

SPR4106 – Syntax and semantics in formal terms

Lecture II: Functional structure

12 February 2015

What's in an f-structure?

- The functional structure is where LFG models *grammatical functions*
- Grammatical functions do not correspond one to one to either c-structure or semantics (thematic roles), so they need a separate representation
- The functional structure is modelled in an *attribute-value matrix*, i.e. a set of attributes with certain values
- The attributes are either grammatical functions (subject, object etc.), or syntactically relevant features such as *tense*, *number*, *gender* etc.
- Three types of values:
 - linguistic “atoms” such as *plural*, *singular*, *present* etc.
 - new, embedded feature structures
 - semantic symbols, “predicators”

Sample f-structure

PRED	'SEE <SUBJ, OBJ>'				
TENSE	PRESENT				
SUBJ	<table><tr><td>PRED</td><td>'PETER'</td></tr><tr><td>NUMBER</td><td>SG</td></tr></table>	PRED	'PETER'	NUMBER	SG
PRED	'PETER'				
NUMBER	SG				
OBJ	<table><tr><td>PRED</td><td>'MARY'</td></tr><tr><td>NUMBER</td><td>SG</td></tr></table>	PRED	'MARY'	NUMBER	SG
PRED	'MARY'				
NUMBER	SG				

- The value of PRED is always a *semantic symbol*
- Syntactic features such as TENSE and NUMBER take atomic values such as PRESENT
- The values of grammatical functions are feature structures

Features

- There is a “received set” of grammatical functions in LFG, but no corresponding received set of features that are not grammatical functions
- The usual suspects are traditional features like TENSE, NUMBER, GENDER, DEF; there are also less typical features like PCASE
- Will vary with the morphological resources of the language, but should *not* be equated with morphological features
 - Morphological features are only present in the f-structure if they are *syntactically relevant*
 - Words can “speak about their environment”: contribute features to other f-structures than their own (agreement)

Semantic symbols

- Semantically contentful words are represented with semantic symbols

saw 'SEE <SUBJ, OBJ> '
John 'JOHN'
him 'PRO'
rains 'RAIN < > SUBJ'

- the semantic symbol includes a representation of the meaning (conventionally in English)
- words that require arguments also list these
- semantic arguments are listed inside angular brackets, non-thematic (purely syntactic) arguments outside
- unique to each instance!

Argument functions

SUBJ	subject
OBJ	object
OBJ2	second object (NB: sometimes called OBJ _θ)
OBL _θ	oblique
COMP	complement clause
POSS	(certain) arguments of nouns

Non-argument functions

FOCUS	focus
TOPIC	topic
ADJ	adjunct

Classifying grammatical functions

- In stating generalizations it is often useful to refer to certain classifications of GFs
- SUBJ, OBJ and OBJ₂ are collectively known as core functions or term functions
- Many grammatical process are sensitive to the functional hierarchy
 - SUBJ > OBJ > OBJ₂ > OBL_θ
- TOPIC, FOCUS and sometimes SUBJ are referred to as grammaticized **discourse functions** or **overlay functions**

Subcategorization

- Our VP rule $\rightarrow V DP DP PP^* (IP|CP)$ would seem to allow the following sentences (given that all nodes are optional)
 - (1) I donated a book to the library.
 - (2) *I donated to the library.
 - (3) *I donated the university a book to the library.
- The c-structures are all well-formed. Instead, the ungrammaticality is accounted for at f-structure

(In)completeness

PRED	'DONATE <SUBJ, OBJ, OBL _{goal} OBJ>'				
SUBJ	["I"]				
OBL _{goal}	<table><tr><td>PCASE</td><td>OBL_{goal}</td></tr><tr><td>OBJ</td><td>["THE LIBRARY"]</td></tr></table>	PCASE	OBL _{goal}	OBJ	["THE LIBRARY"]
PCASE	OBL _{goal}				
OBJ	["THE LIBRARY"]				

(4) *I donated to the library.

- All argument functions specified in the value of the PRED feature must be present in the local f-structure

(In)coherence

PRED	'DONATE <SUBJ, OBJ, OBL _{goal} OBJ>'				
SUBJ	["I"]				
OBJ	["A BOOK"]				
OBJ ₂	["THE UNIVERSITY"]				
OBL _{goal}	<table><tr><td>PCASE</td><td>OBL_{goal}</td></tr><tr><td>OBJ</td><td>["THE LIBRARY"]</td></tr></table>	PCASE	OBL _{goal}	OBJ	["THE LIBRARY"]
PCASE	OBL _{goal}				
OBJ	["THE LIBRARY"]				

(5) *I donated the university a book to the library.

- All argument functions in an f-structure must be selected by their local PRED

Grammaticality

PRED	‘DONATE <SUBJ, OBJ, OBL _{goal} OBJ>’
SUBJ	[“I”]
OBJ	[“A BOOK”]
OBL _{goal}	[PCASE OBL _{goal} OBJ [“THE LIBRARY”]]

Expletives

- Some lexical items – e.g. expletives (*there, it*) and idiom chunks (*keep the tabs on*) – are meaningless; they do not provide a PRED-value
- Some predicates, e.g. *rain*, require purely syntactic (non-thematic) arguments
- We modify completeness and coherence to account for these

Completeness All argument functions specified in the value of the PRED feature must be present in the local f-structure. All functions that receive a thematic role must have a PRED feature.

Coherence All argument functions in an f-structure must be selected by their local PRED. Any argument function that has its own PRED feature must be assigned a thematic role

What goes wrong here?

(6) *We rain

[PRED	'RAIN < >SUBJ']
	SUBJ	PRED	'PRO'	
		NUMBER	PLURAL	
	PERSON	1		

- Incoherent!

What goes wrong here?

(7) *I donated there to the library

PRED	'DONATE <SUBJ, OBJ, OBL _{goal} OBJ>'			
SUBJ	PRED	'PRO'		
	PERSON	1		
	NUMBER	PL		
OBJ	PERSON	3		
OBL _{goal}	PCASE	OBL _{goal}		
	OBJ	PRED	'LIBRARY'	
		DEF	+	
		PERSON	3	

Extended coherence

- What about discourse functions and adjuncts?
- There something wrong with these:
 - *It that came rained.
 - *The man who I saw the woman crossed the street.
- Intuitively, meaningless items cannot be modified, and discourse functions (*who*) must not “dangle”

Extended coherence

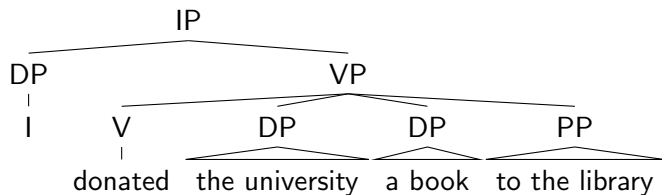
All functions in an f-structure must be incorporated into the semantics. Argument functions are subject to the Coherence condition. Discourse functions must be identified with arguments or adjuncts. Adjuncts must be in f-structures containing PREDs.

Uniqueness

- Every attribute has a single value.
- This falls out of the formal setup of LFG
- So we disallow f-structures with, say, two different tense values or two different objects
- What about adjuncts?

Relating c- and f-structures

(8) *I donated the university a book to the library.



PRED	'DONATE <SUBJ, OBJ, OBL _{goal} OBJ>'
SUBJ	["I"]
OBJ	["A BOOK"]
OBJ2	["THE UNIVERSITY"]
OBL _{goal}	[PCASE OBL _{goal} OBJ ["THE LIBRARY"]]

Relating c- and f-structure

- The c-structure is well-formed by the phrase structure rules and the first f-structure is well-formed by the principles we just saw, and yet something is clearly wrong.
- Informally, the c-structure and the f-structure do not *correspond* in the required way
- The second, incoherent f-structure is intuitively the correct correspondent to the c-structure
- Intuitively, the f-structure that corresponds to the c-structure is the one that contains all the information in the c-structure (and nothing more)
- Technically, we will say that a c-structure and an f-structure correspond iff the f-structure is the minimal solution to the f-description offered by the c-structure

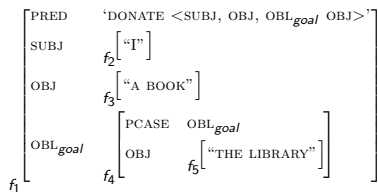
F-structures as functions

$$f_1 \left[\begin{array}{l} \text{PRED} \quad \text{'DONATE <SUBJ, OBJ, OBL}_{goal} \text{ OBJ>'} \\ \text{SUBJ} \quad f_2 \left[\text{"I"} \right] \\ \text{OBJ} \quad f_3 \left[\text{"A BOOK"} \right] \\ \text{OBL}_{goal} \quad f_4 \left[\begin{array}{l} \text{PCASE} \quad \text{OBL}_{goal} \\ \text{OBJ} \quad f_5 \left[\text{"THE LIBRARY"} \right] \end{array} \right] \end{array} \right]$$

- $f_1(\text{SUBJ}) = f_2$, or in LFG notation $(f_1 \text{ SUBJ}) = f_2$
- $(f_1 \text{ OBJ}) = f_3$, $(f_1 \text{ OBL}_{goal}) = f_4$, $(f_4 \text{ OBJ}) = f_5$
- $(f_1 \text{ PRED}) = \text{'donate <SUBJ, OBJ, OBL}_{goal} \text{ OBJ>'}$
- $(f_4 \text{ PCASE}) = \text{OBL}_{goal}$

F-descriptions

- Equations such as $(f_1 \text{ SUBJ}) = f_2$ are known as functional descriptions
- We extracted f-descriptions from the f-structure, but it works the other way around too: we can build an f-structure from the f-description



- $(f_1 \text{ SUBJ}) = f_2$
- $(f_1 \text{ OBJ}) = f_3$
- $(f_1 \text{ OBL}_{goal}) = f_4$
- $(f_4 \text{ OBJ}) = f_5$
- $(f_1 \text{ PRED}) = \text{'donate <SUBJ, OBJ, OBL}_{goal} \text{ OBJ >'}$
- $(f_4 \text{ PCASE}) = \text{OBL}_{goal}$

The f-structure contains all and only the information in the f-descriptions

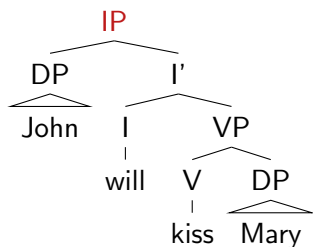
An aside: Identification

- We have several statements about f_1 , e.g.
 - $(f_1 \text{ SUBJ}) = f_2$
 - $(f_1 \text{ OBJ}) = f_3$
- We could also state this in the following way:
 - $(f_1 \text{ SUBJ}) = f_2$
 - $(f_5 \text{ OBJ}) = f_3$
 - $f_1 = f_5$
- The minimal solution remains the same, because the labels aren't essential

$$f_1, f_5 \begin{bmatrix} \text{SUBJ} & f_2 \begin{bmatrix} \end{bmatrix} \\ \text{OBJ} & f_3 \begin{bmatrix} \end{bmatrix} \end{bmatrix}$$

Unification

- Why would we want to do such a thing?
- Syntactic information can arise in different places in the c-structure
- We want to be able to *unify* this information in a single f-structure



- The IP node “knows” that *John* is the subject
- The VP node “knows” that *Mary* is the object
- The I node “knows” that the tense is future
- The V node “knows” that PRED is ‘kiss <SUBJ, OBJ >’
- We want to *unify* this information

Unification II

- The unification of two f-structures A and B is the f-structure C such that it contains all attribute value-pairs of from A and B
- So we collect all features from both f-structures
- If there are duplicated attribute-value pairs, that is not a problem
- If there are conflicting values for the same attribute, the result will fail uniqueness and hence not be a licit f-structure

Unification III

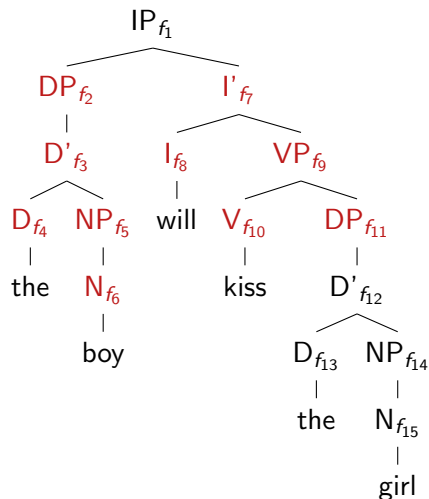
$$\begin{bmatrix} \text{NUMBER} & \text{SG} \\ \text{PERSON} & 3 \end{bmatrix} + \begin{bmatrix} \text{GENDER} & \text{FEM} \end{bmatrix} = \begin{bmatrix} \text{NUMBER} & \text{SG} \\ \text{PERSON} & 3 \\ \text{GENDER} & \text{FEM} \end{bmatrix}$$

$$\begin{bmatrix} \text{NUMBER} & \text{SG} \\ \text{PERSON} & 3 \\ \text{GENDER} & \text{FEM} \end{bmatrix} + \begin{bmatrix} \text{GENDER} & \text{FEM} \end{bmatrix} = \begin{bmatrix} \text{NUMBER} & \text{SG} \\ \text{PERSON} & 3 \\ \text{GENDER} & \text{FEM} \end{bmatrix}$$

$$\begin{bmatrix} \text{NUMBER} & \text{SG} \\ \text{PERSON} & 3 \\ \text{GENDER} & \text{FEM} \end{bmatrix} + \begin{bmatrix} \text{GENDER} & \text{MASC} \end{bmatrix} = \begin{bmatrix} \text{NUMBER} & \text{SG} \\ \text{PERSON} & 3 \\ \text{GENDER} & \text{FEM} \\ \text{GENDER} & \text{MASC} \end{bmatrix}$$

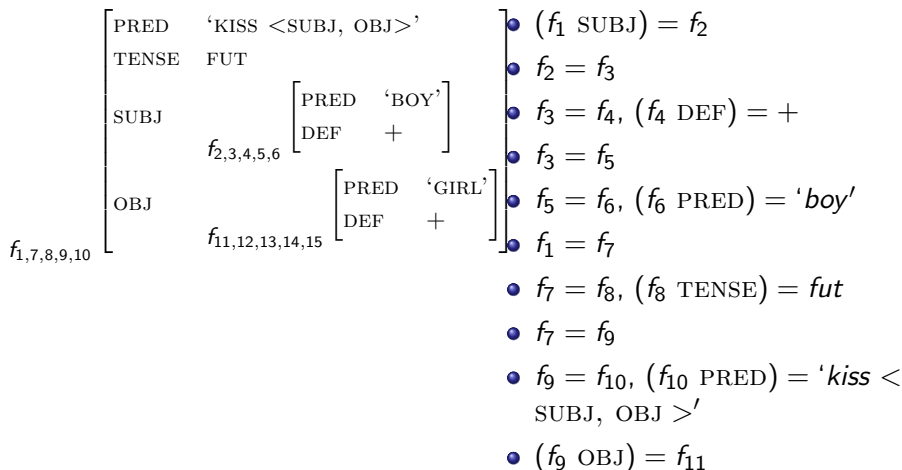
$$\begin{bmatrix} \text{PRED} & \text{'RAIN < > SUBJ'} \\ \text{PRED} & \text{'RAIN < > SUBJ'}_1 \\ \text{PRED} & \text{'RAIN < > SUBJ'}_2 \end{bmatrix} + \begin{bmatrix} \text{PRED} & \text{'RAIN < > SUBJ'} \end{bmatrix} =$$

Back to the c-/f-structure mapping

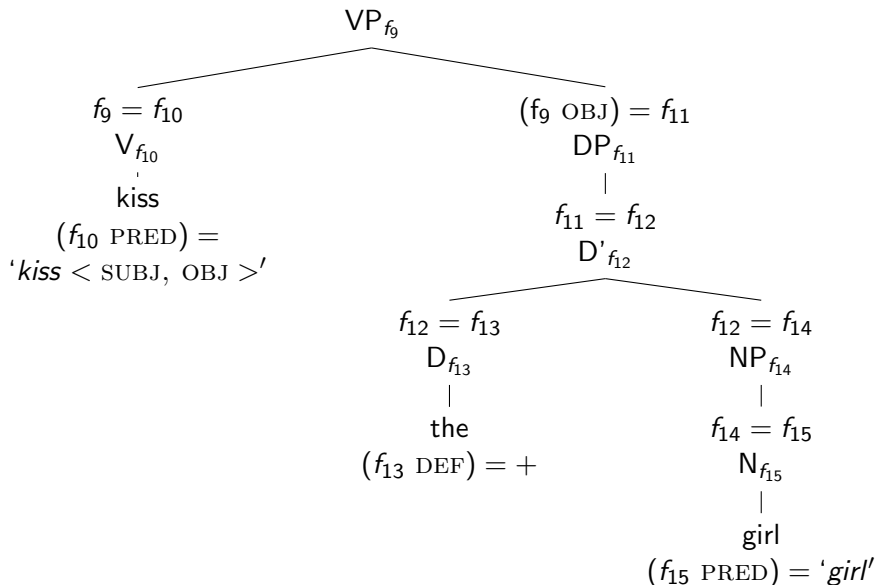


- $(f_1 \text{ SUBJ}) = f_2$
- $f_2 = f_3$
- $f_3 = f_4, (f_4 \text{ DEF}) = +$
- $f_3 = f_5$
- $f_5 = f_6, (f_6 \text{ PRED}) = \text{'boy'}$
- $f_1 = f_7$
- $f_7 = f_8, (f_8 \text{ TENSE}) = \text{fut}$
- $f_7 = f_9$
- $f_9 = f_{10}, (f_{10} \text{ PRED}) = \text{'kiss < SUBJ, OBJ >'}$
- $(f_9 \text{ OBJ}) = f_{11}$

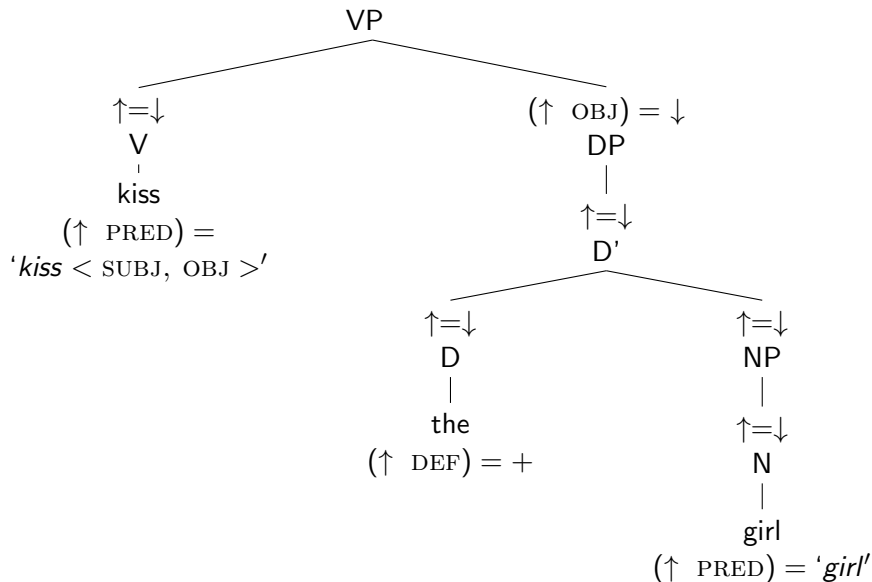
The minimal solution



The tree revisited



Introducing metavariables

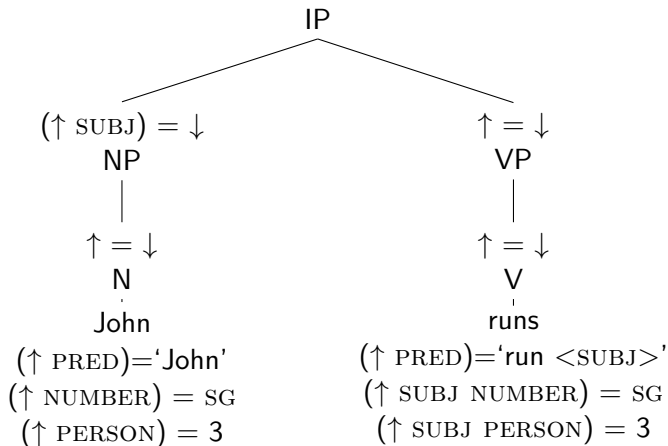


Designators

- \downarrow and \uparrow are metavariables referring to the f-structure of the current node and of the mother of the current node respectively
- We can form complex designators or “paths” through the f-structure
 - $(\uparrow\text{SUBJ}) \equiv$ my mother’s f-structure’s subject
 - $(\uparrow\text{COMP SUBJ}) \equiv$ my mother’s f-structure’s complement’s subject
 - $(\uparrow\text{GF}^*) \equiv$ an f-structure arbitrarily embedded under my mother’s f-structure
- We can go the other way (“outside-in”):
 - $(\text{SUBJ } \uparrow) \equiv$ the f-structure that my mother is the subject of
 - $((\text{SUBJ } \uparrow) \text{OBJ}) \equiv$ the object of the f-structure that my mother is the subject of

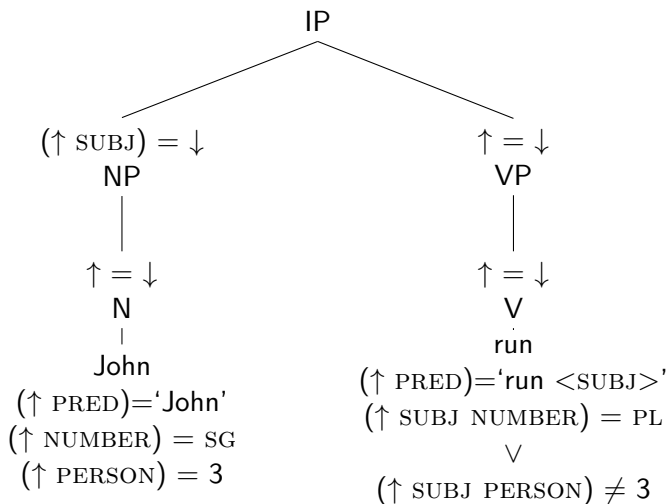
Agreement

- How can we capture agreement with lexical information?



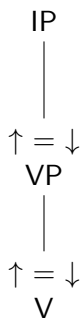
Agreement

- This one is ungrammatical - how can we capture that?



Prodrop

- This one is grammatical in Italian - how can we capture that?



canta

(↑ PRED) = 'sing <SUBJ>'

(↑ SUBJ NUMBER) = SG

(↑ SUBJ PERSON) = 3

((↑ SUBJ PRED) = 'pro')

Annotated phrase structure rules

Functional maximal projections

CP → XP C'
(↑ FOCUS) = ↓ ↑ = ↓

IP → (DP|CP|PP) I'
(↑ SUBJ) = ↓ ↑ = ↓

DP → DP D'
(↑ POSS) = ↓ ↑ = ↓
(↑ DEF) = +
(↓ CASE) =_c GEN

Functional single-bar projections

$C' \rightarrow C \quad IP$
 $\uparrow = \downarrow \quad \uparrow = \downarrow$

$I' \rightarrow I \quad VP$
 $\uparrow = \downarrow \quad \uparrow = \downarrow$

$D' \rightarrow D \quad NP$
 $\uparrow = \downarrow \quad \uparrow = \downarrow$

Lexical phrases

VP	→	V	DP*	PP*	(IP CP)
		$\uparrow = \downarrow$	$(\uparrow \text{ OBJ}) = \downarrow \vee$ $(\uparrow \text{ OBJ2}) = \downarrow$	$(\uparrow \text{ OBL}_{\theta} = \downarrow)$	$(\uparrow \text{ COMP}) = \downarrow$
PP	→	P	DP	PP	IP
		$\uparrow = \downarrow$	$(\uparrow \text{ OBJ}) = \downarrow$	$(\uparrow \text{ OBL}_{\theta} = \downarrow)$	$(\uparrow \text{ COMP}) = \downarrow$
NP	→	N	PP*	CP	
		$\uparrow = \downarrow$	$(\uparrow \text{ OBL}_{\theta} = \downarrow)$	$(\uparrow \text{ COMP}) = \downarrow$	
AP	→	A	PP	(IP CP)	
		$\uparrow = \downarrow$	$(\uparrow \text{ OBL}_{\theta} = \downarrow)$	$(\uparrow \text{ COMP}) = \downarrow$	

Constraining equations

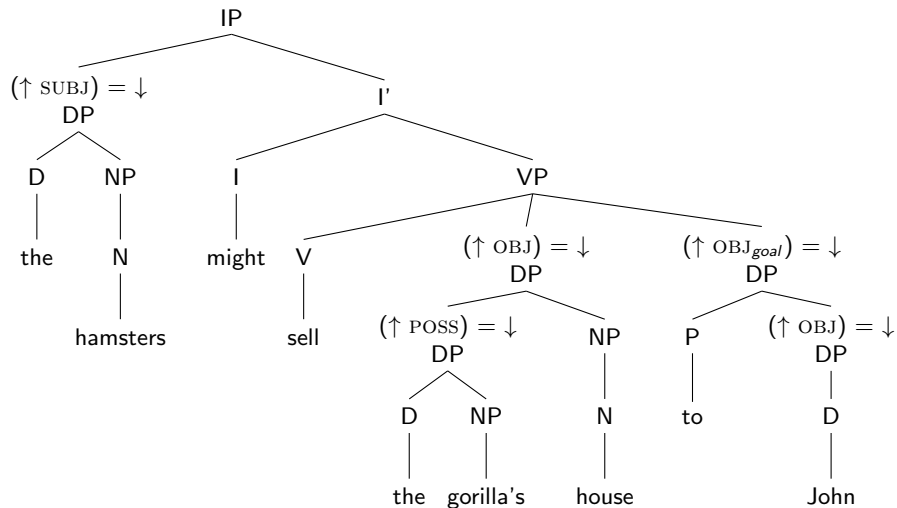
- We have already seen a couple of so-called constraining equations
 - $\text{PERSON} \neq 3$
 - $\text{CASE} =_c \text{GEN}$
- *Existential equations* are another type
 - $(\uparrow \text{TENSE})$ means TENSE should have *some* value
 - $\neg(\uparrow \text{TENSE})$ means TENSE should have *not* have *any* value
 - Useful for capturing the selectional restrictions of the complementizers *to* and *that*

Exercises: English (adapted from Falk 2001)

(9) The hamsters might sell the gorilla's house to John

the	D	DEF = +
hamsters	N	(↑ PRED) = 'hamster' (↑ NUMBER) = PL
might	I	(↑ TENSE) = PRES (↑ MOOD) = POSSIBILITY (↑ VFORM) = _c INF
sell	V	(↑ PRED) = 'sell <SUBJ, OBJ, OBL _θ >' (↑ VFORM) = INF
gorilla's	N	(↑ PRED) = 'gorilla' (↑ NUMBER) = SG (↑ CASE) = GEN (POSS ↑)
house	N	(↑ PRED) = 'house <POSS>' (↑ NUMBER) = SG
to	P	(↑ PRED) = 'to <OBJ>'
John	D	(↑ PRED) = 'John' (↑ NUMBER) = SG

The solution: c-structure



The solution: f-structure

SUBJ	<table style="border-collapse: collapse;"> <tr> <td style="padding-right: 10px;">DEF</td> <td style="padding-left: 10px;">+</td> </tr> <tr> <td style="padding-right: 10px;">PRED</td> <td style="padding-left: 10px;">‘HAMSTER’</td> </tr> <tr> <td style="padding-right: 10px;">NUMBER</td> <td style="padding-left: 10px;">PL</td> </tr> </table>	DEF	+	PRED	‘HAMSTER’	NUMBER	PL								
DEF	+														
PRED	‘HAMSTER’														
NUMBER	PL														
TENSE	PRES														
MOOD	POSSIBILITY														
PRED	‘SELL <SUBJ, OBJ, OBL _{goal} >’														
VFORM	INF														
OBJ	<table style="border-collapse: collapse;"> <tr> <td style="padding-right: 10px;">PRED</td> <td style="padding-left: 10px;">HOUSE</td> </tr> <tr> <td style="padding-right: 10px;">DEF</td> <td style="padding-left: 10px;">+</td> </tr> <tr> <td style="padding-right: 10px;">NUMBER</td> <td style="padding-left: 10px;">SG</td> </tr> <tr> <td style="padding-right: 10px;">POSS</td> <td style="padding-left: 10px;"> <table style="border-collapse: collapse;"> <tr> <td style="padding-right: 10px;">DEF</td> <td style="padding-left: 10px;">+</td> </tr> <tr> <td style="padding-right: 10px;">NUMBER</td> <td style="padding-left: 10px;">SG</td> </tr> <tr> <td style="padding-right: 10px;">CASE</td> <td style="padding-left: 10px;">GEN</td> </tr> </table> </td> </tr> </table>	PRED	HOUSE	DEF	+	NUMBER	SG	POSS	<table style="border-collapse: collapse;"> <tr> <td style="padding-right: 10px;">DEF</td> <td style="padding-left: 10px;">+</td> </tr> <tr> <td style="padding-right: 10px;">NUMBER</td> <td style="padding-left: 10px;">SG</td> </tr> <tr> <td style="padding-right: 10px;">CASE</td> <td style="padding-left: 10px;">GEN</td> </tr> </table>	DEF	+	NUMBER	SG	CASE	GEN
PRED	HOUSE														
DEF	+														
NUMBER	SG														
POSS	<table style="border-collapse: collapse;"> <tr> <td style="padding-right: 10px;">DEF</td> <td style="padding-left: 10px;">+</td> </tr> <tr> <td style="padding-right: 10px;">NUMBER</td> <td style="padding-left: 10px;">SG</td> </tr> <tr> <td style="padding-right: 10px;">CASE</td> <td style="padding-left: 10px;">GEN</td> </tr> </table>	DEF	+	NUMBER	SG	CASE	GEN								
DEF	+														
NUMBER	SG														
CASE	GEN														
OBL _{goal}	<table style="border-collapse: collapse;"> <tr> <td style="padding-right: 10px;">PRED</td> <td style="padding-left: 10px;">‘TO <OBJ>’</td> </tr> <tr> <td style="padding-right: 10px;">OBJ</td> <td style="padding-left: 10px;"> <table style="border-collapse: collapse;"> <tr> <td style="padding-right: 10px;">PRED</td> <td style="padding-left: 10px;">‘JOHN’</td> </tr> </table> </td> </tr> </table>	PRED	‘TO <OBJ>’	OBJ	<table style="border-collapse: collapse;"> <tr> <td style="padding-right: 10px;">PRED</td> <td style="padding-left: 10px;">‘JOHN’</td> </tr> </table>	PRED	‘JOHN’								
PRED	‘TO <OBJ>’														
OBJ	<table style="border-collapse: collapse;"> <tr> <td style="padding-right: 10px;">PRED</td> <td style="padding-left: 10px;">‘JOHN’</td> </tr> </table>	PRED	‘JOHN’												
PRED	‘JOHN’														

Exercises: Warlpiri (adapted from Bresnan 2001)

- (10) Kurdu-jarra-rlu wita-jarra-rlu ka-pala maliki
child-DUAL-ERG small-DUAL-ERG PRES-DUAL dog-ABS
wajilipi-nyi
chase-NONPAST
'The two small children are chasing the dog.'
- (11) Kurdu-jarra-rlu ka-pala maliki wajilipi-nyi
child-DUAL-ERG PRES-DUAL dog-ABS chase-NONPAST
wita-jarra-rlu
small-DUAL-ERG
'The two small children are chasing the dog.'
- (12) Maliki ka-pala kurdu-jarra-rlu wajilipi-nyi
dog-ABS PRES-DUAL child-DUAL-ERG chase-NONPAST
wita-jarra-rlu
small-DUAL-ERG
'The two small children are chasing the dog.'

Exercises: Warlpiri (adapted from Bresnan 2001)

S	→	X	(Aux)	X*	(where X = NP, V)
		(↑(SUBJ OBJ)) = ↓	↑=↓	(↑(SUBJ OBJ))	
NP	→	N*			
		↑=↓			
<i>kurdu-jarra-rlu</i>		N	(↑ PRED) = 'child'		
			(↑ NUM) = 'dual'		
			(↑ CASE) = 'erg'		
<i>maliki</i>		N	(↑ PRED) = 'dog'		
			(↑ NUM) = 'sg'		
			(↑ CASE) = 'abs'		
<i>wita-jarra-rlu</i>		N	(↑ ADJ PRED) = 'small'		
			(↑ NUM) = 'dual'		
			(↑ CASE) = 'erg'		
<i>wajilipi-nyi</i>		V	(↑ PRED) = 'chase <SUBJ, OBJ>'		
			(↑ TENSE) = NONPAST		
			(↑ SUBJ CASE) = 'dual'		
			(↑ OBJ CASE) = 'abs'		
<i>ka-pala</i>		Aux	(↑ ASPECT) = PRESENT.IMPERFECT		
			(↑ SUBJ NUM = = DUAL		