

# Continuation semantics ii<sup>1</sup>

Patrick D. Elliott<sup>2</sup> & Martin Hackl<sup>3</sup>

February 10, 2020

<sup>1</sup> 24.979: Topics in semantics

*Getting high: Scope, projection, and evaluation order*

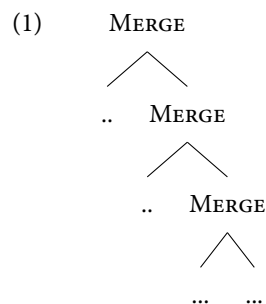
<sup>2</sup> pdell@mit.edu

<sup>3</sup> 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.<sup>4</sup>



<sup>4</sup> 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 at least to Montague (1973), is often described as *direct compositionality*.<sup>5</sup>

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

<sup>5</sup> Although direct compositionality is often associated with frameworks such as variable free semantics and Combinatory Categorical Grammar, it's in principle independent. See, e.g., Kobele (2006) for an explicit formalization of a directly compositional minimalist grammar.

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.

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

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

$$(3) \quad (\mathbb{x}, x) * (y, y) := ([\mathbb{x} \ y], x \ S \ y)$$

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:

$$(4) \quad (\mathbb{x}, x)^\dagger := (\mathbb{x}, x^\dagger)$$

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:

$$(5) \quad a^\dagger := \frac{[]}{a}$$

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.

$$(6) \quad \llbracket \text{everyone} \rrbracket := \lambda k . \forall x [k \ x]$$

$$(7) \quad \llbracket \text{everyone} \rrbracket^\dagger = \lambda l . l (\lambda k . \forall x [k \ x])$$

#### Question

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

Going back to tower notation, lifting a tower adds a trivial third story:<sup>6</sup> 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).

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.

<sup>6</sup> In fact, via successive application of LIFT, we can generate an  $n$ –story tower.

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

### Question

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

$$(9) \quad \frac{\left( \frac{f []}{g []} \right)}{x}$$

$$(10) \quad \frac{f []}{\left( \frac{g []}{x} \right)}$$

The extra ingredient we're going to need, is the ability to sandwich an empty story into the *middle* of our tower, pushing the quantificational part of the meaning to the very top. This is *internal lift* ( $\Uparrow$ ).<sup>7</sup>

(12) *Internal lift* (def.)

- a.  $(\Uparrow) : K_t a \rightarrow K_t (K_t a)$
- b.  $m^\Uparrow := \lambda k . m (\lambda x . k x^\uparrow)$

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.)

$$\left( \frac{f []}{x} \right)^\Uparrow := \frac{f []}{\frac{[]}{x}}$$

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

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

<sup>7</sup> 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 though 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  $S$ .

$$(11) \quad (\text{LIFT}^\uparrow) S \frac{f []}{x} = \frac{f []}{\frac{[]}{x}}$$

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.

$\forall > \exists$

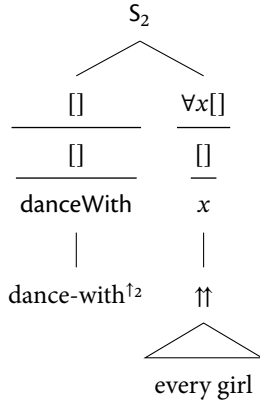
Before we proceed, we need to generalize LIFT and SFA to three-story towers.<sup>8</sup>

$$(16) \quad x^{\uparrow 2} := \frac{\frac{[]}{x}}{[]}$$

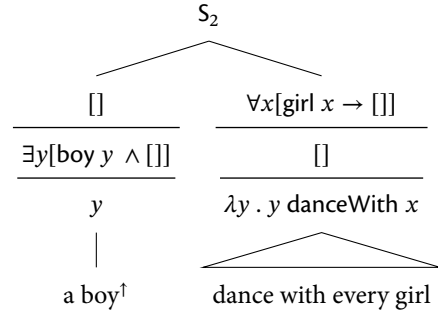
$$(17) \quad \frac{f []}{m} S_2 \frac{g []}{n} := \frac{f (g [])}{m S n}$$

<sup>8</sup> 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: externally lift *a boy*



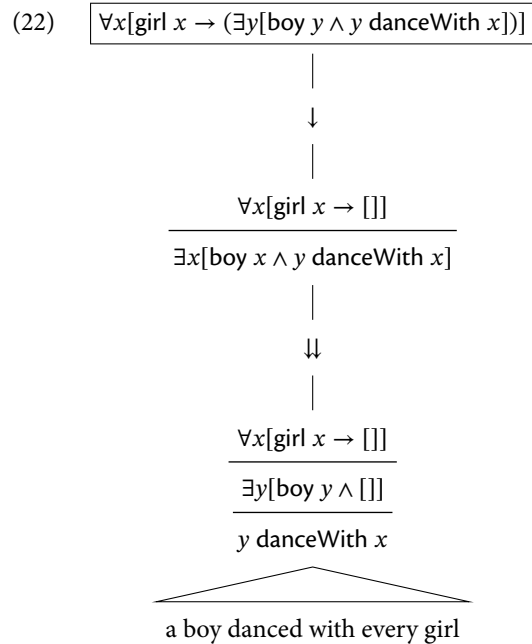
What we're left with now is a 3-story tower with the universal on the top story and the existential on the middle story. We can collapse the tower by first collapsing the bottom two stories, and then collapsing the result. In order to do this, we'll first define *internal lower*.<sup>9</sup>

$$(21) \quad \text{Internal lower (def.)} \quad \left( \frac{f []}{\frac{g []}{p}} \right)^{\Downarrow} := \frac{f []}{\left( \frac{g []}{x} \right)^{\Downarrow}}$$

<sup>9</sup> 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:

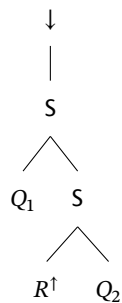
$$(20) \quad m^{\Downarrow} \equiv (1)^{\uparrow} S m$$

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

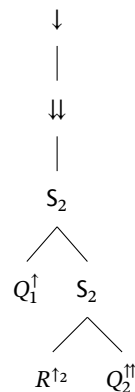


Great! We've shown how to achieve quantifier scope ambiguities using our new framework. Let's look at the derivations again side-by-side.

(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  $\Downarrow$ .
- In order to derive an inverse scope reading, what was *crucial* was the availability of *internal lift*; the remaining operations,  $S_2$ ,  $\uparrow_2$ ,  $\Downarrow$  only functioned to

message 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.<sup>10</sup>

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 in 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 to an ordinary tower in a different way. Let's call it *join*:<sup>11</sup>

$$(25) \text{ join (def.)} \quad m^\mu := \lambda k . m (\lambda c . c k) \quad \mu : K_t (K_t a) \rightarrow K_t a$$

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

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

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).

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

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*

<sup>10</sup> 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*.

<sup>11</sup> *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.

*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.  $\neg > \Box > \exists$

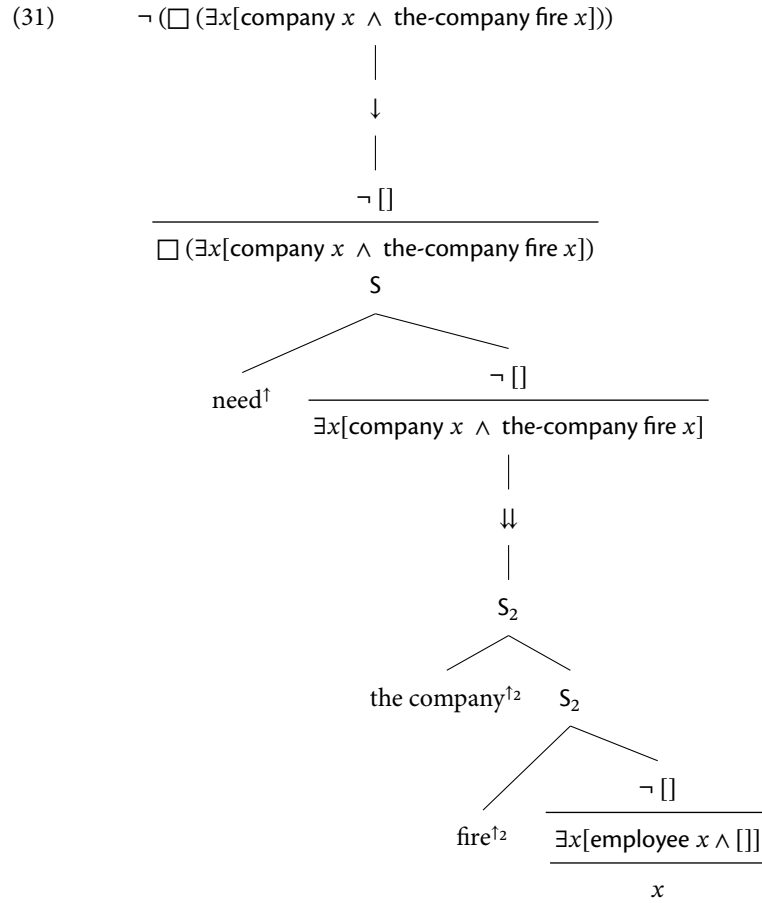
With continuation semantics, we can understand this data as providing support for the idea that expressions can denote three-story towers (something not excluded by, e.g., Heim & Kratzer 1998 in any case).

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

Tower version:

(30)  $\llbracket \text{no employees} \rrbracket := \frac{\frac{\neg []}{\exists x [\text{employee } x \wedge []]}}{x}$

We get the split scope reading by doing internal lower first below the modal, and then external lower above the modal.



One interesting property of split scope readings is that it is essential that we be able to *lower* a 3-story tower in two distinct steps – *internal lower* followed by *lower*.

### 3 Scope islands and obligatory evaluation

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

<sup>12</sup> See, e.g., Danvy & Filinski 1992 and Wadler 1994.

He proposes the following definition:

(32) *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  $\kappa_t$  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*.<sup>13</sup>

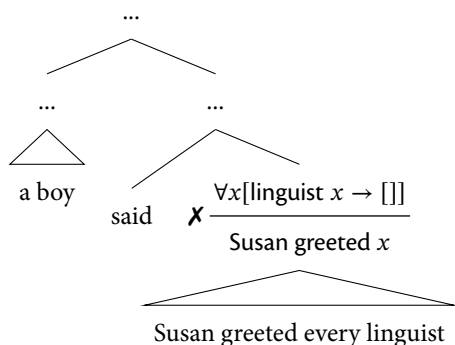
How does this work in practice? A great deal of ink has been spilled arguing that, e.g., a finite clause is a scope island.

(33) A boy said that Susan greeted every linguist.  $\exists > \forall; \forall > \exists$

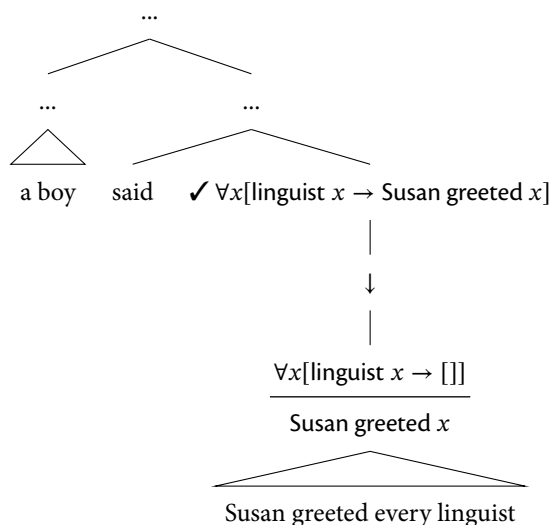
The derivation of the embedded clause proceeds as usual via lift and SEA.

<sup>13</sup> Exploring this parallel in greater depth could make for an interesting term paper topic.

(34) Scope island with an unevaluated type



(35) 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?<sup>14</sup>
- 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?<sup>15</sup>

<sup>14</sup> The answer to this question may ultimately be yes, in my view.

<sup>15</sup> 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 con/dis-junction

The flexibility of *and* and *or* has been discussed at length by, e.g., Partee & Rooth (1983), Winter (2001), among others.<sup>16</sup>

- (36) Lan and some woman arrived.
- (37) Howie sneezed and/or coughed.
- (38) Lan kissed and/or hugged Irene.

<sup>16</sup> We saw an example of this when we motivated LIFT.

Unlike other expressions we've seen so far, we can characterize *and* and *or* as expressions that takes two continuized values as arguments.<sup>17</sup>

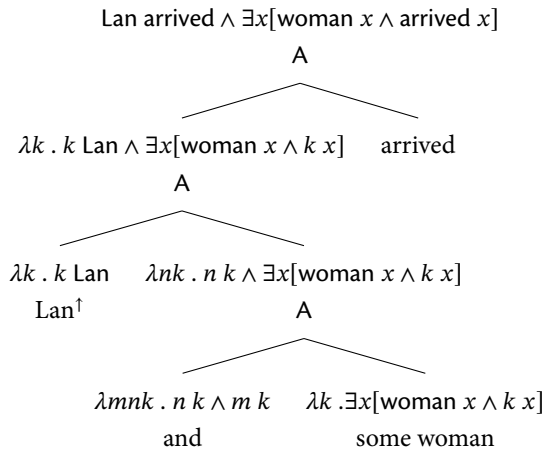
<sup>17</sup> Note that since the continuation argument  $k$  occurs more than once in the function body, we can no longer abbreviate the flat lambda-expression using a tower.

- (39) a.  $m \text{ and } n := \lambda k . m \ k \wedge n \ k$                        $\text{and} : K_t \ a \rightarrow K_t \ a \rightarrow K_t \ a$   
 b.  $m \text{ or } n := \lambda k . m \ k \vee n \ k$                        $\text{or} : K_t \ a \rightarrow K_t \ a \rightarrow K_t \ a$

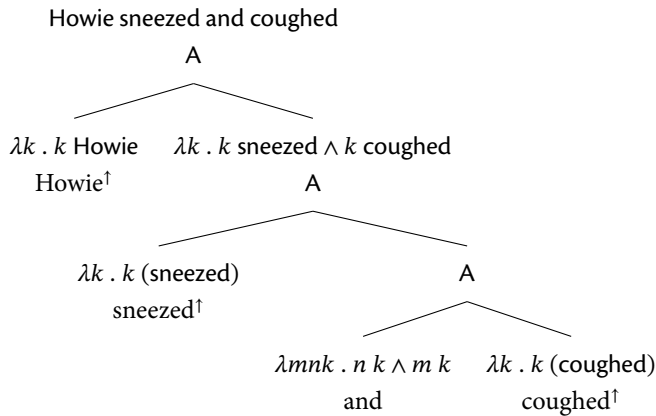
The intuition here is as follows: *and* wants as its arguments things that are *guaranteed* to give back truth values at some future stage of computation.

This accounts for the basic cases discussed by Partee & Rooth (1983), with a single (polymorphic) entry for *and*, as illustrated below:<sup>18</sup>

- (40) Lan and some woman arrived.



- (41) Howie sneezed and coughed.



#### 4.1 The scope of con/dis-junction

Both conjunction and disjunction exhibit “scope” ambiguities. This is illustrated below for conjunction:

<sup>18</sup> Unlike other lexical entries we’ve seen so far, *and* is *lexically specified* as seeking continuized arguments. As such, if a value isn’t already typed as an instantiation of  $K_t \ a$ , it must be lifted.

In the derivations below, you can observe that, unlike the cases we’ve encountered so far, *and* composes with its arguments via Function Application (FA) rather than sFA.

(42) You're not allowed to dance and sing.

a. *You're not allowed to dance and you're not allowed to sing*

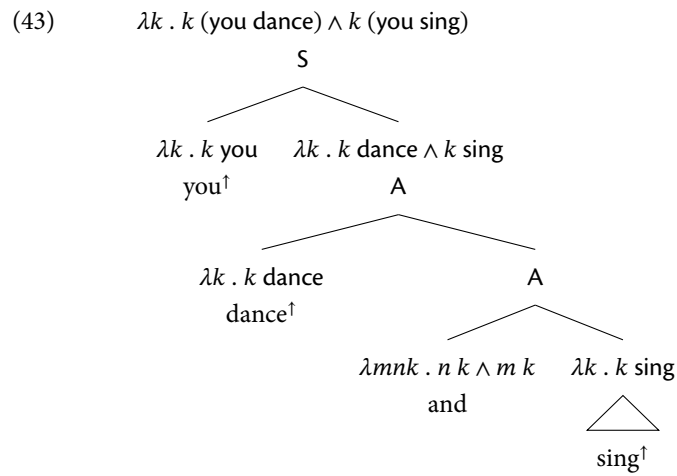
$\wedge > \neg > \Diamond$

b. *You're not allowed to both dance and sing (at the same time)*

$\neg > \Diamond > \wedge$

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

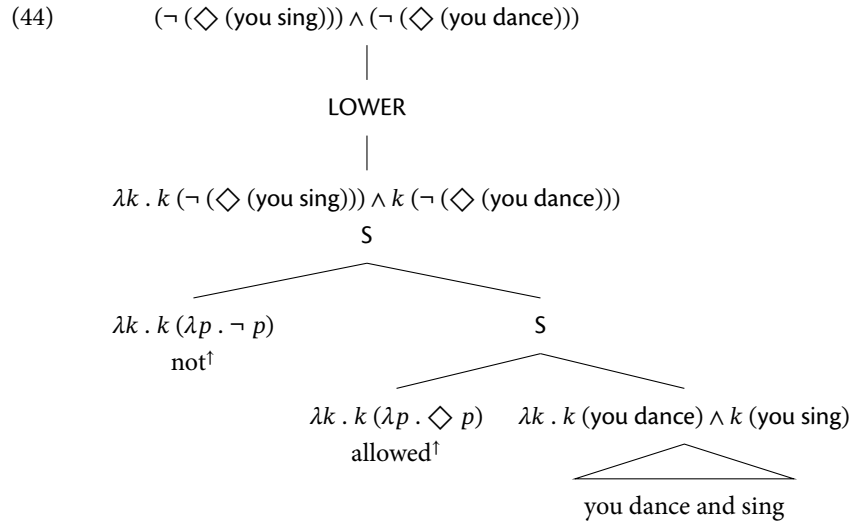
Let's first compute the semantic value of the prejacent of the modal:



If we LOWER immediately we're just going to get a proposition, which *allowed* will take as its argument, deriving the narrow scope reading.<sup>19</sup>

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

<sup>19</sup> Instead of composing via S and lowering, we could equivalently compose lifted *you* with its complement via A, since it's a subject.



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

(45) John isn’t allowed to claim [that you sing and dance].  $\text{X} \wedge > \neg > \Diamond$

#### 4.2 Split scope with conjunction

Hirsch (2017) claims that conjunction reduction is necessary in order to derive the “split scope” reading of the following sentence:

(46) John refused to visit any city in Europe and any city in Asia.

Hirsch’s analysis:

(47) John [<sub>VP</sub> refused to visit any city in Europe ] and  
       [<sub>VP</sub> refused to visit any city in Asia ]

Here, I’m going to argue that we can get the split scope reading of (46) using just the machinery we’ve already introduced. Concretely:

- Our entry for *and*.
- The free availability of *external lift*.

First of all, each conjunct is going to be *externally lifted*.

$$(48) \quad \llbracket \text{any city in Europe} \rrbracket^\dagger = \lambda l . l (\lambda k . \exists x [\text{city-europe } x \wedge k x])$$

and, in our view, is polymorphic – it’s looking for two arguments of type  $K_t a$ . This means it can compose the externally lifted DPs, returning the following meaning:

$$(49) \quad \lambda l . l \left( \frac{\exists x [\text{europe-city } x \wedge []]}{x} \right) \wedge l \left( \frac{\exists x [\text{asia-city } x \wedge []]}{x} \right) : K_t (K_t a)$$

Remember that  $S_2$  is just like  $S$ , only it does  $S$  of the wrapped values. Composition can proceed via  $\uparrow_2$  and  $S_2$  up to the prejacent of *refuse*.

To get the scope of the quantifiers right, we do *internal lower* just before we compose with *refuse*.

$$(50) \quad \lambda l . l (\text{refuse } \exists x [\text{europe-city } x \wedge \text{pro visit } x]) \wedge l (\text{refuse } \exists x [\text{europe-city } x \wedge \text{pro visit } x])$$

S

$$\text{refuse}^\dagger \quad \lambda l . l (\exists x [\text{europe-city } x \wedge \text{pro visit } x]) \wedge l (\exists x [\text{europe-city } x \wedge \text{pro visit } x])$$

|

⇕

|

$$\lambda l . l \left( \frac{\exists x [\text{europe-city } x \wedge []]}{\text{pro visit } x} \right) \wedge l \left( \frac{\exists x [\text{asia-city } x \wedge []]}{\text{pro visit } x} \right)$$

PRO visit any city in Europe  
and any city in Asia

Lowering the result gives us...the split-scope reading:

$$(51) \quad \text{refuse } \exists x [\text{europe-city } x \wedge \text{pro visit } x \wedge \text{refuse } \exists x [\text{asia-city } x \wedge \text{pro visit } x]]$$

|

LOWER

|

$$\lambda l . l (\text{refuse } \exists x [\text{europe-city } x \wedge \text{pro visit } x]) \wedge l (\text{refuse } \exists x [\text{europe-city } x \wedge \text{pro visit } x])$$

So, continuations can get split-scope readings of conjunction straightforwardly.

This is a notable result.<sup>20</sup>

Note that this theory of split-scope coordination explains the lack of split-scope reading in the following example, from Partee & Rooth (1983).

- (52) John hopes [that some company will hire a maid and a cook].  
 ✗ John hopes that some company will hire a maid,  
 and John hopes that some company will hire a cook.

This is simply because the embedded finite clause is a scope island, and therefore the scope of *and* is trapped.

## 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 Massachusetts: The MIT Press.
- Danvy, Oliver & Andrzej 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.
- Hirsch, Aron. 2017. A case for conjunction reduction. unpublished manuscript. Massachusetts Institute of Technology.
- Kobele, Gregory. 2006. *Generating copies - An investigation into structural identity in language and grammar*. UCLA dissertation.
- Montague, Richard. 1973. The Proper Treatment of Quantification in Ordinary English. In K. J. J. Hintikka, J. M. E. Moravcsik & P. Suppes (eds.), *Approaches to Natural Language: Proceedings of the 1970 Stanford Workshop on Grammar and Semantics* (Synthese Library), 221–242. Dordrecht: Springer Netherlands.
- Partee, Barbara & Mats Rooth. 1983. Generalized conjunction and type ambiguity. In, Reprint 2012, 361–383. Berlin, Boston: De Gruyter.
- Wadler, Philip. 1994. Monads and composable continuations. *LISP and Symbolic Computation* 7(1). 39–55.
- Winter, Yoad. 2001. *Flexibility principles in boolean semantics: The interpretation of coordination, plurality, and scope in natural language* (Current Studies in Linguistics 37). Cambridge Massachusetts: The MIT Press. 297 pp.

<sup>20</sup> To my knowledge, this is a novel observation.

Hirsch doesn't consider an analysis of this data in terms of continuations, but does entertain a similar analysis involving lifting the quantificational conjuncts, alongside QR with higher-type traces. Hirsch (2017) rejects this analysis on the basis that it allows for the quantifiers to be syntactically above the intensional verb, while semantically reconstructing below it – a configuration which has been argued to be ruled out. Note that an analysis in terms of continuations doesn't face this objection – at no point in the analysis did we need to invoke covert movement.