(34) dass Jan [Piet [Marie t_3] t_2] [schwimmen$_3$ lassen$_2$ sah]
that Jan Piet Marie swim make saw
'that Jan saw Piet make Marie swim'

Constructions with clusters that have verbs from 4 different clauses are not acceptable in either Dutch or German. Since the dependencies in the German examples are nested, and the dependencies in the Dutch examples are crossing, the idea that they would have similar psychological complexity is rather surprising, but was confirmed in an interesting study done by Bach, Brown and Marslen-Wilson (1986). At two levels of embedding, no difference in complexity at all was found. Three

[16] Den Besten and Rutten (1989), Hoeksema (1988), Haegeman and van Reimsdijk (1986), Den Besten and Edmondson (1983).

levels of embedding was significantly more complex in both languages, but here the center embedded structures actually seemed to be slightly more intelligible. It would be interesting to account for this slight difference (see §8, below), but the more striking result is the similar jump in complexity when we add a third verb. (32) predicts this jump in complexity if we assume that each of these verbal elements assigns an agent θ-role to an NP outside of the complex. Actually, following recent theoretical proposals, it is preferable to assume that the θ-roles are assigned from the original, embedded positions of the verbs, and that the complex is formed by a cyclic verb-raising operation. In this case, each element is related by a chain of head-movements to its embedded position, and the observed bound would be predicted by an assumption which could be formulated roughly along the following lines:

(35) The acceptability of any constituent degrades quickly when it is related to other positions by more than two or three head-movement chains.

Since head-movement analyses of these Dutch constructions are still not well developed, it is not clear that (35) is properly formulated. The problem is that, in many recent theories, it is not clear how the relation between the moved constituent and its original position is represented.[17] (35) suggests that these positions are related by a chain, but the identification of the relevant chains is complicated by the fact that recent theories often assume that when a head moves to another head position, the indices of those two positions must be identified in order to allow for "proper government."[18] So when two verbs move from embedded clauses to the matrix clause, the usual representation is such as to suggest that there is only one chain along which all the elements have been "amalgamated," and it is not clear how the origins of each verb are indicated.

---

[17] There are more fundamental challenges to these head-movement accounts as well. For example, a quite different account of these constructions has been proposed by Kroch and Santorini (1991). On their alternative account, the clause-final sequences of verbs in Dutch and German do not form a constituent. Rather, in Dutch the embedded verbs are extraposed, adjoining to the main clause, forming structures like,

(c) dat [[[Jan [Piet [Marie t_3] t_2] zag] laten$_2$] zwimmen$_3$], (Dutch)

while in German, the embedded verbs can stay in place:

(d) dass [Jan [Piet [Marie schwimmen] lassen] sah] (German)

Even on this alternative analysis, our proposals predict the observed limitations. Notice that in (c), if the minimal indicated boundary containing *Jan...zag* is regarded as a real constituent boundary, the length of the sequence is still restricted by (32), since that principle applies to all constituents, not just X0 constituents. The German construction, on the other hand, becomes a center-embedding construction. As indicated in the introduction, this is ruled out by an elaboration of our connectivity bound, as we will discuss in detail in §6.

[18] See, e.g., Chomsky (1986), Pollock (1989).

A related problem comes from the fact that the V V V sequences that we see in the examples above are actually sequences of 6 or 9 or more elements, because each verb has associated tense and subject agreement. In fact, each main verb in the final cluster may be associated with an overt and phonologically distinct auxiliary verb, in which case we could actually see additional elements in an acceptable cluster, though all of these elements will have come from just three separate clauses. Depending on what the right account of these structures is, we may need to distinguish various kinds head-movement chains, for example to distinguish the chains of main verbs that θ-mark elements in their underlying positions.

4.2. Limitations on other verbal compounds

Although there is no interesting bound on the complexity of N compounds in English, as we observed in (31), there are interesting bounds on other types of compounds found in other languages. We considered Germanic auxiliary verb compounds in the previous section, but some languages allow a more productive verbal compounding that has rather different characteristics. Consider, for example, the "resultative" V-V compounds in Chinese.[19] In Chinese, we find compounds like the following from Li (1990):

(36) Daiyu jiao-dong-le Xiangling zheishou shi.
Daiyu teach-understand-ASP Xiangling this poem
'Daiyu taught Xiangling this poem and as a result Xiangling understood it'

(37) Baochai dai-gei Miaoyu yixie cha.
Baochai carry-give Miaoyu some tea
'Baochai brought Miaoyu some tea'

(38) Daiyu ku-zou-le henduo keren.
Daiyu cry-leave-ASP many guest

[19] I distinguish these V-V compounds from the "serial verb" constructions of West African languages like Kwawu Akan and Carribean creoles (Jansen, Koopman and Muysken, 1978; Baker, 1989), since the verbal elements in these constructions do not appear to form a constituent. For example, as Jansen et al. point out, though in the Carribean creole Sranan you can say:

(e) A waka go
he walk go
'he walked in a different direction'

various syntactic tests show that *waka go* is not a constituent here. For example, material can intervene between the two verbs, and although either of the two verbs can be fronted, they cannot be fronted together. The V-V sequences in Hindi (Hook, 1974) come in many varieties, but at least for many of them, the situation there seems similar. That is, it is not clear that they really form X0 compounds.

'Daiyu cried so much that many guests left'

It is interesting to note that the complexity of these resultative verbal compounds is limited, more like the morphological causatives and verb clusters, and less like noun compounding in English.[20] Consider the following examples:[21]

(39) ?? Baoyu ba Daiyu qi-hui-dao-le jia.
Baoyu OBJ Daiyu anger-return-arrive-ASP home
'Baoyu angered Daiyu with the result that Daiyu went back and arrived home'

(40) ?? Daiyu ku-zou-diao-le keren.
Daiyu cry-leave-away-ASP guest
'Daiyu cried so that the guest left and was away'

(41) * Baoyu ba Daiyu da-ku-zou-diao-le.
Baoyu OBJ Daiyu beat-cry-leave-away-ASP
'Baoyu beat Daiyu so that Daiyu cried and left and was away'

If we assume that these elements of these compounds increase the connective complexity, we predict the unacceptability of these longer compounds.[22] With some thought, even the structures marked unacceptable here are intelligible, but this is no surprise, and may be the result of some sort of "rescuing" heuristics that are at least partially non-linguistic (see §8).

4.3. The complexity of verbal clitic complexes in Romance

Clitics in Romance are sometimes thought to form V0 complexes by adjoining to the verb (e.g., Kayne, 1975, §2; 1989). This proposal is motivated in part by evidence of movement-like relations

---

[20] Chinese has other constructions that involve verb sequences that are not so awkward, such as

(f) Ta hui xiang qu tan gangqin
He will want go play piano
'He will want to go to play the piano'

However, as the gloss would suggest, the verbs do not form a compound here. We just have a sequence of verbs embedded in a control structure.

[21] These examples and judgements are from Andi Wu.

[22] Li (1990) proposes a lexical source for these compounds and allows various kinds of merging of the θ-grids of the component verbs. Since the result of adding a new verb to a compound with such a merging operation need not have any higher valency or connective complexity than the original verb, it is not clear why compounding should be bounded on this account. Against clausal sources for the elements of these compounds, Li suggests that resultative clauses are adjuncts, and as such it should not be possible to move the verb from the resultative up to the main verb by head movement. However, it is not clear that these resultatives are adjuncts. Notice that when we have full clauses as resultatives, it is not possible to have two of them.

between clitics and the argument positions of the verb. The complexity of these constructions appears to be strictly bounded by grammatical requirements, but they are complicated enough that it is worth considering whether they provide counterexamples to our claims.

It is clear that argument clitics stand in significant relations to other structural positions. In languages like standard French, for example, an object clitic prohibits the appearance of an overt object. This relation between the clitic and the object has some similarity to movement relations, but in languages like Spanish the clitic can be doubled. That is, there can be both an object clitic and an overt object. Even in these doubled constructions though, there is clear evidence of grammatical relation between the clitic position and the doubling phrase. For example, when a clitic is doubled, the clitic must agree in number, person, and gender with the doubling object (Jaeggli, 1981, pp12ff). In the present context, it is interesting to consider how many clitics can be associated with a single verb. In French, we get constructions like the following (more causative constructions!):[23]

(42) ? Jean fera detourner Marie du droit chemin a Pierre a Paris.
Jean will-make detour Marie of-the the right road to Pierre to Paris
'Jean will have Pierre detour Marie from the right road to Paris.'

(43) ? Jean le lui y en fera detourner
Jean him to-her there of-it will-make return
'Jean will make him detour her from it that way'

(44) ?? Jean te le lui y en fera detourner.
Jean you him to-her there of-it will-make return
'For your interest, Jean will make him detour her from it that way'

The *te* in the last example is the "ethical dative," which signifies something like "for your interest." It is presumably not related to any argument of the verb or even to any position inside the VP. So in both of these last examples we have just 4 clitics associated with other VP positions. In French, this seems to be the maximum number.[24]

It is interesting to note that the relations these clitics bear to other structural positions are not all of the same kind. Sportiche (1992) observes, for example, that the genitive clitic *en* licenses

[23] Thanks to Dominique Sportiche for these examples.

[24] We can also have a negative clitic, and a subject pronoun clitic, but it is not plausible that these are part of the verbal complex. This is argued by Kayne (1975), Sportiche (1992), and others. Also see the brief remarks on these clitics below.

parasitic gaps and does not allow stranding of an associated quantifier, while dative clitics do not license parasitic gaps and do allow stranding of an associated quantifier:

(45) Marie en_1 a presente [le frere t_1] [a la soeur e_1]
Marie of-him introduced the brother to the sister
Marie introduced his brother to his sister

(46) * Marie en a tous lu le livre t_1(, de ces auteurs)
Marie of-them has all read the book (of these authors)
Marie has read their book (of these authors)

(47) * Marie $leur_1$ offrira des bonbons t_1 [sans demander e_1]
Marie to-them will-offer some candies without speak
Marie will offer candies to them without speaking to them

(48) Marie $leur_1$ offrira tous t_1 des bonbons
Marie to-them will-offer all some candies
Marie will offer all of them some candies

In (45), the genitive clitic *en* can bind both the position t_1 in the direct object, and also the parasitic gap e_1 inside of the prepositional phrase. In (46), we see that the quantifier *tous* cannot be stranded in a position related to the genitive clitic. The examples (47-48) show that dative clitics have the opposite properties. Sportiche (1992) suggests that these differences can be explained if we assume that clitic positions are related to argument positions by phrasal movement relations. In particular, he proposes that each clitic is the head of a projection, and that it must be possible to move the relevant constituents of the VP into the specifiers of the corresponding clitic. The observed differences among the clitics can then be subsumed under the generalizations that only $\bar{A}$-movements license parasitic gaps (Chomsky, 1986), and that quantifiers cannot be stranded in A positions (Sportiche, 1988). We simply make the independently supported assumptions that the genitive clitic is related to an $\bar{A}$-position (inside of an argument of the verb) by an $\bar{A}$-movement, while dative clitics are related to A-positions by A-movements. Sportiche (1992) argues that accusative clitics are also related to their arguments by $\bar{A}$-movements, offering a different explanation of why we do not find them licensing parasitic gaps.

If this account is on the right track, then there is a significant split among relations between clitics and verbal positions: some involve A-movements, while others involve $\bar{A}$-movements. This suggests that the connective complexity of structures with clitic-V0 compounds respect the bound on phrasal connectivity suggested in (1), namely,

(49) When more than two or three elements are extracted from a constituent by Ā-movement, or by A-movement, or by head movement, structures become unacceptable.

That is, assuming that the clitics we have been considering essentially involve Ā-movement and A-movement relations to positions internal to the VP, this bound is respected. Assuming at least that the specifier of the genitive and accusative clitics are Ā-positions, we have at most three A-movements (subject, dative, locative) in the slightly awkward (43), and two Ā-movement relations (genitive, accusative). This fits with our bound (1).[25]

This account, according to which there is a movement of (possibly empty) argument NPs to the specifier positions of the corresponding clitics, weakens the argument for the view that the verbal clitics are adjoined to the original verb V0 to form a huge adjunction structure. Movement-like properties can be attributed to the relationship between the specifiers of the clitics and argument positions. The clitics could simply cliticize to the verb in the phonology in suitable conditions, without forming any single constituent in the syntax, like subject clitic pronouns and the negative clitic *ne*.[26]

To conclude, our observations of even the most complex sorts of X0 constituents show that their relationships to other elements in linguistic structure are severely limited. There are many other grammatical relations involving heads that we have not considered, but the finite connectivity bound (30) seems quite secure for X0 elements. And, at least for X0 constituents, our much stronger bounds on θ-marking, A-movement, Ā-movement, and head-movement relations are, I think, currently tenable though in need of much more careful assessment. We turn now to a brief consideration of the connectivity of phrasal constituents.

---

[25] Notice that even if this particular proposal does not work, it is plausible that we will still be able to respect our connectivity bound, because the evidence is very good that the relations between the different clitics and corresponding argument positions are of different kinds.

[26] This is not the view Sportiche (1992) adopts, though. He argues for phrasal movements to the specifiers of the clitics, and head movements to form a large complex of the verb and clitics up to but not including subject and negative clitics. On this view, assuming that the clitic-V0 complexes themselves are formed by head-movement, and that causative restructuring also involves head-movement, we seem to have 5-head movement relations involving a single complex in (43): [le [lui [y [en [fera [detourner]]]]]]. There is still a finite bound on the number of head-movement relations that involve this complex, but the bound is larger than the two or three allowed by (35). In this case, to preserve our bounds in something like their current form, we might want to consider distinguishing different kinds of head-movement chains, as discussed in §4.1, above.

5. XP connectivity bounds

Phrases are involved in many different kinds of grammatical relations, but as in the case of X0 constituents, to support our finite connectivity claim (30), we will focus on relations that are potentially increased by recursive operations, since these are most likely to yield high degrees of connectivity. Here we consider just XP movement relations and anaphora.

5.1. Ā-movements

The total number of movements out of any constituent is severely limited. Let's distinguish movements of a phrase to an argument position, A-movements, from movements to a non-argument position, Ā-movements. In terms of these two basic types of movements, we have offered the conjecture (49), repeated here:

(49) When more than two or three phrases are extracted from a constituent by Ā-movement, or by A-movement, structures become unacceptable.

Although English allows multiple questions when one or more wh-phrases are left *in situ*, it only marginally allows Ā-extractions of more than one wh-phrase (here we use bold to indicate stress on the *in situ* wh-constituent):[27]

(50) Who_1 did you ask [t_1 to fix the car **how**]?

(51) ? Who_1 did you ask how_2 [t_1 to fix the car t_2]?

Other languages are much more liberal. For example, Mahajan (1990) points out that in Hindi, there are cases where wh-phrases cannot be left *in situ*, but must be extracted:

(52) * [rām-ne kahā ki kOn kis-ko māregā]
Ram-erg said that who whom hit-FUT
'Who did Ram say will hit who?'

(53) * kɔn_1 [rām-ne kahā ki t_1 kis-ko māregā]
who Ram-erg said that whom hit-FUT
'Who did Ram say will hit who?'

(54) kɔn_1 kis-ko_2 [rām-ne kahā ki t_1 t_2 māregā]
who whom Ram-erg said that hit-FUT

[27] Notice that we will want to say that we have one wh-extraction from the clause in the following structure, even though there are 4 traces:

(g) Who does he feed t_1 and wash t_1 and clothe t_1 and support t_1?

'Who did Ram say will hit who?'

Though multiple extractions are allowed and sometimes even required in this language, constructions get progressively more awkward when more than two or three wh-elements are extracted:[28]

(55) ? kis-ne$_1$ kɛse$_2$ [rām-ne kahā ki t_1 t_2 gaRīThīk kī]
who how Ram-erg said that car fixed
'Who did Ram say fixed the car how?'

(56) ??? kis-ne$_1$ kis-ko$_2$ kɛse$_3$ [rām kahā hE ki t_1 t_2 t_3 mārā hogā]
who whom how Ram said that killed
'Who did Ram say that killed whom how?'

(57) ? kis-ne$_1$ kis-ko$_2$ kyā [rām-ne sītā-ko batāyā kī t_1 t_2 t_3 diyā hogā]
who whom what Ram-erg Sita told that gave-FUT
'Did Ram tell Sita who gave what to whom?'[29]

Furthermore, we can show that the movements are contributing to the awkwardness in these constructions by exhibiting a pair in which the movements are optional, where the wh-constituents can be left *in situ*:

(58) ? rām jāntā hɛ ki kis-ne kis-ko kyā kyɔn diyā
Ram knows be that who whom what why gave
Ram knows who gave whom what why

(59) ???? kis-ne$_1$ kis-ko$_2$ kyā$_3$ kyɔn$_4$ [rām-ne kahā ki t_1 t_2 t_3 t_4 diyā hogā]
who whom what why Ram-erg said that gave-FUT
Ram said who gave whom what why

These limitations Hindi provide further support for (49), though they remind us that, beyond the bound of two or three relations of each type, unacceptability increases at different rates depending on the construction. We return to this point in §8.

5.2. A-movements

Movements like passive and NP-raising in English are typically classified as A-movements, that is, as movements to an argument position. Unlike Ā-movements, A-movements are typically clause-bound, and so their contribution to connective complexity is immediately limited by clausal complexity. In fact, Sportiche's proposal about Romance clitics (discussed in §4.3) apparently takes us

[28] Example (55), like the previous Hindi examples, is from Mahajan (1990, §3). Thanks to Anoop Mahajan for providing the additional examples below, with careful comparative acceptability judgements.

[29] This Hindi sentence can be interpreted either as a multiple question, or as a multiple indirect question where the wh-elements have been topicalized.

to the limit here, since on that view, all arguments of a verb (including the external, subject argument) are related to VP-internal positions, and all of the arguments can be moved out of the VP to the specifiers of clitic positions at LF. The analysis is not uncontroversial, but if the movement of the genitive and accusative phrases are Ā-movements, we will remain under the proposed bound.

5.3. Anaphoric relations

The landmark work of Berwick and Weinberg (1984) argued that since pronoun-antecedent relations apparently extend well beyond the local domain needed to make structural decisions, and since a speaker's determination of these relations is apparently based on inferences from general background knowledge, it is plausible that these relations are not computed by the same mechanism that builds syntactic structure. This argument continues to be persuasive, and so we do not assume that pronoun-antecedent relations are subject to the same sorts of low finite bounds that apply to other types of linguistic relations.[30] That is, we assume that there is a finite bound on the number of pronoun-antecedent relations that any acceptable expression can involve, but it does not seem to be an interestingly low, roughly binary or ternary bound of the sort we have with other dependencies. As for other binding relations, we have already discussed antecedent-trace relations, and we have suggested that these are bounded in number according to the type of chain involved. The binding of English reflexive pronouns and other overt anaphors is similarly local, so we predict that it conforms to the same low bound.

6. A hypothesis about syntactic processing

As observed in the introduction, the connectivity bound (1) for linguistic constituents does not predict the well-known unacceptability of deeply center embedded sentences. Notice that in a simple constituent structure tree, with no other linguistic relations, every constituent is connected to the rest of the structure by the constituency relation represented by the arc from the constituent to its parent, and perhaps also by an (immediate) precedence relation to a sister node.[31] Unbounded

[30] We will not consider in any detail the processing issues associated with computing pronoun-antecedent relations, but it is worth noting one attractive option which is actually part of the strategy Berwick and Weinberg proposed. Since pronouns are required to be free in their own "binding domain," we could assume that in structural processing each pronoun is associated with some representation of the set of local elements to which it cannot be bound. The extra step of selecting an appropriate binding is not assumed to be part of structural processing. This is the idea adopted by Giorgi et al. (1990).

[31] Whether linear precedence is relevant in syntax is a controversial matter (e.g., Marantz, 1989, p114f), as is the nature of precedence constraints. However, in the derivation trees of context free grammars, it is natural to regard

crossing dependencies cannot be defined in context free grammars because of this limitation: they require arbitrarily higher connectivity. However, context free grammars can define languages with arbitrarily deep center embedding: this does not entail any additional connectivity, intuitively, because the embedded constituents do not stand in the way of any of the grammatical relations of higher constituents. In English center embedded clauses, too, we presumably have somewhat greater connectivity than in context free languages, but it need not increase with depth of embedding. For example, in the following examples there is some modification relation between each embedded clause and the N′ phrase it modifies, but we do not have increasing numbers of connections to elements in the matrix clause when we add a second or third center embedded clause:

(60) The house [(that) the malt lay in] was built by Jack
(61) * The house [(that) the malt [(that) the rat ate] lay in] was built by Jack

Why should center embedding be subject to such similar bounds as the other constructions we have surveyed? As in the other cases, when we have two clauses as in (60), the construction is fine; with three clauses as in (61), the construction is extremely awkward or impossible. To subsume this fact, we simply assume that not only completed constituents, but also the partial constituents constructed by the parser, are subject to our connectivity bound. This was our proposition (2), repeated here:

(2) Parsing involves the explicit recognition of all significant grammatical relations, at all levels of representation, and partial structures constructed by the human parser respect the same connectivity bounds as completed constituents.

For the moment, we will be rather vague about what the partial structures are like, exactly. We can think of them as sets of nodes, or as descriptions of some sort, or as sequences or stacks of subtrees. The relevant parameter is the number of relations connecting nodes in any partial structure to nodes that are not in that partial structure.[32]

---

every elementary subtree as corresponding to a production, with the children ordered exactly as in the right hand side of the production.

[32] Notice that we have also been rather vague about what the completed linguistic representations are. For present purposes, we can think of them as something like the S-structures of recent transformational grammars, but I prefer the idea that the relevant structure is something like logical form (LF). On most approaches, this level of representation is usually the most elaborated, indicating grammatical relations most completely. So if we say that the parser computes all relations, all levels of representation, these could all be encoded in just a slight augmentation of LF structure, where the augmentation indicates where various elements were at the time of "spell-out," at S-structure (Wu, 1992). With this assumption, the connectivity of an element could be directly assessed on that structure.

In any case, if we assume that parsing proceeds from the beginning of a sentence to the end, and that parsing never involves the prediction of overt, lexical X0 elements, then after seeing the first three NPs in our example (61), the parser presumably has a structure with something like the following form:

(62) [The house (that) [the malt (that)]]

This structure has already has two NPs, each of which must be case- and θ-marked by constituents that have not been found yet, and so we are already at the bound. If, as the parse continues, we get to

(63) [The house (that) [the malt [(that) the rat]]]

we now have three θ-relations connecting the partial structure to constituents outside of the structure, and we correctly predict that the structure is extremely awkward or impossible. We get this prediction no matter how much of the syntactic structure is built for these elements, since whatever the structure is, it will somehow involve relations to three θ-assigning verbs, three V0 elements, which have not been found yet. For example, even if we build VPs for each of the subject we have found, and assume that these VPs compositionally assign θ-roles to the subjects, we will still have three outstanding θ-relations between these VPs and the verbs that have not been found yet. Alternatively, if we assume that all subjects originate VP-internally, we will have 3 outstanding A-movement relations.

We should quickly survey a range of related constructions at least briefly, to support the claim that the proposed bound is always respected in acceptable structures, and to get an idea of the predictive power of the proposal. As noted above, our proposals provide only sufficient conditions for unacceptability; they do not purport to explain all unacceptability! Recall that the proposal is that when two relations of the same kind extend beyond a given partial structure, there is no problem as far as connectivity goes, though there may well be other problems with the structure. However, when there are three relations of the same kind, we predict that the structure should be more awkward, and four of the same kind should be unacceptable. This is what we found with Quechua causatives, and now we find a similar bound is respected in the partial structures of center-embedded relative clauses. Let's turn now to sentential subjects with sentential subjects, as in

(64) [That the rat ate the malt] surprised Jack.
(65) * [That [that the rat ate the malt] surprised Jack] bothered Mary.

If we suppose that the matrix clause of a declarative English sentences is a CP with an empty complementizer which selects (and perhaps θ-marks) an IP complement, then when the parser gets the partial structure:

(66) $[_{CP}\ [_C\ e]\ [_{CP}\ [_C$ that$]\ [_{CP}\ [_C$ that$]\ ]\ ]\ ]$

it already has three relations of the same type connecting it to IPs that are not built. If, on the other hand, we build the IP nodes, matters only become worse, since now we have:

(67) $[_{CP}\ [_C\ e]\ [_{IP}\ [_{CP}\ [_C$ that$]\ [_{IP}\ [_{CP}\ [_C$ that$]$ IP $]\ ]\ ]\ ]\ ]$

where each of the three indicated IPs is related by constituency to a subject and to an I' constituent. So this structure has at least 4 dominance relations connecting it to the completed tree. We leave a careful treatment of constituency relations to the next section, but as we will see, 4 dominance relations is too many.

Finally, notice that whereas it is unacceptable to have two center embedded relative clauses, and it is unacceptable to have two embedded sentential subjects, two relative clauses in object position is not so bad:

(68) ?? She loves the house [(that) the malt [(that) the rat ate] lay in].
(69) ?* She loves the house [(that) the malt [(that) the rat [(that) the cat chased] ate] lay in].

We can allow for this, since after we have parsed

(70) She loves the house [(that) the malt [(that) the rat]]

we have two, not three NPs that lack θ-roles, because *loves* has already θ-marked its object. This explains the contrast with superficially similar (63).

Finally, it has often been observed that a relative clause inside a sentential subject is not as bad as a sentential subject inside a sentential subject, nor is it as bad as a relative clause inside a relative clause in a subject NP. Consider, for example,

(71) ? [That the rat [who the cat chased] lives in the house] surprised Jack.
(72) * [That the rat [who the cat [who the dog bit] chased] lives in the house] surprised Jack.

Here, the parser presumably builds a partial structure like the following,

(73) $[_{CP}$ That the rat [who the cat]]]

In this structure, we seem to have two NPs that need θ-roles, and also a CP subject that needs a θ-role, so why isn't the structure as awkward as the structure beginning with (63)? This question is easily answered because the two structures are very different, as has long been recognized. One prominent tradition has argued, for example, that sentential subjects are not really in subject position at all, but are rather in topic position, where the phrase can escape being case-marked. In this $\bar{A}$ topic position they somehow license an empty subject (perhaps as a trace of topicalization), but they cannot be regarded as elements of case-marked chains.[33] This explains why we cannot get sentential subjects in many of the places where NP subjects can occur:

(74) * Is that the rat lives in the house likely?
(75) * John doubts (that) that the rat lives in the house is likely.

Now we can explain the relative acceptability of (71), because topics do not need to be θ-marked in the regular way, as we can see in other constructions more commonly treated as having topics:

(76) That the rat is in the house, Mary could not doubt (it) for a minute.
(77) That the rat is in the house, who could possibly doubt (that)?

So the partial structure (73) has only two NPs with θ-relations to elements outside of that structure, and is within our bounds.[34]

7. A previous account: reviewing the support for our typology of relations

The proposals made above generalize and unify many earlier ideas, but they are particularly close to Gibson's (1991, §8) theory of processing overload. The proposal made there can be described roughly as follows:

(78) In every partial structure Γ built by the parser, each node whose θ-marker is not yet unambiguously specified and included in Γ imposes a processing load *int*.
Similarly, each node in Γ which θ-marks a constituent XP such that no N, V, A or P in XP is yet included in Γ, imposes a processing load *int*.
And each node which can receive a θ-role and which selects a constituent XP such that no N, V, A or P in XP is yet included in Γ imposes an additional processing load *tt*.
With this typology of parsing loads, partial structures imposing a load greater than K are unacceptable, where

33 E.g., Emonds (1976, §IV.2), Koster (1978), Stowell (1981, pp152ff), Safir (1985, §3.4), Emonds (1985, pp314ff).

34 Notice that even if the particular analysis of sentential subjects proposed here is ultimately refuted, the prospects for maintaining the proposed bound still look good because of the abundant evidence that the relations between sentential subjects and the rest of the sentence are different in character from the relations between simple NP subjects and the rest of the sentence.

$$5 * int > K$$

$$4 * int \leq K$$

$$(4 * int) + tt > K$$

$$(3 * int) + (2 * tt) \leq K$$

We will not discuss this theory in detail, but it is worth comparing its main features with the present proposal. First, and most importantly, Gibson shows that the bound K is respected in a wide range of structures, and yet K is very low! Furthermore the bound involves relations between the partial structure and elements that have not been found yet. These two most salient features are preserved in the present proposal. The difference which matters is in the typology of relevant relations: our typology stands in a closer relation to linguistic theory; it is a finer classification which (as will be briefly argued below) provides a more accurate account of acceptability bounds; and when the bound is relativized to this typology it has a natural functional motivation, as will be argued in §8, below.

Gibson accounts for the difficulty of center-embedded relative clauses in the following way. In the partial structure,

(79) [The house O_1 (that) [the malt O_2 (that)]]

there are already two lexical noun phrases that need to receive θ-roles, imposing a load of $2 * int$. In addition, according to recent theories about relative clause structure, there are two empty operators which, in the completed structure (61), get associated with the θ-roles of positions in the relative clauses, increasing the load by $2 * int$. Gibson also claims that the second complementizer in this structure, which is either the overt *that* or empty, imposes a processing load because it needs to be related to an IP none of which has been found yet, imposing a load of *tt*. So, assessing only these two types of relations *int* and *tt*, he concludes that $(4 * int) + tt > K$.

Our proposal, on the other hand, involves the use of a finer classification of linguistic relations, as we have already seen. The two empty operators in (79) are actually related by $\bar{A}$-movement to the rest of the structure, and so we have two $\bar{A}$-movement relations, and two θ-marking relations (or, if the subject NPs originate VP-internally, two raising, A-movement relations). If we assume that complementizers select their complements by subcategorization, we have this relation in addition. Distinguishing these linguistic relations with a linguistically typology, rather than classifying them all as one of two types *int* or *tt*, we can still explain the awkwardness of the completed structure, as we saw in the discussion of (63) above. But our typology provides different, and better results on a wide range of constructions. Let's quickly review our account to see why this is so.

Consider again the restriction on morphological causatives. If we assume that a morphological causative relates an agent, the causer, to a theme, which is typically either a clause or an NP denoting an event, then what is wrong with a Quechua sentence like (13), repeated here:

(13) * Řiku-chi-chi-chi-ni
see-make-make-make-1s

If we assume a lexical source for this verb, then it might be regarded as having three or four agents, one of which is the subject, one of which is in direct object position, and the remaining objects (at most one of which can be overt) could be expressed in adjuncts. The agreement morphology indicates that the highest agent is first person singular. It is not clear how Gibson's proposal would rule this out, since 4 unconnected θ-relations are presumed to be acceptable. Some elaboration of that theory might be made to fit the facts, but it is not clear how it should go. Alternatively, if the verbal elements of the complex come from embedded clauses, it is again unclear how to rule out the sentence. Then we have 3 verbs each related by head-movement to embedded positions, and θ-roles were of course assigned from those positions. Perhaps Gibson would treat the movement relations as unconnected lexical requirements, but again it is not clear how the story would go, especially since Quechua is SOV, and so if the structure is built from left to right, the empty embedded clauses would presumably have to be built before the complex verb was inserted.

Similar remarks apply to the complexity bounds on Chinese verbal compounds and to verb clusters in Dutch. Gibson (1991, §8.3) discusses verb clusters, assuming that elements of these clusters assign θ-roles.[35] While he agrees that clusters of verbs from 4 clauses are very marginal, he claims that the crucial transition to unacceptability occurs in the step to 5 verbs. (Structures with more than 5 unconnected θ-relations of this kind are just pruned from consideration, on his account.) This view does not have anything to say about the surprising jump in complexity between 2 and 3 that was demonstrated and discussed by Bach, Brown and Marslen-Wilson (1986), nor is the extreme awkwardness of 4 verb clusters predicted. This view also does not fit well with recent

[35] Pritchett (1992, p157n7) seems to suggest that the bounds that Gibson uses to block center-embedding in English are too low to account for the sequence of subject and object NPs in Germanic languages, but here we make the slightly different claim that Gibson counts processing load with respect to too coarse a typology of relations. By conflating wh-movement relations involving empty operators with θ-marking of the subject, he gets a bound that is too high to account for the limitations in causatives or Dutch verb clusters.

theories which say that the overt verbal positions in these structures are not the ones responsible for direct θ-assignment; rather, the positions overtly filled are related by movements to the θ-marking positions.

A more interesting problem arises with verbal clitics in Romance. Consider again the slightly awkward (43), repeated here:

(43) ? Jean le lui y en fera detourner
Jean him to-her there of-it will-make return
'Jean will make him detour her from it that way'

We suggested that the 4 clitics here were related by 3 A-movements and 1 $\bar{A}$-movement to NP positions inside VP. Gibson would have a hard time allowing this structure, though, since at the time when the speaker has processed the subject and these 4 clitics, there are presumably 5 θ-relations unconnected, one for the subject and one for each clitic:

(80) [[Jean] [le lui y en]]

Here, a finer typology of relations can come to the rescue, as discussed above.

A similar problem for Gibson's theory arises for phrasal movements. Consider again (57), repeated here:

(57) ? kis-ne$_1$ kis-ko$_2$ kyā [rām-ne sītā-ko batāyā kī t_1 t_2 t_3 diyā hogā]
who whom what Ram-erg Sita told that gave-FUT
'Did Ram tell Sita who gave what to whom?'

This is awkward, as we predict since it has 3 wh-extractions, but Gibson's theory predicts that it should be completely unacceptable, since at the stage where the speaker has heard only

(81) kis-ne$_1$ kis-ko$_2$ kyā [rām-ne sītā-ko]

there are 5 NPs without θ-roles.

The following section defends an important extension to our proposal that further distinguishes it from Gibson's: the inclusion of constituency among the bounded relations. With this step, a new conception of human parsing begins to become plausible.

8. Constituency and frugal parsers

In light of hypothesis (2), one dependency worth considering more carefully is constituency. It would be appealing to assume that this, most familiar relation is bounded like the others we have discussed. In fact, it appears that it can be bounded as follows:

(82) No acceptable structure is related by more than two or three *immediately dominates* relations to other structural elements, or by more than two or three *is immediately dominated by* relations.

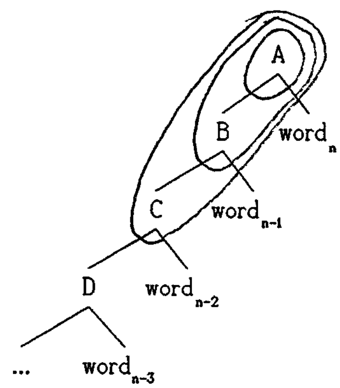
Let's assume that human constituent structures are at most binary branching, following Kayne (1983), Larson (1988; 1990; 1991) and many others. Then (82) is trivially satisfied by any completed tree of phrase structure constituency relations. Certain partial structures are ruled out, though. This kind of view has been considered before, and so we can go rather quickly to the heart of the issue.

First of all, it is well known that purely top-down or bottom-up parsing strategies produce partial structures with unbounded constituency connections, even on acceptable sentences.[36] For example, any parser that builds structure in the standard top-down order quickly go beyond the frugal bounds on left branching structures, even if the only relation we consider is constituency, and yet left branching structures are not subject to interesting bounds of the sort we have been discussing, as we see in:

(83) [[Jack's] house]
[[[Jack's] mother's] house]
[[[[Jack's] mother's] sister's] house]
[[[[[Jack's] mother's] sister's] dog's] house]
...

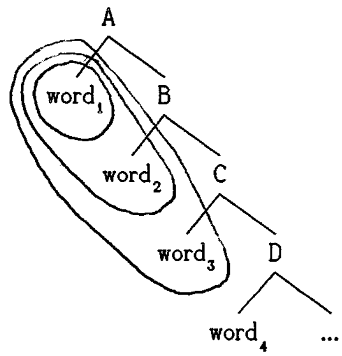
The following diagram shows the how a top-down parser builds structure in parsing a left branching construction:

36 For recent discussions of this point, and comparisons with other recent ideas about complexity, see Gibson (1991) and Abney and Johnson (1991).



We see that after building the root node A, we have two elements to find; after building node B, we have three constituency connections; and so on. A traditional stack-based implementation of a top-down context free parser remembers these constituency relations by keeping a record of them in a stack, but notice that this problem arises just from the order in which structure is built. The problem does not depend at all on whether the parser uses chart- or stack-based structure sharing, whether or not it pursues a single parse at a time or many in parallel, whether it backtracks or is deterministic, whether or not it uses lookahead. The point is not specific to context-free parsing either. For example, it applies in exactly the same way to top-down stack-based implementations of tree adjoining grammars of the sort discussed by Joshi (1989) and Vijay-Shanker (1987, §3.2); these devices similarly posit partial structures with unbounded connectivity in parsing left-branching structure.

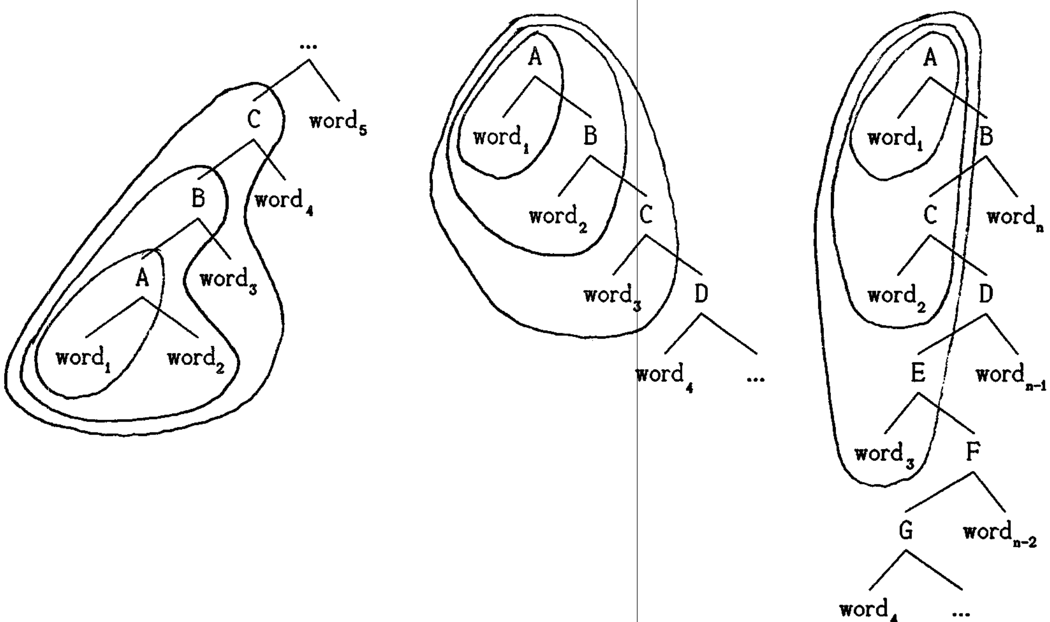
Similarly, parsers using the standard bottom-up order of structure building posit unboundedly connected partial constituents on right-branching structures, even though there is no interesting bound on the depth of right-branching allowed in human languages. The problem for bottom-up parsers can be depicted with a diagram like the following, showing the order in which structure is built:



Again, we see that, processing the input from left-to-right, the number of constituency arcs extending beyond the structure to the right increases with every word. After accepting $word_1$, one constituency arc relates the incomplete structure to the rest of the tree, but after $word_3$ has been processed, we already have three constituency arcs connecting the partial structure to the rest of the tree. This problem arises whether these parsers backtrack or are deterministic, whether they use lookahead or not, whether they build one structure at a time or not.

Some parsing strategies, though, do not posit unboundedly connected partial structures in parsing purely right and left branching binary trees. For example, the operation of a certain kind of left-corner parser can be depicted as follows:[37]

[37] Abney and Johnson (1991, p245n) note that standard left-corner parsers do not work this way, but that a simple modification reduces the stack requirements. The modification simply (nondeterministically) merges the parent P of a left corner L to a predicted category Q already on the stack, before the sisters of L have been completed, so long as the parent category P and the next predicted category Q match. The left-corner parser described in Gibson (1991, §3.3) does not include this modification. These issues, and particularly the relation between connectivity and that stack requirements of conventional parsers, are treated in more detail in Stabler (1992a).



We can be more precise about the behavior of this parsing strategy if we define the depth of center embedding as follows:

Node n in tree T has **ce-depth** 0 iff there is no factorization (cut) $\alpha n \beta$ through T, where neither α nor β is the empty sequence of categories, and

Node n has **ce-depth** k+1 iff the most deeply center-embedded ancestor n' of n such that the subtree with root n' has a factorization $\alpha n \beta$ where α and β are non-empty is such that the ce-depth of n' is k.

In these terms, we can see that the ce-depth of the nodes in the left and middle trees, above, never exceeds 1. But in the rightmost tree we can calculate the ce-depths as follows:

The nodes labeled A, $word_1$, B and $word_n$ have ce-depth 0.

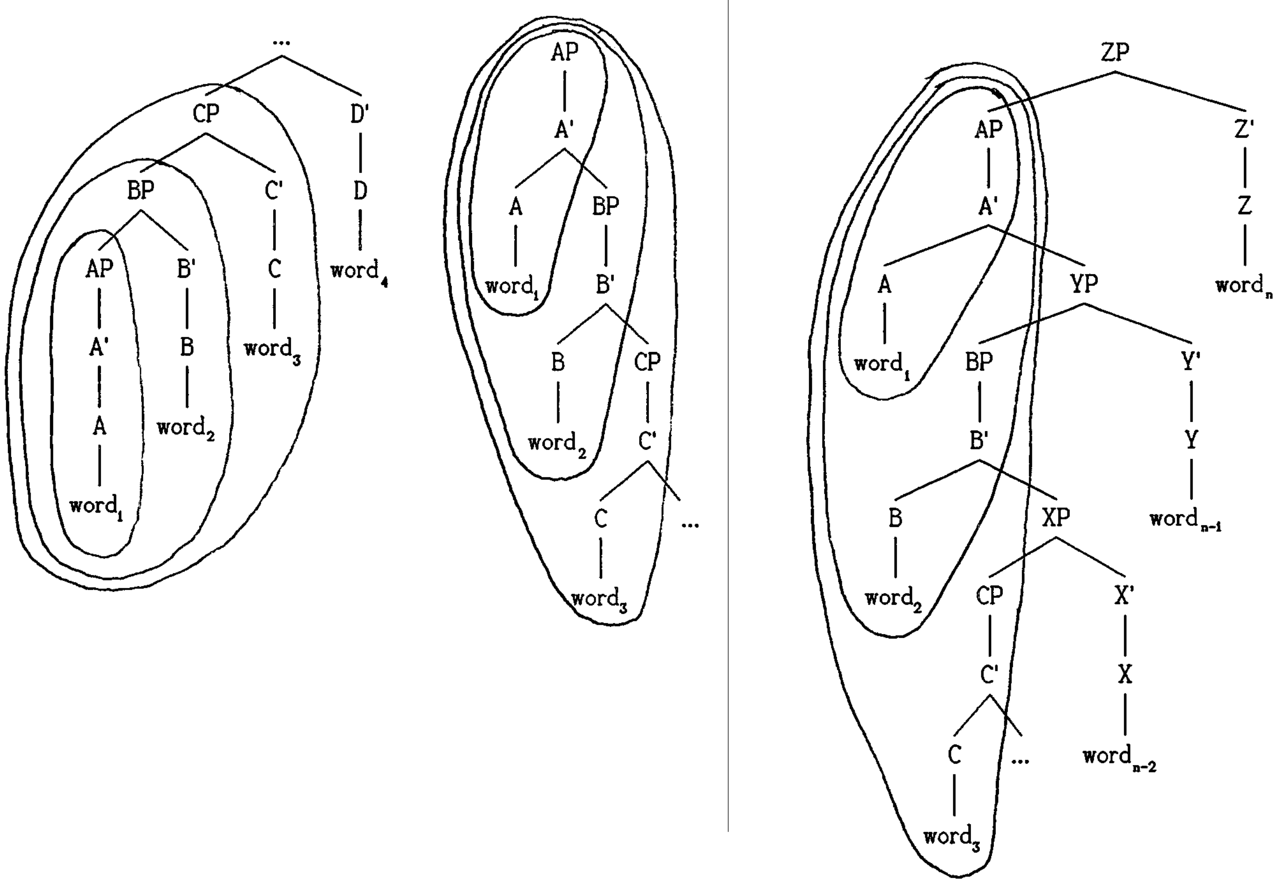
C has ce-depth 1 because $word_1$ C $word_n$ is a factorization of the subtree with root A.

Nodes labeled $word_2$, D and $word_{n-1}$ also have ce-depth 1.

> Looking at items that are center-embedded under C, we see that E, $word_3$ and F have ce-depth 2.

The closed curves in the diagrams above show the partial structures constructed by our left-corner parsing strategy. The connectivity of these partial structures can be determined simply by counting the arcs that cross the curves. So it is easy to see that when our left-corner parser builds the nodes with ce-depth 2, shown in the rightmost diagram, the partial structure has 5 arcs extending across the boundary of the indicated partial structure to the right: 3 *immediately dominates* arcs and 2 *is immediately dominated by* arcs. Clearly, as ce-depth increases, the numbers of these relations connecting the partial structure to the whole will increase.

Head-driven parsers similarly build structure from left-to-right, projecting structure upward, but instead of projecting just a parent and siblings of each left corner, each lexical element X0 projects all the way to the phrasal level XP. These parsers also require few constituency relations between the partial structures and the rest of the tree, except when there is center embedding:



Again, we see that connectedness of the partial structures indicated by the closed curves is bounded in the left and right branching structures, but increases in the center-embedded structure. In the rightmost structure shown above, we can calculate:

The nodes labeled $word_1$, A, A′, AP, ZP, Z′, Z and $word_n$ have ce-depth 0.

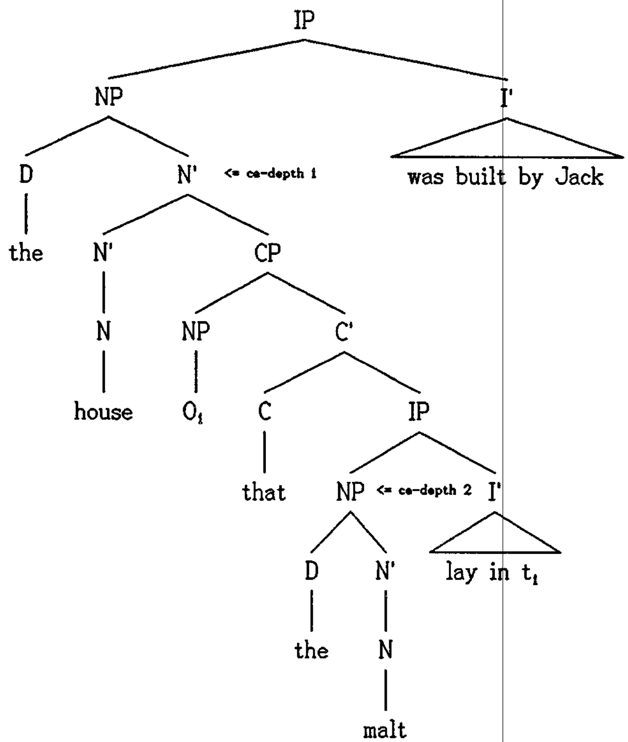
YP has ce-depth 1 because $word_1$ YP $word_n$ is a factorization of the subtree with root XP.

Nodes labeled $word_2$, B, B′, BP, Y′, Y and $word_{n-1}$ also have ce-depth 1.

Looking now at items that are center-embedded under YP, we see that $word_3$, C, C′, CP, XP, X′, X and $word_{n-2}$ have ce-depth 2.

So it is easy to see that when our head-driven parser builds the nodes of CP with ce-depth 2, shown in the rightmost diagram, the partial structure has 6 arcs extending across the boundary of the indicated partial structure to the right (or 5 if C has no complements): 3 *immediately dominates* arcs and 3 *is immediately dominated by* arcs. And clearly, as ce-depth increases, so will the connectedness of the partial structure.

So now let's turn to the constituent structures of human languages, considering the constituency connections of partial structures. We have just seen that there is no interesting bound at all if the human parser is purely top-down or purely bottom-up, but if we adopt a mixed strategy such as we find in left-corner and head-driven parsing, we may find a bound. These strategies only require highly connected partial structures on center-embedded constructions, and as we saw in the previous section, center-embedding seems to be bounded in human languages. Since both left-corner and head-driven strategies propose partial structures exceeding the bound (82) only beyond ce-depth 2, this fits with our account of center-embedded relative clauses of English. Consider the following structure, for example, in which only constituency relations are indicated:



If we add another center-embedded relative clause, the structure becomes unacceptable.

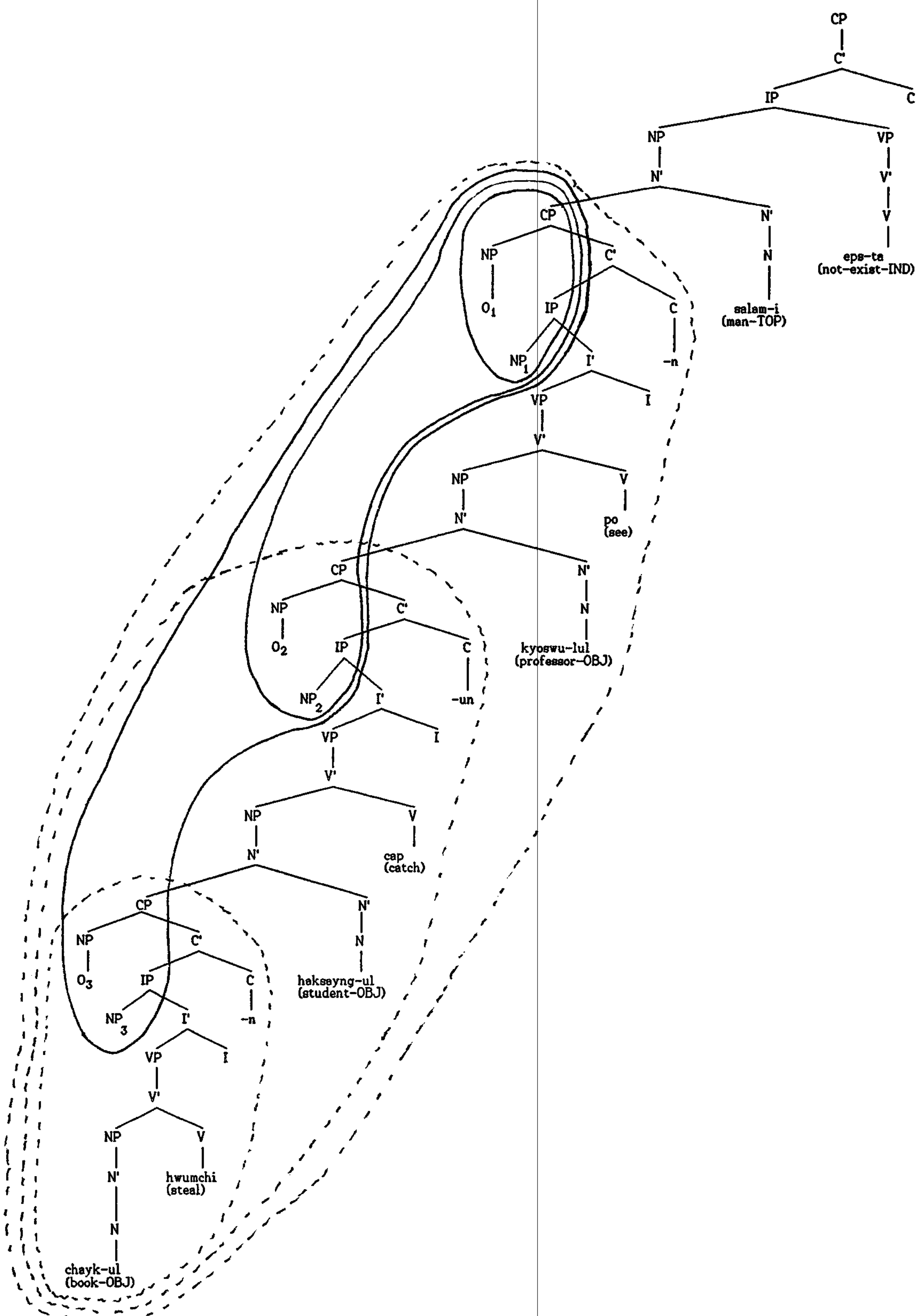
When we look at other constructions, though, matters are not so simple. We observed above that some constructions, such as (71), two center-embedded clauses can be acceptable. This seems to pose a problem for a low constituency connectedness bound like (82). We find even more challenging forms of the same problem in certain relative clause constructions from other languages. For example, in the following Korean constructions, we see that three embedded relative clauses is not too bad! The jump in complexity that produces a really unacceptable structure happens when we embed a fourth clause:[38]

(84) ? [[[Chayk-ul hwumchi-n] haksayng-ul cap-un] kyoswu-lul po-n] salam-i eps-ta.
book-OBJ steal-CMP student-OBJ catch-CMP professor-OBJ see-CMP man-TOP not-exist-IND
'There is no man who saw the professor who caught the student who stole the book'

[38] Thanks to Seungho Nam for discussion of these structures and for judgements (which were also checked with a number of other Korean speakers), and to Wu (1988) for drawing my attention to these constructions. Wu (1988) and Chao (1991) discuss essentially identical structures in Japanese, structures which raise the same issues about constituency bounds for partial structures.

(85) ?* [[[[Chayksang-wuy-ey iss-tan] chayk-ul hwumchi-n] haksayng-ul cap-un] kyoswu-lul po-n] salam-i eps-ta.
desk-on-top-of-at exist-CMP book-OBJ steal-CMP student-OBJ catch-CMP professor-OBJ see-CMP man-TOP not-exist-IND
'There is no man who saw the professor who caught the student who stole the book which is on the desk'

Presented with the bracketing just shown, the structure (84) may appear to be left-branching, and hence unproblematic for our constituency bounds. But most recent analyses of these constructions suggest that there are empty categories in the subject positions of the embedded clauses, in which case we have quite deep center-embedding in this acceptable structure. The following tree for (84) shows that the most deeply embedded verb has a ce-depth of 7:



Standard left-corner and head driven strategies will, in successful parses of this structure, produce partial structures that are connected by quite a few constituency relations, as indicated by the closed curves drawn with solid lines in the figure above.

Rather than taking the acceptable structures shown above as showing that partial structures can be connected by so many constituency relations, we reject the idea that the human parser uses a standard left-corner or head-driven strategy. In fact, there are independent grounds for rejecting those strategies. They both build structure from left-to-right, triggered in bottom-up fashion by the elements that we find at the leaves of our constituent structure. But this idea is hopeless once we allow empty constituents that can occur as left-corners or heads of XPs that have not yet been predicted. That is, the search space for these procedures becomes intractable, simply because any number of empty elements can be found anywhere in the input string.[39] We propose abandoning the idea that structures are built just from left-to-right, with the leftmost leaf triggering the first partial structure, in favor of the idea that structure building is triggered by overt elements. In other words, empty elements are projected top-down, while overt elements trigger the building of structure bottom-up. This idea solves both the search problem and the connectivity problem, though it forces us to abandon left-corner and head-driven strategies. It allows us to construct partial structures as indicated by the closed curves drawn with dotted lines in the figure above.

So let's tentatively accept the bound on constituency connectedness (82) which began this section, and consider what kind of parsing strategy this dictates. A certain natural conception of the human parser now emerges. In the first place, the preceding discussion sets the stage for the claim that <u>all</u> linguistic dependencies are bounded with respect to a natural typology, as the claims (1) and (2) suggest. So the natural idea is that the human parser is **frugal** in the sense that it has great difficulty pursuing parses which involve more than two or three dependencies of any given type. We have seen that standard bottom-up, top-down, left-corner and head-driven strategies are not frugal even on acceptable sentences, but it is easy to imagine one that is. We could get appropriate behavior by keeping track of what sorts of information gets put into the stack, bounding it appropriately. But of course we need more than just a stack if we are going to handle crossing

[39] Notice that the search problem is not be removed by adding an oracle which allows us to restrict structure-building operations with top-down information about the identity of the leftmost constituent we are looking for. (The use of such an oracle is discussed in, for example, Pereira and Shieber, 1987, §6 where it is called "linking.") Nor is the search problem removed by adding k-element lookahead to either the left-corner or the head-driven algorithms, since no particular number of terminal elements k would allow us to see far enough ahead into the string to know how many clauses with empty subjects to initiate. We could, perhaps, get enough information if we are allowed to look ahead over constituents that have already been built, as in the LR(k,t), LRRL(k) and similar strategies (Szymanski and Williams, 1976; Nozohoor-Farshi, 1986; Marcus, 1980; Berwick and Weinberg, 1984). But what these "non-canonical" lookahead strategies are doing is changing the order in which structures are built, as we recommend here.

dependency relations, relations of the sort found in the Dutch verb clusters discussed above, for example. One idea is that the constituency relations are recorded in a stack, while other types of dependencies, such as those induced by head-movements, are available for satisfaction at any point in the parser. But if we are going to separate our representations of these two types of information, why not separate all the types of information. Then we do not need any stack or other unbounded memory for frugal parsing, since we only need to remember two or three dependencies of any given type, where there are only finitely many types.[40] So one possible implementation could proceed along the following lines. Imagine that for every type of dependency (where the relevant typology is the one mentioned in (1) and (2)), there are just two memory cells. Then, when a third dependency of that kind is needed, a certain "rescue" operation is possible, where this rescue operation itself has finitely bounded resources at its disposal, and so it allows us, with more or less difficulty, to handle a third or sometimes even a fourth dependency of any given type, depending on the type of the dependencies, whether they are crossing or nested, etc. A rigorous development of such a parsing strategy goes beyond the scope of this paper (see Stabler, 1992a), but something along these lines now appears plausible.

9. Some consequences and conclusions

It is interesting to note that, if it were not for the fact that claims (1) and (2) involve "acceptability" rather than "grammaticality," they would conflict with the widely discussed claims of Barton, Berwick and Ristad (1987) and Ristad (1990) to the effect that grammatical structures exhibit the sort of "nonlocal information flow" characteristic of intractable problems in the complexity class called "NP." For example, the satisfiability of a 3-CNF formula of the propositional calculus is generally assumed to be intractable, and we accordingly find that there is no finite upper bound on the amount of information we need about a subformula to decide whether the whole formula is satisfiable. What is not usually emphasized is that this kind of information flow is not found in any acceptable sentences in any human language. The intractability arguments obviously require assuming that language is infinite, and so the sentences involved in these arguments go beyond reasonable acceptability bounds in the relatively uninteresting respect that some of the sentences are too long for any human to understand. What is more interesting is that the constructions used in these arguments also involve unbounded violations of our connectivity bounds. In short,

[40] Actually, if we assume that our structures contain indices of various kinds, as is usual in recent transformational theory, a certain kind of unboundedness could creep in here, but the actual identity of any index never determines structure-building operations. See Stabler (1992a) for more discussion.

although the problems defined by the grammarian may require unbounded nonlocal information flow of the sort we find in NP hard problems, our bound immediately implies that the language people use does not involve any such problems, even when we abstract away from bounds on the length of acceptable sentences.[41]

Perhaps the traditional delimitation of "grammatical structures" should be reexamined. It is not clear why (1) should not be regarded as a grammatical constraint. Shieber (1985) warns against using restrictions like (1) on the set of grammatical sentences, saying "Down this path lies tyranny. Acceptance of this argument opens the way to proofs of natural languages as regular, nay, finite." But this is a mistake. We have not had to sacrifice our grip on the nature of linguistic structure at all, let alone give up on it completely as we would if we treated the acceptable structures as just a finite, listed set. In fact, the bound (1) is stated in explicitly grammatical terms, and it is not clear why it should not be regarded as a grammatical constraint. Chomsky (1956) rejects the idea that an arbitrary bound on sentence length should be treated as part of the grammar. This still seems right: it is plausible that whatever bounds there are on length follow not from the human representation of language, but from other, independent cognitive limitations. The complexity bound (1) is not like the length bound, though. It is explicitly linguistic, and while it may not be exactly right, the evidence for extremely low connectivity bounds is overwhelming.

It is no surprise that structures acceptable to finite creatures like us exhibit finite connectivity, but it is surprising how low the finite bounds are! (1) is a (slightly squishy) binary connectedness hypothesis for acceptable structures! And (2) is the similar hypothesis for the partial structures constructed in parsing. If it turns out to be possible to integrate more evidence into the the preliminary case for these claims that has been made here, it will be of considerable interest to pursue in detail the idea that they are functionally explained by the "frugal" nature of the human parser. The natural, interestingly low, and linguistically motivated bounds seem, at least

---

41 The famous example *Buffalo buffalo buffalo buffalo buffalo*, is mentioned by Barton et al. (1987), but this example is not representative of a hard processing problem, and it does not involve high connectivity. It is hard for other reasons. The word *buffalo* is a transitive verb, a common noun denoting a kind of animal, and a proper name denoting a city, so in fact most grammars would treat all sentences $buffalo^k$ for $k \geq 2$ as grammatical. The parsing problem for $buffalo^k$ is tractable though – the constituent structures of these sentences are easily represented in a context free grammar and can be parsed efficiently by Earley's algorithm, for example. And the connective complexity of all of these sentences is very low, staying under a fixed finite bound. The lesson of these sentences has nothing to do with real intractability; rather, these sentences show us that people apparently have trouble understanding highly ambiguous sentences in the absence of semantic cues that could guide disambiguation. To get intractability, sentences with unboundedly higher connective complexity must be used.

at first blush, to fit with such a conception of how the human parser might work. Indeed, even before getting clear about natural bounds, the imperative of explaining nearly real time parsing of acceptable sentences with the finitely bounded human "hardware" has prompted a good deal of interest in finite state approximations to more complex languages (Pereira, 1990) and connectionist implementations of such models (Moisl, 1992; Servan-Schrieber et al., 1991). But the resource demands of frugal parsers are restricted along linguistically motivated dimensions, and the claim is that their finite resources suffice to handle the full structural complexity of acceptable strings, as characterized by grammarians who had no ulterior computational motives in mind, but only the goal of a clear understanding of linguistic structure.

Acknowledgements. This work is part of a long-term project to find reasonable ways to bound the resource demands of problems like those formulated in the axiomatic system of Stabler (1992). The idea proposed here, which in effect imposes bounds according to a typology of relations, will be described in more computational detail in Stabler (1992a). I benefitted from discussions about various aspects of this work with Steve Abney, Alex Alsina, Ted Gibson, Mark Johnson, Aravind Joshi, Ed Keenan, Stuart Shieber, Dominique Sportiche, Tim Stowell, K. Vijay-Shanker, and especially Andi Wu. Thanks also to my audiences at Rutgers, Princeton and the University of Arizona where various early versions of this material were presented. Finally, this project would have been impossible without the generous assistance of the language specialists noted in the text.

References.

Abney, S.P. and Johnson, M. (1991) Memory requirements and local ambiguities of parsing strategies. *J. of Psycholinguistic Research 20*(3): 233-249.

Alsina, A. (1992) On the argument structure of causatives. *Linguistic Inquiry 23*(4): 517-556.

Bach, E., Brown, C., and Marslen-Wilson, W. (1986) Crossed and nested dependencies in German and Dutch: A psycholinguistic study. *Language and Cognitive Processes 1*: 249-262.

Baker, M.C. (1988) *Incorporation: A Theory of Grammatical Function Changing.* Chicago: University of Chicago Press.

Baker, M.C. (1989) Object sharing and projection in serial verb constructions. *Linguistic Inquiry 20*(4): 513-553.

Barton, G.E., Jr., Berwick, R.C., and Ristad, E.S. (1987) *Computational Complexity and Natural Language.* Cambridge, Massachusetts: Bradford/MIT Press.

Berwick, R. and Weinberg, A. (1984) *The Grammatical Basis of Linguistic Performance.* Cambridge, Massachusetts: MIT Press.

Besten, H. den, and Edmondson, J. (1983) The verbal complex in continental West Germanic. In W. Abraham, ed., *On the Formal Syntax of West Germania.* Amsterdam: Benjamins.

Besten, H. den, and Rutten, J. (1989) On verb raising, extraposition, and free word order in Dutch. In D. Jaspers, W. Klooster, Y. Putseys, and P. Seuren, eds., *Sentential Complementation and the Lexicon.* Dordrecht: Foris.

Bresnan, J. and Moshi, L. (1990) Object asymmetries in comparative Bantu syntax. *Linguistic Inquiry 21*(2): 147-186.

Chao, Hsi-chia (1991) Jump in difficulty and incomplete clause hypothesis. Manuscript, University of Massachusetts, Amherst.

Chomsky, N. (1956) Three models for the description of language. *IRE Transactions on Information Theory, IT-2*: 113-124.

Chomsky, N. (1965) *Aspects of the Theory of Syntax.* Cambridge, Massachusetts: MIT Press.

Chomsky, N. (1986) *Barriers.* Cambridge, Massachusetts: MIT Press.

Cole, P. (1985) *Imbabura Quechua.* London: Croom Helm.

Comrie, B. (1976) The syntax of causative constructions: cross-language similarities and divergences. In M. Shibatani, ed., *Syntax and Semantics, Volume 6: The Grammar of Causative Constructions.* NY: Academic Press.

Comrie, B. (1981) *Language Universals and Linguistic Typology.* Oxford: Basil Blackwell.

Crocker, M.W. (1992) *A Logical Model of Competence and Performance in the Human Sentence Processor.* Ph.D. Dissertation, University of Edinburgh.

Emonds, J. (1976) *A Transformational Approach to English Syntax.* New York: Academic Press.

Emonds, J. (1985) *A Unified Treatment of Syntactic Categories.* Cinnaminson, New Jersey: Foris Publications.

Fortescue (1984) *West Greenlandic.* London: Croom Helm.

Gazdar, G., Klein, E., Pullum, G. and Sag, I. (1985) *Generalized Phrase Structure Grammar.* Cambridge, Massachusetts: Harvard University Press.

Gibson, E.A.F. (1991) *A Computational Theory of Human Linguistic Processing.* Ph.D. Dissertation, Carnegie-Mellon University.

Giorgi, A., Pianesi, F. and Satta, G. (1990) A computational approach to binding theory. *Procs. of the 13th Int. Conf. on Computational Linguistics, COLING-90*: 120-125.

Givón, T. (1976) Some constraints on Bantu causativization. In M. Shibatani, ed., *Syntax and Semantics, Volume 6: The Grammar of Causative Constructions.* NY: Academic Press.

Gould, R. (1988) *Graph Theory.* Menlo Park, California: Benjamin/Cummings.

Haegeman, L. and van Riemsdijk, H. (1986) Verb projection raising, scope, and the typology of movement rules. *Linguistic Inquiry 17*: 417-466.

Herrero, S.J. and Sánchez de Lozada, F. (1978) *Gramatica Quechua: Estructura del Quechua Boliviano Contemporaneo.* Cochabamba: Editorial Universo.

Hetzron, R. (1969) *The Verbal System of Souther Agaw.* Berkeley: University of California Press.

Hetzron, R. (1976) On the Hungarian causative verb and its syntax. In M. Shibatani, ed., *Syntax and Semantics, Volume 6: The Grammar of Causative Constructions.* NY: Academic Press.

Hoeksema, J. (1988) A constraint on governors in the West Germanic Verb Cluster. In M. Everaert, A. Evers, R. Huybregts and M. Trommelen (eds) *Morphology and Modularity: In Honour of Henk Schultink.* Dordrecht: Foris.

Hook, P.E. (1974) *The Compound Verb in Hindi.* Ann Arbor: Michigan Series in South and Southeast Asian Languages and Linguistics.

Jaeggli, O. (1981) *Topics in Romance Syntax.* Dordrecht: Foris.

Jansen, B., Koopman, H., and Muysken, P. (1978) Serial verbs in the creole languages. *Amsterdam Creole Studies II*: 125-159.

Joshi, A.K. (1989) Processing crossed and nested dependencies: an automaton perspective on the psycholinguistic results. *Language and Cognitive Processes 5*: 1-27.

Kayne, R.S. (1975) *French Syntax: The Transformational Cycle.* Cambridge, Massachusetts: MIT Press.

Kayne, R.S. (1983) *Connectedness and Binary Branching.* Dordrecht: Foris.

Kayne, R.S. (1989) Null subjects and clitic climbing. In O. Jaeggli and K. Safir, eds., *The Null Subject Parameter.* Boston: Kluwer.

Koopman, H. (1984) *The Syntax of Verbs.* Dordrecht: Foris.

Koopman, H. and Sportiche, D. (1992). Subjects. *Lingua ???.*

Koster, J. (1978) Why subject sentences don't exist. In S.J. Keyser, ed., *Recent Transformational Studies in European Languages.* Cambridge, Massachusetts: MIT Press.

Kroch, A.S. and Joshi, A.K. (1985) *The Linguistic Relevance of Tree Adjoining Grammar.* Technical Report, University of Pennsylvania Dept. of Computer and Information Science.

Kroch, A.S. and Santorini, B. (1991) The derived constituent structure of the West Germanic verb raising construction. In R. Friedin (ed.), *Principles and Parameters in Comparative Grammar,* 269-338.

Larson, R.K. (1988) On the double object construction. *Linguistic Inquiry 19*: 335-391.

Larson, R.K. (1990) Double objects revisited: reply to Jackendoff. *Linguistic Inquiry 21*: 589-632.

Larson, R.K. (1991) *Promise* and the theory of control. *Linguistic Inquiry 22*: 103-139.

Li, Y. (1990) On V-V compounds in Chinese. *Natural Language and Linguistic Theory 8*(2): 177-208.

Mahajan, A.K. (1982) *Causative Constructions in the Indian Subcontinent.* Dissertation, Department of Linguistics, University of Delhi.

Mahajan, A.K. (1982) *The A/A-bar Distinction and Movement Theory.* MIT Ph.D. Dissertation.

Marantz, A.P. (1984) *On the Nature of Grammatical Relations.* Cambridge, Massachusetts: MIT Press.

Marantz, A.P. (1989) Clitics and phrase structure. In M.R. Baltin and A.S. Kroch, eds., *Alternative Conceptions of Phrase Structure.* Chicago: University of Chicago Press.

Marcus, M. (1980) *A Theory of Syntactic Recognition for Natural Language.* Cambridge, Massachusetts: MIT Press.

Mohanan, K.P. (1982) Grammatical relations and clause structure in Malayalam. In J. Bresnan, ed., *The Mental Representation of Grammatical Relations.* Cambridge, Massachusetts: MIT Press.

Miller, G.A. and Chomsky, N. (1963) Finitary models of language users. In R.D. Luce, R. Bush and E. Galanter (eds.), *Handbook of Mathematical Psychology, Volume II.* New York: Wiley.

Moisl, H. (1992) Connectionist finite state natural language processing. *Connection Science 4*(2): 67-91.

Muysken, P. (1977) *Syntactic Developments in the Verb Phrase of Ecuadorian Quechua.* Dordrecht: Foris.

Muysken, P. (1988) Affix order and interpretation: Quechua. In M. Everaert, A. Evers, R. Huybregts and M. Trommelen, eds., *Morphology and Modularity.* Dordrecht: Foris.

Nozohoor-Farshi, R. (1986) *LRRL(k) Grammars: A Left to Right Parsing Technique with Reduced Lookaheads.* University of Alberta Ph.D. thesis.

Pereira, F.C.N. (1990) Finite-state approximations of grammars. *Procs. Speech and Natural Language Workshop:* 20-25.

Pereira, F.C.N. and Shieber, S.M. (1987) *Prolog and Natural Language Analysis.* Chicago: Chicago University Press.

Pinker, S. (1984) *Language Learnability and Language Development.* Cambridge, Massachusetts: Harvard University Press.

Pollock, J.-Y. (1989) Verb movement, universal grammar, and the structure of IP. *Linguistic Inquiry 20*: 365-424.

Pritchett, B.L. (1992) *Grammatical Competence and Parsing Performance.* Chicago: University of Chicago Press.

Ristad, E.S. (1990) *A Computational Complexity Thesis for Human Language.* Technical Report, Princeton University Dept. of Computer Science. (Revised version of his MIT Ph.D. thesis.)

Rizzi, L. (1978) A restructuring rule in Italian syntax. In S.J. Keyser, ed., *Recent Transformational Studies in European Languages.* Cambridge, Massachusetts: MIT Press.

Rizzi, L. (1986) On the status of subject clitics in Romance. In O. Jaeggli and C. Silva Corvalan, eds., *Studies in Romance Linguistics.* Dordrecht: Foris.

Safir, K.J. (1985) *Syntactic Chains.* NY: Cambridge University Press.

Schiffman, H. (1976) *A Reference Grammar of Spoken Kannada.* Seattle: University of Washington, Dept. of Asian Languages.

Servan-Schreiber, D., Cleeremans, A. and McClelland, J.L. (1991) Graded state machines. *Machine Learning 7*: 161-193.

Shieber, S.M. (1985) Evidence against the context-freeness of natural language. *Linguistics and Philosophy 8*: 333-343. Reprinted in W.J. Savitch, ed., *The Formal Complexity of Natural Language.* Boston: Reidel.

Shieber, S.M. (1992) *Constraint-based Grammar Formalisms.* Cambridge, Massachusetts: Bradford/MIT Press.

Sportiche, D. (1988) A theory of floating quantifiers and its corollaries for constituent structure. *Linguistic Inquiry 19*(3): 425-449.

Sportiche, D. (1990) *Movement, Agreement and Case.* UCLA manuscript, forthcoming.

Sportiche, D. (1992) Clitic constructions. UCLA ms, forthcoming.

Stabler, E.P., Jr. (1991) Avoid the pedestrian's paradox. In R.C. Berwick, S. Abney, C. Tenny, eds., *Principle-based Parsing: Computation and Psycholinguistics.* Boston: Kluwer.

Stabler, E.P., Jr. (1992) *The Logical Approach to Syntax.* Cambridge, Massachusetts: MIT Press.

Stabler, E.P., Jr. (1992a) Frugal computation of syntactic relations. In preparation.

Stowell, T. (1981) *Origins of Phrase Structure.* Ph.D. Dissertation, Department of Linguistics and Philosophy, Massachusetts Institute of Technology.

Syeed, S.M. (1985) Morphological causatives and the problems of the transformational approach. Bloomington, Indiana: Indiana University Linguistics Club.

Szymanski, T.G. and Williams, J.H. (1976) Noncanonical extensions of bottom-up parsing techniques. *SIAM Journal of Computing 5*: 231-250.

Vijay-Shanker, K. (1987) *A Study of Tree Adjoining Grammars.* Ph.D. Thesis, University of Pennsylvania.

Ward, G., Sproat, R. and McKoon, G. (1991) A pragmatic analysis of so-called anaphoric islands. *Language 67*: 439-473.

Weber, D. (1989) *A Grammar of (Huánuco) Quechua.* UC Publications in Linguistics. Berkeley: University of California Press.

Wu, A. (1988) Jump in difficulty and the two-tree hypothesis in parsing. UCLA M.A. thesis, Dept. of Linguistics.

Wu, A. (1992) A minimalist universal parser. *Working Papers in Syntax and Semantics* (this volume), UCLA Dept. of Linguistics.

Zimmer, K.E. (1976) Some constraints on Turkish causativization. In M. Shibatani, ed., *Syntax and Semantics, Volume 6: The Grammar of Causative Constructions.* NY: Academic Press.

A MINIMALIST UNIVERSAL PARSER

Andi Wu

wu@cognet.ucla.edu

In this paper, I present a principle-based parser[1] which is inspired by Chomsky's (1992) Minimalist Program and Kayne's (1992) idea that cross-linguistic variations in word order result from movement rather than head positions.[2] Traditionally, word order has been regarded as a property of phrase structure. Consequently, parsers for different languages must use different context-free rules. Since Chomsky (1992) and Kayne (1992) appeared, however, a new picture has become conceivable: all languages are identical in terms of phrase structure and the movements that operate on it. Differences in word order depend on whether the movements occur before or after *Spell-Out*, the point where the syntactic representation is fed into *PF* (Phonetic Form). With regard to parsing, this means that, no matter what the language being parsed is, we always build the same X-bar tree and the same movement chains, resulting in the same *LF* (Logical Form) representation. The only decision we have to make in parsing a particular language is where each movable element is to appear in the relevant chain at Spell-Out time. I will show that such decisions can be made on the basis of a small number of parameter values. Implementing these ideas, we get a parser which has the following two properties: (a) it simultaneously constructs the D-structure, S-structure and LF representations, and (b) its algorithm applies universally to languages of any word orders and no parametrization is needed in the parser itself. If the present approach is correct, there will be the possibility of a universal parser and a model of language acquisition where no learning is needed in terms of parsing.

The paper is organized as follows: Section 1 briefly summarizes the problems encountered in the traditional approach where different word orders result from different phrase structure rules. Section 2 spells out the new approach and illustrates how different orders can be derived from movement. Section 3 proposes a set of parameters whose values determine the timing (before or after Spell-Out) of each movement. It also demonstrates how the parameter space thus created accommodates a word order

---

[1]See Johnson 1989, Berwick 1991, Fong 1991, Stabler 1991, etc. for discussions of principle-based parsing.

[2]This idea has been presented in his talks at UCLA (April 1992) and his speech at GLOW 1992, though no published versions of those talks are available.

typology. Section 4 introduces the parser which implements the new ideas. Section 5 points out the issues that need to be addressed in the future.

1. Problems with Varying Phrase Structures

In the traditional view, word order is at least in part a property of phrase structure. Different languages can have different word orders when the phrase structure rules they use are different. In order to parse a language, the parser must be equipped with the phrase structure rules specifically designed for this language. This poses both a learning problem and a parsing problem.

In terms of acquisition, this approach predicts that a child has to acquire or select a different set of phrase structure rules for each different language. The question is whether these rules are learnable. As Fodor (1992) shows, there is even no selection criterion for phrase structure rules that satisfies the Subset Principle. A pure phrase structure grammar is not learnable without negative evidence. To make the selection procedure more feasible, linguists have been trying to constrain the phrase structure rules in natural languages so that the selection can be made from a very small set of "possible rules". In Government and Binding (GB) theory, this attempt has resulted in the current version of X-bar theory (Jackendoff 1977, Stowell 1981, Chomsky 1986, Sportiche 1991, etc.). The X-bar theoretic constraints on phrase structure restricts phrasal projections to a uniform structure where every category has a head, a specifier (possibly empty) and a complement (possibly empty).[3] Cross-linguistic differences in phrase structure now translate into the values of two parameters: the Specifier-Head Parameter and the Complement-Head parameter (cf. Travis 1984, Nyberb 1987, Gibson and Wexler 1992, Wu 1992a). However, this seemingly small parameter space still cannot guarantee the learnability of all languages. As pointed out in Gibson and Wexler (1992), the parameter-setting process may get into a state, called a *local maximum*, where the setting is incorrect (i.e., some input data are not interpretable with the current grammar) but there exist no data from the input that can trigger the correct setting. So even the most restrictive theory of phrase structure currently available fails to solve the learning problem.

In terms of parsing, the variation in phrase structure makes it difficult to design a psychologically viable parsing algorithm which is completely universal. [4] Generally

[3]The number of complements can be more than one, as in a ditransitive sentence. But the Larsonian treatment of VP structure has made it possible to split the VP into several layers with each argument corresponding to a single layer and each layer other than the top one being a complement of the higher VP. In this way, each VP only has one complement.

[4]Some parsing algorithms, such as the Earley parser and chart parsers in general, can handle

speaking, top-down strategies seem optimal for languages that are primarily right-branching while bottom-up strategies are favored when a language is left-branching. In addition, the existence of empty categories requires top-down parsing and the existence of left-recursive rules requires bottom-up parsing. Since top-down and bottom-up strategies are complementary to each other, most parsing models combine the two. But the actual combination has to be tailored to suit different languages. To maintain a uniform parser while accommodating individual differences, many people have proposed to parameterize the parser. For instance, Frazier and Rayner(1988) suggests that whether the matrix S node should be projected at the initial step is parameterized across languages. The value of this parameter is "yes" for a right-branching language like English and "no" for a left-branching language like Japanese. A different kind of parameterization is found in Weinberg (1992), where she argues that, in a head-final language like Japanese, the projection of IP is licensed by the case morphology of the subject, but it is not necessarily the case in a head-initial language like English. These proposals have minimalized the amount of parsing knowledge that has to be learned, yet the basic assumption remains that, in addition to acquiring the grammar, children have to acquire some parsing strategies. Although the learning is minimal, it is nonetheless simpler to have a theory where all the learning takes place in the grammar and everything about parsing is innate. The model to be presented in this paper is potentially capable of achieving this higher goal.

2. An Alternative View of Word Order

All the problems discussed in the previous section come from our assumption that different languages have different phrase structures. None of these problem would exist if a single, invariant X-bar structure is used in all languages. This is the possibility that Kayne (1992) has suggested.[5] He proposed a fixed X-bar structure where the specifier always precedes the head and the complement follows the head. All languages have the same word order at D-structure and all other orders are derived from movement.

The notion that word order variation can be the consequence of movement is a familiar one. However, so far this idea has been used mainly to account for word order variations *within* a language. Since any given language is supposed to have a fixed phrase structure, the only way to get different word orders is through movement.

any phrase structure rules. However, these algorithms are usually not meant to be psychologically motivated. They are computationally feasible but are not designed as models of human parsing.

[5] Wu (1992b) has tried to follow up on this idea.

This is how we explain, for instance, why statements and questions have different word orders in English. So movement can be a source of word order variation at least within a single language. Across languages, we also find it necessary sometimes to derive a certain order from movement. Take the VSO order as an example. If we assume that an object is always the complement of a verb and therefore adjacent to it, this common word order would be impossible no matter how the head parameters are set. The only explanation for this order comes from movement: the verb and the object *are* adjacent to each other at D-structure(DS), but the verb has moved to a higher position at S-structure (SS) (cf. Chomsky and Lasnik 1991, Chomsky 1992). The question now is whether *all* cross-linguistic differences in word order can be accounted for in this way. In other words, is movement not only a necessary but also a sufficient source of word order variation across languages?

The answer to this question was "No" before the appearance of Chomsky's (1992) Minimalist Program. As we know, syntactic movements are all motivated by UG (universal grammar) constraints. Furthermore, the Principle of Economy (Chomsky 1991) requires that no movement occur unless motivated by a UG constraint. Therefore, movement takes place *if and only if* it is the only means to satisfy a UG constraint. Let us call a UG constraint that forces movement a "Movement-Forcing UG Constraint" (MFUGC for short). Before Chomsky proposed the Minimalist Program, some MFUGC were S-structure requirements and some were LF requirements. Therefore, some movements had to take place in the derivation between D-structure and S-structure and some between S-structure and LF. Within this framework, the movement story is deficient. Suppose all languages have the same word order at D-structure, which is what Kayne suggests. To get different word orders in different languages, which are properties of S-structure, we have to assume one of the following:

(i) Different movements take place in different languages.

(ii) The same set of movements applies in all languages, but the movements that occur between D-structure and S-structure (i.e. those that are visible) vary from language to language.[6]

Assumption (i) is unappealing for theoretical reasons. Since UG constraints are universal, the movements motivated by MFUGCs must be the same. To say that different languages have different movements is equivalent to saying that different grammatical principles operate in different languages. Assumption (ii) seems more plausible. In

[6]This is a familiar idea which has been assumed in Huang (1982), Stowell (1983), May (1985), etc. as well as Chomsky (1991) and Chomsky (1992).

fact, this is the standard explanation for the question of why wh-phrases go clause-initial in English but remain in situ in Chinese (Huang 1982). However, this is possible only because the MFUGC which forces wh-movement (namely the scope requirement) is an LF requirement whose satisfaction at S-structure is optional. If most MFUGCs must apply at S-structure, as is the case in the traditional model, this option will be unavailable in most cases.

In the recently proposed Minimalist Program framework (Chomsky 1992), however, the scenario in assumption (ii) has become a realistic one. This new model eliminates D-structure and S-structure, reducing the levels of representation to PF and LF only. As a result, LF is now the only level at which the MFUGCs must be satisfied. In this new model, syntactic derivation starts from Lexical Projection, going through GT operations (Generalized Transformations which map individual projection trees to a single tree) and Move-α (which maps one tree to another), finally ending in LF when all movements motivated by MFUGCs have taken place. Lexical Projection and GT operations must occur before *Spell-Out* (SO), the point where the syntactic representation is fed to the PF component. This is probably due to a PF constraint which requires that the input to PF be single trees[7]. Once the projections have been mapped to a single tree which satisfies the X-bar constraints, Spell-Out may occur any time during the derivation, regardless of how many MFUGC-related movements have taken place. The surface form of a sentence will therefore vary depending what has moved before SO. This model is still being developed and many questions remain unanswered. It is still debatable, for example, whether the elimination of D-structure and S-structure is justified or whether they are eliminated at all. But this makes no difference to the word order issue I am pursuing here. Let us keep the terms D-structure and S-structure and let the former refer to the representation resulting from lexical projection and GT operations and the latter to the representation that is "spelled out". The important thing is that no UG constraint is associated with S-structure any more. All MFUGCs, including the Case Filter which forces NP-movement and Lasnik's Filter (Lasnik 1981) which forces verb movement, are now LF requirements. They must be satisfied by LF, but that does not prevent the possibility that they may be satisfied earlier the derivation, such as at SS or SO. (I will use the terms S-structure(SS) and Spell-Out(SO) more or less interchangeably.) As far as UG is concerned, whether a given MFUGC is satisfied at SS does not matter, as long as it is satisfied at LF. However, individual languages may differ as to what MFUGCs must be satisfied before Spell-Out. Consequently, their word

[7]One possible reason why this should be so is that each intonational pattern has to be mapped onto a single tree.

orders may vary according to how many pre-SO movements have occurred. In sum, we now have the following picture: all languages have the same D-structure and the same LF. Furthermore, they go through the same movements to satisfy the same set of UG requirements. However, whether a given movement takes place before or after Spell-Out is a choice that has to be made for each individual language. In the next section, I will propose a way to parameterize the choices and use a concrete model to illustrate the relationship between parameter values and word order variations.

3. S-Parameters and Word Orders

3.1. The S-parameter

To formalize the option of whether a movement is to occur before or after Spell-Out, I propose that an *S-Parameter* be associated with each MFUGC.[8] This parameter has two values: 1 if an MFUGC must be satisfied at S-structure and 0 if it must not be satisfied at S-structure. When the S-parameter of an MFUGC is set to 1, the movement forced by this requirement must occur before Spell-Out and thus be visible; otherwise the movement has to occur after Spell-Out and the constituent to be moved will stay in situ at S-structure. I assume that the S-parameter is a set of parameters rather than a single parameter and the number of S-parameters in this set depends on the number of MFUGCs we assume. Different languages can choose to assign 1 or 0 to different subsets of S-parameters. As a result, the set of movements that occur before Spell-Out can vary from language to language and the word order will vary accordingly.

3.2. Syntactic assumptions

The actual correspondence between S-parameter values and word order depends on the syntactic model we assume. It is determined by the following three things:

- The D-structure. This is the "pre-movement" structure created by lexical projection and GT operation. In our theory, this structure is identical across languages.
- The LF-structure. This is the "post-movement" structure where all MFUGCs have been satisfied. This structure is again identical across languages.
- The movements. These are the operations that map D-structures into LF structures. These are again universal, except that the timing of those movements

[8] In Wu (1992b) this parameter is called the *procrastination parameter.*

can vary from language to language.

These three things are not explicitly stated in UG. They are derived from the following components of UG:

- The rules of Lexical Projection which determine the phrasal projections of different categories.
- The rules of GT operation which determine how the phrasal trees produced by Lexical Projection are put together into a single tree.
- The LF constraints (i.e. the MFUGCs) which determine the LF structures which in turn determine the movements that are required.

I will use a particular version of those rules and constraints to illustrate how the parameter space I have just proposed accommodates word order variations. The version to be used is not meant to be *the* correct version. It is chosen simply for the purpose of illustration. Furthermore, it will be a simplified version. I only want to illustrate the derivation of *basic word orders* here, with my attention limited to simple declarative sentences with a subject, an object and a verb. Therefore, only those parts of UG which are relevant to this simple structure will be touched upon. The approach illustrated here is potentially applicable to more complicated word order phenomena, but a full discussion of these issues is beyond the scope of this paper.

3.2.1. Lexical projection

Given any category X, the tree projected from the head of this category is always the one in Figure 1.

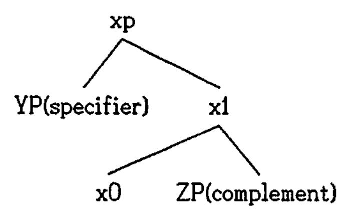


Figure 1

I will call this tree an *elementary X-bar tree.* This structure is not subject to parametric variation: the specifier (Spec) is always on the left and the complement (Comp) on the right. The categories of YP and ZP are selected by X. A YP can be in the Spec position of an XP only if there is a rule such as

specifier(X,Y)

which says "YP is a possible specifier of XP". Similarly, a ZP can be in the complement position of an XP only if we have the following rule:

complement(X,Z)

which says "ZP is a possible complement of XP.

The set of rules to be used in the illustration is given below:

(1)

specifier(c,X) (i)
specifier(agr-s,n) (ii)
specifier(t,ϵ) (iii)
specifier(agr-o,n) (iv)
specifier(v,n) (v)

complement(c,agr-s) (vi)
complement(agr-s,t) (vii)
complement(t,agr-o) (viii)
complement(agr-o,o) (ix)
complement(v,v) (x)
complement(v,ϵ) (xi)

"X" stands for an unspecified category and ϵ means there is no specifier and complement.

At the time when an elementary X-bar tree is projected, the Spec and Comp positions are empty. They are just place-holders.

3.2.2. GT operations

A generalized transformation (GT) takes two trees and substitutes one of the trees for an empty Spec or Comp position in the other tree. The substitution is possible only if the XP being substituted has the same category as the specifier or complement. Figures 2 and 3 illustrate the two cases of GT operation. In Figure 2, YP is substituted into the complement position of XP. In Figure 3, XP is substituted into the specifier position of YP.

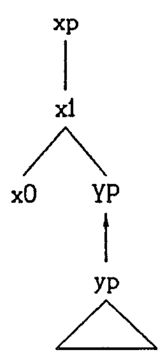


Figure 2

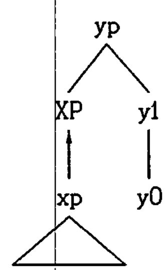


Figure 3

The GT operation is applied recursively and, if a sentence is grammatical, the operation should eventually result in a single tree.

For a simple transitive sentence, Lexical Projection and GT operations result in the D-structure in Figure 4.

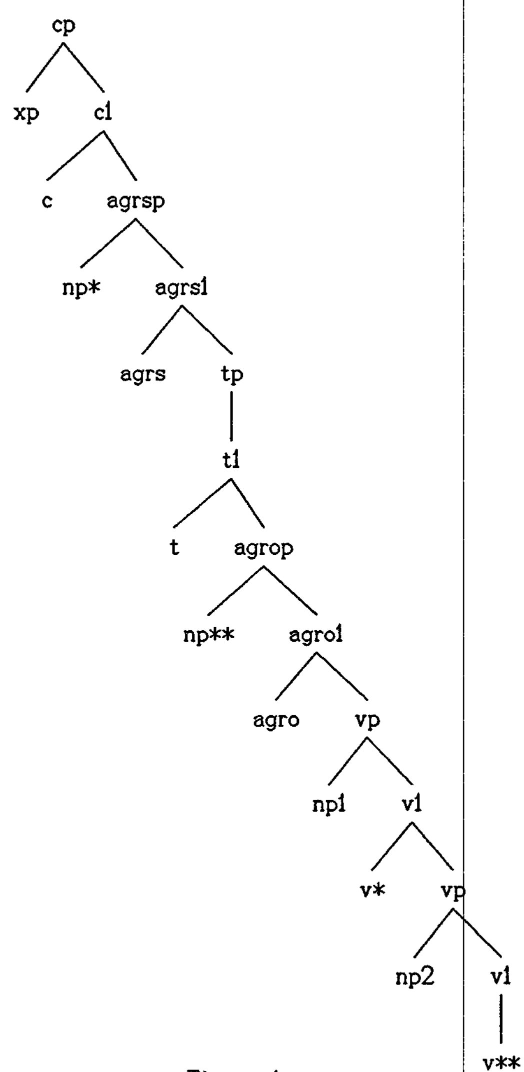


Figure 4

This is basically the structure assumed in Chomsky (1992, p10). The VP internal structure is a pseudo Larsonian one (cf. Larson 1990) where the number of VP layers corresponds to the number of theta roles the verb assigns. I follow Sportiche (1990) in assuming that, in every layer, the argument of the verb is in the Spec position, so that theta-role assignment uniformly takes place in a Spec-Head configuration. Furthermore, no NP is assigned case VP-internally, so all NPs have to move for case reasons. To assign a theta-role to each of the arguments, the verb has to move through all layers of VP (moving from V** to V* in the present case). I assume that this movement always takes place before Spell-Out as part of the VP projection.[9] At D-structure, therefore, the subject is in NP1, the object in NP2, and the verb in V*.

3.2.3. Movement-forcing UG constraints

In the traditional version of Principles and Parameters theory, movements are often forced for morphological reasons: nouns move to get their case morphology and verbs move to get their tense or agreement morphology. In the Minimalist Program, however, this morphological motivation no longer exists. In this model, nouns and verbs are "drawn from the lexicon with all of their morphological features, including case and ϕ-features" (Chomsky 1992, p41). The purpose of movement has changed from feature-assignment to *feature-checking.* Viewed from this new perspective, all MFUGCs are "checking" requirements. They force movements because the checking requires some particular configuration relations. I assume that a Spec-Head configuration is required for all of the following: case-assignment, theta-role assignment, as well as the checking of agreement features. This requirement forces XP movements. The checking of verbal features (*eg.* tense) requires a sisterhood relation. To have a tense feature checked, for instance, the verb has to adjoin to T. This requirement forces head movements or X0 movements in some cases.

As a first approximation, I propose the following set of MFUGCs.

A. CHECK-AGREEMENT(O): the verb must have its ϕ-features checked for object-verb agreement. This requirement is satisfied iff the verb is currently adjoined to AgrO or has been adjoined to AgrO during a movement.

B. CHECK-TENSE: the verb must have its tense morphology checked. This requirement is satisfied iff the verb is currently adjoined to T or has been adjoined to T during a movement.

C. CHECK-AGREEMENT(S): the verb must have its ϕ-features checked for subject-verb agreement. This requirement is satisfied iff the verb is currently adjoined to AgrS or has been adjoined to AgrS during a movement.

D. CHECK-C: the verb must have some features checked in C.[10] This requirement is

[9] This has to be done before Spell-Out because it determines how many layers of VP should be projected. In other words, the theta-assigning movement is relevant to Lexical Projection and GT operations while the other movements do not.

[10] The motivation for this movement is less clear than the previous ones. Various explanations have been proposed. For example, Stowell (1983) proposes that tense is located in C in some languages and the verb moves there to get T-marked. Koopman (1992) suggests that the verb has to move to C

satisfied iff the verb is currently adjoined to C or has been adjoined to C during a movement.

E. CHECK-CASE(S): the subject NP must have its case morphology checked. This requirement is satisfied iff the subject NP has moved to or through the Spec of AgrS-P.

F. CHECK-CASE(O): the object NP must have its case morphology checked. This requirement is satisfied iff the object NP has moved to or through the Spec of AgrO-P.

G. CHECK-TOPIC: the NP which is the topic of a sentence must be in agreement with C. This requirement is satisfied iff the topic NP has moved to the Spec of CP.[11]

From now on, I will use the letters A, B, C, D, E, F, G to refer to the respective S-parameters of the 7 MFUGCs introduced above. In addition, I will use the letters a, b, c, d, e, f, g to refer to the respective movements forced by those MFUGCs.

The movements forced by these MFUGCs are illustrated in Figure 5.

to "close a predicate". But one reason might be the following: C contains features which determine, among other things, whether a sentence is a statement or a question. These features are related to the intonation assigned to a sentence. We may suppose that the verb has to move to C to have its prosodic features checked.

[11]I assume that every sentence has a topic and the topic feature is checked in C. The movement required here is found overtly in verb-second (V2) languages and some other languages where the Spec of CP must be filled.

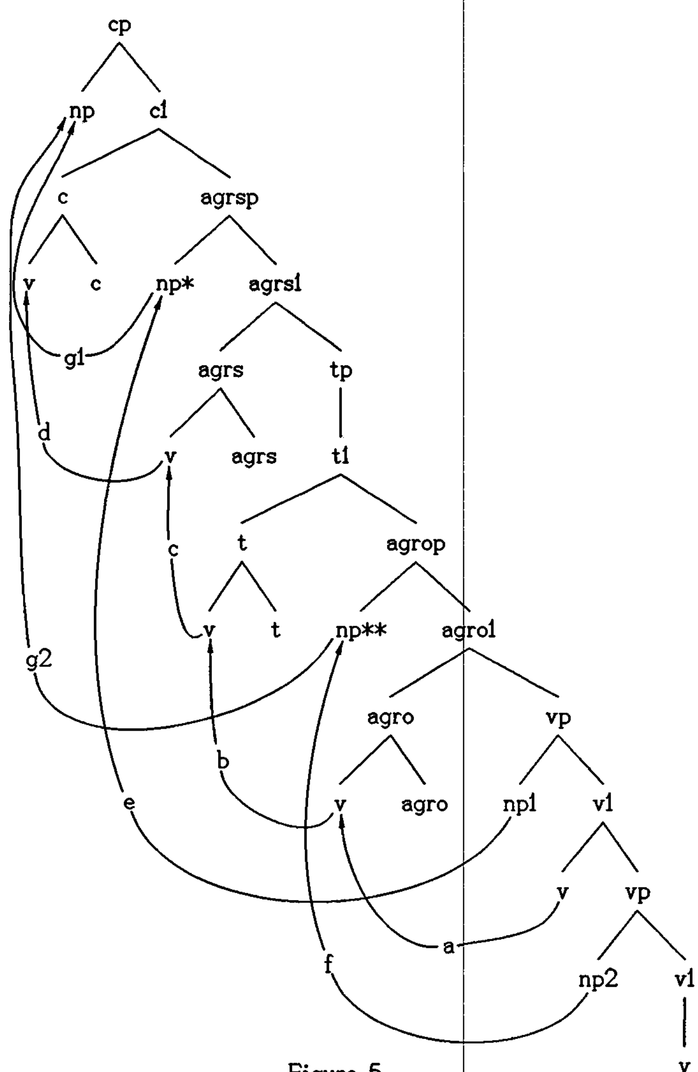


Figure 5

The movements found here are of three types: head movement (a, b, c and d), A-movement(e and f) and A-bar movement ($g1$ and $g2$). For the sake of simplicity, the intermediate positions involved in these movements are not shown. We notice that CHECK-TOPIC can force two different movements: g1 and g2. The former moves the subject NP into the Spec of CP and the latter the object NP. However, these two movements are exclusive of each other, i.e. only one of the two can occur in a particular sentence. We also notice that these two movements starts from the positions where the cases of those NPs are checked. This means an NP must have its case checked before moving to its topic position.

The head movement observed in Figure 5 is different from the standard one. I propose this one because it is more consistent with the Minimalist Program. Now that the verb comes from the lexicon together with its morphology, the function verb movement performs has changed from "picking up" the morphology to merely checking the morphology. So the verb no longer has to be amalgamated with the heads it moves to. It adjoins to another head simply to satisfy the sisterhood requirement for feature checking. After the adjunction and the feature checking, the verb is still an independent unit and can move further up by itself if necessary. The resulting head movement chain is just a V-chain and it does not have to be coindexed with the heads it passes by. Of course, this kind of head movement only applies to the "checking" movements where a lexical head moves to non-lexical head. A different kind of head movement (probably the standard one) is needed to account for V-incorporation, for instance, where a verb moves to another with the result that the two verbs are merged into one. This involves the movement of a lexical head to another lexical head and I consider it a different type of head movement.

3.3. The parameter space

Now we consider the parameter space of the S-parameters. Associating a binary-valued S-parameter to each of the 7 MFUGCs, we have a parameter space of 2^7 settings. However, due to the dependencies between parameter values, not all those settings are syntactically possible. The dependencies are derived from the following three syntactic constraints assumed in this experimental model.

- The Head Movement Constraint. This constraint requires that no intermediate head be skipped during head movement. For a verb to move from its VP-internal position and get adjoined to C, it must get adjoined successively to AgrO, T and AgrS. We have assumed that the checking requirement for a verbal feature is satisfied iff the verb is adjoined or *has been adjoined* to the position where the feature is checked. This means that, if the verb has moved to C and satisfied the CHECK-C requirement, CHECK-AGREEMENT-O, CHECK-T and CHECK-AGREEMENT(S) must also have been satisfied. If CHECK-C is satisfied before Spell-Out, the other three must also be satisfied before Spell-Out. Consequently, setting D to 1 requires that A, B and C also be set to 1. So there is a transitive implicational relationship between the values of A, B, C and D: given the order here, if one of them is set to 1, then ones that precede it must also be set to 1.

- The constraint that the Spec of CP cannot be filled unless a verb has moved to C.[12] Therefore, if CHECK-TOPIC is satisfied before Spell-Out, CHECK-C must also be satisfied before Spell-Out. So setting G to 1 implies setting D to 1 as well. Since D cannot be set to 1 unless A, B and C are also set to 1, the assignment of 1 to G implies the assignment of 1 to A, B, C and D.

- The constraint that an NP must have its case checked before moving to Spec of CP. This means that G cannot be set to 1 unless E and F are set to 1.

It turns out that, with all the value dependencies assumed here, there are only 21 settings that are syntactically possible. These settings and their corresponding word orders are given in (2).

(2)

#	*Values of S-parameters*							*Word Order*
	A	B	C	D	E	F	G	
1	0	0	0	0	0	0	0	SVO
2	0	0	0	0	1	0	0	SVO
3	1	0	0	0	1	0	0	SVO
4	1	1	0	0	1	0	0	SVO
5	1	1	0	0	1	1	0	SVO
6	1	1	1	0	1	0	0	SVO
7	1	1	1	0	1	1	0	SVO
8	0	0	0	0	1	1	0	SOV
9	1	0	0	0	1	1	0	SOV
10	1	0	0	0	0	0	0	VSO
11	1	1	0	0	0	0	0	VSO
12	1	1	1	0	0	0	0	VSO
13	1	1	1	1	0	0	0	VSO
14	1	1	1	1	1	0	0	VSO
15	1	1	1	1	1	1	0	VSO
16	1	1	1	1	1	1	1	V2
17	1	1	0	0	0	1	0	VOS
18	1	1	1	0	0	1	0	VOS
19	1	1	1	1	0	1	0	VOS
20	0	0	0	0	0	1	0	OSV
21	1	0	0	0	0	1	0	OVS

As we can see, all the possible basic word orders are accommodated in the parameter space. The number of settings corresponding to each order is different, but that

[12]This phenomenon has been observed in V2 languages. It is also found in English where the verb has to move to C whenever the Spec of CP is filled.

does not necessarily mean that the word orders compatible with a greater number of settings are more numerously found in the world's languages. Some of those settings may not actually occur.

I will use the settings in #4, #8, #14 and #16 to illustrate the relations between word order and S-parameter values.

#4: This setting requires that CHECK-AGRO, CHECK-T, CHECK-CASE(S) and CHECK-CASE(O) be satisfied before Spell-Out. This causes the movements *a*, *b* and *e* to take place before Spell-Out and thus be visible. At S-structure, therefore, the subject is in Spec of AgrS-P, the verb in T, and the object in situ. This setting may well be the one for English, in which case the sentence *John loves Mary* will have the S-structure in (3):

(3) $[_{cp}\ [_{c1}\ C\ [_{agrsp}\ \mathbf{John}\ [_{agrs1}\ AgrS\ [_{tp}\ [_{t}\ \mathbf{loves}_j]\ [_{agrop}\ [_{agro1}\ [_{agro}\ e_j]\ [_{vp}\ e_i\ e_j\ \mathbf{Mary}]]]]]]]]$

#8: This setting requires that CHECK-CASE(S) and CHECK-CASE(O) be satisfied before Spell-Out. The overt movements will thus be *e* and *f*. At S-structure, the subject is in Spec of AgrS-P, the object in Spec of AgrO-P and the verb in situ. This is a possible setting for Japanese. In that case, the Japanese sentence in (4a) will have the S-structure in (4b).

(4a) *Taroo-ga* *Hanako-o* *mi-ta*
Taroo-nom Hanako-acc see-past
'Taroo saw Hanako'

(4b) $[_{cp}\ [_{c1}\ C\ [_{agrs}\ \mathbf{Taroo-ga}_i\ [_{agrs1}\ AgrS\ [_{tp}\ T\ [_{agrop}\ \mathbf{Hanako-o}_j\ [_{agro1}\ AgrO\ [_{vp}\ e_i\ \mathbf{mi-ta}\ e_j\ ]]]]]]]$

#14: This setting requires that CHECK-AGRO, CHECK-T, CHECK-AGRS, CHECK-C and CHECK-CASE(S) be satisfied before Spell-Out. The movements *a*, *b*, *c*, *d* and *e* will thus be overt. At S-structure, the verb is in C, the subject in Spec of AgrS-P, and the object in situ. This is a possible structure for a VSO language like Berber. For instance, the Berber sentence in (5a) might have the representation in (5b) spelled out.

(5a) *i-annay* *urba* *tarbatt*
he-saw boy girl
'The boy saw the girl'

(5b) $[_{cp}\ [_{c1}\ [_{c}\mathbf{i-annay}_i]\ [_{agrsp}\ \mathbf{urba}_j\ [_{agrs1}\ [_{agrs}\ e_i[_{tp}\ [_{t}\ e_i][_{agrop}\ [_{agro1}\ [_{a}gro\ e_i]\ [_{vp}\ e_j\ e_i\ \mathbf{tarbatt}]]]]]]]]]$

#16: This setting requires that every MFUGC be satisfied before Spell-Out. Every movement is overt and the word order of the S-structure representation is V2: either SVO or OVS depending on which NP is being topicalized. Let us take German as an example. The sentences in (6a) and (7a) both mean 'Hans loves Bettina', but the topicalized NP is the subject in (6a) but the object in (6a). The structure of these sentences are in (6b) and (7b).

(6a) *Hans liebt Bettina*
Hans loves Bettina
'Hans loves Bettina'

(6b) $[_{cp}$ **Hans**$_i$ $[_{c1}$ $[_c$ **liebt**$_j]$ $[_{agrsp}$ e_i $[_{agrs1}$ $[_{agrs}$ $e_j[_{tp}$ $[_t$ $e_j][_{agrop}$ **Bettina**$_k$ $[_{agro1}$ $[_a gro\ e_j]$ $[_{vp}$ $e_i\ e_j\ e_k]]]]]]]]$

(7a) *Bettina liebt Hans*
Hans loves Bettina
'Hans loves Bettina'

(7b) $[_{cp}$ **Bettina**$_k$ $[_{c1}$ $[_c$ **liebt**$_j]$ $[_{agrsp}$ **Hans**$_i$ $[_{agrs1}$ $[_{agrs}$ $e_j[_{tp}$ $[_t$ $e_j][_{agrop}$ e_k $[_{agro1}$ $[_a gro\ e_j]$ $[_{vp}$ $e_i\ e_j\ e_k]]]]]]]]$

4. A Universal Parser

In this section, I present a parsing algorithm which implements the new approach to word order we have discussed so far. It should be pointed out that the parser to be presented is still in its elementary form. It is not intended to be a full-power parser which can handle any syntactic phenomena, though potentially it can be. The purpose of presentation here is not to provide a piece of ready-to-use software, but to introduce a new concept of parsing. The syntactic phenomena to be handled by this parser will be limited to those which have been covered in the minimal model we have discussed so far. In other words, it only parses simple declarative sentences with a subject, a verb, and an optional object. But it is capable of processing such simple sentences in *all* languages. It is a small parser but a universal one. The syntactic phenomena which are not meant to be covered here include, but not limited to, the following:

- Deletion, insertion and merge. Every chain built by the parser is expected to contain exactly one lexical constituent. By lexical constituent I mean a constituent which is actually pronounced. Furthermore, a NP chain is assigned exactly one theta role and a V chain contains exactly one verb.

- Reanalysis and other syntactic operations which involve movements of the 'yoyo' type. Every movement is a strict raising. Lowering is prohibited.

- Quantifier Raising. This operation is required by a LF requirement concerning scope representation. It should be represented at LF, but since we have not treated it as a component in our simple system, the present parser will not be responsible for it.

- Complex sentences. We are only concerned with the main clause here. In some languages, such as German, main clauses and embedded clauses have different word orders. This phenomenon can be handled in an augmented and more sophisticated version of the theory, but that is not our immediate concern.
- Adjuncts and modifiers. The present parser is mainly concerned with the relations between the head, the specifier and the complement. The only adjunction structures involved here are (a) adjunction of an X0 to another X0 in head movement, and (b) the adjunction sites which serve as intermediate landing sites in XP movement. It is not very difficult to extend the parser so that it can handle modifiers and other adjuncts, but I have avoided doing so in order to concentrate on the main issue.

This present parser is different from most existing parsers in two important respects:

- It parses D-structure, S-structure and LF at the same time. Most existing parsers are not concerned with LF. The parse tree they produce is either an S-structure representation or a combination of D-structure and S-structure representations. This works well in the pre-Minimalist framework because most UG constraints apply at D-structure and S-structure. As long as issues like the representation of scope ambiguity are not at stake, LF can be put aside without affecting the "grammaticality judgment" of the parser. In the Minimalist framework, however, LF has become much more important, since most "grammaticality judgments" are now made at LF. Therefore, parsing the LF structure has become a virtual necessity. I will show how to represent DS, SS and LF in a single parse tree.
- It builds the same X-bar tree no matter what language is being parsed. This is because (1) Lexical Projection and GT operations are performed uniformly so that the D-structure is identical in all languages, and (2) the MFUGCs are universal so that the movements and resulting LF representation are identical across languages. The only structure that can vary from language to language is S-structure. As we will see, this variation affects the position of a lexical constituent in the tree but not the overall X-bar structure.

4.1. Representing DS, SS, and LF in a Single Tree

The parser produces a parse tree which merges the DS, SS, and LF representations. The X-bar structures at these three levels of representation are essentially the same, so the X-bar tree produced by the parser is simultaneously a DS tree, an SS tree and an LF tree. In my simplified syntactic model, the only changes that can be made in the tree geometry during the derivation from DS to LF is the addition of adjunction sites. For instance, when the verb moves to a functional category to get certain features checked, it must be adjoined to the head of that functional category. The adjunction structure will be represented in the parse tree, with the understanding that the additional structure is in the LF representation but not in the DS representation. Whether it is part of S-structure depends on how far the pre-SO movements go.

The major difference between DS, SS and LF is the representation of movement chains. At LF, all possible movements have taken place and all the chains are complete. Therefore, in order for LF to be represented, the parse tree must be basically an LF tree where all the chains are fully represented. Each of those chains contains exactly one lexical constituent and its position in the chain differs at the three levels of representation. At LF, the lexical constituent is always at the "head" of some chain.[13] The verb is in C which is the head of a V0 chain. The NP which has gone through both an A-movement and an A-bar movement is in Spec of CP, which is the head of an A-bar chain, and all the other NPs are in their case positions which are heads of A chains. At DS, nothing has moved and every lexical constituent is at the "tail" of some chain: verbs and nouns are all in their VP-internal positions. In short, we can identify the LF and DS positions of each lexical constituent once the chain it appears in is built. Since the chains to be built are identical across languages, the LF and DS positions are fixed. We know where they are even if they are not explicitly represented. The S-structure position of a lexical constituent, however, is subject to parametric variation. It can be at either the "head" or "tail" of some chain depending on what movement has occurred before Spell-Out in a given language. We cannot tell where it is unless it is explicitly represented. Furthermore, the SS positions determine the order in which the lexical constituents are pronounced. It is those positions which identify the word order of a language. Therefore, only the SS positions of lexical constituents will be explicitly represented in the parse tree, with the assumption that the DS and LF positions are predictable. Take the structure in Figure 6 as an example. This is supposed to be a parse tree for the English sentence

[13] Following the standard terminology, I use "head" to refer to the highest link of a chain where the movement terminates and "tail" to refer to the lowest link of a chain where the movement starts.

John loves Mary where *John* is the topic.[14] The parameter setting assumed here is #4 where A, B and E are set to 1 while C, D, F and G are set to 0.

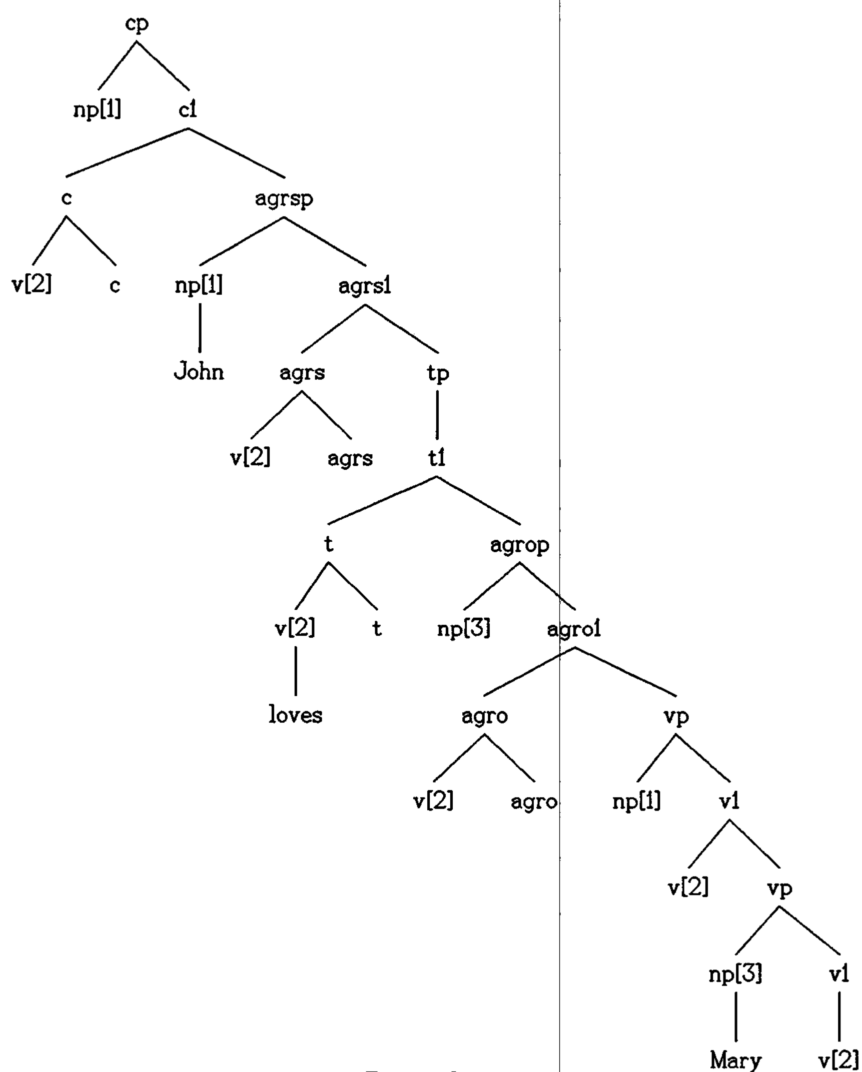


Figure 6

The nodes that are in the same chain are coindexed in this tree. The lexical constituents (*John*, *loves* and *Mary*) are in their S-structure positions. The terminals that contain no lexical constituents are empty (i.e. they have features but no phonetic content). The surface word order found here is SVO. This tree also tells us the DS and LF positions of each lexical head. For instance, the DS position of *loves* is in V* and its LF position is in C.

[14]In a non-topicalized sentence of English, the subject is assumed to be the default topic.

4.2. Parsing procedures that deal with universals

In this section, we discuss the parsing procedures that are universal. These procedures are responsible for construction of DS and LF representations. Since the trees built here are identical across languages, the parsing algorithm used to build these trees can be universal. The algorithm to be introduced here combines top-down and head-driven procedures. We choose this algorithm because it parallels the computational system described in the Minimalist Program very closely.

The basic operations in this algorithm are PROJECT, ATTACH and CHAIN-FORMATION.

4.2.1 Project

The PROJECT operation builds the maximal projection of each category. It corresponds to the computational system in the Minimalist Program which "selects an item α from the lexicon and projects it to an X-bar structure". (Chomsky (1992) p.30) Given a category X, it starts from X0, projects to X1, leaving an attachment site for YP on the right if the rule "complement(X,Y)" exists, and then projects further up to XP (X2) where an attachment site for ZP is built on the left if the rule "specifier(X,Z)" exists. This operation results in the elementary X-bar tree in Figure 1 which is the output of Lexical Projection. (The attachment sites are represented in upper-case letters.) When a sentence is parsed, PROJECT is typically applied more than once, which produces an array of elementary X-bar trees.

Each attachment site of XP consists of a set of features which are selected by the head of XP. For instance, the Spec of AgrS-P contains case and agreement features selected by AgrS. This is where case assignment and Head-Spec agreement are checked. The NP to be attached to this site must have features which are compatible or unifiable with the features of this attachment site.

An XP can be projected from either an empty or non-empty head. A non-empty head is an X0 which contains a lexical constituent or has a non-empty head adjoined to it.

The projection of XP from an empty head X0 must be licensed by two things:

- There must be an attachment site for XP in the current tree. In other words, a category can be projected from an empty head only if it has been predicted by the present structure.

- X0 must be licensed by a chain: either X0 is a link in a chain or it has a link of a chain adjoined to it. An example of the second case is the head of T. It is not in a chain but it can have a link of the V chain adjoined to it.

4.2.2 Attach

The ATTACH operation connects the array of elementary X-bar trees into a single tree. It corresponds to the GT operation in the Minimalist Program. Given two elementary X-bar trees XP and YP with XP preceding (i.e. projected before) YP, it takes one of the following actions.

- If XP has an attachment site for YP on the right, YP is substituted into this attachment site. In this case, YP is attached to XP as a complement. This is what has been illustrated in Figure 2.
- If YP has an attachment site for XP on the left, XP is substituted into this attachment site. In this case, XP is attached to YP as a specifier. This is what has been illustrated in Figure 3.
- Otherwise, the two trees remain unattached to each other.

This operation applies to each adjacent pair of elementary X-bar trees from left to right until no more action can be taken. The operation succeeds if the result is a single tree and fails otherwise.

Given the the rules in (1), we will never find a situation where XP and YP have attachment sites for each other. Either XP is a specifier of YP or YP is a complement of XP, but not both. Therefore, only one kind of GT operation is possible for each pair of trees.

The substitution involved in the ATTACH operation is a unificational process. When a tree *xp* is substituted into to an attachment site *XP*, all the features in *xp* and *XP* become unified. Attachment is not possible if *xp* and *XP* have conflicting features.

4.2.3 Chain-Formation

The operation that corresponds to *move-α* in the Minimalist Program is CHAIN-FORMATION. Since the chains being built are LF chains which are universal, this operation is again performed in a uniform way. The parser keeps three lists for the three types of chains being built: head chains, A chains, A-bar chains. The members of these lists are nodes and they represent chains which are yet to be completed. In other words, they serve as potential antecedents for the current node. As the PROJECT and ATTACH operations proceed from left to right, the lists are updated whenever a new X0 or XP nodes are built. The updating is done in one of the following ways.

- Initializing a chain. This happens when we encounter a node which finds no antecedent in the current chain lists.[15] This indicates the beginning of a new chain, so a copy of this node is added to one of the chain lists: to the head chain if the current node is an X0, to the A-chain list if it is an XP in an A position, and to the A'-chain list if it is an XP in an A-bar position.

- Passing on a chain. This happens when the node we encounter has an antecedent in the current chain lists but the node is not in a D-structure position.[16] This means that the current node is part of an on-going chain but the tail of this chain has not been reached. In this case, we take the antecedent node off the list, unify it with the current node, and place a copy of the unified node back in the list. The new node in the list is different from the original one in that it contains additional information from the current node.

- Terminating a chain. Normally, a chain terminates when it reaches a node in a DS position. In this case, the antecedent node is taken off the list, unified with the current node, but no copy of the unified node is put back into the list again. A chain can also terminate (prematurely) when a barrier is crossed.[17] This means the chains cannot be passed on even though they are still incomplete. In this case, the chain list will become empty no matter how many nodes it used to contain. This happens, however, only when the input sentence is ungrammatical.

4.3. Parsing procedures that handle language variation

So far I have described the universal operations of the parser. The parse tree they build contain the DS and LF representations which are universal, but there is no information about S-structure. In this section, I discuss the parsing procedures which respond to parametric variations. These procedures add the S-structure representation to the parse tree. It is those procedures that are responsible for the acceptance or rejection of a certain word order.

[15] An X0 element is only allowed to look for an antecedent in the head chain list. An XP in an A-bar position can only find an antecedent in the A-bar chain list. An XP in an A position, however, can look for an antecedent in either the A chain list or A-bar chain list. The latter happens when the XP in question is an NP which has ungergone A-bar movement from its case position to Spec of CP.

[16] The DS position of an NP is the position where it is assigned a theta-role, and the DS position of a verb is its VP internal position.

[17] How the barriers are implemented will be discussed in a different paper.

We have seen that the lexical constituents are expected to appear in different chain positions in different languages. The particular positions expected in a particular language are determined by the S-parameter values of this language. Intuitively, what the parser has to do is look at the S-parameters and decide where the lexical constituents are required to appear at S-structure. The algorithm which puts this idea into practice is the following:

> For each terminal node in the tree, decide on the basis of S-parameter values whether it must contain a lexical constituent or not. Continue if the requirement is satisfied and fail otherwise.

If a terminal node is expected to contain a lexical constituent, we must find one in the input string and attach it to this node. If a terminal node is not expected to contain a lexical constituent, it must be empty. However, being empty is not enough. An empty node (whether XP or X0) must be licensed by a chain which contains a lexical constituent. It is licensed if it is a link in the chain or it has a link of the chain adjoined to it.

Let us examine two nodes, an XP and an X0, to illustrate the algorithm.

The XP node to be examined will be Spec of AgrS-P and the X0 will be T. The Spec of AgrS-P must contain a lexical constituent if E is set to 1 and G to 0. It must be empty when both E and F are set to 0.[18] In this case, a copy of this node will be put in the A-chain list. In a grammatical sentence, this node will eventually be unified with a node which does contain a lexical constituent. When both E and G are set to 1, there are two possibilities. If the subject NP moves to Spec of CP, Spec of AgrS-P will be empty. If the object NP moves to CP, however, the Spec of AgrS must be non-empty. What the parser does at this time is the following. If the A'-chain list contains a node which is unifiable with the current node, (which means the object NP has not moved to Spec of CP,) the Spec of AgrS-P must be empty. The subject NP has already moved to Spec of CP. Otherwise, the subject NP cannot have moved and it must be in Spec of AgrS-P.

The X0 node for illustration is T. It must contain a lexical constituent (i.e. have a non-empty V adjoined to it) if B is set to 1 and C is set to 0. This means the verb has moved exactly to T before Spell-Out. T must be empty if B is set 0 (the verb has not moved to T) or C is set to 1 (the verb has moved through T to a higher position). In both cases, the parser will take a V node from the head chain list, adjoin it to T for feature checking, and put the checked V node back in the head chain list.

[18]This implies G is also set to 0.

4.4. An Example

In this section, I show how the above procedures work together by going through the parse of a Japanese sentence. Before the demonstration, however, we have to look at one more detail of the parsing algorithm which concerns how the parsing starts. It is assumed that, no matter what the input string is, the parsing of a sentence always starts by projecting CP. In the present system, every complete sentence constitutes a CP. Therefore a CP is predicted whenever a sentence is parsed. In this sense, CP is an original attachment site which tells us what structure we are going to build. Once CP is built, other predictions will follow and projections from empty heads will not be a problem.

The Japanese sentence to be parsed is the one in (4a), repeated here as (8).

(8) *Taroo-ga* *Hanako-o* *mi-ta*
Taroo-nom Hanako-acc see-past
'Taroo saw Hanako'

We again suppose that the S-parameters are assigned the values in #8 where E and F are set to 1 while the others are set to 0. The parse will be described in 6 steps, with the elementary X-bar tree projected at each step corresponding the numbered boxes in Figure 7. The attachment sites where those elementary X-bar trees are joined together are in the intersections of those boxes.

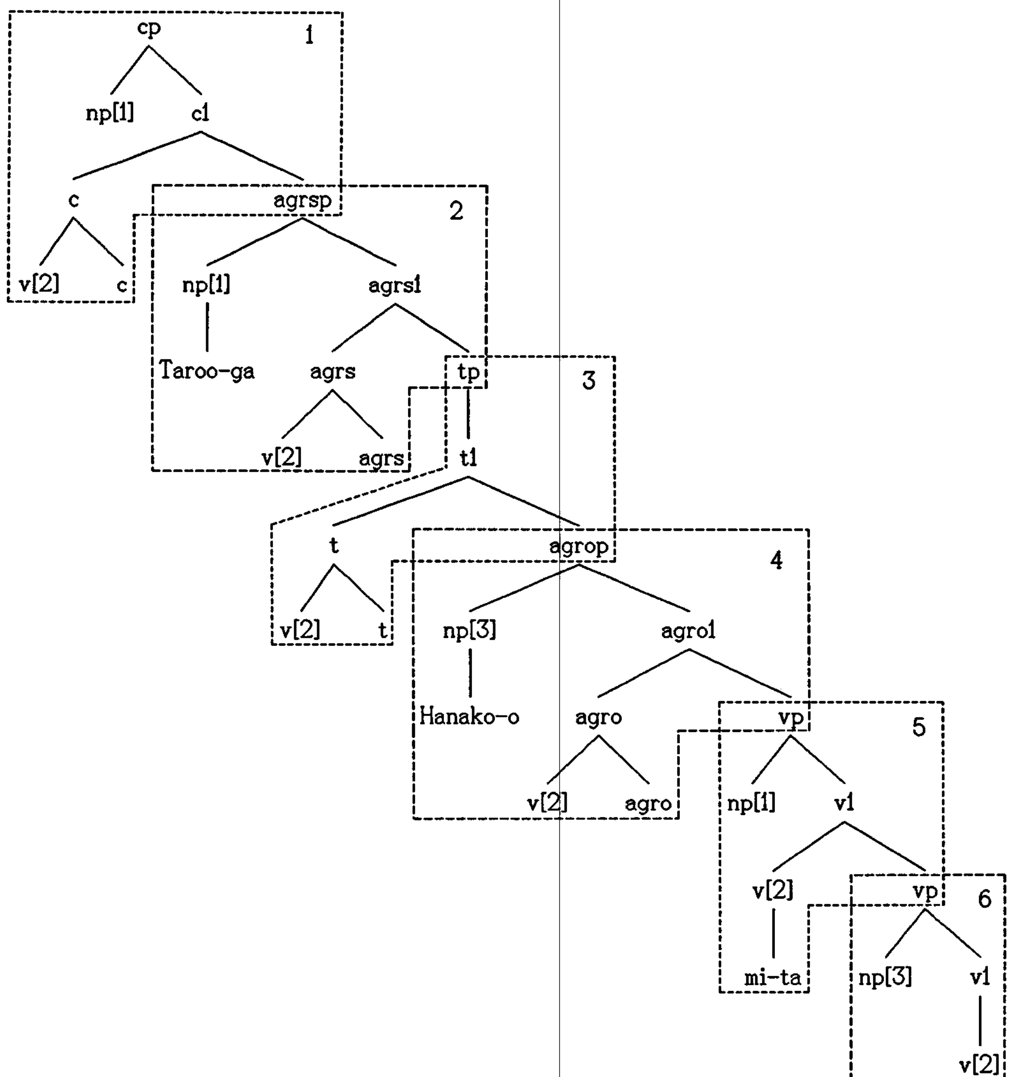


Figure 7

Step 1. Project CP. This CP is projected from an empty head. The projection is licensed by the initial CP attachment site just mentioned. The empty head is licensed by the V adjoined to it. Since D is set to 0, the V must be empty. A copy of the V node is then put in the head chain list (initializing a chain).

This CP tree here is looking for an NP specifier[19] on the left and an AgrS-P

[19]Actually, the Spec of CP has not yet had its category specified at this point. NP is put there instead of XP because in this sentence it is eventually specified as an NP.

complement on the right. Since no other tree has been built before this tree, the Spec position must be empty. This is licensed by the fact G is set 0 (Spec of CP not filled). So we just put a copy of this node in the A'-chain list.

Step 2. Since the complement attachment site tells us that an AgrS-P is expected, we now project the AgrS-P tree in Box 2. The head of AgrS is licensed by the adjoining V. This V must also be empty, as C is set to 0 (no overt verb movement to AgrS). We pick up the V node from the head chain list, unify it with the present V, and put a copy of the unified V back in the list (passing a chain). The Spec of AgrS-P, however, must not be empty, because E is set to 1. So we look for a lexical constituent in the string and found *Taroo-ga* to be the NP we are looking for. The NP headed by *Taroo-ga* is then attached to the Spec position. If *Taroo-ga* is the topic of this sentence, we will pick up the NP node from the A'-chain list and unify it with the current NP[20] Then the unified node is put in the A-chain list, Spec of AgrS-P being an A position.

Now we have two elementary X-bar trees: a CP tree and an AgrS-P tree. Since CP has an AgrS-P attachment site on the right, the AgrS-P tree is attached there. This results in the tree in the union of Box 1 and Box 2.

Step 3. The attachment site which is still open is TP. This TP is again projected from an empty head which is licensed by the V adjoined to 1. As B is set to 0 (no verb overt movement to T), the V must again be empty. The chain-passing is done in the usual way. There being no specifier to look for, TP is attached right away, which results in the tree in Boxes 1, 2 and 3.

Step 4. TP has left an attachment site for AgrO-P. The projection of AgrO-P is similar to that of AgrS-P. F being set to 1, the Spec of AgrO-P must be non-empty. We find the NP *Hanako-o* in the string and attach it to the Spec position. The A-chain list now contains an NP node but it cannot serve as an antecedent for the present NP, for they have different case features. So the new NP node is simply added to the A-chain list and a new chain is started. After this is done, AgrO-P is attached and the whole tree now extends through Boxes 1, 2, 3 and 4.

Step 5. The complement attachment site of AgrO-P tells us that a VP is expected. So we project this VP. Since A is set to 0 (no verb movement at all), a lexical

[20]A node can be unified with another node only if the two nodes have the same category. They must not contain features which are contradictory to each other. After the unification, they will have the same set of features and the same index.

verb must reside in its head. Looking at the string, we find mi-ta which is then attached to V0. Meanwhile, we pick up the V node from the head chain list and unify it with the present node. By now the V-chain has finally got a lexical head. The current VP has a Spec attachment site for NP. Since E is set to 1, this NP must be empty. This empty NP must be licensed by an antecedent in the A-chain list. There are two nodes in the list, but only one (the subject NP) can serve as the antecedent. We pick up this node and unify the current NP with it. The unified node is not put back in the list, however. The Spec of VP is a DS position[21] and the chain terminates here.

Step 6. The verb *mi-ta* is transitive one and assigns two theta roles. After assigning the agent role to the subject NP chain, it has one more role to assign. Therefore an additional VP is expected, represented by the complement attachment site in the higher VP. The additional VP is projected from an empty head which is licensed by being a link in the V chain. It is unified with the V node in the head chain list and the chain terminates. The Spec of this VP is also empty (F being set to 1) and it is licensed by the remaining NP node in the A-chain list. After unifying the two NP nodes, we terminate this second NP chain. By now the parse is complete. We have a single tree and the input string is empty.

5. Concluding Remarks

The present approach to parsing is still in its infant stage. The syntactic model used in this paper is an extremely simplified one and the parsing algorithm is still very rudimentary. An obvious question is how far this approach can go when 'real' sentences are being dealt with. There are two sub-questions here. One is whether the new approach to word order can account for the full range of word order phenomenon in natural languages. In particular, can we replace the head parameter and still get all the word order variation originally explained by the values of head parameters. The other question is whether the parsing algorithm will be feasible when more complex structures are to parsed.

No definite answer can be given to these questions at this stage, but the approach seems to be a promising one. The prospect of replacing head parameters is especially good. We have already derived many orders which the head parameters are intended to cover. We have also accounted for some orders which the head parameter has failed

[21]The parser knows this because this is the position where the NP chain is assigned a theta role.

to explain. There are some additional orders which can be explained by both the head-parameter or S-parameter. Whether a language has preposition or postposition can be explained by the S-parameter value of an additional MFUGC: the complement NP of a P has to move to the Spec of PP to have its case checked. We have a preposition if the movement is after Spell-out and a postposition otherwise. Whether an NP precedes or follows the determiner can be explained in a similar fashion: the NP must move to the Spec of DP for agreement reasons and the NP will precede the determiner if the movement takes place before Spell-Out. The full range of possibilities are yet to be explored.

The prospect for parsing is less clear. This is partly because many parsing problems depend on the syntactic theory for solution. No full parser can be built unless we have full syntax. For example, we have not discussed how the positions of relative clauses are to be represented in the new model. Therefore, we currently do not know what will happen when these modifiers enter the picture. For this reason, I have not addressed the parsing issue in Frazier and Rayner (1988), since their conclusions rest crucially on relative constructions. As far as left-branching is concerned, however, the present algorithm should have no problem, because it is basically head-driven. As has been pointed out in Pritchett (1991), the head-driven parsing algorithm can handle left-branching and right-branching structures equally well. However, we cannot see the whole picture until the present model has been developed further.

Bibliography

Berwick, R.C. (1991) Principles of principle-based parsing. In R.C. Berwick et al (eds.) *Principle-Based parsing: Computation and Psycholinguistics.* Kluwer Academic Publishers, the Netherlands.

Chomsky, N. (1986) *Barriers*, MIT Press, Cambridge, Massachusetts.

Chomsky, N. (1991) Some notes on the economy of derivation and representation. In R. Freidin, ed., *Principles and Parameters in Comparative Grammar.* MIT Press, Cambridge, Massachusetts.

Chomsky, N. and H. Lasnik (1991) Principles and parameters theory", to appear in J. Jacobs, A. von Stechow, W. Sternefeld, and T. Vennemann, eds., *Syntax: An International Handbook of Contemporary Research*, Walter de Gruyter, Berlin.

Chomsky, N. (1992) A minimalist program for linguistic theory. *MIT Occasional Papers in Linguisitics* Number 1.

Fodor, J. (1992) Learnability of phrase structure grammars. To appear in R. Levine (ed.) *Formal Grammar: Theory and Implementation*, Vancouver Studies in Cognitive Science, University of British Columbia Press.

Fong. S. (1991) The computational implementation of principle-based parsers. In R.C. Berwick et al (eds.) *Principle-Based parsing: Computation and Psycholinguistics.* Kluwer Academic Publishers, the Netherlands.

Frazier, L and K. Rayner (1988) Parameterizing the language processing system: left-vs.right-branching within and across languages. In J. Hawkins (ed) *Explaining Language Universals.* Basil Blackwell, New York.

Gibson, E. and K. Wexler (1992) Parameter setting, triggers and V2", Paper presented at GLOW 1992.

Huang, C.-T. J. (1982) *Logical Relations in Chinese and the Theory of Grammar*, Doctoral dissertation, MIT, Cambridge, Massachusetts.

Jackendoff, R. (1977) *X-bar Syntax*, MIT Press, Cambridge, Massachusetts.

Johnson, M. (1989) Parsing as deduction: the use of knowledge of language. *Journal of Psycholinguistic Research 18(1)*: 105–128.

Kayne, R. (1992) Talks at GLOW 1992 and UCLA.

Koopman, H. (1992) Licensing Heads. UCLA manuscript.

Larson, R.K. (1990) Double objects revisited: reply to Jackendoff. *Linguistic Inquiry 21*: 589–632.

Lasnik, H. (1981) Restricting the theory of transformation: a case study. In N. Hornstein and D. Lightfoot (eds.) *Explanation in Linguistics.* Longmans, London.

May, R. (1985) *Logical Form.* The MIT Press, Cambridge, Massachusetts.

Nyberg, E. (1990) A limited non-deterministic parameter-setting model", *Proceedings of the Twenty-First North East Linguistic Society Conference*, McGill University, Montreal, Quebec. Pritchett, B. (1991) Head Position and Parsing Ambiguity. *Journal of Psycholinguistic Research 20(3)*, 251–279.

Sportiche, D. (1990) *Movement, Agreement and Case.* UCLA manuscript.

Stabler, E.P., Jr. (1992) *The Logical Approach to Syntax.* The MIT Press, Cambridge, Massachusetts.

Stowell, T. (1981) *Origins of Phrase Structure.* Doctoral dissertation, MIT, Cambridge, Massachusetts.

Stowell, T. (1983) Subjects across categories. *Linguistic Review 2*: 561-570.

Travis, L. (1984) *Parameters and Effects of Word Order Variation*, Doctoral dissertation, MIT, Cambridge, Massachusetts.

Weinberg, A (1992) Parameterising the parser: a grammar-based theory of garden paths. Paper presented at GLOW 1992.

Wu, A (1992a) A computational approach to 'intake' ordering in syntactic acquisition", Paper presented at the Eleventh West Coast Conference on Formal Linguistics, UCLA, Los Angeles.

Wu, A (1992b) Acquiring word order without word-order parameters? To appear in *UCLA Working Papers in Psycholinguistics.*