Continuation semantics i[1]

Patrick D. Elliott[2] *& Martin Hackl*[3]

February 7, 2020

[1] 24.979: Topics in semantics

Getting high: Scope, projection, and evaluation order

[2] `pdell@mit.edu`

[3] `hackl@mit.edu`

1 *A note on syntax*

So far, I've been a little shy about saying explicitly what we're assuming here about syntax, and what we're assuming about the syntax-semantics mapping.

I'll assume a derivational theory, according to which structures are built-up via successive application of Merge.[4]

(1) Merge
- ..
- Merge
  - ..
  - Merge
    - ...
    - ...

[4] I'll often use "syntactic structure speak" when talking about trees. This is harmless, since they can always be interpreted as the graph of a syntactic derivation, especially since trees encode both structure and order.

I'll furthermore adopt the hypothesis that the syntactic derivation proceeds in lockstep with the semantic computation. This conjecture, which goes back to MONTAGUE is often described as *direct compositionality*.[5]

cite montague

[5] Although direct compositionality is often strongly associated with Combinatory Categorial Grammar, it's in principle independent. See, e.g., Kobele (2006) for an explicit formalization of a directly compositional minimalist grammar.

Minimally, the formatives must be *tuples* consisting of phonological features and semantic features: $(\texttt{phon}, \texttt{sem}, \ldots)$. Semantic features could be cashed out as model theoretic objects, or perhaps as expressions of the simply typed lambda calculus.

Heretically, I'll assume that part of what Merge does is concatenate phonological features. This is because Merge is just an instruction for combining formatives. On the semantic side, it typically does function application.

(2) $(\mathbb{x}, x) * (\mathbb{y}, y) := ([\mathbb{x}\ \mathbb{y}], x \mathsf{A}\, y)$

It can also do concatenation of phonological features, plus Scopal Function Application (sfa) of semantic values (whence the left-to-right bias of S).

(3) $(\mathbb{x}, x) * (\mathbb{y}, y) \coloneqq ([\mathbb{x}\ \mathbb{y}], x \mathsf{S}\ y)$

I've define Lift as a purely semantic operation – this is to be taken as shorthand for an operation on a formative that only effects the semantic value:

(4) $(\mathbb{x}, x)^{\uparrow} \coloneqq (\mathbb{x}, x^{\uparrow})$

This constitutes the basics of the system laid out in Elliott 2019.

1.1 Deriving inverse scope

Recall that LIFT is a *polymorphic function* – it lifts a value into a trivial tower:

(5) $a^{\uparrow} \coloneqq \dfrac{[]}{a}$

Since LIFT is polymorphic, in principle it can apply to any kind of value – even a tower! Let's flip back to lambda notation to see what happens.

(6) $[\![\text{everyone}]\!] \coloneqq \lambda k\ .\ \forall x[k\ x]$

(7) $[\![\text{everyone}]\!]^{\uparrow} = \lambda l\ .\ l\ (\lambda k\ .\ \forall x[k\ x])$

> Question
>
> What is the *type* of lifted *everyone*?

Going back to tower notation, lifting a tower adds a trivial third story:[6] Following Charlow (2014), when we apply LIFT to a tower, we'll describe the operation as *external lift* (although, it's worth bearing in mind that this is really just our original LIFT function).

[6] In fact, via successive application of LIFT, we can generate an n−story tower.

One important thing to note is that, when we externally lift a tower, the quantificational part of the meaning always remains on the same story relative to the bottom floor. Intuitively, this reflects the fact that, ultimately, LIFT alone isn't going to be enough to derive quantifier scope ambiguities.

(8) $\left(\dfrac{\forall x[\,]}{x}\right)^{\uparrow} = \dfrac{[\,]}{\dfrac{\forall x[\,]}{x}}$

Question

Which (if any) of the following bracketings make sense for a three-story tower:

(9) $\dfrac{\left(\dfrac{f\,[\,]}{g\,[\,]}\right)}{x}$

(10) $\dfrac{f\,[\,]}{\left(\dfrac{g\,[\,]}{x}\right)}$

The extra ingredient we're going to need, is the ability to sandwhich an empty story into the *middle* of our tower, pushing the quantificational part of the meaning to the very top. This is *internal lift* ($\uparrow\uparrow$).[7]

[7] I can tell what you're thinking: "seriously? Another *darn* type-shifter? How many of these are we going to need?!". Don't worry, I got you. Even thought we've defined internal lift here as a primitive operation, it actually just follows from our existing machinery. Concretely, *internal lift* is really just *lifted* LIFT (so many lifts!). Lifted LIFT applies to its argument via S.

(11) $(\mathsf{LIFT}^{\uparrow})\ \mathsf{S}\ \dfrac{f\,[\,]}{x} = \dfrac{f\,[\,]}{\dfrac{[\,]}{x}}$

(12) *Internal lift* (def.)

a. $(\uparrow\uparrow) : \mathsf{K_t\ a} \to \mathsf{K_t\ (K_t\ a)}$

b. $m^{\uparrow\uparrow} := \lambda k\,.\,m\,(\lambda x\,.\,k\,x^{\uparrow})$

It's much easier to see what internal lift is doing by using the tower notation. We can also handily compare its effects to those of *external* lift.

(13) *Internal lift* (tower ver.)

$$\left(\dfrac{f\,[\,]}{x}\right)^{\uparrow\uparrow} := \dfrac{f\,[\,]}{\dfrac{[\,]}{x}}$$

(14) *External lift* (tower ver.)

$$\left(\dfrac{f\,[\,]}{x}\right)^{\uparrow} := \dfrac{[\,]}{\dfrac{f\,[\,]}{x}}$$

Armed with *internal* and *external* lifting operations, we now have everything we need to derive inverse scope. We'll start with a simple example (15).

The trick is: we *internally* lift the quantifier that is destined to take wide scope.

(15) A boy danced with every girl. $\forall > \exists$

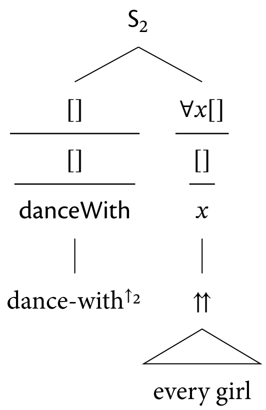
Before we proceed, we need to generalize LIFT and SFA to three-story towers.[8]

(16) $x^{\uparrow_2} := \dfrac{[]}{\dfrac{[]}{x}}$

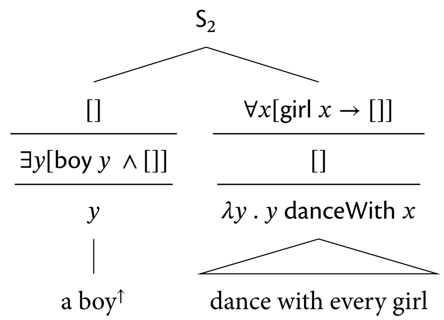
(17) $\dfrac{f\,[]}{m} \; \mathsf{S}_2 \; \dfrac{g\,[]}{n} := \dfrac{f\,(g\,[])}{m \; \mathsf{S} \; n}$

[8] Before you get worries about expanding our set of primitive operations, notice that *3-story lift* is just ordinary lift applied twice. *3-story* SFA is just SFA, but where the bottom story combines via S not A. In fact, we can generalize these operations to n−story towers.

(18) Step 1: internally lift *every girl*



(19) Step 2: *ex*ternally lift *a boy*



What we're left with now is a 3-story tower with the universal on the top story and the existential on the middle story. We can collapse the tower by first collapsing the bottom two stories, and then collapsing the result. In order to do this, we'll first define *internal lower*.[9]

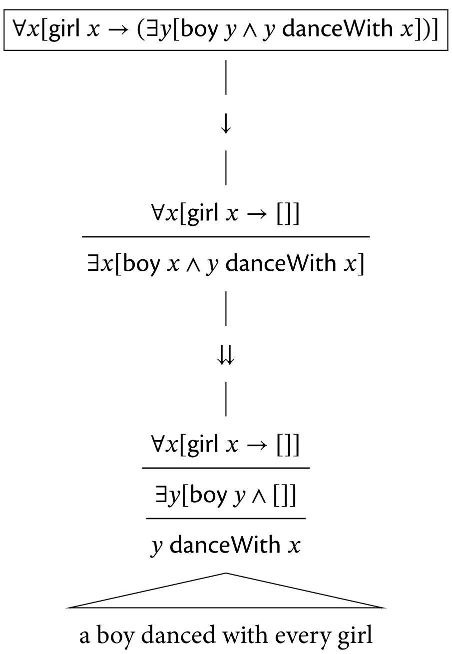
(21) *Internal lower* (def.) $\left(\dfrac{f\,[]}{\dfrac{g\,[]}{p}}\right)^{\downdownarrows} := \dfrac{f\,[]}{\left(\dfrac{g\,[]}{x}\right)^{\downarrow}}$

[9] Let's again address the issue of expanding our set of primitive operations (in what is becoming something of a theme). Internal lower is just lifted lower, applying via S. In other words:

(20) $m^{\downdownarrows} \equiv (\downarrow)^{\uparrow} \; \mathsf{S} \; m$

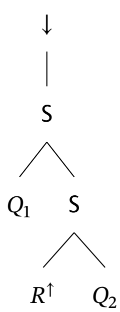
Now we can collapse the tower by doing *internal lower*, followed by *lower*:

(22)

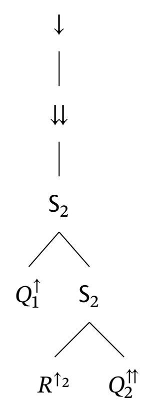


Great! We've shown how to achieve quantifier scope ambiguities using our new framework. Let's look at the derivations again side-by-side.

(23) Surface scope (schematic derivation)



(24) Inverse scope (schematic derivation)



There's a couple of interesting things to note here:

- The inverse scope derivation involves more applications of our type-shifting operations – this becomes especially clear if we decompose the complex operations S_2, $\uparrow_2$, $\uparrow\uparrow$, and $\downdownarrows$.
- In order to derive an inverse scope reading, what was *crucial* was the availability of *internal lift*; the remaining operations, S_2, $\uparrow_2$, $\downdownarrows$ only functioned to

massage composition for three-story towers.

On the latter point, it's tempting to conjecture that in, e.g., German, Japanese and other languages which "wear their LF on their sleeve", the semantic correlate of *scrambling* is *internal lift*, whereas in scope-flexible languages such as English, internal lift is a freely available operation.[10]

[10] To make sense of this, we would of course need to say something more concrete about the relationship between syntax and semantics. For an attempt at marrying continuations to a standard, minimalist syntactic component, see my manuscript *Movement as higher-order structure building*.

If we adopt some version of the *derivational complexity hypothesis*, we also predict that inverse scope readings should take longer to process than surface scope readings. This is something Martin may discuss is a couple of weeks time.

It's worth mentioning, incidentally, that although we collapsed the resulting three-story tower via internal lower followed by lower, we can also define an operation that collapses a three-story tower two an ordinary tower in a different way. Let's call it *join*:[11]

[11] Join for three-story towers corresponds directly to the *join* function associated with the continuation monad. For more on continuations from a categorical perspective, see the first appendix.

(25) *join* (def.)
$m^{\mu} \coloneqq \lambda k \,.\, m\,(\lambda c \,.\, c\, k)$ $\quad\quad \mu : \mathtt{K_t}\ (\mathtt{K_t}\ \mathtt{a}) \rightarrow \mathtt{K_t}\ \mathtt{a}$

In tower terms, join takes a three-story tower and sequences quantifiers from top to bottom:

(26) $$\left(\frac{f\,[]}{\dfrac{g\,[]}{x}}\right)^{\mu} = \frac{f\,(g\,[])}{x}$$

Doing internal lower on a three-story tower followed by lower is equivalent to doing join on a three-story tower followed by lower (as an exercise, convince yourself of this). However, there's may be a good empirical reason for having internal lower as a distinct operation (and since it's just lifted lower, it "comes for free" in a certain sense).

(27) Daniele wants a boy to dance with every girl. $\quad \forall > \mathsf{want} > \exists$

Arguably, (27) can be true if for every girl x, Daniele has the following desire: *a boy dances with y*. This is the reading on which *every boy* scopes over the intensional verb, and *a boy* scopes below it.

If we have *internal lower*, getting this is easy. We *internally lift every girl* and *externally lift a boy*. Before we reach the intensional verb, we fix the scope of *a* *boy* by doing internal lower. Now we have an ordinary tower, and we can defer fixing the scope of *every girl* via *lower* until after the intensional verb.

If we only have *join* then the scope of *a boy* and *every girl* may vary amongst themselves, but they should either both scope below *want* or both scope above *want*.

2 Split scope

In the first p-set, I asked you how to think about analyzing split scope of non-upward-monotone quantifiers:

(28) The company need fire no employees. $\neg > \Box > \exists$

With continuation semantics, we can understand this data as providing support for the idea that expressions can denote three-story towers (something not excluded by, e.g., Heim & Kratzer 1998 in any case).

(29) $[\![\text{no employees}]\!] := \lambda k \,.\, \neg\, k\, (\lambda l \,.\, \exists x[\mathsf{employee}\ x \wedge l\ x])$

3 Scope islands and obligatory evaluation

Inspired by research on *delimited control* in computer science[12], Charlow (2014) develops an interesting take on scope islands couched in terms of continuations.

[12] See, e.g., Danvy & Filinski 1992 and Wadler 1994.

He proposes the following definition:

(30) *Scope islands* (def.)
A *scope island* is a constituent that is subject to *obligatory evaluation*.
(Charlow 2014: p. 90)

By *obligatory evaluation*, we mean that every continuation argument *must* be saturated before semantic computation can proceed. In other words, a scope island is a constituent where, if we have something of type $\mathtt{K_t\ a}$, we cannot proceed.

One way of thinking about this, is that the presence of an unsaturated continuation argument means that there is some computation that is being deferred until later. Scope islands are constituents at which evaluation is *forced*. As noted by Charlow, this idea bears an intriguing similarity to Chomsky's notion of a *phase*.[13]

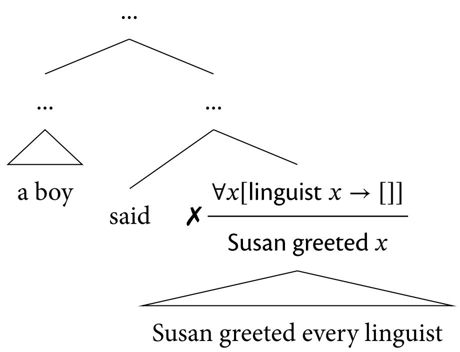
[13] Exploring this parallel in greater depth could make for an interesting term paper topic.

How does this work in practice? A great deal of ink has been spilled arguing that, e.g., a finite clause is a scope island.

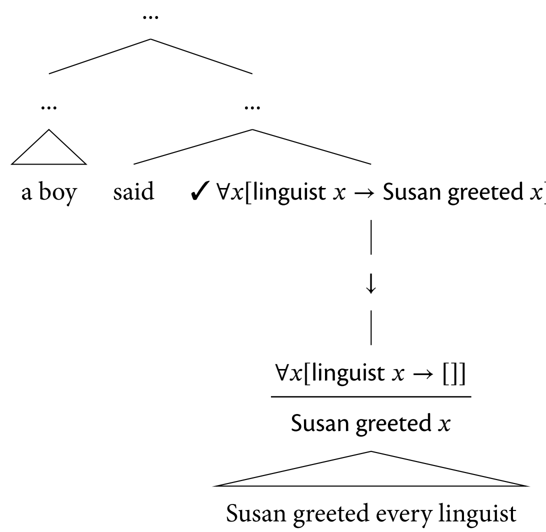
(31) A boy said $\overbrace{\boxed{\text{that Susan greeted every linguist}}}^{\text{scope island}}$. $\exists > \forall$; ✗ $\forall > \exists$

The derivation of the embedded clause proceeds as usual via lift and SFA.

(32) Scope island with an unevaluated type



(33) Scope island with an evaluated type



This story leaves a lot of questions unanswered of course:

- Is this just a recapitulation of a representational constraint on quantifier raising?[14]
- Can we give a principled story about islands for overt movement using similar mechanisms? What explains the difference between overt movement and scope taking with respect to locality?[15]

[14] The answer to this question may ultimate be *yes*, in my view.

[15] If we want to give a more general account of phases using this mechanism, we need to give an account of overt movement in terms of continuations, too. See my unpublished ms. *Movement as higher-order structure building* for progress in this direction.

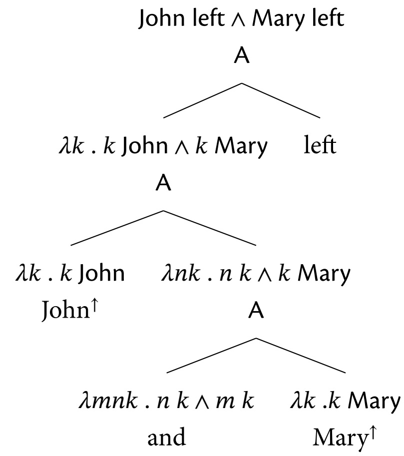
4 *Generalized coordination*

Unlike other expressions we've seen so far, we can characterize *and* as something that takes two continuized values as arguments.

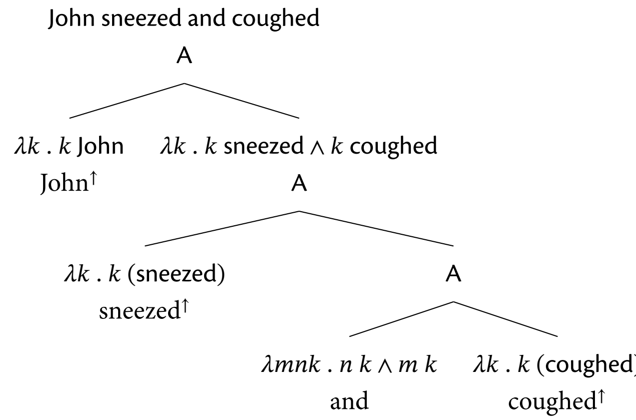
(34) $m \text{ and } n := \lambda k \,.\, m\, k \wedge n\, k$

The intuition here is as follows: *and* wants as its arguments things that are *guaranteed* to give back truth values at some future stage of computation.

(35)



(36)



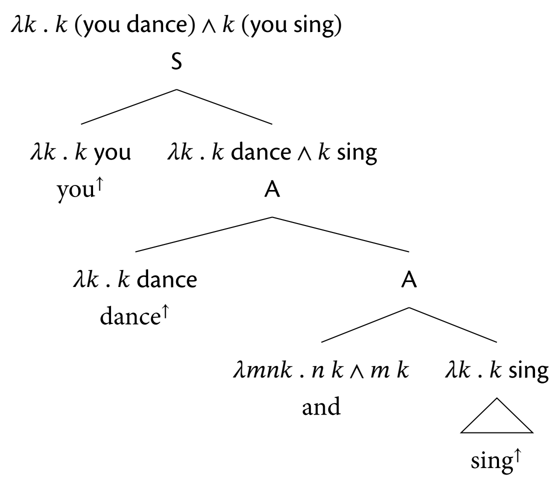
Conjunction exhibits "scope" ambiguities:

(37) You're not allowed to dance and sing. $\neg > \Diamond > \wedge; \wedge > \neg > \Diamond$

We can account for the wide/narrow scope ambiguity as a matter of where we LOWER.

Let's first compute the semantic value of the prejacent of the modal:

(38)

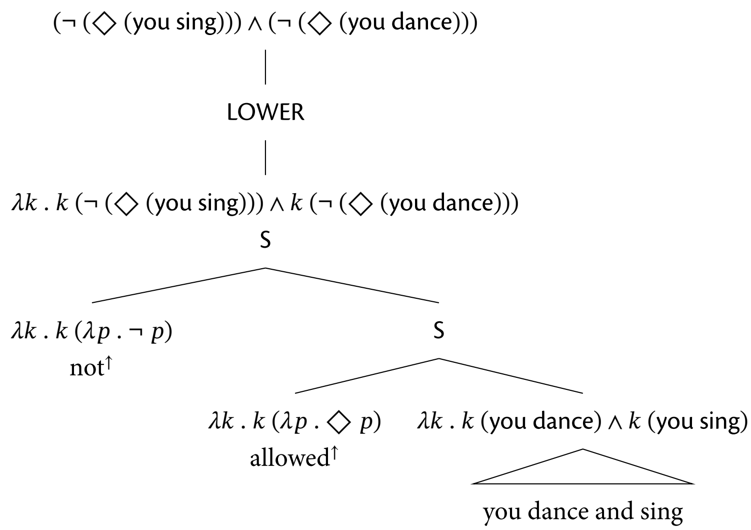


If we LOWER immediately we're just going to get a proposition, which allowed will take as its argument, deriving the narrow scope reading.[16]

[16] Instead of composing via S and lowering, we could equivalently compose lifted *you* with its complement via A, since it's a subject.

The "wide scope" reading is more interesting. We can simply defer lowering, and compose the prejacent with lifted *allowed* via S.

(39)



We make the nice prediction that "wide scope" readings of conjunction should be subject to scope islands:

(40) John isn't allowed to claim [that you sing and dance]. ✗ $\wedge > \neg > \Diamond$

References

Charlow, Simon. 2014. *On the semantics of exceptional scope*.

Chomsky, Noam. 2001. Derivation by phase. In Kenneth L. Hale & Michael J. Kenstowicz (eds.), *Ken hale: A life in language* (Current Studies in Linguistics 36). Cambridge Massachussetts: The MIT Press.

Danvy, Oliver & Andrzex Filinski. 1992. Representing Control: a Study of the CPS Transformation. *Mathematical Structures in Computer Science* 2(4). 361–391.

Elliott, Patrick D. 2019. Overt movement as higher-order structure building. unpublished manuscript. Leibniz-Zentrum Allgemeine Sprachwissenschaft.

Heim, Irene & Angelika Kratzer. 1998. *Semantics in generative grammar* (Blackwell textbooks in linguistics 13). Malden, MA: Blackwell. 324 pp.

Kobele, Gregory. 2006. *Generating copies - An investigation into structural identity in language and grammar*. UCLA dissertation.

Wadler, Philip. 1994. Monads and composable continuations. *LISP and Symbolic Computation* 7(1). 39–55.