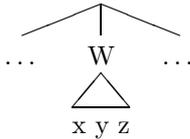


Constituency

Doug Arnold
doug@essex.ac.uk

Suppose we have a string $\dots xyz \dots$, how can we establish whether xyz is a constituent (i.e. syntactic unit); i.e. whether the representation of $\dots xyz \dots$ should be:

$\dots [w \ x \ y \ z] \dots$



1 Diagnostics

- Movement
- Substitution
- Ellipsis
- Intrusion
- Coordination
- Theory Internal:
 - Semantics
 - Subcategorization
 - Internal Structure
- Parsing
- etc.

1.1 Movement

Constituents can be moved around (constituents are strings of words or categories that may appear in different environments) – “we see xyz in a number of different positions... , hence...”;

- (1) a. This is a very common example.
b. *A very common example* though this may be...
c. **Common example* though this may be a very

But **beware**:

- (2) **Very common example* though this may be a...

Very common example may still be a constituent: there may be other reasons why it cannot move.

1.2 Substitution

One constituent can often be substituted for another – “We can substitute B, which is a constituent, for xyz , hence...”;

- (3) a. Kim spoke to a very nice student.
b. Kim spoke to what/who?
- (4) a. Kim spoke to this very nice student.

- b. Sam spoke to that *one*.

Beware:

- (5) a. *Paul will* sleep in the garage.
- b. *Tramps* sleep in the garage.

does not show that *Paul will* is a constituent.

We look for (a) substitution in a number of environments, and (b) substitution by something ‘similar’ (e.g. a pro-form).

1.3 Ellipsis

Constituents can be omitted – “*xyz* can be omitted in sentences like . . . , hence . . .”;

- (6) a. This is a very common example.
- b. That may be Δ too.
- (7) a. Sam ate some fish.
- b. Sam ate Δ
- (8) a. Sam likes the blue Chinese vase with flowers more than the green Δ .

1.4 Intrusion

Constituents resist intrusion – “. . . *xyz*. . . and *xyz* are ungrammatical, but . . . *wxyz*. . . (etc.) is grammatical, hence . . .”;

- (9) a. This sentence provides a very good example.
(I suppose) (and why not?)

1.5 Coordination

Constituents can be coordinated: “*xyz and uvw* and *xyz or uvw* are grammatical, hence . . .”;

- (10) a. Sam ran *on the grass* and *on the gravel*.
- b. *Sam rang *up his mother* and *up his sister*.

Beware: ‘Right-Node-Raising’

- (11) a. Sam will Δ and Kim might Δ *go to Paris*.
- b. Sam admired this Δ and Kim wanted to buy Δ *that book*.
- c. A very tall Δ and very stupid Δ *person*

1.6 Theory Internal

1.6.1 Semantics

Constituents are interpreted as units – “*xyz* is interpreted as a unit, hence . . .”;

- (12) a. I’ll tell Sam [that you saw her *yesterday*].
- b. I told Sam [that you will see her] *yesterday*.

1.6.2 Subcategorization

Heads subcategorize only sisters – “ x subcategorizes y and z , hence...”;

- (13) a. Sam is *fond of the job*.
 b. Sam is *happy about the job*.

Structural Parallels

“ xyz has the internal structure of XP, which suggests it is an XP, hence a constituent.”

- (14) a. [_S [_{NP} Sam] [_{VP} likes Kim]].
 b. I believe [_S [_{NP} Sam] [_{VP} likes Kim]].
 c. I expect [_S [_{NP} Sam] [_{VP} to like Kim]].

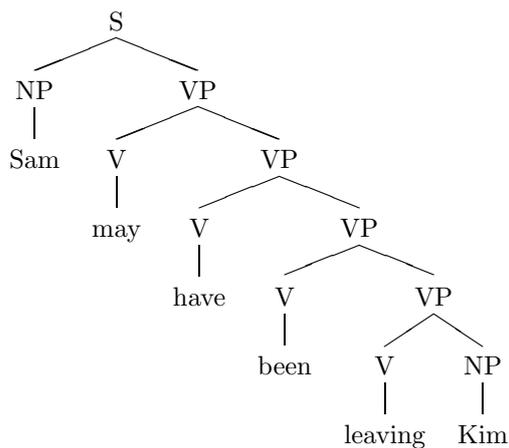
1.6.3 Capturing Generalizations...

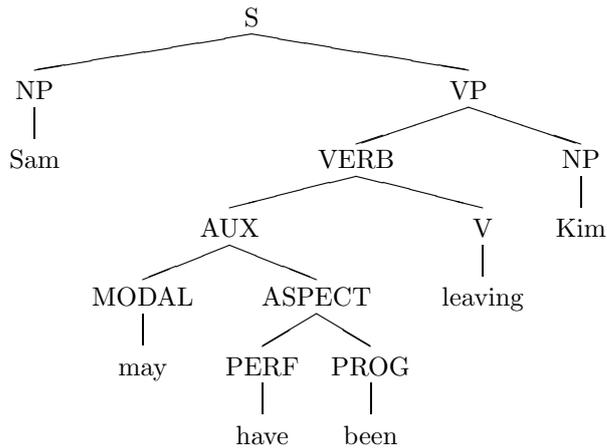
1.7 Parsing

The constituent structure determines how the parsing process divides into subtasks. e.g.

- (15) a. NP → DETP \bar{N}
 (16) a. VP_{<>} → V
 b. VP_{<XP,...>} → VP_{<...>} XP
 (17) a. VP_{<>} → V
 b. VP_{<NP>} → V NP
 c. VP_{<NP,PP>} → V NP PP etc
 (18) VP_{<...>} → V ...

2 Example: Aux, Modals, Main verbs



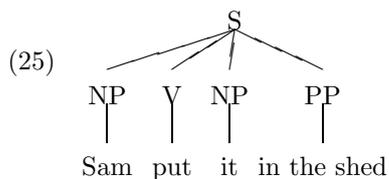


- (19) a. A: What might she have been doing?
 b. B: Leaving Kim.
 c. B: Been leaving Kim.
 d. B: Have been leaving Kim.
- (20) a. Sam may have been leaving Kim...
 b. or leaving Sandy.
 c. or been leaving Sandy.
 d. or have been leaving Sandy.
- (21) a. Sam — may — have — been — leaving Sandy.
 b. — for example —
- (22) a. Bev thinks Sam may have been leaving Kim...
 b. and so he may.
 c. and so he may have.
 d. and so he may have been.
- (23) a. A: Do you think Sam may have been leaving Kim?
 b. B: Yes, she may have been Δ (leaving Kim)
 c. B: Yes, she may have Δ (been leaving Kim)
 d. B: Yes, she may Δ (have been leaving Kim)

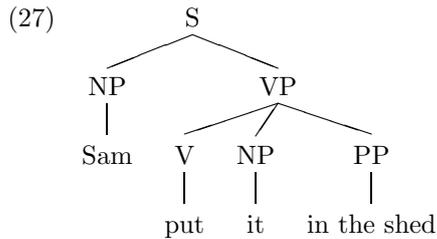
3 Example: VP in English

“Is there a VP in English?” or “What is the structure of S?”

- (24) a. Sam put the car in the shed.
 b. Sam saw Kim.
 c. etc.



- (26) $S \rightarrow NP V NP \dots$



- (28) a. $S \rightarrow NP VP \dots$
 b. $VP \rightarrow V NP \dots$

3.1 Movement

I thought Sam might *put the car in the shed* and ...

- (29) a. put the car in the shed he did.
 b. *put the car he did in the shed.

3.2 Ellipsis

- (30) a. and (so) he did Δ
 b. and (so) he did Δ in the shed.

3.3 Coordination

- (31) a. Sam put the car in the shed and left.
 b. Sam parked the car and put the bike in the shed.
 c. *Sam parked the car and put the bike.
 d. *Sam parked the car and put.

3.4 Substitution

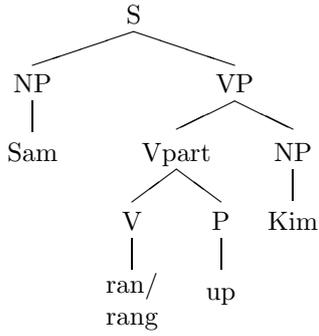
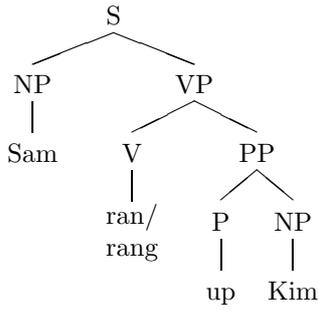
- (32) a. Sam will put the car in the shed.
 b. Sam will *what*?
 c. *Sam will *what* in the shed?

3.5 Subcategorization

- (33) a. put the car in/on/under the shed.
 b. put the car there.
 c. *put under the car.
 d. *put the car.
 e. *put the car the shed.
 f. *put the car from the shed.

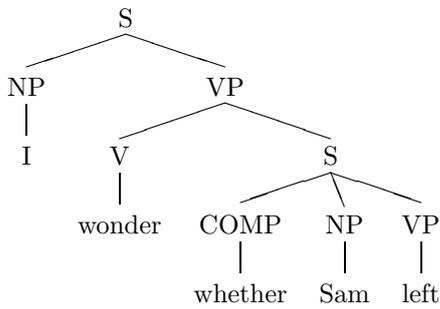
4 Example: prepositional vs particle verbs

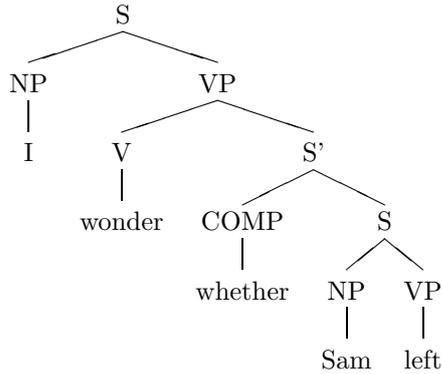
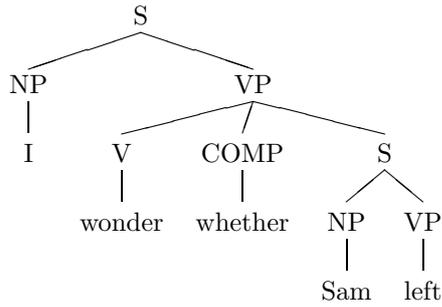
- (34) a. Sam ran up Everest.
 b. Sam rang up Kim.



- (35) a. *Sam rang up Kim and up Sandy.
 b. *Up Kim Sam rang.
 c. Sam + rang * up + Kim.
 (and why not)
 d. Kim was rung up by someone strange.
 e. Sam rang Kim up.
- (36) a. Sam ran up Everest and up K2.
 b. Up Everest Sam ran.
 c. Sam + ran + up ? Everest
 (and why not)
 d. ?Everest was run up by someone strange.
 e. *Sam ran Everest up.

5 Example: Complementizers





- (37) a. Whether Sam left, I sometimes wonder.
 b. Sam left, I sometimes wonder whether.
- (38) a. I wonder whether Sam left or whether Sam stayed.
 b. I wonder whether Sam left or Sam stayed.
- (39) a. You wonder whether *what*?
 b. You wonder *what*?
- (40) a. I wonder whether Δ
 b. I wonder Δ

6 Practical Application

Practically, this matters:

1. because getting the constituency right simplifies the description (grammar);
2. constituency determines the structure of the parsing process into subtasks.

6.1 Movement

- (41) a. and *leave* [_S he did].
 b. and *see Kim* [_S he did].
 c. and *put the car in the shed* [_S he did].

No VP:

- (42) a. $S \rightarrow V S$
 b. $S \rightarrow V NP S$
 c. $S \rightarrow V NP PP S$ and similarly for other VP expansions....

With VP:

- (43) a. $S \rightarrow VP S$

6.2 Ellipsis

- (44) a. (and (so)) he did Δ

No VP:

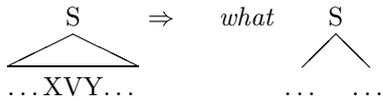
- (45) a. $S \rightarrow NP V_{AUX} (V)$
b. $S \rightarrow NP V_{AUX} (V) (NP)$
c. $S \rightarrow NP V_{AUX} (V) (NP) (PP)$ etc.

With VP:

- (46) a. $S \rightarrow NP (VP)$

6.3 Substitution

No VP: No PS account possible:

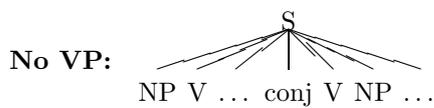


With VP:

- (47) a. $VP \rightarrow \text{what}$

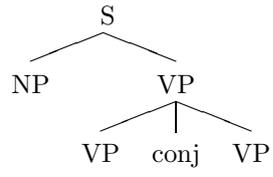
6.4 Coordination

- (48) a. Sam arrived and saw Kim.
b. Sam arrived and and put the car in the shed.
c. Sam saw Kim and put the car in the shed.



- (49) a. $S \rightarrow NP V \text{ conj } V$
b. $S \rightarrow NP V \text{ conj } V NP$
c. $S \rightarrow NP V \text{ conj } V NP PP$
d. $S \rightarrow NP V NP \text{ conj } V$
e. $S \rightarrow NP V NP \text{ conj } V NP$
f. $S \rightarrow NP V NP \text{ conj } V NP PP$
g. $S \rightarrow NP V NP PP \text{ conj } V$
h. $S \rightarrow NP V NP PP \text{ conj } V NP$
i. $S \rightarrow NP V NP PP \text{ conj } V NP PP$

With VP:

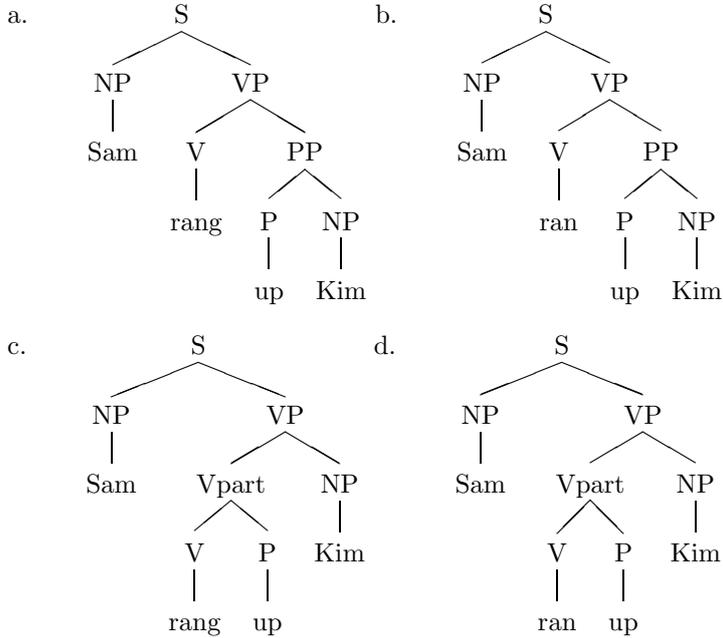


- (50) a. $S \rightarrow NP VP$
b. $VP \rightarrow VP \text{ and } VP$

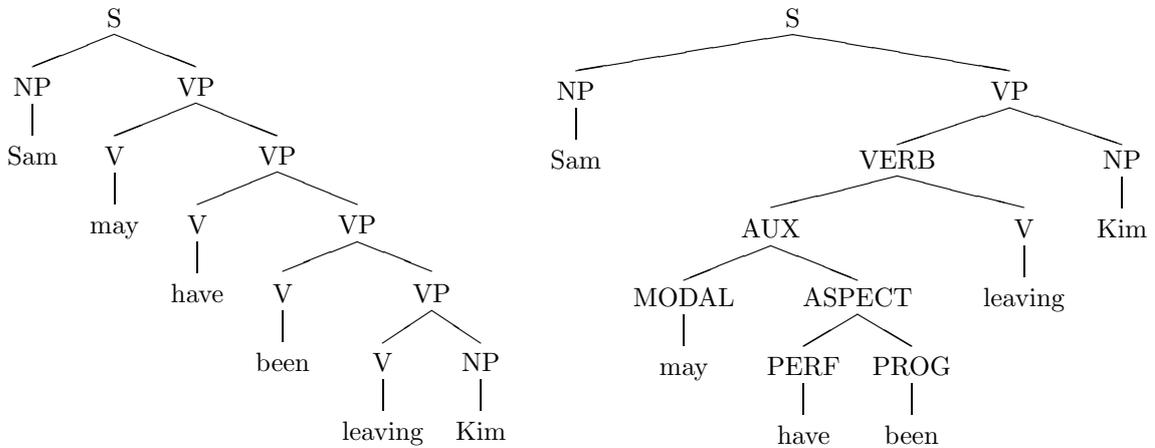
7 Appendix: Trees

For easier comparison, here are some of the trees printed together.

Particle vs Prepositional Verbs



Auxiliary Verbs



Complementizers

