How do syntactic structures emerge?
Communicative approach from dynamic semantics

Ivan Rygaev
Laboratory of computational linguistics
Institute for information transmission problems RAS, Moscow
irygaev@gmail.com

Open workshop of Institute of linguistics RSUH, September 29, 2022

Communication

- Communication is a primary and main function of language
 - Information transmission from the speaker to the hearer
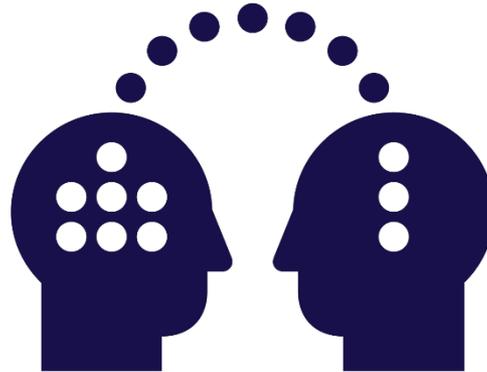
Information transmission

- New information should be linked to the old one:

Ivanov's salary is 50 000 rub.

```
update employees
set salary = 50000
where name = 'Ivanov'
[where id = 2837]
```

id	name	salary	...
...			
2837	Ivanov	50 000	...
...			
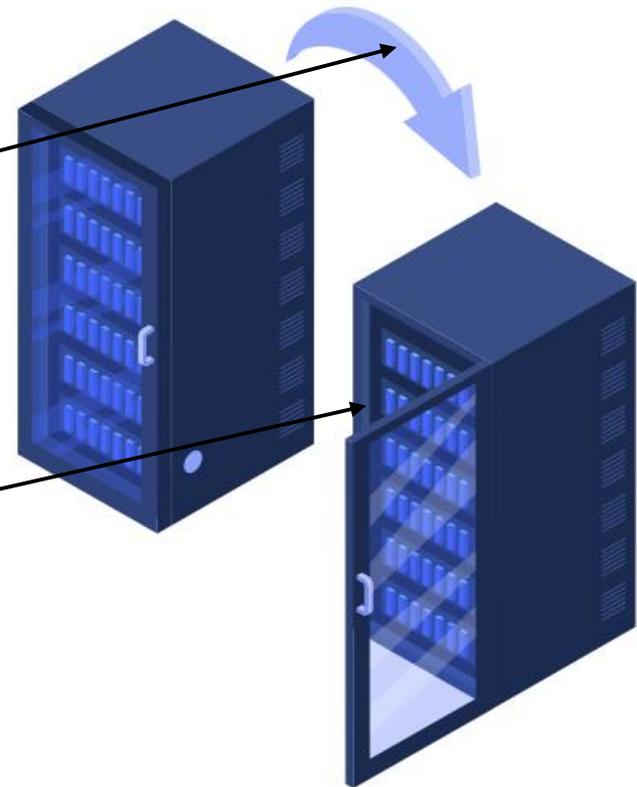
Linguistic communication

- New information should be linked to the old one in the hearer's mind. One part of the sentence serves to find a mental file, the other – to update it.

Ivanov's salary is 50 000 rub.

x
name (x, 'Ivanov')
salary (x, 50000)
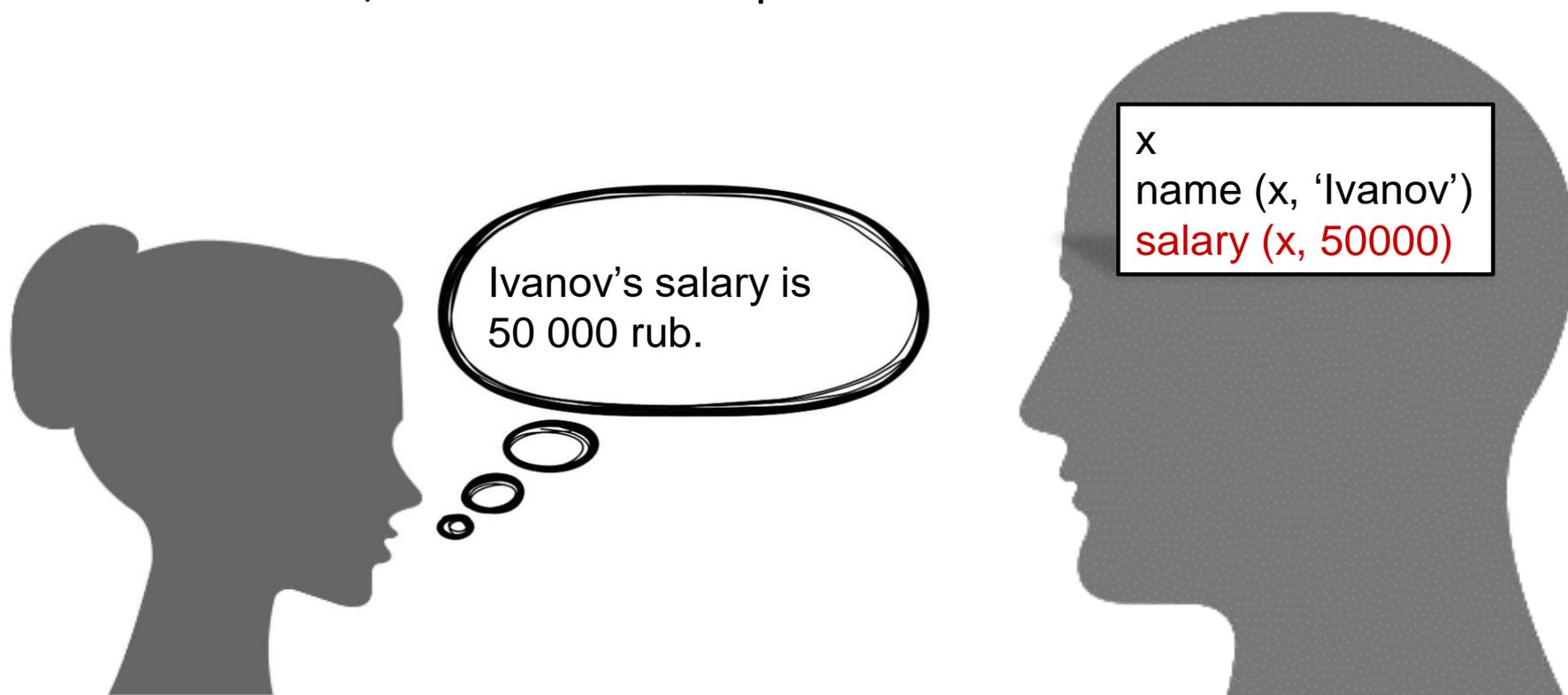
File Change Semantics

- File of cards metaphor:
  - *"B's task is to construct and update a file which, at any point in the conversation, contains all the information that A has conveyed up to that point"* (Heim 1982:178)

- A card corresponds to a *discourse* referent

- The meaning of noun phrases is procedural
  - *"For every indefinite, start a new card; for every definite, update a suitable old card"* (Heim 1982:179)

- Sentence meaning is a context change potential
 - It the entire file, not a sentence, which has truth conditions

File Change Semantics

- *A woman was bitten by a dog*
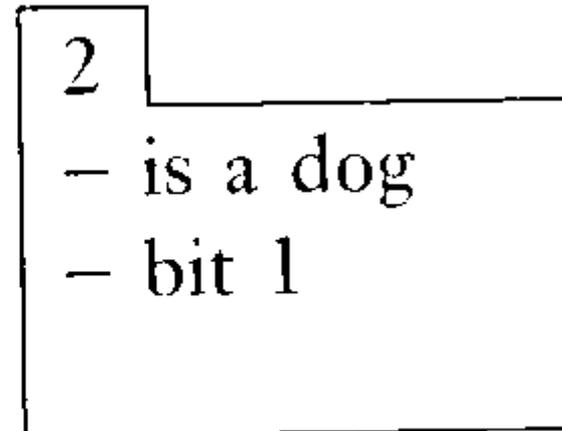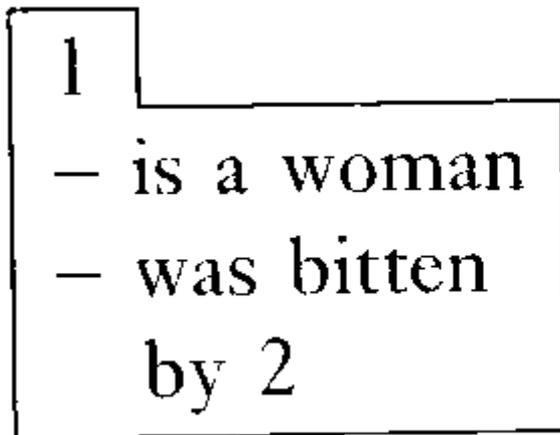
File Change Semantics

- *A woman was bitten by a dog*
- *She hit it*
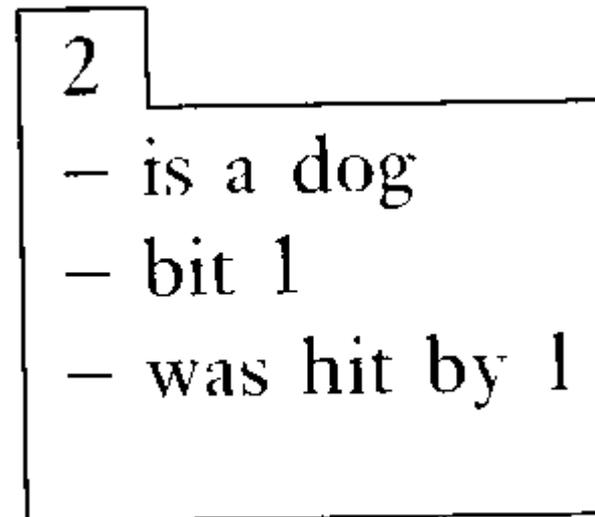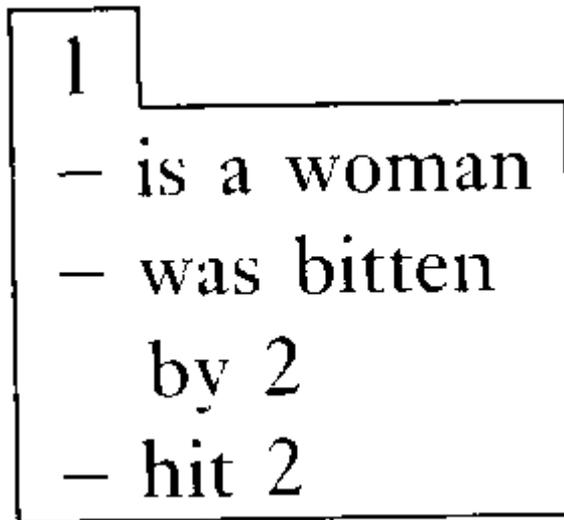
File Change Semantics

- *A woman was bitten by a dog*
- *She hit it*
- *It jumped over the fence*

```
1
– is a woman
– was bitten
  by 2
– hit 2
```

```
2
– is a dog
– bit 1
– was hit by 1
– jumped over 3
```

```
3
– is a fence
– was jumped
  over by 2
```

Discourse representation theory

- *A woman was bitten by a dog*

$$\begin{array}{|l|} \hline x, y \\ \hline \text{woman } (x) \\ \text{dog } (y) \\ \text{bit } (y, x) \\ \hline \end{array}$$

- DRS (discourse representation structure) consists of:
  - Discourse referents (markers, variables)
  - Conditions (properties, predicates)
- DRS is a whole discourse representation (Kamp 1981)

Anaphora resolutions

- *A woman was bitten by a dog*

- *She hit it*

x, y, u, v
woman (x)
dog (y)
bit (y, x)
hit (u, v)
she (u)
it (v)

x, y, u, v
woman (x)
dog (y)
bit (y, x)
hit (u, v)
$u = x$
$v = y$

x, y
woman (x)
dog (y)
bit (y, x)
hit (x, y)

Complex DRSs: implication

- *If a farmer owns a donkey he beats it*
- *Every farmer who owns a donkey beats it*

Complex DRSs: negation

- *John owns no donkey*

- *John does not own a donkey*

$$
\begin{array}{|l|}
\hline
x \\
\hline
\text{John } (x) \\
\\
\neg \quad
\begin{array}{|l|}
\hline
y \\
\hline
\text{donkey } (y) \\
\text{own } (x, y) \\
\hline
\end{array} \\
\\
\hline
\end{array}
$$
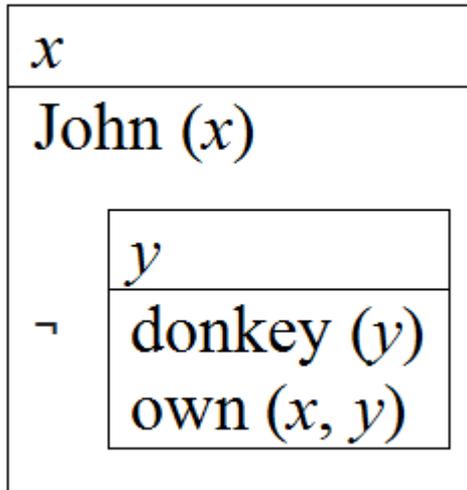
Complex DRSs: disjunction

- *John owns a donkey or a horse*

$$
\boxed{
\begin{array}{l}
x, y \\
\hline
\text{John } (x) \\
\text{owns } (x, y) \\
\\
\boxed{\begin{array}{l} \\ \hline \text{donkey } (y) \end{array}} \lor \boxed{\begin{array}{l} \\ \hline \text{horse } (y) \end{array}}
\end{array}
}
$$
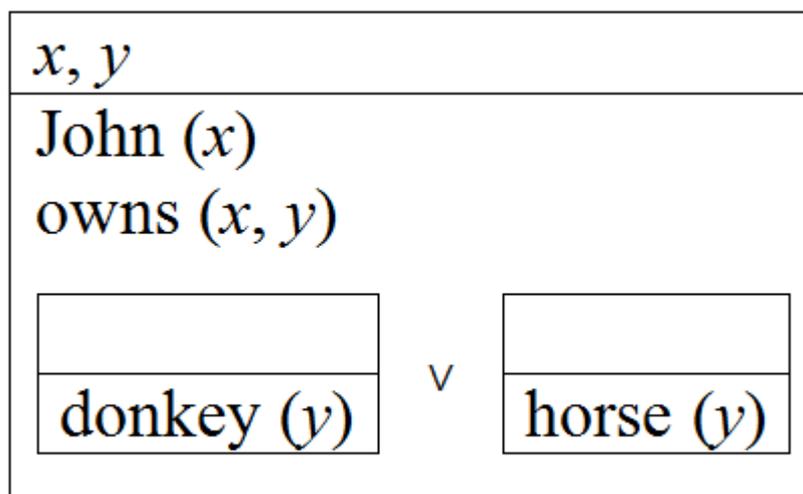
Subordination and accessibility

- DRS B is subordinate to A if (informally):
  - B is embedded into A or
  - 'A => B' is a condition in some other DR

- Accessibility
  - Discourse referent from DRS A is accessible to an (anaphoric) discourse referent in DRS B, just in case B is subordinate to A

Subordination and accessibility

- *If a farmer owns a donkey, he feeds it*

$$\boxed{\begin{array}{|l|} \hline x,\ y \\ \hline \text{farmer } (x) \\ \text{donkey } (y) \\ \text{own } (x,\ y) \\ \hline \end{array}} \Rightarrow \boxed{\begin{array}{|l|} \hline u,\ v \\ \hline \text{he } (u) \\ \text{it } (v) \\ \text{feed } (u,\ v) \\ \boldsymbol{u = x} \\ \boldsymbol{v = y} \\ \hline \end{array}}$$
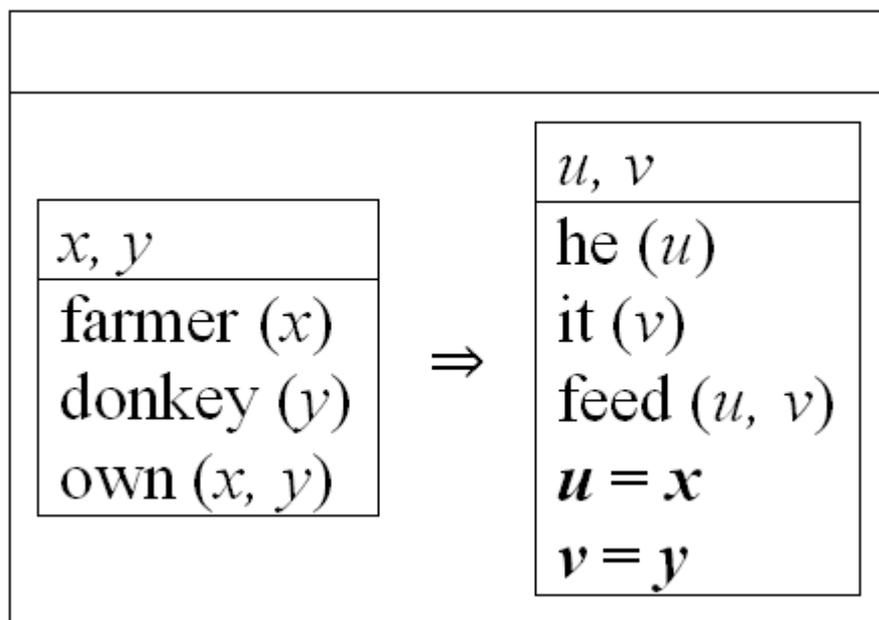
- *x* and *y are accessible* for *u* and *v*, since they are located in a superordinate DRS

Subordination and accessibility

- *Every farmer owns a donkey. *It is grey.*
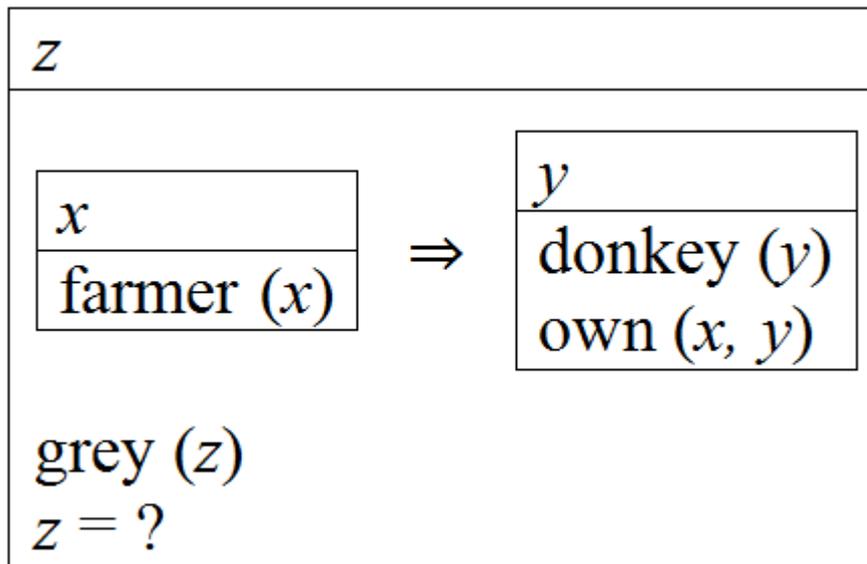
- Neither x, nor y is accessible for z, since they are located in subordinate DRSs

DRT advantages

- Allows the scope of (top level) NPs to be extended indefinitely

- Explains binding of anaphoric pronouns which are not syntactically bound

- Explains impossibility of anaphoric links where the antecedent is inaccessible

Presupposition

- Presupposition is an information which the speaker linguistically marks as taken for granted
  - i.e. already known by the audience
  - i.e. constituting a part of the common ground

Presupposition triggers

- Definite descriptions
  - *The king of France is bald*
  - *> There is a king of France*

- Complements of factive verbs
  - *John knows that the Earth is flat*
  - *> The Earth is flat*

- Clefts
  - *It was John who killed the butcher*
  - *> Somebody killed the butcher*

- Adverbs *even, too, again*, etc.

Presupposition and negation

- Negation does not affect presupposition
- If an affirmative sentence carries a presupposition
  - *The king of France is bald*
  - *> There is a king of France*
- Then its negative counterpart carries the same presupposition
  - *The king of France is **not** bald*
  - *> There is a king of France*

Presupposition projection

- Presuppositions also normally survive under other logical operators:
  - *If Fred has stopped beating Zelda, then Fred no longer resents Zelda's infidelity*
  - *> Fred has been beating Zelda*
  - *> Zelda has been unfaithful*
- And in other complex sentences:
  - *Bill does not know that all of Jack's children are bald*
  - *> All of Jack's children are bald*
  - *> Jack has children*

Presupposition projection

- Sometimes presuppositions seem to disappear in complex sentences:

  – *If Jack has children, then all of **Jack's children** are bald*

  – *Jack has children and all of **Jack's children** are bald*

  – *Either Jack has no children of all of **Jack's children** are bald*

- Presupposition projection problem:

  – Explain in which cases presuppositions disappear and why

  – Determine the presuppositions of a complex sentence out of presuppositions of its parts

Presupposition as anaphora

- Rob van der Sandt (1992) proposed that presupposition and anaphora is essentially the same phenomenon:

  - *Theo has a little rabbit and **his rabbit** is grey*
  - *Theo has a little rabbit and **it** is grey*
  - *If Theo has a rabbit, **his rabbit** is grey*
  - *If Theo has a rabbit, **it** is grey*

Parallels

- Presupposition
  - *Jack has children and all of **Jack's children** are bald*
  - *If Jack has children, then all of **Jack's children** are bald*
  - *Either Jack has no children or all of **Jack's children** are bald*

- Anaphora
  - *John owns a donkey. He beats **it**.*
  - *If John owns a donkey, he beats **it***
  - *Either John does not own a donkey or he beats **it***

Parallels

- VP-anaphora:

  - *If someone solved the problem it was Julius who **{solved it/did}***

  - *If Harry stopped smoking, John **{stopped/did}** too.*

- Sentential anaphora:

  - *If John is ill, Mary regrets **{that/that he is ill}***

  - *If John died, he did see his children before **{that/he did/he died}***

Presupposition as anaphora

- Presuppositions are just anaphors
  - Can be treated by the same mechanism as anaphora resolution
- But unlike pronouns they contain descriptive content
  - They have internal structure that must be represented
  - They can be accommodated – if there is no antecedent found then the information can be just added to the DRS

Presupposition projection in DRT

- First a separate sentence DRS (preliminary DRS) is built and only after that it is merged into the main DRS

- Anaphoric elements are encoded separately in a DRS
  - They are processed only after the sentence DRS is merged into the main DRS
  - In addition to discourse referents and conditions there is now an A-structure – a set of presuppositional A-DRSs
  - Presuppositional A-DRS can have its own A-structure, i.e. they can be embedded into one another

# Binding

- *John has a cat.* *His cat purrs*

$$y, x$$

John (y)
cat (x)
owns (y, x)

purr (x)

$$x$$

cat (x)
owns (y, x)

$$y_{masc}$$

$$u, v$$

John (u)
cat (v)
owns (u, v)

purr (x)

$$x$$

cat (x)
owns (y, x)

$$y_{masc}$$

Binding

- *John has a cat.* His cat purrs

u, v
John (u)
cat (v)
owns (u, v)
purr (x)
x
cat (x)
owns $(\boldsymbol{u}, x)$

u, v
John (u)
cat (v)
owns (u, v)
purr (v)

Accommodation

- *John's cat purrs*
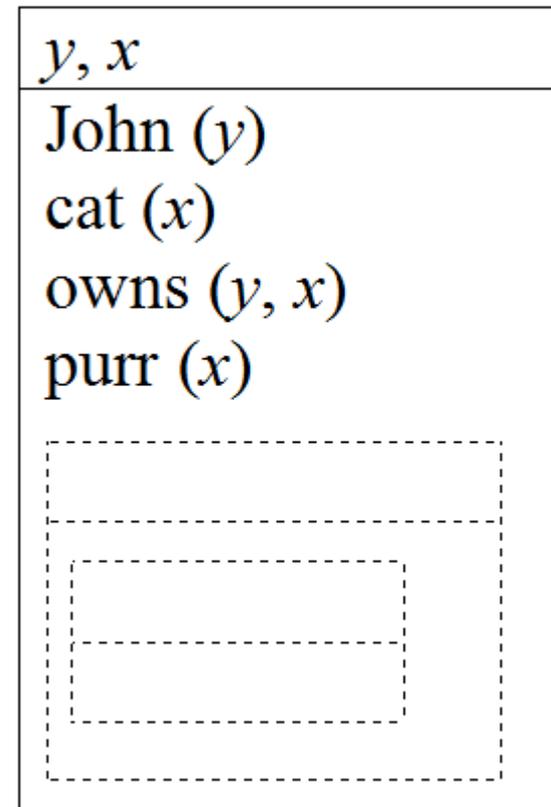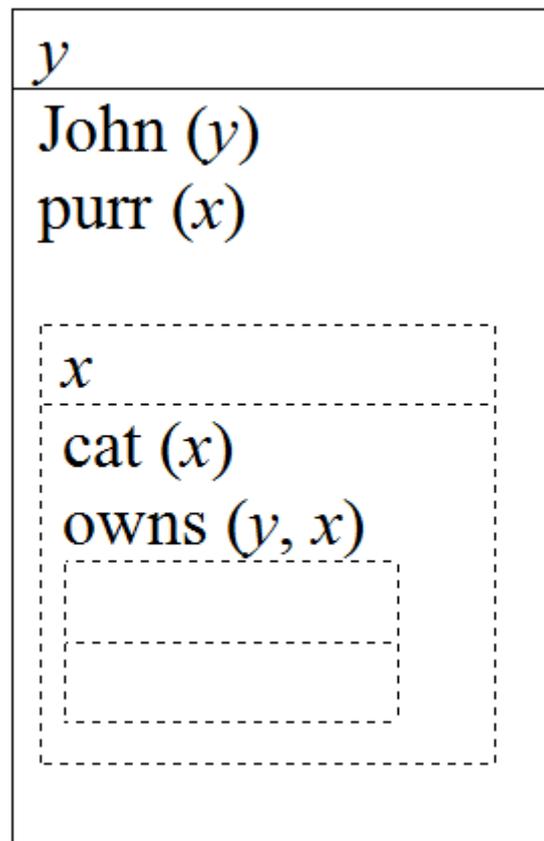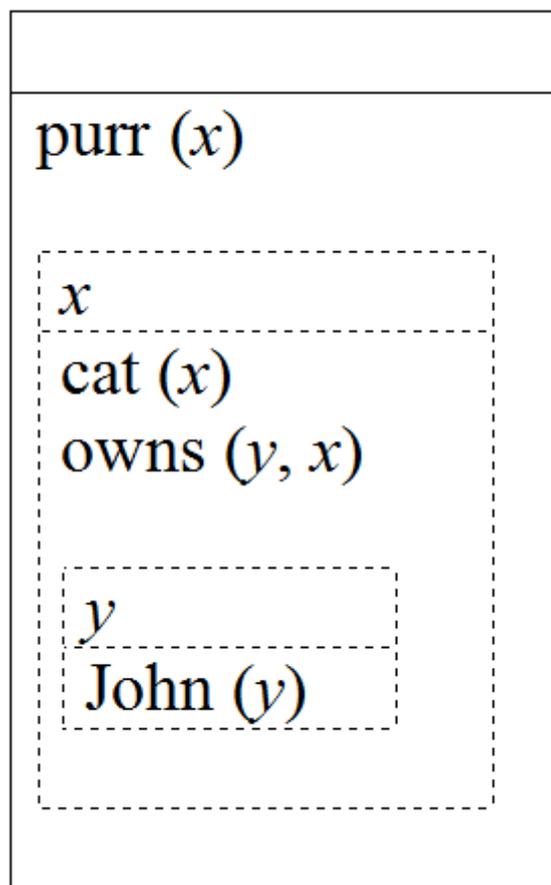
	y	y, x
purr (x)	John (y)	John (y)
	purr (x)	cat (x)
		owns (y, x)
		purr (x)

First box:
- purr (x)
  - x
    - cat (x)
    - owns (y, x)
      - y
        - John (y)

Second box:
- y
- John (y)
- purr (x)
  - x
    - cat (x)
    - owns (y, x)

Third box:
- y, x
- John (y)
- cat (x)
- owns (y, x)
- purr (x)

Presupposition disappears

- *If John has a child, his child is happy*

$$y$$
$$\text{child\_of}\,(y, x)$$

$$x$$
$$\text{John}\,(x)$$

$$\Rightarrow$$

$$\text{happy}\,(z)$$

$$z$$
$$\text{child\_of}\,(z, w)$$

$$w_{masc}$$
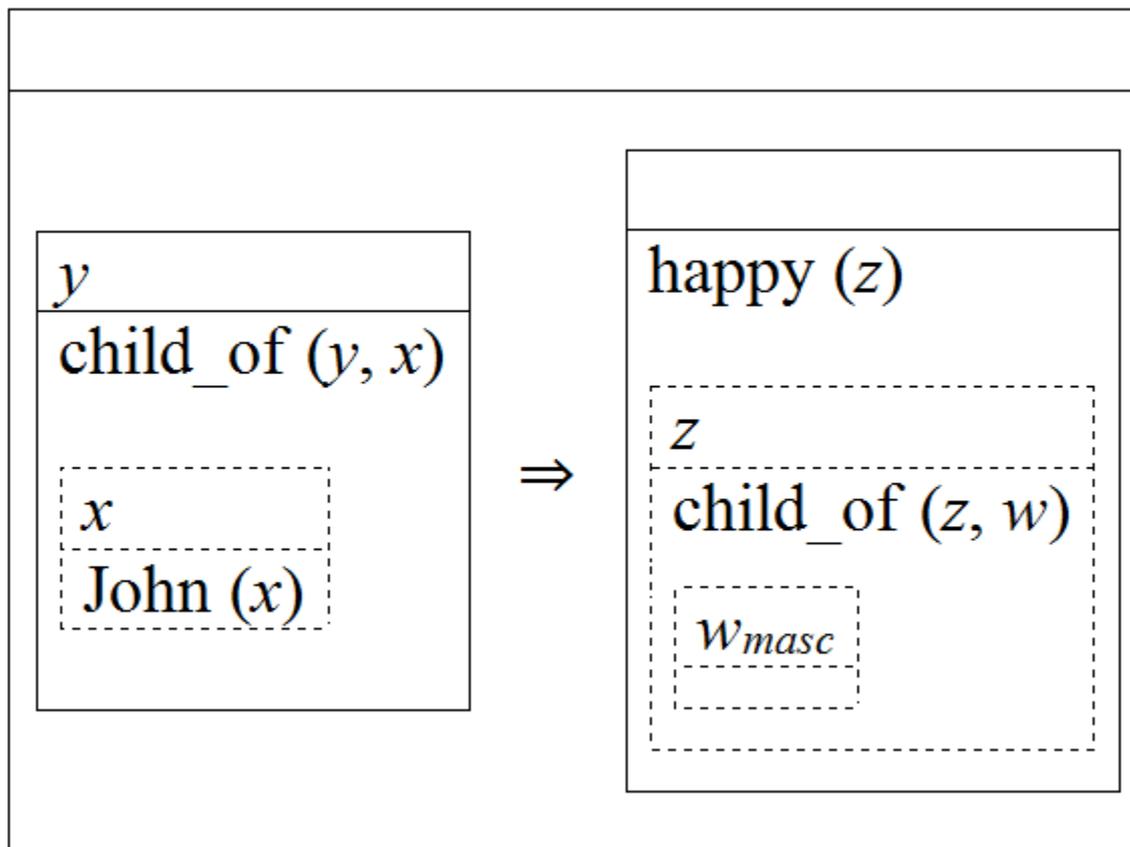
Presupposition disappears

- *If John has a child, his child is happy*
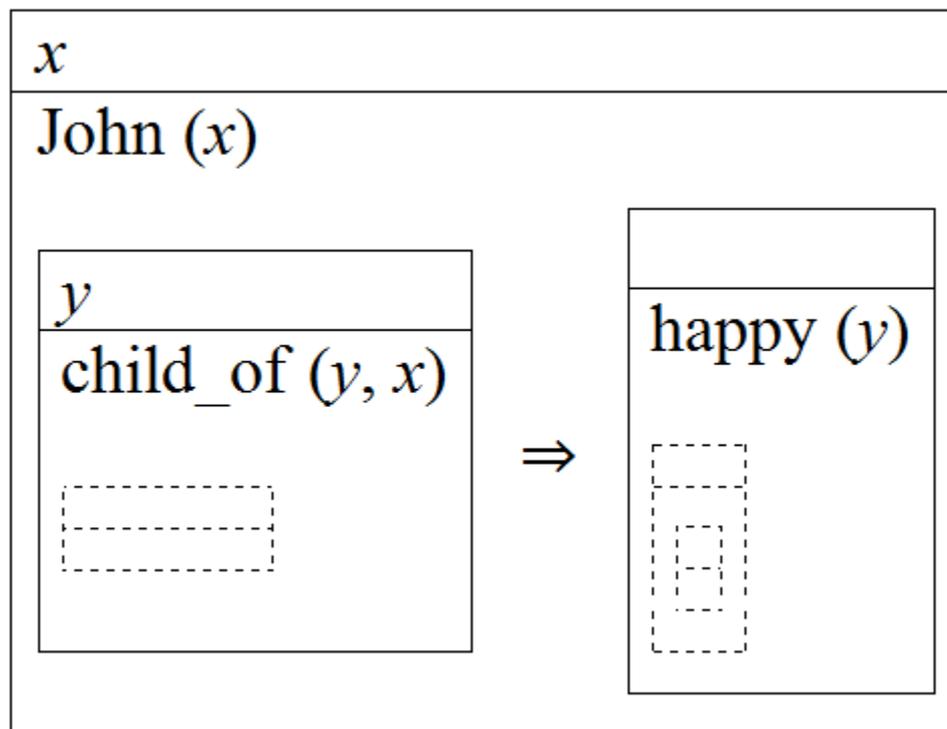
Constraints on resolution

- General algorithm:
  - A-DRS goes up through hierarchy looking for an antecedent
  - If not found, tries to accommodate at the top DRS
  - If failed – goes back down to accommodate where possible

- Constraints on resolution:
  - Main DRS must stay consistent
  - New main DRS must be informative (must not be entailed by the previous one)
  - No subordinate DRS must contradict a superordinate one
  - A subordinate DRS must informative relative to a superordinate one

Specific indefinites

- Specific indefinites (van Geenhoven 1998)
  - Similar to presuppositions
  - They are interpreted not in the place they appear
  - But somewhere higher in the structure
  - They are normally accommodated rather than bound
- Are they a special type of presupposition?
- *Peter intends to visit **a museum** every day*
  - *Has at least three different interpretations*
  - *Depending on the level where 'a museum' is interpreted*

Backgrounding (Geurts 2010)

- Specific indefinites are not presuppositions
  - Accommodation is a repair strategy
  - It would be strange to use it normally as specifics do
- Different types of backgrounding:
  - Presuppositions
  - Specific indefinites
  - "Parenthetical" constructions
- The Buoyancy Principle:
 - Backgrounded material tends to float up towards the main DRS.

DRSs as instructions

- We can use A-DRSs for all backgrounded expressions

- But they have to be marked with their function

  - Propositional A-DRS serves to *find* a discourse referent

  - Specificity A-DRS – to *create* a new discourse referent

  - "Parenthetical" A-DRS – to *update* an existing one

  - Such A-DRSs symbolize instructions for the hearer to update his mental database

- Now we can call them B-DRSs (backgrounded DRS)

- Main DRS is an instruction to update the topic referent

DRSs as instructions

- *Bill saw a certain picture of John, a friend of mine*

update b:

see (b, p)

find b: b
bill (b)

create p: p
picture (p)
of (p, j)

update j: i
friend (i, j)

find j: j
john (j)

---
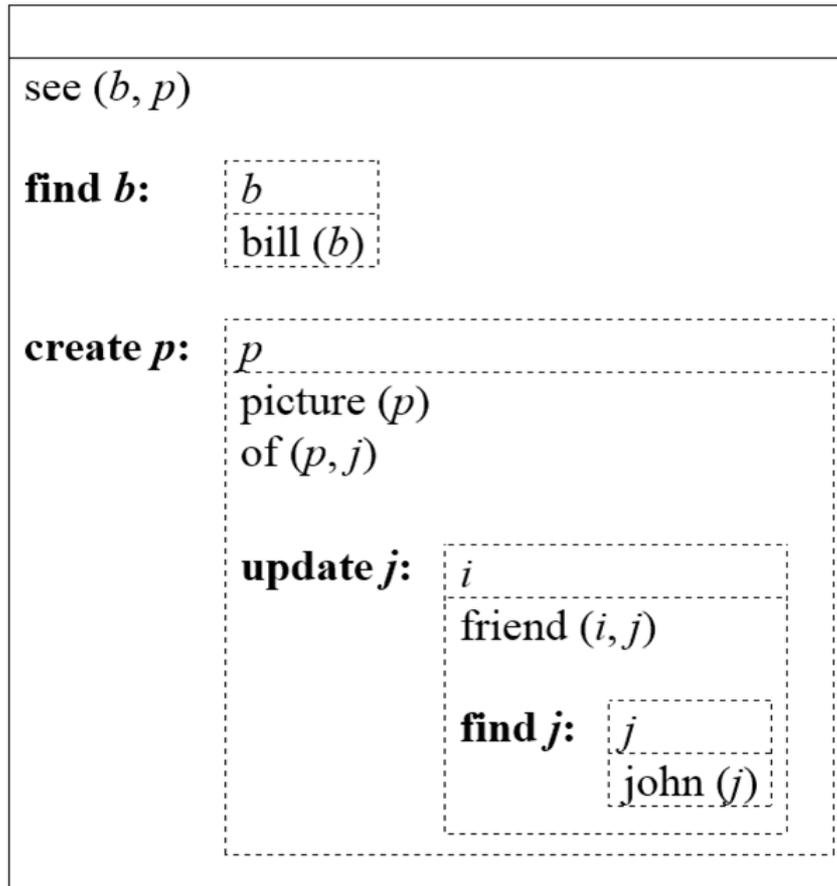
b, p, j, i
bill (b)
see (b, p)
picture (p)
of (p, j)
john (j)
friend (i, j)

Syntax

- The structure of a Preliminary DRS:
  - Each backgrounded constituent corresponds to a B-DRS
  - B-DRS hierarchy forms a tree
  - The tree mirrors the syntactic tree of the sentence

- Sentence production
  - The speaker intends to convey information to the hearer
  - He splits his mental Proper DRS into a set of instructions to find, create or update mental referents in the hearers' mind
  - Being dependent on one another they form a tree
  - The tree is then realized as a syntactic tree of the sentence

Utterances as programs

- Two steps of NLU (Davies & Isard 1972)
  - Compilation
  - Execution
  - Understanding an utterance vs carrying it out

- In our model
  - Compiling instructions = building a Preliminary DRS
  - Executing instructions = resolving B-DRSs to obtain a Proper DRS

- A book is not a knowledge base
  - It is a script to create the knowledge base

Two layers of representation

- Preliminary DRS – sentence representation
    - A sequence of instructions
    - Is completely context-independent
    - But nevertheless is context-sensitive
    - Captures information structure
    - Reflects the syntax on the level of semantics
    - Can serve well as an interlingua for translation
- Proper DRS – mental representation
    - Captures truth conditions

Syntactic islands

- Islands
  - Syntactic constructions which contain an element that cannot be extracted out of it

- Non-island example
  - *Bill saw [a picture of John]*
  - ***Who** did Bill see [a picture of ____]?*

- Island example
  - *Bill saw [the picture of John]*
  - **Who** did Bill see [the picture of ____]?*

Syntactic islands

- More island examples (Newmeyer 2016)
  - *∗**What** did you take a class from [the chef that created ___]?*
  - *∗**What** did you eat [beans and ___]?*
  - ***What** [that Mary solved ___] is likely?*
  - *∗**What** were you happy [because John bought ___]?*
  - ***Whose** did Sue borrow [___ pencil]?*
- Main approaches explaining islands:
  - Syntactic
  - Resource-based
  - Communicative

Communicative approach

- Islands appear because of the clash in the communicative structure of the sentence

- Semantic dominance (Erteschik-Shir 1973)

  – A constituent is semantically dominant if it is not presuppositional and does not have a referent in context

- Condition for extraction

  – An element can be extracted only out of those constituents which can be treated as dominant in some context

"This is a lie" test

- A test for dominance

- *Tom said: "Bill thinks that they're gonna win."*
- *which is a lie — he doesn't.*
- *which is a lie — they're not.*

- *Tom said "John rejoices that they came to the party."*
- *which is a lie — he doesn't.*
- ** which is a lie — they didn't.*

Goldberg 2006

- Backgrounded constituents are islands (BCI)
- An experiment (Ambridge & Goldberg 2008)

Ungrammaticality of extraction
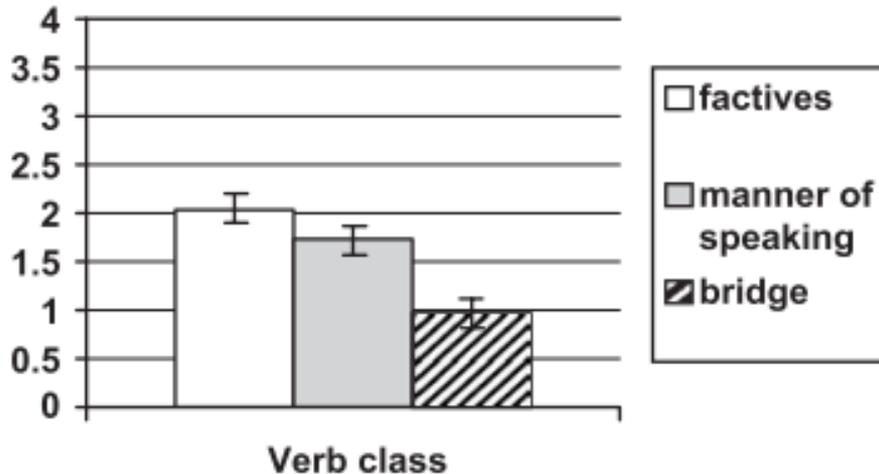
Focality
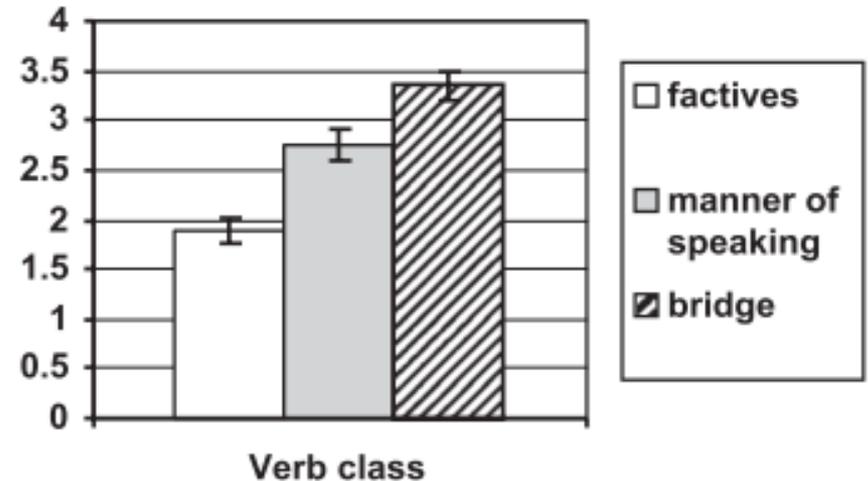
p < 0.001

correlation = 0.83

Explanation

- Why are backgrounded constituents islands?

- The answer
  - Each backgrounded constituent is a separate instruction
  - It is executed separately
  - All discourse referents it depends on must have already been found or created by other instructions
  - If that is not the case (e. g. there are vicious circles in the instruction dependencies) the set in not executable
  - Hence the sentence is not interpretable

Non-island example

- *Bill saw [a picture of John]*
- ***Who*** *did Bill see [a picture of ___]?*

Island example

- *Bill saw [the picture of John]*

- ***Who** did Bill see [the picture of __]?*

Relevance violation

- *Bill [raised a son and planted a tree]*

- ***Who** did Bill [raise __ and planted a tree]?*

Conclusions

- Proposed a unified account of backgrounded meaning within the DRT framework

  - Presupposition

  - Specific indefinites

  - "Parenthetical" constructions

- Proposed how the syntactic tree of the sentence arises out of knowledge in our mind

- Proposed an explanation why backgrounded constituents are syntactic islands

References

1. Ambridge, B., & Goldberg, A. E. (2008). The island status of clausal complements: Evidence in favor of an information structure explanation.

2. Davies, D. J. M., & Isard, S. D. (1972). Utterances as programs. *Machine intelligence, 7*, 325-339.

3. Erteschik-Shir N. 1973. On the nature of island constraints. PhD thesis, Dep. Foreign Lang. Linguist., MIT, Cambridge. 175 pp.

4. Geurts, B. (2010). Specific Indefinites, Presupposition and Scope. *Presuppositions and Discourse: Essays Offered to Hans Kamp*, 21, 125.

5. Goldberg AE. 2006. Constructions at Work: The Nature of Generalization in Language. Oxford, UK: Oxford Univ. Press

References

6. Heim, I. (1982). The semantics of definite and indefinite noun phrases.

7. Kamp, H. (1981). A theory of truth and semantic representation. Methods in the Study of Language Representation.

8. Newmeyer, F. J. (2016). Nonsyntactic explanations of island constraints. *Annual Review of Linguistics, 2*, 187-210.

9. van der Sandt, R. A. (1992). Presupposition projection as anaphora resolution. Journal of semantics, 9(4), 333-377.

10. van Geenhoven, V. (1998). Semantic incorporation and indefinite descriptions: *Semantic and syntactic aspects of noun incorporation in West Greenlandic*. CSLI publications.

Thanks for your attention! Questions?