PART 6. SEMANTIC PARSING OF THE VERB CLUSTER IN DUTCH AND GERMAN

6.1 Basics of functional type theory, semantics of serial verbs via function composition.

Basics of functional type theory (ignoring intensionality)

This is rigorously done in Advanced Semantics (warning : the text below is horribly sloppy on blurring object language-meta language)

Types

e is the type of individuals, t is the type of truth-values <a,b> is the type of functions from a-entities into b-entities.

<e,t> is the type of functions from individuals into truth values = one-place properties.

Functional application

If α is of type $\langle a, b \rangle$ and β is of type a, then $(\alpha(\beta))$ is of type b. $(\langle a, b \rangle + a \rightarrow b)$

PURR	is of type <e,t></e,t>	one place property
Ronya	is of type e	individual
(PURR(Ronya))	is of type t	truth value

Curried functions

EAT[Pat, Pap]two-pl	ace relation between individuals
curried:	
EAT	is of type <e,<et>></e,<et>
	function from individuals into one place properties
functional application	:
(EAT(pap))	is of type <e,t>, one place property</e,t>
	the property that you have if you eat pap
Functional application	n:
((EAT(pap)) (Pat)	is of type t
	Pat has the property that you have if you eat pap.

Relational notation: if R is of type $\langle b, \langle a, c \rangle \rangle$ and β of type b and α of type a, then R(α,β) is relational notation for ((R(β))(α))

EAT(Pat, pap) is relational notation for ((EAT(pap)) (Pat))

Functional abstraction

If x is a variable of type a and β is of type b, then $\lambda x.\beta$ is of type $\langle a,b \rangle$

 $\lambda x.\beta$ is the function that maps all a-entities d onto the interpretation of β [x:d], the interpretation of β that sets the interpretation of variable x to d.

lambda abstraction allows us to define complex functions.

 $\lambda x_n \dots x_1 \beta (x_1, \dots, x_n)$

The lambda prefix $\lambda x_n \dots x_1$ indicates the arguments that go into the function in the order $x_n \dots x_1$.

 $\beta(x_1,...,x_n)$ is the description of the function, it tells you what the function does. In particular,

the order in $\beta(x_1,...,x_n)$ indicates the argument (or thematic) structure (who does what to whom).

Example:



eat first combines with the object *pap*, and the result *eat pap* combines with the subject *Pat*:

eat $\rightarrow \lambda y \lambda x. EAT(x,y)$ the relation that holds between x and y if x eats y. ($\lambda y \lambda x. EAT(x,y)$ (pap)) (($\lambda y \lambda x. EAT(x,y)$ (pap)) (Pat))

 λ -conversion:

 $(\lambda x.\beta (\alpha)) = \beta[\alpha/x]$ the result of replacing every variable x free in β by α , if no variable is bound in $\beta[\alpha/x]$ that was free in β

 $(\lambda y \lambda x. EAT(x, y) (pap)) = \lambda x. EAT(x, pap)$

 $((\lambda y \lambda x. EAT(x, y) (pap)) (Pat)) = (\lambda x. EAT(x, pap) (Pat))$

 $(\lambda x.EAT(x, pap) (Pat)) = EAT(Pat, pap)$

Higher order abstraction

Example: attributive adjectives:

smart $\rightarrow \lambda P \lambda x. P(x) \land SMART(x)$ of type <<e,t>,<e,t>>,

P is a variable of type <e,t>, a variable over properties.

smart denotes a function from one place properties into one place properties

maps property Q onto the property that you have if you have Q and you are smart.



smart cat = $(\lambda P\lambda x.P(x) \land SMART(x) (CAT))$

 λ -conversion:

$$(\lambda P\lambda x.P(x) \land SMART(x) (CAT)) = \lambda x.CAT(x) \land SMART(x)$$

The property that you have if you are a cat and you are smart. Small clause analysis for help/let/see... auxiliary verbs:



Semantics matching the syntax exactly (simplifying by ignoring intensionality)

let $\rightarrow \lambda p \lambda x.LET(x,p)$

p a variable over propositions, for simplicity here identified with sentences (type t, this is incorrect, but will do for pour purposes here).

Semantics of the small clause:	((EAT(pap)) (Pat))
	= EAT(Pat, pap)
Semantics of the sentence:	(LET(EAT(Pat, pap)) (Sam))
	= LET(Sam, EAT(Pat, pap))

Sam stands in the LET relation to the proposition that Pat eats pap. Lexical semantics of LET: LET(x,p) x allows p to happen (where p is under x's control, etc.)

Semantics:



two place relation between an individual and a proposition

Non-matching semantics

let $\rightarrow \lambda P \lambda y \lambda x. LET(x, P(y))$

Syntactically *eat* takes a small clause complement and a higher subject. Semantically *eat* accesses the subject of the small class, the predicate of the small clause and the higher subject, i.e. is a three-place relation, it applies to the the small clause predicate, the small clause subject and the higher subject.

So:

Small clause predicate: $eat pap \rightarrow (EAT(pap))$

let eat pap \rightarrow ($\lambda P \lambda y \lambda x.LET(x, P(y))$ (**EAT(pap**))

 λ -conversion:

 $\lambda y \lambda x.LET(x,(EAT(pap))(y)))$ Relational notation:

 $\lambda y \lambda x. LET(x, EAT(y, pap))$

let Pat eat pap \rightarrow ($\lambda y \lambda x.LET(x, EAT(y, pap))$ (**Pat**))

 λ -conversion: $\lambda x.LET(x, EAT(Pat, pap))$

Sam let Pat eat pap \rightarrow (λ **x**.LET(**x**, EAT(Pat, pap)) (**Sam**))

 λ -conversion: LET(**Sam**, EAT(Pat, pap))

lexical semantics: LET(x, P(y)): x allows it to happen that y has P.

So: Sam lets it happen that Pat eats her pap.

-The small clause syntax is for a variety of syntactic reasons better than a syntax mirroring the semantics.

-The 3-place relational semanrics is for a variety of semantic reasons better than a semantics mirroring the syntax.

Hence: mismatches between syntax and semantics.

(Other examples: Landman 2000, 2004 on numericals in nominal versus argument or predicate position, 2016 on measure versus classifier readings).

Semantic argument: the semantics given here allows a very elegant semantics for the verb cluster in terms of function composition.

function composition:

If f is a function of type $\langle a,b \rangle$ and g a function of type $\langle b,c \rangle$ then g \circ f is a function of type $\langle a,c \rangle$ ($\langle a,b \rangle + \langle b,c \rangle \rightarrow \langle a,c \rangle$)

 $g \circ f = \lambda x.g(f(x))$

Apply f to variable x of type a:	f(x)	is of type b
Apply g to the result:	g(f(x))	is of type c
Abstract over x:	$\lambda x.g(f(x))$	is of type <a,b></a,b>

 $\lambda x.g(f(x))$ The function that maps every x of type a onto the result of applying f to x and

then g to the result.

generalized function composition:

Resolve a type mimatch through function composition:

You want to compose function g with function f.

So apply f to a variable x. But f(x) isn't of the right type to be fed into g. Solution: continue to apply f to variables (i.e. apply f(x) to a variable y, etc. until the types match.

Apply g to the result and abstract over all variables you have applies f to:

 $COMP[g,f] = \lambda x_n \dots \lambda x_1 g(f(x_1, \dots, x_n))$

(where the types of $x_1...x_n$ can be anything and $f(x_1,...x_n)$ should be understood curried:

 $(\dots((f(x_n))(x_{n-1}))\dots(x_1))$)

Fact:

Let \mathbb{R}^n be an n-place relation between individuals (of type <e,...,<e,t>> with n e's). Let **LET** = $\lambda P \lambda y \lambda x.LET(x,P(y))$

Then: COMP[LET, Rⁿ] is an n+1 place relation between individuals

Proof:

We calculate COMP[LET, Rⁿ] Step 1: apply Rⁿ to variables $x_2,...,x_n$ to bring the type down to <e,t>, the type of variable P: Rⁿ($x_2,...,x_n$) is a one place predicate. Step 2: apply LET to Rⁿ($x_2,...,x_n$): ($\lambda P \lambda y \lambda x.LET(x, P(y))$ (Rⁿ($x_2,...,x_n$)) λ -conversion: $\lambda y \lambda x.LET(x, R^n(x_2,...,x_n) (y))$ Relational notation: $\lambda y \lambda x.LET(x, R^n(y,x_2,...,x_n))$ Step 2: abstract over variables $x_2,...,x_n$: $\lambda x_n...\lambda x_2 \lambda y \lambda x.LET(x, R^n(y,x_2,...,x_n))$

With this we understand what happens in the verb cluster:

COMP[LET ,EAT]	$(1-1) = \Gamma \wedge T(-1-1)$	01	
Apply to a variable z:	λ uλv.EAT(v,u) (z) λ uλv.EAT(v,u) (z)	2-place	relation
λ -conversion:	$\lambda v.EAT(v, \mathbf{z})$	1-place	property
Apply LET : λ-conversion:	$(\lambda P \lambda y \lambda x.LET(x, P(y)))$ λyλx.LET(x, λv.EAT	(λ v.EA (v,z) (y)	T(v,z))
λ-conversion:	λyλx.LET(x, λv .EAT(λyλx.LET(x, EAT(y ,z	(v,z) (y) ()))
abstract over z:	λzλyλx.LET(x, EAT(y	y,z))	3-place relation

So given the semantics for *let*, function composition in the verb cluster of *let* and two place relation *eat* forms three-place relation *let eat*.

-In **phrasal domains** (IP (S), VP), the basic **meaning composition** operation is function-argument application.

-In **lexical domains** (like V), the basic **meaning composition** operation is function composition. (Hoeksema 19??)

Claim: The mismatch assumption and the composition assumption are the only assumptions that need to be made to get the semantics of the serial verb cluster come out right.

Example: (ignoring the modal *zal*)



Predicted Semantics:

LET = $\lambda P \lambda y \lambda x$. **LET**(x, P(y)) **EAT** = $\lambda z \lambda y$. **EAT**(y,z)

LET ° **EAT** = $\lambda P \lambda y \lambda x$. **LET**(x, P(y)) ° $\lambda z \lambda y$. **EAT**(y,z) =

 λz [λΡλγλx. LET(x, P(y))] ($\lambda z \lambda xy. EAT(y,z)$ (z)) = λz [λΡλγλx. LET(x, P(y))] ($\lambda y. EAT(y,z)$) = λz [$\lambda y \lambda x. LET(x, \lambda y. EAT(y,z)$ (y))] =

 $\lambda z \lambda y \lambda x.$ **LET**(x, **EAT**(y,z)) x lets: y eat z

Conclusion:

2 place verb	+ LET	\rightarrow	3-place verb
λzλy. $EAT(y,z)$	LET		$\lambda z \lambda y \lambda x.$ LET(x, EAT(y,z))
1 2			1 23

Next:

HELP \circ (LET \circ EAT) =

 λ P λ x λ u. **HELP**(u, P(x)) ° λ z λ y λ x. **LET**(x, **EAT**(y,z)) =

 $\lambda z \lambda y$ [$\lambda P \lambda x \lambda u$. **HELP**(u, P(x))] ($\lambda z \lambda y \lambda x$. **LET**(x, **EAT**(y,z)) (y,z)) = $\lambda z \lambda y$ [$\lambda P \lambda x \lambda u$. **HELP**(u, **P**(x))] (λx . **LET**(x, **EAT**(y,z))) = $\lambda z \lambda y \lambda x \lambda u$. **HELP**(u, **LET**(x, **EAT**(y,z))) $\begin{array}{cccc} 3 \text{ place verb} & + \text{ HELP} \rightarrow & \text{4-place verb} \\ \lambda z \lambda y \lambda x. \text{ LET}(x, \text{ EAT}(y, z)) & & \lambda z \lambda y \lambda x \lambda u. \text{ HELP}(u, \text{ LET}(x, \text{ EAT}(y, z))) \\ 1 & 2 & 3 & 1 & 2 & 3 & 4 \end{array}$

So: helpen laten eten $\rightarrow \lambda z \lambda y \lambda x \lambda u$. HELP(u, LET(x, EAT(y,z))) u helps x; x lets y, y eats z.

-The basic composition operation in phrasal domains is **function-argument** application:

Type theory: The left-right order in the λ -prefix represents the order of application.

-The meaning of the V *helpen laten eten* applies to the meaning of *haar pap* (λz) -the result applies to the meaning of *Pat*, (λy), -the result to the meaning of *Sam* (λx), -the result to the meaning of *Kim* (λu), giving, the correct meaning for the sentence:

APPLY[, Kim]
APPLY[, Sam]
APPLY[, Pat]
APPLY[HE	LP ° (LET ° EAT), Her Porridge]

=

HELP(Kim, LET(Sam, EAT(Pat, Her Porridge)))

The semantics proposed gets the meanings right within a framework of standard assumptions about the semantic composition operations applicable in different domains (composition and application).

Given the meanings of the verbs that enter into the serial verb, composition has the effect of:

n-PLACE SERIAL VERB FORMATION:

Let α be one of LET, HELP, SEE, HEAR ,... Let β be an n-1 place relation, then $\alpha \circ \beta$ is an n-place relation.

6.2. Semantic Parsing

Bach et. al. 1986 performed a cross-linguistic experiment.

The idea of the experiment was the following: Dutch and German are similar enough to be able to compare the speed of processing of the Dutch verb cluster by native speakers of Dutch with that of the German verb cluster by native speakers of German. What Bach et. al. measured for Dutch and for German was the following.

Let's use the numbers 0,1,2,3,... for 0 embeddings, 1 embedding, 2 embeddings,... as indicated below:

German:

0	Jan weißt daß wi	r d	as Ha	us hat	en wo	llen m	alen
1		Hans				/	lassen _l
2		Hans	Peter		/	helfe	n lassen
3		Hans	Peter	Marie	sehen	helfe	n lassen
		NP ₁	NP ₂	NP ₃	V ₃	V_2	V_1

Center embedded dependencies.

Dutch:

2		
0	Jan weet dat we het huis hebben willen	n schilderen
1	Hans	laten
2	Hans Peter	laten helpen
3	Hans Peter Marie	laten helpen zien
	NP ₁ NP ₂ NP ₃	$V_1 V_2 V_3$
	l l	

Cross serial dependencies.

Bach et. al. measured, in terms of processing time, in each language, how much longer type 1 sentences take to process than type 0, how much longer type 2 than type 1, etc. And then they compared the figures they got for Dutch and for German. They found that indeed type 0 in Dutch and in German (and in English) take about the same time (which forms the basis for comparison). Interestingly enough they found the following:

Systematically Dutch speakers process n Dutch embeddings FASTER than German speakers process n German embeddings.

I want to suggest a possible explanation of this result in terms of **semantic parsing**. (For an alternative explanation, see Joshi 1989.)

Let's first explain the idea of semantic parsing.

We have an input string which is read symbol by symbol from left to right:

dat Jan Marie kust [that Jan kisses Marie]

In semantic parsing the task is to come up with a semantic interpretation of the sentence. And we use the **types** of the input expressions and the type assignment of the grammar to do that online. Basically the idea is to find the values for the variables introduced in the parsing process, or equivalently, eliminate the variables. The parse is done when all variables are eliminated.

We indicate in boldface where we are in the parse:

Step 1	<i>dat Jan Marie</i> Semantics: Task: find the	φ φ value of φ.	$[\phi \in VAR_t, a \text{ sentential variable}]$
Step 2	dat Jan Marie Semantics:	φ kust $\varphi = P^{1}(j)$	$[P^1 \in VAR_{\langle e,t \rangle}, a \text{ one-place predicate variable}]$
	Task: find the	value of P^1 .	-
Step 3	dat Jan Marie Semantics:	e^{kust} $\varphi = P^{1}(j)$ $P^{1} = P^{2}(m)$	$[P^2 \in VAR_{\langle e, \langle e, t \rangle}$, a two place predicate variable]
This means th	at we can elimi Semantics:	nate P^1 : $\varphi = P^2(j,m)$	
	Task: find the	e value of P^2 .	
Step 4	dat Jan Marie Semantics:	$kust \varphi = P^{2}(j,m) P^{2} = KISS$	
We eliminate	p ² .		

We eliminate P²: Semantics:

We eliminate φ:

Semantics: KISS(j,m)

Done.

Applying the very same strategy in the verb cluster, we get for Dutch (and for German) the following partial parse: (LF stands for: 'look for the value of')

 $\varphi = KISS(j,m)$

So at the point where we reach the verb cluster, the parser is looking for a **four-place** relation.

We are at the following stage in Dutch:

dat Kim Sam Pat haar pap **zal** helpen laten eten. **SEMANTICS**: P⁴(k,s,p,pap) **LOOK FOR**: P⁴

In German we are, similarly at the following stage:

daβ Kim Sam Pat ihr Brei **essen** lassen helfen wird. **SEMANTICS**: $P^4(k,s,p,pap)$ **LOOK FOR**: P^4

In both cases, we are looking for a **four-place relation** P^4 and we rely on function composition to find it.

Let's argue the German case first. We are looking for a four place relation. But *essen* is a two-place relation. So we are stuck. At this point we **start a store in which we build a four place relation**:

daβ Kim Sam Pat ihr Brei **essen** lassen helfen wird. **SEMANTICS**: $φ = P^4(k,s,p,pap)$ **LF**: P^4 **STORE: EAT** (a two place relation)

We continue and at the next step we apply function composition in the store:

daβ Kim Sam Pat ihr Brei essen **lassen** helfen wird. **SEMANTICS**: $φ = P^4(k,s,p,pap)$ **LF**: P^4 **STORE: LET** o **EAT** (a three-place relation) **LET** o **EAT** =

We continue to *helfen* and again do function composition on the store:

daβ Kim Sam Pat ihr Brei essen lassen **helfen** wird. **SEMANTICS**: $φ = P^4(k,s,p,pap)$ **LF**: P^4 **STORE: HELP** o (**LET** o **EAT**) (a four-place relation)

HELP $o(LET \circ EAT) =$ $\lambda u \lambda z \lambda y \lambda x.HELP(x,LET(y,EAT(z,u)))$

 $\lambda z \lambda y \lambda x. LET(x, EAT(y, z))$

We have now a four-place relation in store, but the parse continues inside V, so we continue with function composition:

daβ Kim Sam Pat ihr Brei essen lassen helfen wird. **SEMANTICS**: $φ = P^4(k,s,p,pap)$ **LF**: P^4 **STORE: WILL o (HELP** o (**LET** o **EAT**)) (a four-place relation)

WILL o (HELP o (LET o EAT)) = $\lambda u \lambda z \lambda y \lambda x$. WILL(HELP(x,LET(y,EAT(z,u))))

This completes the parse in the V domain, we match the store and P^4 :

daβ Kim Sam Pat ihr Brei essen lassen helfen wird. **SEMANTICS**: $φ = P^4(k,s,p,pap)$ **LF**: $P^4 = WILL o (HELP o (LET o EAT))$

We eliminate variables and get the correct parse:

daβ Kim Sam Pat ihr Brei essen lassen helfen wird. **SEMANTICS**: WILL(HELP(k,LET(s,EAT(p,pap)))) **DONE**

In Dutch we have exactly the same option as in German, we can create a store:

dat Kim Sam Pat haar pap **zal** helpen laten eten. **SEMANTICS**: P⁴(k,s,p,pap) **LOOK FOR**: P⁴ **STORE: WILL**

We continue in the store with function composition:

dat Kim Sam Pat haar pap zal **helpen** laten eten. **SEMANTICS**: P⁴(k,s,p,pap) **LOOK FOR**: P⁴ **STORE: WILL 0 HELP**

> **WILL 0 HELP =** $\lambda P^{1} \lambda y \lambda x. WILL(HELP(x, P^{1}(y)))$

We continue with *laten*:

dat Kim Sam Pat haar pap zal helpen **laten** eten. **SEMANTICS**: P⁴(k,s,p,pap) **LOOK FOR**: P⁴ **STORE:** (WILL o HELP) o LET (WILL o HELP) o LET = $\lambda P^1 \lambda z \lambda y \lambda x$. WILL(HELP(x,LET(y,P¹(z))))

We compose with *eten*:

dat Kim Sam Pat haar pap zal helpen laten **eten**.

SEMANTICS: P⁴(k,s,p,pap) **LOOK FOR**: P⁴ **STORE: ((WILL 0 HELP) 0 LET) 0 EAT**

((WILL o HELP) o LET) o EAT = $\lambda u \lambda z \lambda y \lambda x$.WILL(HELP(x,LET(y,EAT(z,u))))

We are done in the V-domain, we have the same relation in store as in German, so we match, eliminate variables and are done:

dat Kim Sam Pat haar pap zal helpen laten **eten**. **SEMANTICS**: WILL(HELP(k,LET(s,EAT(p,pap)))) **DONE**

Thus far, there is no difference between the Dutch and the German case. The difference comes in with the following observation:

Dutch allows a straightforward alternative parsing strategy that does not involve a store at all.

We go back to the point where we switched from functional application to function composition:

dat Kim Sam Pat haar pap zal helpen laten eten.

SEMANTICS: $P^4(k,s,p,pap)$ **LOOK FOR**: P^4

We continue the parse by introducing search variable Q^4 and compose as follows:

```
dat Kim Sam Pat haar pap zal helpen laten eten.

SEMANTICS: P^4(k,s,p,pap)

P^4 = WILL \text{ o } Q^4

LOOK FOR: P^4
```

So we get:

dat Kim Sam Pat haar pap zal **helpen** laten eten. **SEMANTICS**: [WILL o Q⁴] (k,s,p,pap) LOOK FOR: Q⁴

And we continue by introducing a search variable P^3 and compose as follows:

```
dat Kim Sam Pat haar pap zal helpen laten eten.

SEMANTICS: [WILL o (HELP o P<sup>3</sup>)] (k,s,p,pap)

LOOK FOR: P<sup>3</sup>
```

We continue by introducing search variable P^2 and compose similarly:

dat Kim Sam Pat haar pap zal helpen laten eten.

SEMANTICS: [WIIL o (HELP o (LET o P^2))] (k,s,p,pap) LOOK FOR: P^2

At this point we reach the end of the V and we resolve:

dat Kim Sam Pat haar pap zal helpen laten **eten**. **SEMANTICS:** [WIIL o (HELP o (LET o EAT))] (k,s,p,pap)

The result is the same:

dat Kim Sam Pat haar pap zal helpen laten **eten**. **SEMANTICS**: WILL(HELP(k,LET(s,EAT(p,pap)))) **DONE**

Thus, on this strategy, we just continue to compose on. In this parse, we do not use a store, and we only need to introduce **six search** variables all in all: φ , P¹,P²,P³,P⁴, Q⁴ of five different types.

Now, composition is a powerful mechanism, so it shouldn't come as a surprise that also for German we can find a direct parse that doesn't rely on the store.

The parse in German can continue directly as follows:

P⁴(k,s,p,pap) LF: P⁴

 $(R^3 \circ EAT)(k,s,p,pap)$ where R^3 is a variable of type <<e,t>,<e,<e,<e,t>>>> LF: R^3

 $(R^2 o (LET o EAT))(k,s,p,pap)$ where R^2 is a variable of type <<e,t>,<e,<e,t>>> LF: R^2

```
(R^1 \circ (HELP \circ (LET \circ EAT)))(k,s,p,pap) where R^1 is a variable of type <<e,t>,<e,t>> LF: R^1
```

(WILL o (HELP o (LET o EAT)))(k,s,p,pap)

which is:

WILL(HELP(k,LET(s,EAT(p,pap))))

This parse takes as many steps as the Dutch parse, and doesn't use a store either. It differs from the Dutch parse, though, in that **it introduces more search variables than the Dutch parse**: **eight search variables all in all:** ϕ , P¹,P²,P³,P⁴,R³,R²,R¹ of **eight different types**.

If we make the plausible hypotheses that **using a store is costly**, and that **using more search variables of more different type is costly** it follows that both parsing

strategies potentially available in German are **more costly** than the fastest strategy available in Dutch. And this is what Bach et. al. 1986 found.