Linguistics 819: Seminar on TAG and CCG

Alexander Williams, 9 April 2008

Introduction to Combinatory Categorial Grammar

Steedman and Dowty

1 The biggest picture

- Grammatical operations apply only to constituents.

  This goal is often forsaken, particularly in accounts of coordination that involve movement or deletion of nonconstituents.

2 Principles of the categories

- Lexicalism

(1) The Principle of Lexical Head Government
Both bounded and unbounded syntactic dependencies are specified by the lexical syntactic type of their head.

- But unbounded dependencies are not projected from special lexical entries, as in TAG.

(2) The Principle [*sic*; really, *Policy*] of Head Categorial Uniqueness
A single nondisjunctive lexical category for the head of a given construction specifies both the bounded and the unbounded dependencies that arise when those complements are displaced under relativization, coordination, and the like.

- Syntactic and semantic types go together

(3) The Principle of Categorial Type Transparency
For a given language, the semantic type of the interpretation together with a number of language-specific directional parameter settings uniquely determines the syntactic category of a category.

(4) The inverse of Type Transparency
For any category, the semantic type is a function of the syntactic type.

- "CCG categories are therefore reminiscent of the lexical items of [everyone you know]. CCG differs only in the way in which these predicate-argument relations are 'projected' by the combinatory rules of syntax." (38)

3 Principles of combination

- Semantically, combination adds no substantive content. Same goes for 'unary rules,' like Type Raising. Only application, abstraction, and place-holder variables to abstract over.

(5) The Principle of Combinatory Type Transparency
All syntactic combinatory rules are type-transparent versions of one of a small number simple semantic operations over functions.

a. Application: $X/Y : f \;\; Y : a \;\; \Rightarrow \;\; X : fa$

b. Composition: $X/Y : f \;\; Y/Z : g \;\; \Rightarrow_{\mathrm{B}} \;\; X/Z : \lambda x.f(gx)$

c. Type Raising: $X : x \;\; \Rightarrow_{\mathrm{T}} \;\; T\backslash(T/X) : \lambda f.fx$

d. $\vdots$

- CCG is strictly concatenative

(6) The Principle of Adjacency
Combinatory rules may only apply to finitely many phonologically realized and string-adjacent entitites.

- Respect the slash!

(7) The Principle of Consistency
All syntactic combinatory rules must be consistent with the directionality of the principal function [i.e. the function among the inputs that determines the *range* of the output]

(8) $*\, X\backslash Y \;\; Y \;\; \Rightarrow \;\; X$

- Pass the slash!

If its $\ldots/Z$ in the input, and Z is in the output,

then its $\ldots/Z$ in the output. Likewise for $\ldots\backslash Z$.

(9) The Principle of Inheritance
If the category that results from the application of a combinatory rule is a function category, then the slash defining directionality for a given argument in that category will be **the same** as the one(s) defining directionality for the corresponding argument(s) in the input function(s).

(10) $*\, X/Y \;\; Y/Z \;\; \Rightarrow X\backslash Z$

- Pursuant to these principles, we have only FOUR possible instances of Composition, for instance. (You'll see the others below.)

3.1 Combinators

- We can define the semantic effect of the combinatory operations without variables and lambdas, by means of **combinators**: operations on functions.

(11) a. Composition: B
$((\mathrm{B}f)g)x \equiv f(g(x))$, thus: $(\mathrm{B}f)g \equiv \lambda x.f(gx)$

b. Type Lifting: T
$\mathrm{T}xf \equiv fx$, thus: $\mathrm{T}x \equiv \lambda f.fx$

3.2 More on Composition

- Besides Forward Composition, (12), we have the (also Lambek-provable) case of Backwards (Harmonic) Composition, (13).

(12) $X/Y \;\; Y/Z \;\; \Rightarrow_{\mathrm{B}} \;\; X/Z$

(13) $Y\backslash Z \;\; X\backslash Y \;\; \Rightarrow_{\mathrm{B}} \;\; X\backslash Z$

In a minute we'll see the role this plays in coordination.

- As we'll see later, Steedman will ban **Forward Crossing Composition** from English, (14), in order to rule out subject-extraction, (15)

(14) $X/Y \;\; Y\backslash Z \;\; \Rightarrow_{\mathrm{B}} \;\; X\backslash Z$

(15) * Who do you think that likes CCG?

But he will allow *limited* uses of **Backward Crossing Composition**, (16), in order to account for licit cases of nonperipheral extraction and extraposition.

(16) $Y/Z \;\; X\backslash Z \;\; \Rightarrow_{\mathrm{B}} \;\; X\backslash Z$

(17) I will buy today (and sell tomorrow) every copy of *LGB* in the bookstore.

- Steedman introduces a generalization of Composition, which amounts to abstraction over n arguments, for an n bounded by grammatical facts (probably $n \leq 4$).

(18) $X/Y : f \qquad (Y/Z)/\$_1 : \ldots \lambda z.gz \ldots \qquad \Rightarrow_{\mathrm{B}^n} \quad (X/Z)/\$_1 : \lambda z.f(gz \ldots)$

3.3 More on Type Raising

- Type Raising is restricted to its **order-preserving** instances:

(19) a. $X \Rightarrow_{\mathrm{T}} T\backslash(T/X)$
b. $X \Rightarrow_{\mathrm{T}} T/(T\backslash X)$

- Besides this, it is also **only permitted** to produce functions over functions, T/X or $T\backslash X$, that are **otherwise attested** as **basic categories** in the language.

  Thus, because English has no basic SOV verbs, type $(S\backslash NP)\backslash NP$, the following instance of Type Raising is illegal:

(20) Illicit instance of Type Raising in English
$NP \Rightarrow_{\mathrm{T}} (S\backslash NP)/((S\backslash NP)\backslash NP)$

3.4 Substitution

- Curry's combinator S has the effect of sharing an argument x between two functions f and g, and then applying fx to gx.

(21) $((\mathrm{S}f)g)x \equiv fx(gx)$
Thus: $\mathrm{S}fg \equiv \lambda x fx(gx)$

  You can think about this as:

  1. Like Composition, just abstracting over an argument both in the lower and in the higher function.
  2. Exactly like Coordination of f and g, except that what combines f and g semantically is not $\wedge$ but function application.

- Anna Szabolcsi motivated the linguistic use of S for parasitic gaps:

(22) a. file without reading
b. S(without reading)(file)
c. $\lambda x.(\mathit{without}(\mathit{reading}))(x)[\mathit{file}(x)]$

- Notice that Substitution in this instance is backwards: the higher-order function, *without reading*, is to the right of the lower-order function, *file*.

  More importantly, it is **crossed**, or **disharmonic**: the higher order function f swallows a function g to its *left*, abstracting over an argument that g wants to its *right*.

  (23) Backward crossed Substitution
  $Y/Z \;\; (X\backslash Y)/Z \;\; \Rightarrow_{\mathrm{S}} \;\; X/Z$

  (24) I will file without reading.

	I will	file	without	reading	
0	S/VP	VP/NP	$(VP\backslash VP)/VPing$	$VPing/NP$	Fwds Comp
1	S/VP	VP/NP	$(VP\backslash VP)/NP$		Bwd Xed Sub
2	S/VP	VP/NP			Fwds Comp
3	S/NP				

- Crucially, Y in this formulation of the Substitution rule is further restricted to ranging only VP-like categories. This is necessary to rule out examples like (25).

  (25) * a nice with a girl = 'a nice girl with a girl'

	a	nice	with a	girl	
0	NP/N	N/N	$(N\backslash N)/N$	N	Bwd Xed Sub
1	NP/N	N/N		N	FA
2	NP/N	N			FA
3	NP				

  Compare: "There is no extraction of NP from DP." Any worse?

- Restricting the domain of crossed rules is necessary, to restrict generative capacity.

  1. Disharmonic composition can result in a kind of discontinuity. A function might find its argument *past* a higher-order function that swallows it.
  2. No disharmonic rules are provable in the Lambek calculus.
  3. And if disharmonic rules *are* let in to the Lambek calculus, the result is *permutation closure.*
  4. As we'll see, it is the inclusion of Crossed rules in CCG that boost its generative capacity past CF.

  So if the same permutation closure is to be avoided in CCG, the application of crossed rules will have to limited to applying only to certain categories, *finite in number.*

4 Coordination

4.1 The rule of coordination

- Rule of coordination: likes coordinate with likes

(26) Coordination ($\langle\Phi_n\rangle$)
$X : g \quad CONJ : b \quad X : f \quad \Rightarrow_{\Phi_n} \quad X : \lambda \ldots b(f \ldots)(g \ldots)$

Semantics:

Supply to the interpretations of the two coordinates, f and g, as many variables $\langle\ldots\rangle$ as you need to get an expression of the type for which the coordinating function b is defined; apply b to the two coordinate interpretations; and abstract over $\langle\ldots\rangle$.

(27) Examples:

a. $[(S\backslash NP) : \lambda x[Px]] \quad [and : \wedge] \quad [(S\backslash NP) : \lambda y[Qy]]$
$\Rightarrow_{\Phi} \quad (S\backslash NP) : \lambda z[\wedge(Pz)(Qz)]$

b. $((S\backslash NP)/NP)/NP : \lambda z\lambda y\lambda x[Pxyz] \quad [and : \wedge] \quad [((S\backslash NP)/NP)/NP : \lambda s\lambda r\lambda q[Qqrs]]$
$\Rightarrow_{\Phi} \quad ((S\backslash NP)/NP)/NP : \lambda c\lambda b\lambda a[\wedge(Pabc)(Qabc)]$

(28) The Φ combinator

a. $\Phi^0 bxy \equiv bxy$
b. $\Phi^1 bfg \equiv \lambda y.b(fx)(gy)$
c. $\Phi^2 bfg \equiv \lambda y\lambda x.b(fyx)(gyx)$
d. etc.

- Since all rules in CCG are strictly concatenative, it is only adjacent sequences that can combine in any way. Thus you cannot get semantic coordination of "met" with "might marry" in (29).

(29) Anna $[_{(S\backslash NP)/NP}$ met $]$ Manny, and $[_{(S\backslash NP)/NP}$ might marry $]$.

In contrast, a CG that allows operations like WRAP would have some difficulty ruling out something like (29) via an fully general principle.

A theory with movement and deletion will have similar challenges. (Why not delete "Manny" from the right conjunct? Why not do covert movement of "Manny" to license a gap in the right conjunct?)

4.2 Basic cases

- The "noncanonical constituents" created by Composition and Type Raising can be coordinated, resulting in "Right Node Raising"

(30) New Hampshire selected, but Maryland rejected, Junior Senator Hillary Rodham Clinton.

	NH	elected	but	Md	rejected	Clinton	
0	NP	$(S\backslash NP)/NP$	$CONJ$	NP	$(S\backslash NP)/NP$	NP	TR
1	$S/(S\backslash NP)$	$(S\backslash NP)/NP$	$CONJ$	$S/(S\backslash NP)$	$(S\backslash NP)/NP$	NP	Fwds Comp
2	S/NP		$CONJ$	S/NP		NP	Coord
3	S/NP					NP	FA
4	S						

4.3 Fancy cases: (Steedman 1985 and) Dowty 1988

- The central case of TR is of NPs lifting to GQs. Steedman, like Lambek, talks about this as analogous to Case.

  But we can Type Raise anything (within the bounds of the principles above), and this allows even more "noncanonical constituency."

(31) John eats rice quickly and beans slowly.

	John	eats	rice	quickly	and	beans	slowly	
0	S/VP	VP/NP	NP	$VP\backslash VP$	&	NP	$VP\backslash VP$	TR
1	S/VP	VP/NP	$VP\backslash(VP/NP)$	$VP\backslash VP$	&	$VP\backslash(VP/NP)$	$VP\backslash VP$	Bwd FC
2	S/VP	VP/NP	$VP\backslash(VP/NP)$		&	$VP\backslash(VP/NP)$		Coord
3	S/VP	VP/NP	$VP\backslash(VP/NP)$					FA
4	S/VP	VP						FA
5	S							

(32) John gave Mary a book and Susan a record.

	John	gave	Mary	a book	and	Susan	a record	
0	NP	$((S\backslash NP)/NP)/NP$	NP	NP	&	NP	NP	Abbrev
1	NP	$DiTrV$	NP	NP	&	NP	NP	TR
2	NP	$DiTrV$	$TrV\backslash DiTV$	$VP\backslash TrV$	&	$TrV\backslash DiTrV$	$VP\backslash TrV$	Unabbrev
3	NP	$VP/NP/NP$	$(VP/NP)\backslash((VP/NP)/NP)$	$VP\backslash(VP/NP)$	&	$TrV\backslash DiTrV$	$VP\backslash TrV$	Bwd FC
4	NP	$DiTrV$	$VP\backslash DiTV$		&	$VP\backslash DiTrV$		Coord
5	NP	$DiTrV$	$VP\backslash DiTV$					FA
6	NP	VP						Unabbrev
7	NP	$S\backslash NP$						FA
8	S							

- Dowty observes that this sort of "left-peripheral deletion" or "left-node raising" or "non-constituent coordation," is not readily assimilated to leftward head-movement.

(33) John gave Mary a book on Monday and a record on Friday.

	John	gave	Mary	a book	on Monday	and	
0	S/VP	$(VP/NP)/NP$	NP	NP	$VP\backslash VP$	&	FA
1	S/VP	VP/NP		NP	$VP\backslash VP$	&	TR
2	S/VP	VP/NP		$VP\backslash(VP/NP)$	$VP\backslash VP$	&	Bwd FC
3	S/VP	VP/NP		$VP\backslash(VP/NP)$		&	FA
4	S/VP	VP				&	FA
5	S					&	

- (34) is a particularly interesting. Here the coordinated sequences includes the object of relatively embedded PP, without its P.

  What is 'shared' by the coordinates, namely "went to," is not obviously a constituent itself (?*"Chicago was gone to by John.") So a theory that either *raises* or *deletes* "went to" from "Detroit on Tuesday" would be performing this operation on a *nonconstituent*.

(34) John went to Chicago on Monday and Detroit on Tuesday.

	John	went	to	Chicago	on Monday	
0	$S/(S\backslash NP)$	VP	$(VP\backslash VP)/NP$	NP	$VP\backslash VP$	TR
1	$S/(S\backslash NP)$	$VP/(VP\backslash VP)$	$(VP\backslash VP)/NP$	NP	$VP\backslash VP$	Fwd FC
2	$S/(S\backslash NP)$	VP/NP		NP	$VP\backslash VP$	TR
3	$S/(S\backslash NP)$	VP/NP		$VP\backslash(VP/NP)$	$VP\backslash VP$	Bwd FC
4	$S/(S\backslash NP)$	VP/NP		$VP\backslash(VP/NP)$		FA
5	$S/(S\backslash NP)$	VP				FA
6	S					

- Dowty also discusses analogous cases inside NP.
  (He also claims that this will go towards explaining the seemingly dissonant agreement in: "This man and woman were both making trouble."

(35) Every man and woman dies.

	every	man	and	woman	dies	
0	$((S\backslash NP)/NP)/N$	N	&	N	$S\backslash NP$	Abbrev
1	GQ/N	N	&	N	$S\backslash NP$	TR
2	GQ/N	$GQ\backslash(GQ/N)$	&	$GQ\backslash(GQ/N)$	$S\backslash NP$	Coord
3	GQ/N	$GQ\backslash(GQ/N)$			$S\backslash NP$	FA
4	GQ				$S\backslash NP$	FA
5	S					

- Notice that in all these derivations, the subject **did not** combine with the verb before the VP-internal coordination was completed. Type Lifting was over the verb (e.g. (S\NP)/NP), not the subject+verb (e.g. S/NP).

(36) Bill gave and Max sold a book to Mary and a record to Susan.

a. (Bill gave and Max sold) = $(S/PP)/NP$

b. (a book to Mary and a record to Susan) = $VP\backslash((VP/PP)/NP)$

c. Can't combine: $[(S/PP)/NP] + [VP\backslash((VP/PP)/NP)]$

1. Dowty's way out is to first lift the type of the verb:

   (37) a. $DatV \mapsto VP/(VP\backslash DatV)$

   b. That is:
   $(VP/PP)/NP \mapsto VP/(VP\backslash((VP/PP)/NP))$

   The subject, S/VP, can then compose into the verb, leaving:

   (38) Bill gave = $S/(VP\backslash((VP/PP)/NP))$

   This can coordinate with Max sold, and then combine with (36b) by FA.

2. An alternative (which seems to make more sense from a L-R processing perspective) is instead to assemble the subject and verb as usual, and then raise the Object over the *resulting type*:

   (39) a. Bill gave, Max Sold $\mapsto$ $(S/PP)/NP$

   b. a book, a record $\mapsto (S/PP)\backslash((S/PP)/NP)$

   c. to Mary, to Susan $\mapsto S\backslash(S/PP)$

   The object NP and PP can then compose by Backwards Harmonic Composition, yielding type: $S\backslash((S/PP)/NP)$. And this can apply to the Subject+Verb sequence by FA to yield a an S.

   Problem is, the Type Raising here seems to run afoul of Steedman's requirement that the domain-term in a category resulting from TR be an attested basic category of the language, which $(S/PP)/NP$ is not.

4.4 The use of Backwards Crossed Composition

- Displacement of nonperipheral constituents is permitted by Backwards Crossing Composition.

(40) I will buy today (and sell tomorrow) every copy of *LGB* in the bookstore.

(41) buy today $\mapsto (S\backslash NP)/NP \;\; (S\backslash NP)\backslash(S\backslash NP) \Rightarrow_B (S\backslash NP)/NP$

We'll see how this has to be restricted next time.