Structure Building and Unification

Colin Phillips
University of Maryland
colin@glue.umd.edu

The Standard View

• There are multiple (related) structure-building systems
  Grammar
A recursive characterization of the structural descriptions of all possible expressions of a language. When we speak of a grammar, we mean the grammar used to generate the language (i.e., the competence grammar). If a language has more than one grammar, the question is which grammar should be chosen in the First Language Acquisition Task (FLAT).

• How are grammatical inflections made? If the parser is a competence grammar for the language, then the parser is a performance system for the language. If the parser is a covering grammar for the language, then the parser is a performance system for the language.

Questions about Unification?

• How do performance-systems recruit the information in the grammar?
• How are grammaticality judgments made? If the parser is a covering grammar for the language, is it a perfectly accurate implementation of the grammar?
• How does learning occur, given that the grammar is ‘hidden’ behind the parser? (error signals in learning are only indirectly related to the changes that they must trigger)

Alternative Approach

(2) A speaker’s grammatical knowledge takes the form of an incremental generative system, which subsumes production, comprehension, and grammar. This is an explanation of why grammars are so different.

Comprehension/Performance

A formal characterization of the structural descriptions of all possible expressions of a language. How do performance-systems recruit the information in the grammar?

Questions about Unification?

• How do performance-systems recruit the information in the grammar?
• How are grammaticality judgments made? If the parser is a covering grammar for the language, is it a perfectly accurate implementation of the grammar?
• How does learning occur, given that the grammar is ‘hidden’ behind the parser? (error signals in learning are only indirectly related to the changes that they must trigger)

Evidence

• In learning, there is more closely related to learning grammar. Processes that are sensitive to structural descriptions more fragmentary performance systems become more proficient. These systems are incremental (analysis-by-synthesis). In learning, better incremental descriptions can be shown to be preferable.

Unification Questions Again

• Relation between grammatical knowledge and the grammatical component of performance systems becomes straightforward, since they are identical.
• Processes underlying grammaticality judgment become more transparent.
• In learning, error signals more closely related to resulting changes.

Evidence

• Is this the right way to think about linguistic knowledge? What evidence can be brought to bear?…the available evidence varies across different domains...

Problems

Predictably, some way in which the parser adds to the grammar is desired.

ND (1999: 320)

-"...parsing routines exist merely for the purpose of performance, etc.

Evidence

- Produces a structural notion of simplicity, force in essential ambiguity

Evidence

- A system for incremental structuring of sentences for speaking.

Evidence

- A system for (possibly) incremental structuring of sentences for speaking.

Evidence

- A system for (possibly) incremental structuring of sentences for speaking.

Evidence

- A system for (possibly) incremental structuring of sentences for speaking.
Some strings that can be coordinated cannot undergo ellipsis:

(1) *Whereas [Mike believes [Wallace did the work that he did]] and [the queen did the work that he did]

Some strings that can be coordinated cannot undergo movement:

(2) *Whereas [Mike believes [Wallace did the work that he did]] and [the queen did the work that he did]

Coordination is a very special operation. Although it does not contradict standard assumptions about constituency, it does create conflicts between syntactic constituency tests: these are opacity effects. Certain constituents are not included in the order in which terminal elements are pronounced.

Incrementality: from left-to-right.

• Differences between tests can be understood if we assume that strings are built up incrementally, from left-to-right.

• Consequences: some tests pick out different sets of strings as constituents, and others do not

Incrementality hypothesis

Sentence structures are built incrementally from left-to-right, i.e. in the order in which terminal elements are pronounced.

Incrementality hypothesis:

Although there are excellent characterizations of the range of constituents identified by individual structural diagnostics, there is no general theory of why different tests apply at different points in a derivation.

Consequences:

Proposal:

• Conflicts between syntactic constituency tests are opacity effects, which arise because different tests apply at different points in a derivation.

• Differences between tests can be understood if we assume that structures are built up incrementally, from left-to-right.

Incrementality hypothesis

Sentence structures are built incrementally from left-to-right, i.e. in the order in which terminal elements are pronounced.

Incrementality hypothesis:

Although there are excellent characterizations of the range of constituents identified by individual structural diagnostics, there is no general theory of why different tests apply at different points in a derivation.

Consequences:

Proposal:

• Conflicts between syntactic constituency tests are opacity effects, which arise because different tests apply at different points in a derivation.

• Differences between tests can be understood if we assume that structures are built up incrementally, from left-to-right.
3

Conflict between movement preferences and binding results is possible because binding preferences are more flexible than is predicted by flexible constituency theories.

ii. Makes it possible to maintain the assumption that syntactic operations manipulate constituents, and that sentences ... than is predicted by flexible constituency theories.

A. Coordination

Predicted to be extremely liberal, because coordinated strings are immediately adjacent; hence no risk of first conjunct being left unexpressed.

Wallace will give Gromit crackers for breakfast and toast for lunch. ... of course, some strings are never constituents:

(9)

a. *The man [who built the rocket has] and [who studied robots designed] a dog.

b. *Wallace will give [his dog] [after Wallace] and [his sheep the milkman]
c. *After Wallace fed [his dog] and [his sheep the milkman], the milkman arrived.

c. After Wallace fed the dog half a dozen and [a handful of crackers for Gromit] a...

B. Coordination vs. Pseudogapping

Predicted to pick out a narrower range of constituents: the string that licenses ellipsis must be a constituent both when it appears before and after the ellipsis site is built & licensed, i.e. constituent must be present beyond completion of sentence which is antecedent of ellipsis.

• Verb–preposition strings may be coordinated because they are coordinated at different stages of the incremental derivation of a sentence.


But they cannot be deleted in elliptical coordination, e.g., a: 'somebody ate lunch, (P)ossilis of dinosaur's

(11)


But they cannot be deleted in elliptical coordination, e.g., a: 'somebody ate lunch, (P)ossilis of dinosaur's

• Verbs–position strings may be coordinated

(10)

a. *The cat slept on the rug in the middle of the living room.

b. *Jonathan slept on the rug in the middle of the living room.

c. *After Jonathan slept on the rug in the middle of the living room, the cat ate the cat food for dinner. ...
4

A. Movement vs. Ellipsis

Prediction: constituents become unavailable to syntactic processes once they have been destroyed.

Both VP-fronting (VPF) and VP-ellipsis (VPE) can move/delete strings beginning at the left-edge of VP.

(17) a. and [give candy to children in libraries on weekends] he did.
b. and [give candy to children in libraries] he did on weekends.
c. and [give candy to children] he did in libraries on weekends.

d. and [give candy to children in libraries on weekends] he did.

(19) a. and [introduce the children to each other] the teacher proceeded to do.
b. and [congratulate everybody on his birthday] he did.

c. John gives candy to children in libraries on weekends and Mary does in urban parks on federal holidays.

Internal to the fronted/elided portion of VP, both VPF and VPE pass tests for right-branching VP-structure (i.e. tests from Barss & Lasnik 1986, Larson 1988 etc.).

(19) a. and [introduce the children to each other] the teacher proceeded to do.
b. and [congratulate everybody on his birthday] he did.

c. John gives candy to children in libraries on weekends and Mary does in urban parks on federal holidays.

(21) a. *John gave books to them on each other's birthdays, and Mary did [on each other's first day of school].

b. *Mary congratulated every boy at his graduation, and Sue did [at his 21st birthday party].

Collective reading: the total time required to read the books was fast

Distributive reading: the time required to read each individual book was fast

b. *Mary read all the books quickly, and John did slowly.
Colin Phillips, Supper @ CUNY, 3/27/01

5

• Explanation: ellipsis must be licensed in second conjunct after adverbial has been added in first conjunct. Creation of right-branching structure in first conjunct destroys the V+NP constituent, which therefore can no longer license ellipsis.

(25)
a. all the books quickly Adv V' VP V
b. all the books quickly Adv V' VP V

Similar effect in Japanese

(26)
a. John-ga isoide dono hon-mo yonda. quickly all books-acc read'John read all the books quickly read, (and) Mary-top slowly did-so'John read all the books quickly, and Mary did slowly.' (both scopes impossible)

Control: Comparative Ellipsis (Heim 1985, Diesing 1992) allows more minimal contrast than VPF/VPE. Adverbial phrase may be present in either or both conjuncts.

(28)
a. John read as many books as Bill did on a Monday.
b. John read as many books as Bill did on a Tuesday.

Prediction: diagnostics of right-branching VPs will succeed when non-elided adverbial is present in just one conjunct (29a-b), and fail when it is present in both conjuncts (29c).

(29)
a. John read as many books as Bill did on a Monday.
b. John read as many books as Bill did on a Tuesday.
c. John read as many books as Bill did on a Tuesday.

Note: the basic loss-of-scope-in-ellipsis facts are not the evidence for incremental structure building: rather it is the contrast between the ellipsis and fronting facts, and the parallel contrast in comparative ellipsis examples.

E. Multiple Dependencies

(32)
a. *Which book did the senator deny [NP the rumor that he wanted __]
b. Which senator __ denied [NP the rumor that he wanted __]
c. Which senator __ denied [NP the rumor that he wanted which book]

• What is surprising about this paradigm is the fact that the short-WD-dependent (right)

where sentence — subject [NP the rumor that he wanted to know which book]

E. Multiple Dependencies (Richards, 1999)

E. Multiple Dependencies

(33)
E. Multiple Dependencies

(34)
a. *Who __ wonders what who bought __
b. John wonders what who bought __
Colin Phillips, Supper @ CUNY, 3/27/01

What is surprising about (a) is that the embedded what can cross over the embedded who. This is impossible in (b).

Richards: this is only possible when there is a well-formed matrix wh-dependency in a higher clause. This implies that ... of the matrix CP precedes the construction of the embedded CP, as is expected in a left-to-right / top-down derivation.

Additional arguments based on expletive-associate dependencies (e.g. There seems to be a man in the room etc.)

Kempson, Meyer-Viol & Gabbay (2001): argue that left-to-right derivations offer an improved treatment of various kinds of anaphoric dependencies, including relativization, weak crossover.


c. Grammaticality Judgment Processes

still largely unexplored

d. Parsing

Main concern is parsing of unambiguous sequences:Quite varied assumptions made in this area:

IF: Parser is a perfect covering grammar
† it's hard to find parsing evidence to distinguish this from the single system approach

IF: Parser is a poor covering grammar, as evidenced by the errors that it makes (older version: Fodor, Bever & Garrett, 1974; newer versions: Townsend & Bever, 2001; Ferreira, Bailey & Ferraro, 2001)
‡ evidence of this kind is more troubling

(i) Incremental Analysis of Parasitic Gaps

Phillips & Wong, 2000)

Problem: parasitic gaps in complex subject NPs pose a look-ahead problem — forming a wh-dependency involves crossing a syntactic island (condition on Extraction Domains: Huang 1982), but the island violation may be rescued at a later point in the sentence.

Fully incremental approach to parsing parasitic gaps (cf. Schneider 1999): form dependencies across islands that may involve a parasitic gap, but do not mark the dependency as completed. This approach to parsing parasitic gaps (cf. Schneider 1999) form dependency across islands that may involve a parasitic gap, but do not mark the dependency as completed. This approach to parsing parasitic gaps (cf. Schneider 1999) form dependency across islands that may involve a parasitic gap, but do not mark the dependency as completed. This approach to parsing parasitic gaps (cf. Schneider 1999) form dependency across islands that may involve a parasitic gap, but do not mark the dependency as completed. This approach to parsing parasitic gaps (cf. Schneider 1999) form dependency across islands that may involve a parasitic gap, but do not mark the dependency as completed. This approach to parsing parasitic gaps (cf. Schneider 1999) form dependency across islands that may involve a parasitic gap, but do not mark the dependency as completed. This approach to parsing parasitic gaps (cf. Schneider 1999) form dependency across islands that may involve a parasitic gap, but do not mark the dependency as completed. This approach to parsing parasitic gaps (cf. Schneider 1999) form dependency across islands that may involve a parasitic gap, but do not mark the dependency as completed. This approach to parsing parasitic gaps (cf. Schneider 1999) form dependency across islands that may involve a parasitic gap, but do not mark the dependency as completed. This approach to parsing parasitic gaps (cf. Schneider 1999) form dependency across islands that may involve a parasitic gap, but do not mark the dependency as completed. This approach to parsing parasitic gaps (cf. Schneider 1999) form dependency across islands that may involve a parasitic gap, but do not mark the dependency as completed. This approach to parsing parasitic gaps (cf. Schneider 1999) form dependency across islands that may involve a parasitic gap, but do not mark the dependency as completed. This approach to parsing parasitic gaps (cf. Schneider 1999) form dependency across islands that may involve a parasitic gap, but do not mark the dependency as completed. This approach to parsing parasitic gaps (cf. Schneider 1999) form dependency across islands that may involve a parasitic gap, but do not mark the dependency as completed. This approach to parsing parasitic gaps (cf. Schneider 1999) form dependency across islands that may involve a parasitic gap, but do not mark the dependency as completed. This approach to parsing parasitic gaps (cf. Schneider 1999) form dependency across islands that may involve a parasitic gap, but do not mark the dependency as completed. This approach to parsing parasitic gaps (cf. Schneider 1999) form dependency across islands that may involve a parasitic gap, but do not mark the dependency as completed. This approach to parsing parasitic gaps (cf. Schneider 1999) form dependency across islands that may involve a parasitic gap, but do not mark the dependency as completed. This approach to parsing parasitic gaps (cf. Schneider 1999) form dependency across islands that may involve a parasitic gap, but do not mark the dependency as completed. This approach to parsing parasitic gaps (cf. Schneider 1999) form dependency across islands that may involve a parasitic gap, but do not mark the dependency as completed. This approach to parsing parasitic gaps (cf. Schneider 1999) form dependency across islands that may involve a parasitic gap, but do not mark the dependency as completed. This approach to parsing parasitic gaps (cf. Schneider 1999) form dependency across islands that may involve a parasitic gap, but do not mark the dependency as completed. This approach to parsing parasitic gaps (cf. Schneider 1999) form dependency across islands that may involve a parasitic gap, but do not mark the dependency as completed. This approach to parsing parasitic gaps (cf. Schneider 1999) form dependency across islands that may involve a parasitic gap, but do not mark the dependency as completed. This approach to parsing parasitic gaps (cf. Schneider 1999) form dependency across islands that may involve a parasitic gap, but do not mark the dependency as completed. This approach to parsing parasitic gaps (cf. Schneider 1999) form dependency across islands that may involve a parasitic gap, but do
(41) No! a. Freedman & Forster 1985 [sentence-matching] NP w/ possessor
   b. Neville et al. 1991 [ERP violation paradigm] NP w/ possessor
   c. Stevenson 1993 [comprehension, matching] complex NP
   e. Studies with marginal results …

- The kinds of island contexts examined in the different studies appear to be fairly good predictors of whether or not the grammatical category (e.g. subject, adjunct, specific NP) will be licensed in the sentence.
- (ii) This is the kind of environment that supports parasitic gaps!

Preliminary: are parasitic gaps acceptable to anybody other than linguists? Yes!

- Grammaticality Judgment Task

RESULTS: sentences with both a parasitic gap and an object gap were rated exactly the same as sentence containing only the object gap (P-gap filled by an overt NP) when the P–gap was in an infinitival clause. P–gap without O-gap rated much worse; combination of P-gap and O-gap rated lower when the P–gap was in an infinitival clause.

SELF-PACED READING STUDY

- Question: will subjects show evidence of actively creating a wh-dependency inside an island in exactly the environments where P-gaps will be licensed?

Materials: 24 sets of 4 conditions, 96 filler sentences; comprehension question after all trials

Note: in the experiment subjects never saw a parasitic gap – they merely saw environments where P-gaps could potentially occur.

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Condition</th>
<th>Reading Time (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. The outspoken environmentalist worked to investigate what the local campaign to preserve the important habitats had actually harmed in the area that the birds once used as a place for resting while flying south. [infinitive, gap]</td>
<td>Finite + Gap</td>
<td>1234</td>
</tr>
<tr>
<td>b. …whether the local campaign to preserve the area that the birds once used as a place for resting while flying south. [finite, no gap]</td>
<td>Finite + No Gap</td>
<td>8765</td>
</tr>
</tbody>
</table>

Critical verb immediately preceding parasitic gap site was always obligatorily transitive:

cheapen, embarrass, remove, nominate, alter, order, aid, preserve, entertain, heal, secure, destroy, rehabilitate, reclaim, liberate, remove, upset, impress, uplift, convict, instruct, astonish, develop, overprice

RESULTS

Comprehension accuracy: 87.8% on comprehension questions (~33 word sentences!)
Analyses exclude trials on which comprehension question was answered incorrectly or was not answered.

Figure 1: Mean Grammaticality Ratings for PG-related sentences

RESULTS: sentences with both a parasitic gap and an object gap were rated exactly the same as sentence containing only the object gap (P-gap filled by an overt NP) when the P–gap was in an infinitival clause. P–gap without O-gap rated much worse; combination of P-gap and O-gap rated lower when the P–gap was in an infinitival clause.

Figure 2: Residual Reading Times for PG Self-paced Reading Experiment
At verb inside complex subject (region 12): Infinitive conditions: significant effect of gap/no-gap: (F1(1,45) = 4.02, p < 0.05; F2(1,35) = 5.01, p < 0.05)

Finite conditions: no difference between gap/no-gap conditions.

- Slowdown in reading times at verb reflects wh-dependency formation, but this only occurs in the infinitive conditions, where a P-gap is possible.
- Wh-dependencies are formed across islands – but in exactly the ways that the grammar allows.

**CONCLUSION:** Structural analysis is fully incremental, and quite sophisticated!

Caveat: this experiment found slowdown due to wh-dependency formation at the verb where the wh-dependency is created, rather than when it needs to be retracted at the following NP. Although this is not a 'filled-gap' effect per se, it is consistent with other results in the literature (e.g. Pickering et al. 1994).

(ii) Monotonic Structure Building (Schneider & Phillips, 2001)

QUESTIONS: (i) how is the grammar searched for grammatical parses; (ii) is reanalysis a last resort? (Fodor & Frazier 1980; cf. Frazier 1990, Frazier & Clifton 1998) or are reanalyses part of the parser's initial search space (Gibson 1991; Stevenson 1998; Fodor & Inoue 1998)? In other words, does the parser attempt to keep to the same derivation?

- Evidence: contrast between strong evidence for locality and strong evidence against.


(a) John said Bill left yesterday.

(b) During his vacations, Wallace likes to swim very much.

(c) John saw the daughter of the lawyer who he had met at the party in the mall.

Although Erica hated the house her father had owned...

Wallace knew the cheese was being stolen from under his nose.

...While Wallace was eating the cheese was being stolen from under his nose.

Wallace knew the cheese was being stolen from under his nose.

Although Erica hated the house her father had owned...

Wallace knew the cheese was being stolen from under his nose.

...While Wallace was eating the cheese was being stolen from under his nose.

Wallace knew the cheese was being stolen from under his nose.

Although Erica hated the house her father had owned...
Colin Phillips, Supper @ CUNY, 3/27/01

PREDICTION:

A combination of locality and easy reanalysis predicts embedded attachment to be easy; reanalysis as a last resort predicts matrix attachment bias. (49)

NP

S

The surprised woman

VP

locked the front door

S'

who

V

discovered

S'

S

himself

NP

S

The surprised woman

VP

locked the front door

S'

who

discovered

S

herself

the drunk man

Low Attachment   High Attachment

Philips: 56 undergraduate students. (7 additional subjects excluded for scoring <75% on comprehension questions in experimental items or <80% on filler items.)

Stimuli:

2 x 2 x 2 design: attachment site (high, low), ambiguity, verb class (strongly NP-biased, neutral or weakly NP-biased). 48 sets of 4 items. Verb classes based on Trueswell et al. (1993), Garnsey (unpublished completion norms)

Strongly NP-biased verbs:

hear, discover, acknowledge, appreciate, warn, understand

Weakly NP-biased verbs:

know, mention, doubt, notice

Disambiguation with himself, herself; gender counterbalanced across high/low sites. All possible antecedents of reflexives are animate NPs. 100 filler items.

High-attachment sentences: low ambiguous – 80%; high ambiguous – 90%

Comprehension question accuracy: low ambiguous – 80%; high ambiguous – 90%

Some examples:

- Embedded, ambiguous:
  The bilingual man who knows the well-traveled woman translated the travel books for himself without any extra help. Was quite impressed with the result.

- Embedded, unambiguous:
  The bilingual man who knows that the well-traveled woman translated the travel books for himself without any extra help. Was quite impressed with the result.

- Matrix, ambiguous:
  The bilingual man who knows the well-traveled woman translated the travel books for himself from a dictionary.

- Matrix, unambiguous:
  The bilingual man who knows her translated the travel books for himself from a dictionary.

Experiments with non-linking verbs were not counterevaluated across high/low sites. All possible antecedents of reflexives are animate NPs. 100 filler items.

Experiment 2: manipulation of statistical verb complement bias

NP-biased verbs,

hear, discover, acknowledge, appreciate, warn, understand

S-complement biased verbs,

claim, believe, suspect

Subjects: 56 English-speaking, 7 attachment species solicited for sentence (75% or more)

High Attachment

S

VP

The surprised woman

locked the front door

Low Attachment

S

VP

The surprised woman

locked the front door

PREDICTION: A combination of locality and easy reanalysis predicts embedded attachment.


Results: NP-complement bias verbs replicate results of Experiment 1. Pattern of difficulty is reversed in S-complement bias verbs.

Conclusion: Structure building is monotonic, i.e. redundancies is a last resort. Syntactic search pursues a simple algorithm – avoid reanalysis – but this is not necessarily the most memory-efficient approach.

Distinguishing Access and Search in Dative Alternations

Question: Can the parser be fooled into creating bogus ambiguities with dative verbs? Much recent parsing work assumes to assume ambiguous with dative verbs. Can the parser be fooled into creating bogus ambiguities with dative verbs? (Kemp, 2000)

(iii) Distinguishing Access and Search in Dative Alternations (Phillips, Edgar & Kabak, 2000)

Other pairs: bring/transport, assign/distribute, buy/obtain, hand/deliver, mail/address, make/create, offer/suggest, show/display...
Control 1: compare reading times to unambiguous conditions with complementizer.
Control 2: add conditions with monotransitive verbs, to gain independent measure of the difficulty of analyzing unambiguous V NP NP sequences.

Materials: 36 sets of 6 conditions, 84 fillers. Fillers: more closely matched in length and complexity than the materials but more closely matched in the number of words, more closely matched in complexity than the materials but not at all in the number of words.

RESULTS: alternator and non-alternator verbs clearly caused in the experiment, but...
Lexical access is not enough to parse successfully – grammatical search must also succeed.

The experiment (unwittingly) created an environment which provided a good deal of support for parsing of V NP NP sequences.

Support for this analysis: separate analysis of first half and second half of experiment shows that reading time profiles for alternator and monotransitive verb classes are stable across the experiment; for dative verbs, however, there is substantial change over the course of the experiment – the slowdown at the relative clause and in the second half of the experiment is most similar to the monotransitive verbs.

More support for this analysis: we ran an additional experiment, which was identical to the previous one, except that we removed the items that may have been helping accurate parsing of V NP NP sequences... containing relative clauses were replaced with sentences lacking relative clauses. Result: practice effects disappeared.

CONCLUSION: grammatical search is a resource-limited process (perhaps in contrast to lexical access); full and immediate access to grammatical information associated with lexical heads does not guarantee successful parsing.

(i) ...but what about real ambiguities? These are a less central concern, since there's not the clear redundancy between the parser-implementation and the grammar. Of course, try to show that notions of 'structural simplicity' govern preferences in selecting between alternative interpretations (ii) rigid constraints in selecting between alternative representations of a single interpretation (i.e. structural economy conditions)

(e) Learning

Consequences largely unexplored

(g) Challenges

• What if the parser does behave stupidly?

• What about head-final languages?

• Need further development of the implemented model

• How could this be implemented in the brain?

References


