PHRASE STRUCTURE PARADOXES: INTRO

PHRASE STRUCTURE AND TYPES OF MISMATCHES

HIERARCHICAL STRUCTURES: TREES

Most if not all disciplines of formal linguistics (syntax, semantics, morphology, phonology) employ the heuristic strategy of segmenting the input strings to be analyzed (sentences, units of meaning, words) into substrings, labeling these substrings and organizing them in a hierarchical order. Tree diagrams provide a convenient mode of representation for graphically depicting this mapping from strings to hierarchically organized substrings. To take an example from morphology, the word *indecipherability* can be broken up into its morphemes, which are in turn combined by the labeled nodes of a tree into larger units (in this case morphemes):

(1) \[
\begin{array}{c}
N \\
A \\
in \quad A \\
V \\
de \quad N \\
cipher
\end{array}
\]

(Spencer 1991: 183)

Nodes fall in two groups: non-terminals (e.g. A, N) and terminals which contain lexical content (*cipher, de, ...*). As can be seen from (1), tree diagrams encode three types of information:

(i) linear order of the nodes (*de* precedes *cipher*)
(ii) hierarchical organization of the nodes
(iii) category labels of the nodes

Different subfields employ different types of diagnostics in order to infer the exact structure of a given string: For instance, evidence that words are indeed organized by means of a hierarchical ordering - as opposed by simple concatenation - comes from the observation that not all logically possible combinations of substrings lead to well-formed outputs. If the structure of the word were flat, as in (2), it would e.g. not be evident why the prefix *in-* can attach to a noun in *in-*{\_}\_N *decipherability*, while it otherwise only combines with adjectives (*incorrect* vs. *intruth*):
In syntax, there is a battery of tests which can be used as a diagnostic for the organization of the syntactic constituents - the phrase structure - of a given sentence. These CONSTITUENCY TESTS aid in establishing the exact shape of a syntactic tree for this sentence.

- Only full constituents may be dislocated by MOVEMENT operations such as topicalization and wh-movement:

(3) **TOPICALIZATION**

a. They moved \([\text{DP this syntactic constituent}]\)

b. \([\text{DP This syntactic constituent}], \text{ they moved}\)

c. *This syntactic, they moved constituent

d. *This constituent, they moved syntactic

e. *This, they moved syntactic constituent

(4) **VP-TOPICALIZATION**

They tried to move something,

a. ...and \([\text{VP move} \ [\text{DP this syntactic constituent}]\], \text{ they did}\)

b. *...and \([\text{VP move} \ [\text{this syntactic}]], \text{ they did constituent}\)

c. *...and \([\text{VP move} \ [\text{this constituent}]\], \text{ they did syntactic}\)

d. *...and \([\text{VP move} \ [\text{syntactic constituent}]], \text{ they did this}\)

e. *...and \([\text{VP move} \ [\text{this}]\], \text{ they did syntactic constituent etc....}\)

(5) **WH-MOVEMENT**

a. They moved \([\text{DP this syntactic constituent}]\)

b. Which constituent did they move \([\text{DP}]\)

c. *Which syntactic did they move constituent

d. *Which constituent did they move syntactic

e. *Which did they move syntactic constituent

- Other tests help to determine the c-command relations in a tree. NEGATIVE POLARITY ITEMS have to be c-commanded by negation or a negative quantifier:

(6) **NEGATIVE POLARITY ITEM (NPI)**

a. John didn’t buy anything

b. Nobody bought anything

c. *Anybody bought nothing

d. *Anybody didn’t buy the book

Similarly, ANAPHORS (reflexive and reciprocal pronouns) have to be locally bound by c-commanding antecedents (as encoded in Principle A of Binding Theory):
(7) a. Sally likes herself
   b. *Herself likes Sally
   c. *We showed [the book [about Sally]] to herself

In general, constituency tests are hypotheses about constituency, which have been arrived at by observational generalizations (i.e. by induction from the facts observed) about phenomena such as movement, anaphoric dependencies, coordination and ellipsis. For instance, the hypothesis that anaphors have to be c-commanded by their antecedents can be used to determine the structure and in particular the c-command relations of other constituents based on the following rationale: if in the string \( \alpha \ldots \beta \ldots \), \( \alpha \) can antecede an anaphor \( \beta \), \( \alpha \) c-commands \( \beta \), resulting in the bracketing (8)a. If \( \beta \) can bind \( \alpha \), this indicates that \( \beta \) c-commands \( \alpha \), yielding the structure (8)b. If both binding relations are attested, it can be inferred that \( \alpha \) and \( \beta \) mutually c-command each other, i.e. the structure is flat, as in (8)c:

(8) a. \[
   \begin{array}{c}
   \alpha \\
   \beta \\
   \end{array}
   \]
   b. \[
   \begin{array}{c}
   \beta \\
   \alpha \\
   \end{array}
   \]
   c. \[
   \begin{array}{c}
   \alpha \\
   \beta \\
   \end{array}
   \]

Tests of this sort have e.g. been employed to refute the claim that the VP in non-configurational (free-word order) languages such as Japanese, Greek and German is flat (see Hoji 1985; Anagnostopoulou & Alexiadou 1998; Webelhuth 1989 among others).

To say that a context meets/fails a test means that this context conforms/fails to conform with the hypothesis which the test is based on. In a wide variety of contexts, application of more than one constituency test yield inconsistent results. Consider e.g. example (9)b. According to the movement test, the structural representation for (9)b should be mainly right-branching, as in (10)a. If, however, the binding test is given priority, sentence (9)b should be assigned the predominantly left-branching tree in (10)b, in which the antecedent c-commands the name:

(9) a. Sally likes [\( \text{DP this picture of herself} \)]
   b. [\( \text{DP Which picture of herself} \)] does Sally like
The observation above might be taken as an indication that the hypotheses underlying the two tests involved are incorrect. But it’s hard to see how a plausible hypothesis about e.g. binding should look like which accommodate for (9)b alongside with the basic phenomena. Let us therefore assume that the tests are in essence correct.

With this in the background and given the widely held assumption that each string can only be assigned a single representation (for a different view see e.g. Pesetsky 1995), we arrive now at a paradox: according to the criteria for structure, the string (9)b should be assigned a left-branching as well as a right-branching representation. It will be Phrase Structure Paradoxes (‘PS-Paradoxes’) of this sort which will be the main concern of this course.

Examples such as (9)b have been extensively discussed in the literature on Reconstruction, which has advanced a number of interesting proposals (Barss 1986; Chomsky 1992; Epstein et al; 1997; Fox 1999; Frey 1989; Heycock 1995; Lebeaux 1990; Lechner 1998; Sternefeld 1997 among others). The core of the problem - conflicting evidence from different types of PS-tests - generalizes however to a much wider range of phenomena. In order to evaluate the full range of the problem, let us therefore address the question first, where exactly PS-Paradoxes are predicted to surface, turning from there to possible strategies for resolving them.

Next:
- Assumptions about organization of the grammar
- Generalizing the problem: what types of PS-Paradoxes does one expect?
- Strategies for resolving PS-Paradoxes
- Case studies

**Organization of the Grammar and Types of PS-Paradoxes**

According to the transformational generative tradition, the grammar is internally organized in discrete components (morphology, phonology, syntax and semantics,...), which themselves may consist of
different levels. The components interact with each other in that the output of one component furnishes the input to another component\(^1\).

Within the Minimalist framework (Chomsky 1995), sentences are generated in syntax, and then submitted to phonology (Phonological Form; ‘PF’) and the semantic component, where they are further processed and assigned a phonetic and semantic representation, respectively. If a component is internally made up of several levels, a similar feeding relation passes on information from one level to the next. In Minimalism, syntactic strings are first processed in the surface syntactic component. At Spell-Out, phonologically relevant information is sent off to PF, while remaining syntactic information is submitted to further (covert) manipulation at Logical Form. Thus, there are two different ‘points of transition’ in the grammar: First, between two different components (e.g. between syntax and semantics). This type of transition is also referred to as the INTERFACE between two components. Second, there are transitions between two different levels (overt syntax and LF, or different levels in derivational models of phonology such as Lexical Phonology (Kiparsky 1982)).

Assume that the mapping from one level/component to the next is transparent in that the tree structure assigned to a string remains the same when it enters the next level/component. On this - somewhat idealistic - view, there are no operations which elide or reorganize the output of one component or level before it enters the next stage of the derivation, and the interfaces as well as the different levels are ISOMORPHIC (Gr. ‘of the same form’).

Assume finally that all levels/components include a certain array of PS-tests. Then, PS-Paradoxes arise whenever two (or more) tests yield conflicting evidence as to the hierarchical tree structure of a string. On the basis of this model it is possible to distinguish three classes of paradoxes:

A. LEVEL INTERNAL PARADOXES
Constituency tests leads to conflicting results as to the factorization of a string within a single level of a component (surface syntax, LF, morphology,...).

B. LEVEL ORDERING PARADOXES
The structure assigned to a string at the output of one level conflicts with the way the string has to parsed in the next level.

C. INTERFACE PARADOXES
A string is assigned different parses in different components of the grammar (e.g. syntax and phonology).

\(^1\)There are also theories which advocate parallel processing of certain components. Parallel Morphology (Borer 1991) postulates that morphology is computed parallel to syntax. Similarly, in Distributed Morphology, syntax operates on morphological feature bundles, which are lexically instantiated at a post-syntactic level (Embick 1997; Halle & Marantz 1993; Noyer 1997; Harley & Noyer 199? among many others).
The taxonomy above does not have theoretical relevance in itself. It only exhausts the logical possibilities where PS-Paradoxes can be found and thereby guides one in locating potentially troubling configurations. Here are some examples of various types of paradoxes from the literature (for an overview see e.g. Kenstowicz 1994; Spencer 1991):

**Phonology-Syntax/LF**

- Evidence that syntactic information is accessible to phonology: Liaison in French (Selkirk 1982)

(11) a. Il y a encore deux après-midi  
   PF: [ ...dœspr...]  
   ‘there are still two afternoons’

   b. Il y a encore deux après lui  
   PF: [ ....dœapr...]  
   ‘there are still two (of them) after him’

(12) a. Il y a eu deux après-midi  
   b. Il y a eu deux après lui

**Clitics**

Clitics are phonologically bracketed with host NP, but syntactically parsed within VP (Spencer 1991):

(13) a. Phonology: [Tom’s] [a linguist]  
    b. Syntax: [NP Tom] [VP ’s a linguist]

Nespor & Vogel (1986: 2): “prosodic constituents built on the basis of information contained in the morphological and syntactic component are not necessarily in a one-to-one relation with any of the constituents of morphology or syntax”

‘unhappier’

- * unhappier means ‘more unhappy’ and not ‘not happier’ (Beard 1991; Pesetsky 1985; Kang 1993; Sproat 1992)

(14) Fred is unhappier than Sally is

   a. Fred is more unhappy than Sally  
      ‘Fred is less happy than Sally’

   b. *It is not the case that Fred is more happy than Sally  
      ‘Fred is as happy as Sally or less happy than Sally’
- *-ier* attaches only to words with maximally two syllables:

(15)  

<table>
<thead>
<tr>
<th>a. tall</th>
<th>b. taller</th>
</tr>
</thead>
<tbody>
<tr>
<td>noble</td>
<td>nobler</td>
</tr>
<tr>
<td>curious</td>
<td>*curiouser</td>
</tr>
<tr>
<td>intelligent</td>
<td>*intelligenter</td>
</tr>
</tbody>
</table>

(16)  

<table>
<thead>
<tr>
<th>a.</th>
<th>b. PHONOLOGY</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>A</td>
<td>er</td>
</tr>
<tr>
<td>un</td>
<td>A</td>
</tr>
<tr>
<td>happy</td>
<td>A</td>
</tr>
<tr>
<td>un happy</td>
<td></td>
</tr>
</tbody>
</table>

---

**Morphology/MorphonoLOGY**

**Stratal Morphology** (Siegel 1979): Affixes are attached in two steps. Phonological rules are sandwiched in between these two levels (Kiparsky 1982; Mohanan 1982):

(17)  

1. Class I affixation  
   (-ity, -ation, -ion, -ive, -al, -ous; con-, de-, in-, en-...)

2. Stress rules

3. Class II affixation  
   (-able, -ize, -ment, -less, -ness, -hood, -ful; non-, un-, de-...)

Roughly, Class I affixes trigger phonological processes such as stress-shift, while Class II affixes are phonologically inert. The theory-internal assumption of level ordering leads to a number of paradoxes, among them:

[[STEM Class II] Class I]  
- *-ity, -ation*, and -*al* are Class I and trigger stress-shift to a preceding heavy syllable:

(18)  

<table>
<thead>
<tr>
<th>a. frágile</th>
<th>b. fragility</th>
</tr>
</thead>
<tbody>
<tr>
<td>augment</td>
<td>augmentation</td>
</tr>
<tr>
<td>párrent</td>
<td>paréntal</td>
</tr>
</tbody>
</table>

- *-able, -ize, -ment* are Class II and do not trigger stress shift:

(19)  

<table>
<thead>
<tr>
<th>a. cóunt</th>
<th>b. cóuntable</th>
</tr>
</thead>
<tbody>
<tr>
<td>standard</td>
<td>standardize</td>
</tr>
<tr>
<td>góvern</td>
<td>góvernment</td>
</tr>
</tbody>
</table>
**PARADOX**: Some Class I suffixes attach to some Class II affixes (Aronoff 1976; Kenstowicz 1994: 234):

(20) a. STEM c. [[STEM CLASS II] CLASS I]
cóunt countability
standard standardization
góvern government

'BUNGRAMMATICALITY'

- *-in* is a Class I prefix, *-un* is a Class II prefix, which is phonologically inert (Siegel 1979):

(21) a. PLACE ASSIMILATION b. NO PLACE ASSIMILATION
convenient eatable
illogical lawful
possible bound

- *-un* only attaches to adjectives:

(22) unhealthy unlawful
unlawful unwoman

**PARADOX**: *-ity* is Class I suffix, and *un-* is Class II. In ungrammaticality, *-ity* should accordingly attach before *-un*. But *-ity* changes category into noun, and *un-* cannot attach to nouns, indicating that *un-* has to be affixed prior to *-ity* (for recent discussion see Benua 1997; McCarthy 1999).

(23) a. CATEGORICAL SELECTION b. PHONOLOGY

\[ N \]
\[ A \] ity
\[ \text{un} \] A grammatical

\[ N \]
\[ \text{un} \] A ity

**COMPOUNDING**

- On the level-ordering hypothesis, derivational morphology precedes compounding. (*school teacher, compound formation, ...*). In numerous cases, though, affixation seems to follow compounding, as can be seen from the interpretation of the output (Fabb 1984; Williams 1981; Selkirk 1982; Sproat 1985):

(24) a. Level Ordering: [nuclear] [physicist]
    b. Interpretation: [nuclear physicist]ist (not: a physicist who is nuclear)

(25) transformational grammarian (not: a grammarian who is transformational)
criminal lawyer (not: a lawyer who is criminal)
Russian teacher (not: a teacher who is Russian)

Similarly, set theoretic, cross-sectional,...
Sometimes, the atoms of the compound are unattested, indicating the bracketing under b:

(26)  
\begin{align*}
\text{a. Level Ordering:} & \quad \text{[four] [legged]} \\
\text{b.} & \quad \text{[[four-leg]ed]} \quad (*\text{four-leg})
\end{align*}

Similar problem for synthetic compounds such as truck driver (*truckdrive) (DiScullio & Williams 1987).

**STRATEGIES FOR RESOLVING PARADOXES**

**DISTRIBUTION ACROSS NON-ISOMORPHIC COMPONENTS**

The components are not mapped in an isomorphic way (i.e. structure is lost). Principles in conflict apply in different components of the grammar (Syntax/LF, semantics, Conceptual Structure; see e.g. Beard 1991; Cullicover & Jackendoff 1995, 1997, 1999 for Binding Theory and the CSC).

**PARALLEL STRUCTURES**

Strings are assigned two distinct representations, which coexist during the course of the derivation. The principles in conflict apply to different structures (Brody 1994; Pesetsky 1995).

**DERIVATIONAL REARRANGEMENT**

Principles at conflict are assumed to apply at distinct levels or stages during the derivation. In addition, trees change their shape either by

- Conspiracy of conventional reordering processes (movement, ellipsis,...) which are independently attested in the grammar (Pesetsky 1995) or
- By means of some other derivational operation (Phillips 1996).

**MOVEMENT SOLUTIONS**

(9)b \[_{\text{DP}} \text{Which picture of } \text{herself} \text{ does Sally like}\]

Returning to the conflict between movement and binding in (9)b, the standard answer to why a dislocated anaphors can satisfy Principle A runs as follows (Chomsky 1992):

(27) **ASSUMPTIONS:**

- (9)b involves movement
- Movement leaves Copies (COPY THEORY; Chomsky 1992)
- Binding Theory applies at LF
- Binding Theory has access to strings inside copies

I.e., trees contain more information than can be retrieved from the surface string alone. In the more abstract representation for (9)b below, Sally c-commands the anaphor herself inside the
Movement can e.g. not be undone if the category which undergoes movement contains a trace and this trace fails to be bound by a c-command antecedent in surface syntax, as in (ii), leading to a violation of the Proper Binding Condition (Fiengo 1972):

(i) ??Who do you wonder [which picture of t_i] Bill bought t_k
(ii) *[Which picture of t_i] do you wonder who_i bought t_k

On discussion of counter-cyclic derivations see e.g. Kitahara (1997), Lasnik & Saito (1992) and Müller (1998).

A movement solution has been proposed for the analysis of some bracketing paradoxes in Pesetsky (1985). Selectional requirements (ungrammaticality) have to be met only at LF:

While most of the morphonological paradoxes proved recalcitrant and have not been given a commonly accepted solution, many other paradoxes are theory-internal in that they arise only on the
assumption of certain set of axioms. The analysis of Exceptional Case Marking (ECM)-constructions in the GB-framework manifests a prototypical case in point.

(30) Sally expected him to like the movie

(31) **ASSUMPTIONS:**
- Verbs assign (internal) \( \theta \)-roles to their complements.
- Verbs case-mark their complements (GB-style).

In ECM contexts, the higher verb (where verb\( \in \{ \text{believed, imagined, reported, considered, allege, discover, figure, know, perceive, recall, observe, reckon, recognize, sense, understand} \} \)) assigns a \( \theta \)-role to its CP-sister node, but assigns case to the subject of the CP. Thus, two categories compete for the same position: the CP, which needs to be \( \theta \)-marked by its sister (\( V^o \)), and the DP, which needs to be case-marked by \( V^o \).

(32) a. Sally expected him to like the movie

b. 
```
  V^o
     CP
    expected
      IP
      DP_{ACC}
```

In GB, ECM necessitated the assumption of S’-deletion (and a loosening of the structural conditions on Case assignment/government Chomsky 1981; 1986; Stowell 1981).

Once Case checking is disassociated from \( \theta \)-role assignment, as in Minimalism, the problem disappears (Lasnik 1993). Again, the resolution of the paradox depends on movement: the subject of the subordinate clause raises overtly to SpecAgrOP to check accusative case:

(33) 
```
  AgrOP
     DP_{ACC}
       AgrO'
         VP
           expect
             CP
               t_i
```

- Evidence for raising comes from Postal’s (1974) observation that ECM-subjects take scope over matrix adjuncts:

(34) **LICENSEING OF ANAPHORIC DEPENDENCIES**
  a. ?The DA \([\text{VP proved the defendants to be guilty]} \text{ during each other’s trials}\]
  b. ?*The DA \([\text{VP proved that the defendants are guilty]} \text{ during each other’s trials}\]
(35) **NPI-LICENSING**  
a. ?The DA \[vp \text{proved none of the defendants to be guilty} \] during any of the trials]  
b. ?*The DA \[vp \text{proved that none of the defendants are guilty} \] during any of the trials]

• Movement should take place in the overt component, since binding relations and NPI-licencing are well-known to not interact with covert movement (Chomsky 1981; Riemsdijk & Williams 1981; for exceptions see Fiengo & May 1994; Fox 1995):

(36) a. Nobody saw anybody  
b. *Anybody saw nobody

(37) *Somebody gave him, [every book that John, liked]

(38) **QUESTION:** How come that the subject of the ECM complement follows, rather than precedes the ECM-verb?

**ANALYSIS:** ECM-raising is overt movement, but lower copy is pronounced (along the lines of ‘overt QR’ of Fox & Nissenbaum 1999).

(For a similar case see also non-obligatory ECM in Hungarian, where a DP can be assigned dative case if this DP is the subject of CP in the dative position; Horvath 1997)

~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

**NEXT:** Selected PS-Paradoxes and their solution. It will be suggested that the movement strategy generalizes to the analysis of these constructions.

**BIBLIOGRAPHY**


Fox, Danny and John Nissenbaum. 1999. Extraposition and Scope: A Case for overt QR. In *WCCFL 18*


