

# Как возникает синтаксическая структура предложения?

Коммуникативный подход от динамической семантики

Иван Рыгаев

Лаборатория компьютерной лингвистики

Институт проблем передачи информации РАН, Москва

[irygaev@gmail.com](mailto:irygaev@gmail.com)

Открытый семинар Института лингвистики РГГУ, 29 сентября 2022 г.

# Коммуникация

- Коммуникация – первичная и основная функция языка
  - Передача информации от говорящего к слушающему

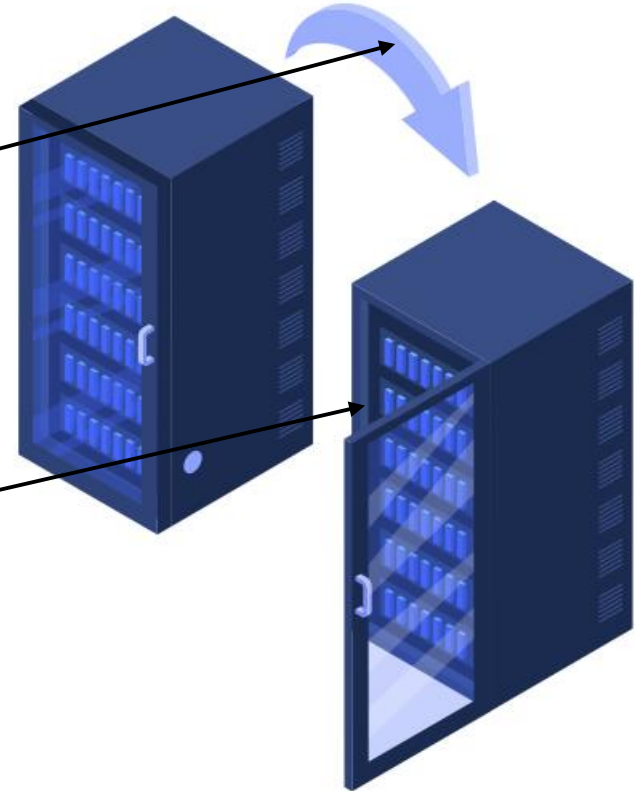


# Передача информации

- Новую информацию надо связать с имеющейся:  
*Зарплата Иванова – 50 000 руб.*

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

id	name	salary	...
...			
2837	Иванов	50 000	...
...			



# Речевая коммуникация

- Новую информацию нужно связать с имеющейся в голове слушающего. Одна часть предложения служит для поиска ментального файла, другая – для изменения

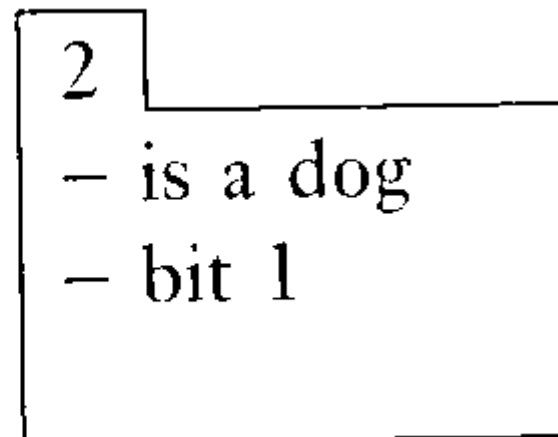
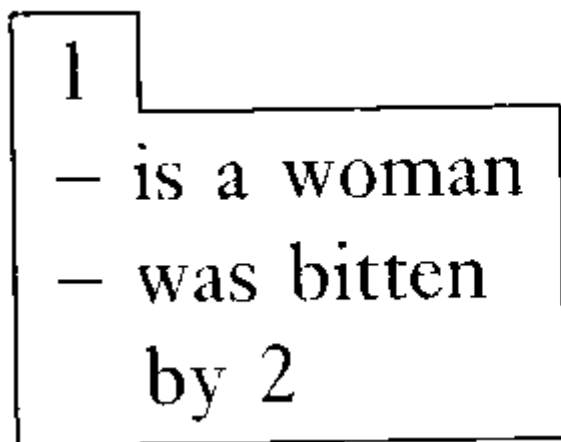


# File Change Semantics

- Метафора картотеки:
  - *“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)
- Карточка соответствует *дискурсивному* референту
- Значение именных групп – процедурное
  - *“For every indefinite, start a new card; for every definite, update a suitable old card”* (Heim 1982:179)
- Значение предложения – изменение контекста
  - Условия истинности имеет картотека, а не предложение

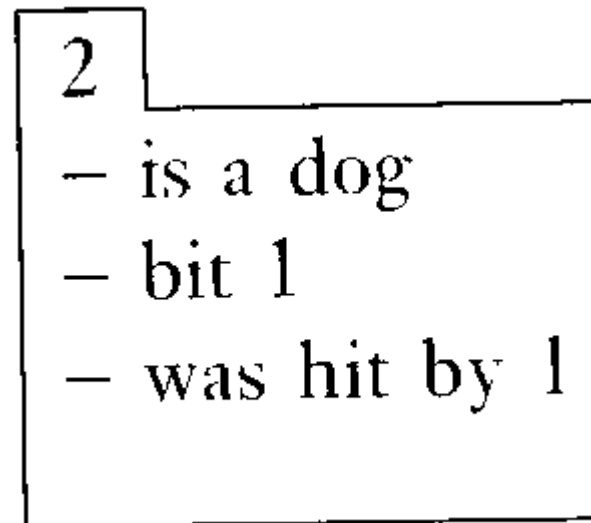
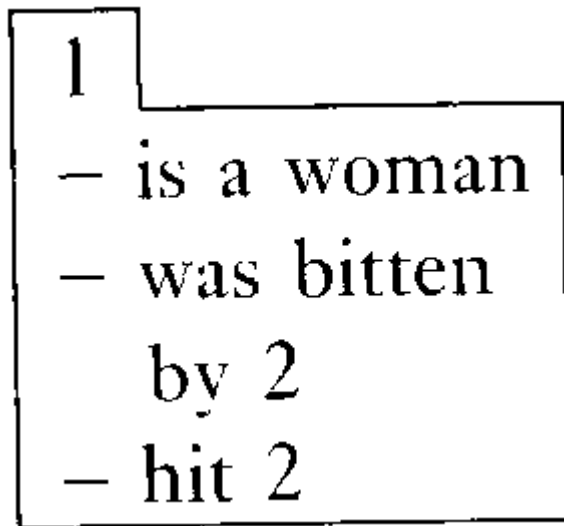
# File Change Semantics

- *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*

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

- DRS (discourse representation structure) состоит из:
  - Дискурсивных референтов (маркеров, переменных)
  - Условий (свойств, предикатов)
- DRS – представление всего дискурса (Kamp 1981)

# Разрешение анафоры

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

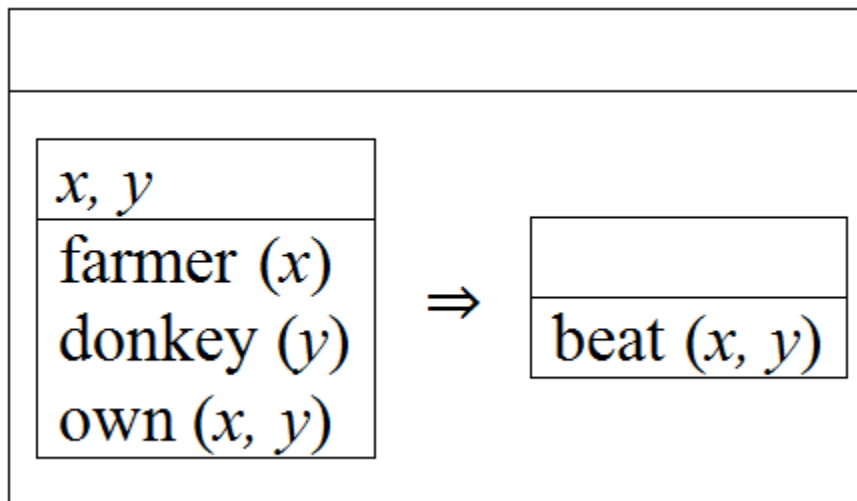
$x, y, u, v$
woman ( $x$ )
dog ( $y$ )
bit ( $y, x$ )
hit ( $u, v$ )
she ( $u$ )
it ( $v$ )

$x, y, u, v$
woman ( $x$ )
dog ( $y$ )
bit ( $y, x$ )
hit ( $u, v$ )
<b><math>u = x</math></b>
<b><math>v = y</math></b>

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

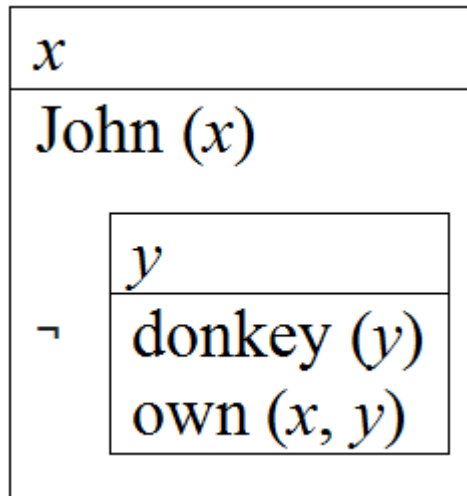
# Комплексные DRS: импликация

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



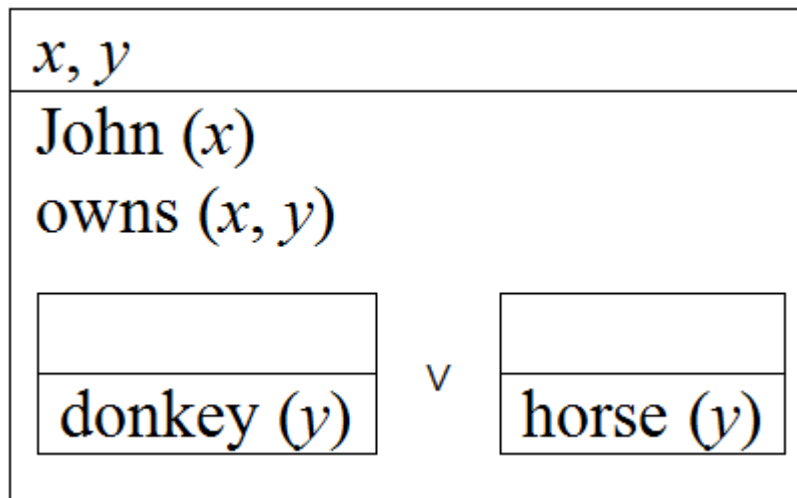
# Комплексные DRS: отрицание

- *John owns no donkey*
- *John does not own a donkey*



# Комплексные DRS: дизъюнкция

- *John owns a donkey or a horse*

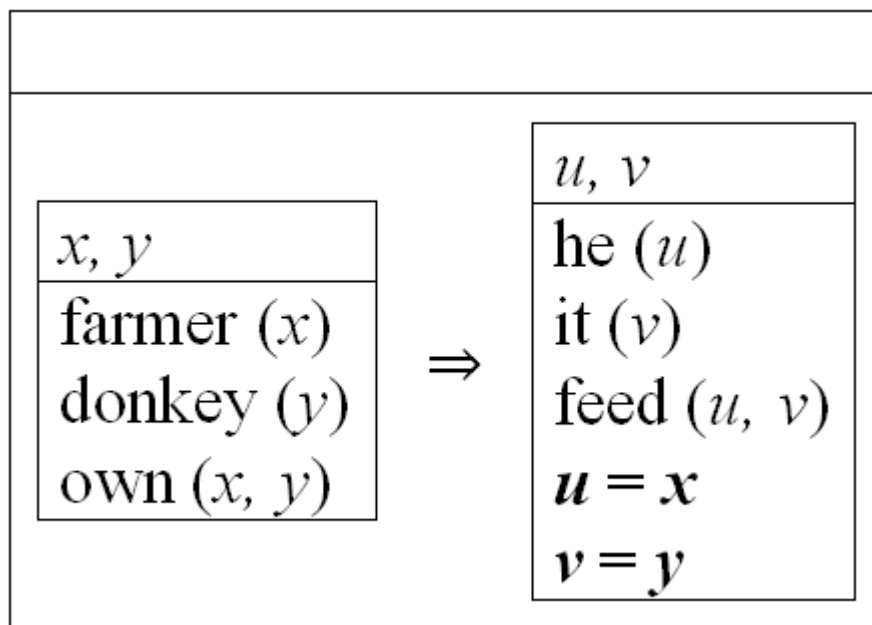


# Подчинение и доступность

- DRS B подчинён A если (неформально):
  - B вложен в A или
  - Где-то имеется импликация  $A \Rightarrow B$
- Доступность
  - Дискурсивный референт из DRS B может (анафорически) ссылаться на дискурсивный референт из DRS A, если B подчинён A.

# Подчинение и доступность

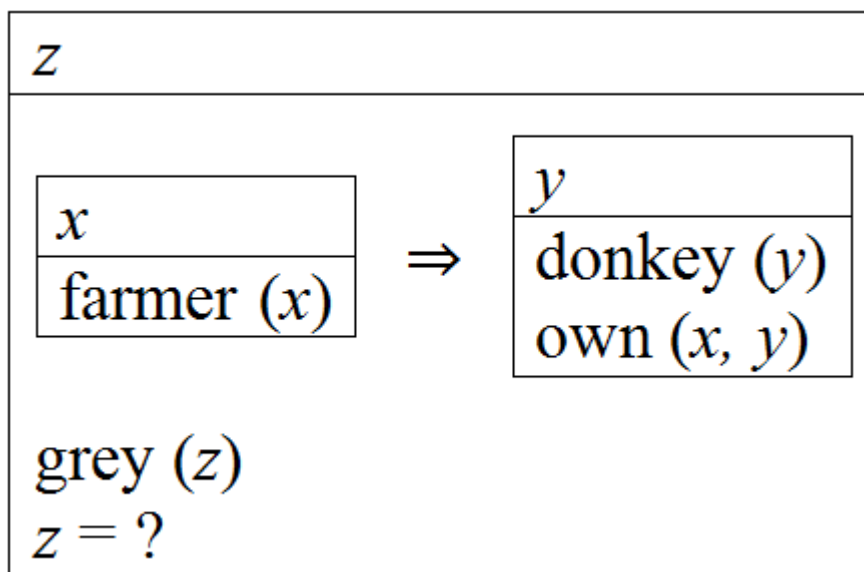
- *If a farmer owns a donkey, he feeds it*



- $x$  и  $y$  доступны для  $u$  и  $v$ , т. к. находятся в вышестоящем DRS

# Подчинение и доступность

- *Every farmer owns a donkey. \*It is grey.*



- Ни  $x$ , ни  $y$  недоступны для  $z$ , т. к. находятся в подчинённых DRS



# Достоинства DRT

- Позволяет расширить сферу действия квантора существования за пределы предложения
- Объясняет связывание анафорических местоимений, которые не связаны синтаксически
- Объясняет невозможность анафорических отсылок, когда антецедент недоступен

# Пресуппозиции

- Пресуппозиция – это информация в предложении, которую говорящий лингвистически помечает как само собой разумеющуюся
  - т. е. как уже известную слушающим
  - т. е. как часть common ground

# Триггеры пресуппозиций

- Определённые именные группы
  - *The king of France is bald*
  - > *There is a king of France*
- Дополнения фактивных глаголов
  - *John knows that the Earth is flat*
  - > *The Earth is flat*
- Клефты
  - *It was John who killed the butcher*
  - > *Somebody killed the butcher*
- Наречия *even, too, again* и т. д.

# Пресуппозиции и отрицание

- Отрицание не влияет на пресуппозиции
- Если утвердительное предложение содержит пресуппозицию
  - *The king of France is bald*
  - > *There is a king of France*
- То и отрицательное – тоже
  - *The king of France is **not** bald*
  - > *There is a king of France*

# Проекция пресуппозиций

- Пресуппозиции также обычно выживают под другими логическими операторами:
  - *If Fred has stopped beating Zelda, then Fred no longer resents Zelda's infidelity*
  - *> Fred has been beating Zelda*
  - *> Zelda has been unfaithful*
- И в других сложных предложениях:
  - *Bill does not know that all of Jack's children are bald*
  - *> All of Jack's children are bald*
  - *> Jack has children*

# Проекция пресуппозиций

- Однако иногда пресуппозиции исчезают в сложных предложениях:
  - *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 or all of **Jack's children** are bald*
- Проблема проекции пресуппозиций:
  - Объяснить, в каких случаях и почему пресуппозиции исчезают
  - Вычислить пресуппозиции сложного предложения из пресуппозиций его частей

# Пресуппозиция как анафора

- Rob van der Sandt (1992) предложил, что пресуппозиция и анафора – это по сути один и тот же феномен:
  - *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*

# Параллели

- Пресуппозиция
  - *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*
- Анафора
  - *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***



# Параллели

- VP-анафора:
  - *If someone solved the problem it was Julius who **{solved it/did}***
  - *If Harry stopped smoking, John **{stopped/did}** too.*
- Сентенциальная анафора:
  - *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}***

# Пресуппозиция как анафора

- Пресуппозиции – это анафоры
  - Могут обрабатываться тем же механизмом разрешения анафоры
- Но они содержат больше информации, чем местоимения, поэтому
  - Имеют внутреннюю структуру, которую надо отразить
  - Могут аккомодироваться – если антецедент не найден, информация может быть просто добавлена в DRS.

# Проекция пресуппозиций в DRT

- Сначала строится отдельный (предварительный) DRS для предложения и только потом он сливается с главным DRS
- Анафорические элементы специально кодируются
  - Они обрабатываются только после того, кого DRS предложения построен и влит в главный
  - DRS предложения содержит дополнительную A-структуру
    - набор т. н. A-DRS пресуппозиций
  - A-DRS могут сами иметь A-структуру, то есть могут быть вложены друг в друга

# СВЯЗЫВАНИЕ

- *John has a cat. His cat purrs*

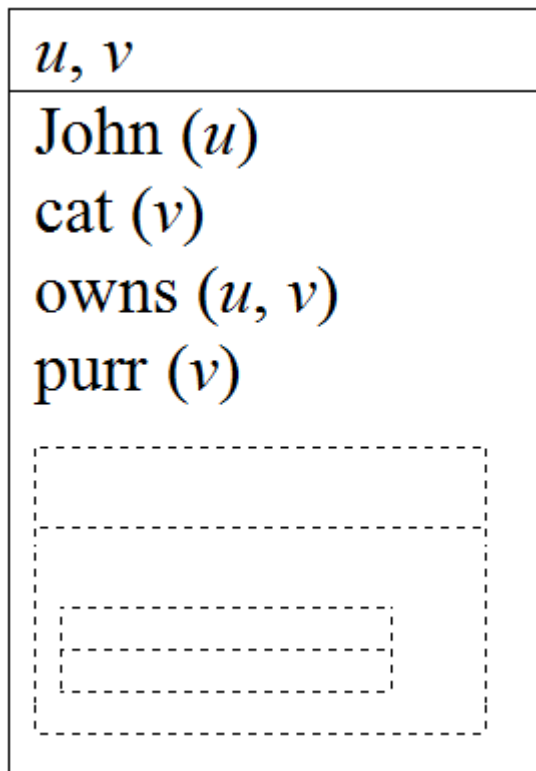
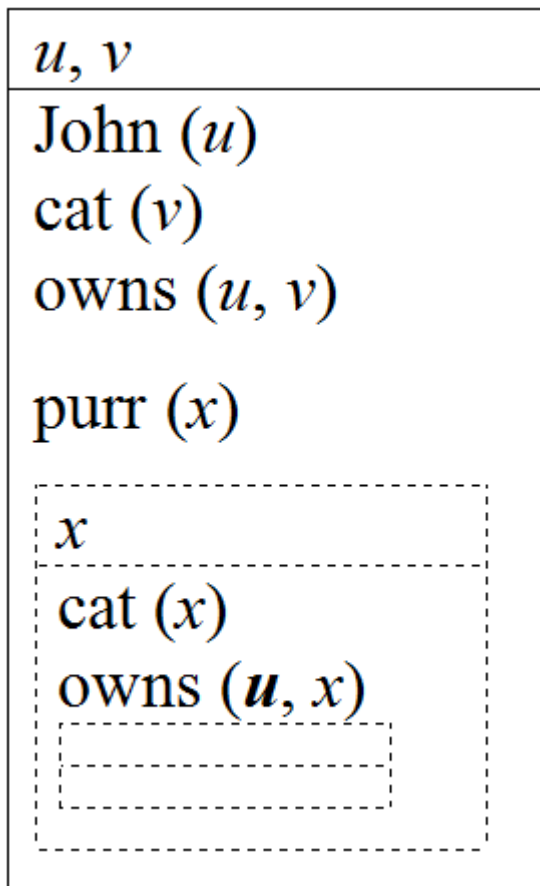
$y, x$
John ( $y$ )
cat ( $x$ )
owns ( $y, x$ )

purrr ( $x$ )					
<table border="1" style="border-style: dashed;"> <tr> <td><math>x</math></td> </tr> <tr> <td>cat (<math>x</math>)</td> </tr> <tr> <td>owns (<math>y, x</math>)</td> </tr> <tr> <td> <table border="1" style="border-style: dashed;"> <tr> <td><math>y_{masc}</math></td> </tr> </table> </td> </tr> </table>	$x$	cat ( $x$ )	owns ( $y, x$ )	<table border="1" style="border-style: dashed;"> <tr> <td><math>y_{masc}</math></td> </tr> </table>	$y_{masc}$
$x$					
cat ( $x$ )					
owns ( $y, x$ )					
<table border="1" style="border-style: dashed;"> <tr> <td><math>y_{masc}</math></td> </tr> </table>	$y_{masc}$				
$y_{masc}$					

$u, v$					
John ( $u$ )					
cat ( $v$ )					
owns ( $u, v$ )					
purrr ( $x$ )					
<table border="1" style="border-style: dashed;"> <tr> <td><math>x</math></td> </tr> <tr> <td>cat (<math>x</math>)</td> </tr> <tr> <td>owns (<math>y, x</math>)</td> </tr> <tr> <td> <table border="1" style="border-style: dashed;"> <tr> <td><math>y_{masc}</math></td> </tr> </table> </td> </tr> </table>	$x$	cat ( $x$ )	owns ( $y, x$ )	<table border="1" style="border-style: dashed;"> <tr> <td><math>y_{masc}</math></td> </tr> </table>	$y_{masc}$
$x$					
cat ( $x$ )					
owns ( $y, x$ )					
<table border="1" style="border-style: dashed;"> <tr> <td><math>y_{masc}</math></td> </tr> </table>	$y_{masc}$				
$y_{masc}$					

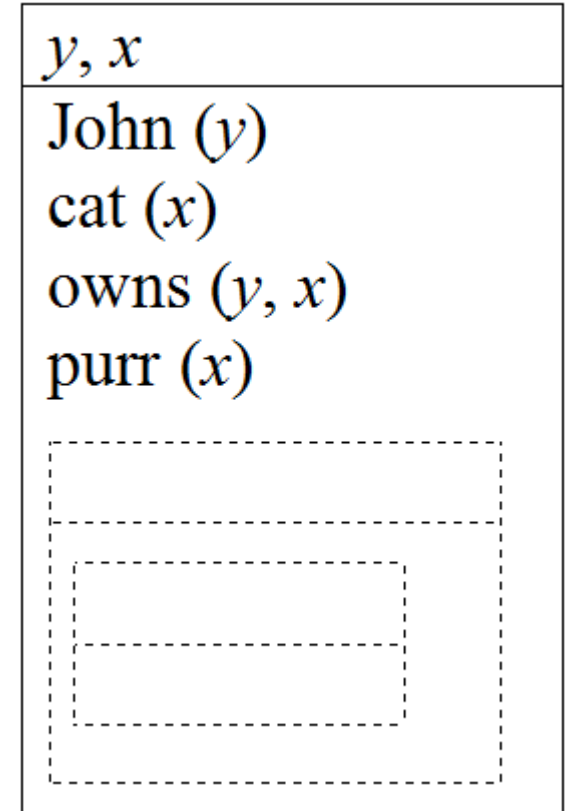
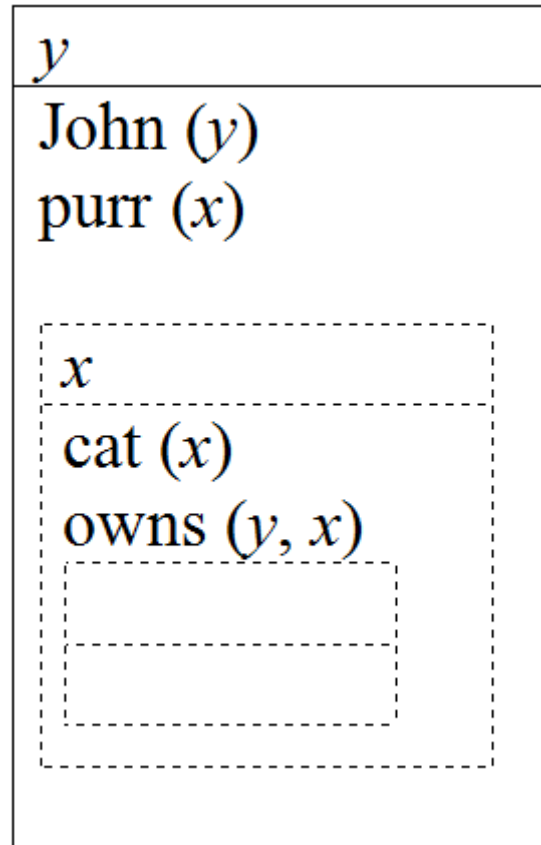
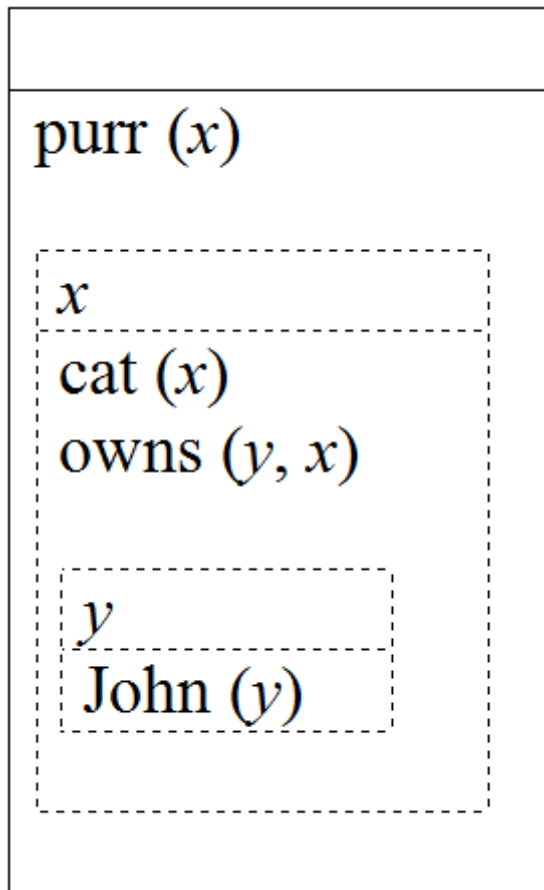
# Связывание

- *John has a cat. His cat purrs*



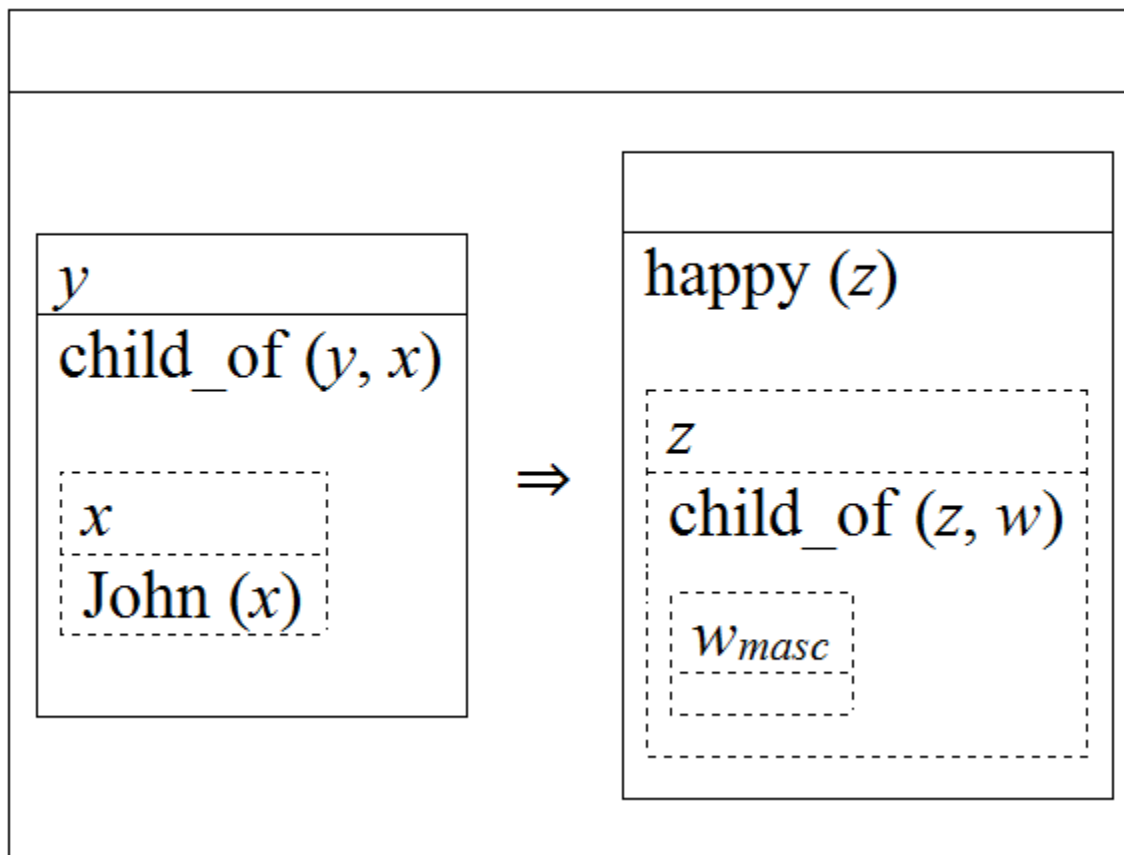
# Аккомодация

- *John's cat purrs*



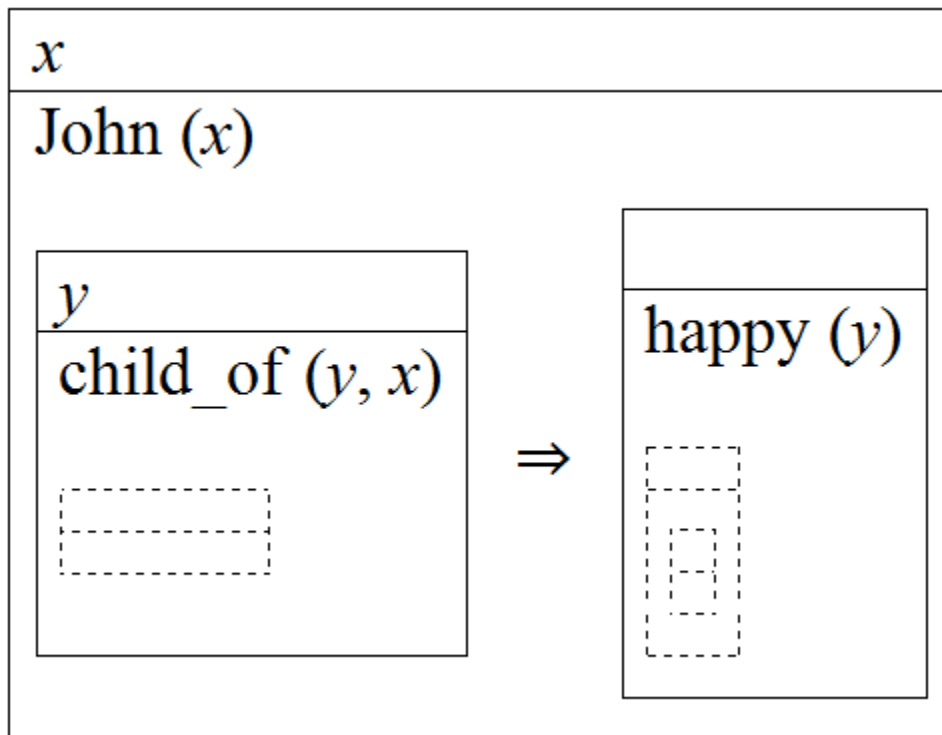
# Исчезновение пресуппозиции

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



# Исчезновение пресуппозиции

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





# Ограничения на разрешение

- Общий алгоритм:
  - A-DRS движется вверх по иерархии в поисках антецедента
  - Если не найдено – пытается аккомодироваться наверху
  - Если не удалось – движется вниз, где удастся
- Ограничения на разрешение:
  - Главный DRS должен оставаться непротиворечивым
  - Новый главный DRS должен быть информативным (не должен логически следовать из предшествующего)
  - Подчинённый DRS не противоречит вышестоящему
  - Подчинённый DRS информативен по отношению к вышестоящему

# Specific indefinites

- Specific indefinites (van Geenhoven 1998)
  - Похожи на пресуппозиции
  - Они интерпретируются не там, где появляются
  - А выше в иерархии
  - Обычно они аккомодируются, а не связываются
- Может это особый тип пресуппозиций?
- *Peter intends to visit **a museum** every day*
  - Предложение имеет как минимум 3 интерпретации
  - Смотря на каком уровне интерпретируется 'a museum'

# Backgrounding (Geurts 2010)

- Specific indefinites – это не пресуппозиции
  - Аккомодация – это запасная стратегия (починка)
  - Было бы странно её использовать как основную
- Разные типы выражений заднего плана:
  - Пресуппозиции
  - Specific indefinites
  - «Скобочные» конструкции
- The Buoyancy Principle (принцип плавучести):
  - Материал заднего плана стремиться «всплыть» наверх в главный DRS

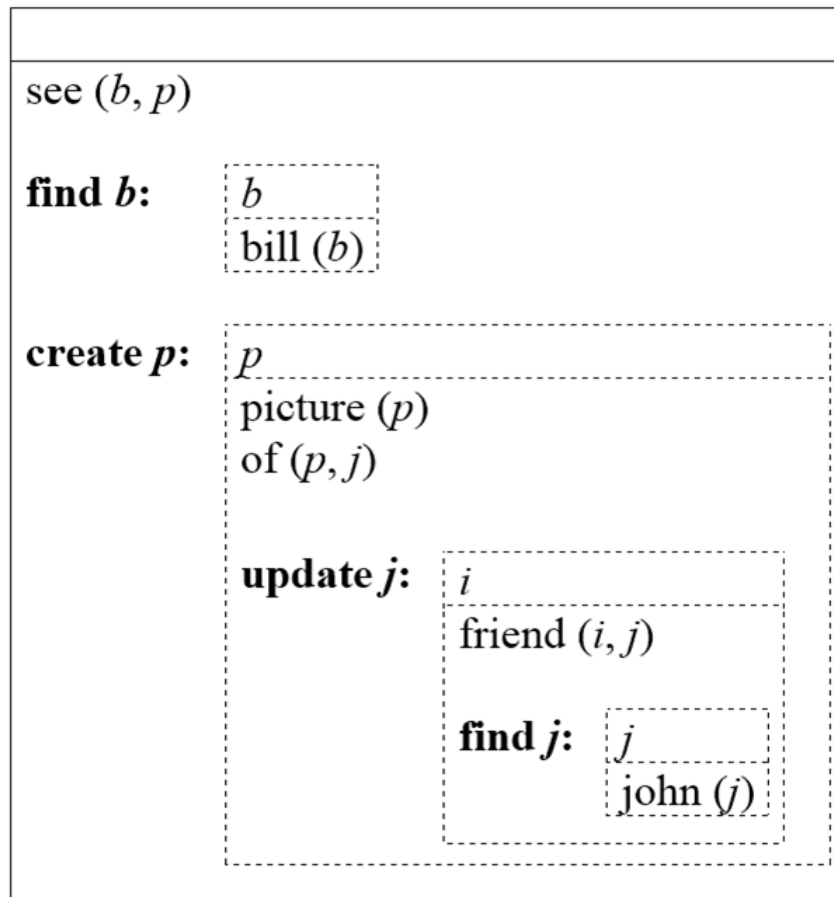
# DRS как инструкции

- A-DRS для всех составляющих заднего плана
- Но надо помечать их тип (функцию)
  - A-DRS пресуппозиций служат для поиска референта
  - Specificity A-DRS – для создания нового референта
  - «Скобочные» A-DRS – для изменения существующего
  - Такие DRS символизируют инструкции по модификации ментальной базы данных слушающего
- Теперь их можно назвать B-DRSs (backgrounded DRS)
- Главный DRS – это инструкция по изменению топикального референта

# DRS как инструкции

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

update *b*:



<i>b</i> , <i>p</i> , <i>j</i> , <i>i</i>
bill ( <i>b</i> )
see ( <i>b</i> , <i>p</i> )
picture ( <i>p</i> )
of ( <i>p</i> , <i>j</i> )
john ( <i>j</i> )
friend ( <i>i</i> , <i>j</i> )

# Синтаксис

- Структура предварительного DRS:
  - Каждая составляющая заднего плана – один B-DRS
  - Иерархия B-DRS образует дерево
  - Это дерево отражает синтаксическое дерево предложения
- Порождение речи
  - Говорящий намеревается сообщить нечто слушающему
  - Он разбивает всю информацию на инструкции по поиску, созданию и изменению ментальных референтов
  - Будучи зависимыми друг от друга, они образуют дерево
  - Это дерево затем реализуется как синтаксическое дерево

# Высказывания как программы

- Два шага в NLU (Davies & Isard 1972)
  - Компиляция
  - Выполнение
  - Понять высказывания vs принять его
- В нашей модели
  - Компиляция = Построение предварительного DRS
  - Выполнение = разрешение B-DRSs, чтобы получить новый главный DRS
- Книга – это не база знаний
  - Это скрипт по созданию этой базы

# Два уровня представления

- Предварительный DRS – представление предложения
  - Набор инструкций
  - Полностью контекстно-независимо
  - Но чувствительно к контексту
  - Отражает коммуникативную структуру
  - Отражает синтаксис на семантическом уровне
  - Хорошая интерлингва для перевода
- Главный DRS – ментальное представление
  - Отражает условия истинности



# Синтаксические острова

- Остров
  - Синтаксическая конструкция, содержащая элемент, который не может быть вынесен за её пределы
- Пример не острова
  - *Bill saw [a picture of John]*
  - **Who** did Bill see [*a picture of \_\_\_\_*]?
- Пример острова
  - *Bill saw [the picture of John]*
  - \***Who** did Bill see [*the picture of \_\_\_\_*]?

# Синтаксические острова

- Ещё примеры островов (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]?*
- Основные подходы к объяснению островов:
  - Синтаксический
  - Ресурсный
  - Коммуникативный

# Коммуникативный подход

- Острова возникают из-за конфликта в коммуникативной структуре предложения
- Сем. доминирование (Erteschik-Shir 1973)
  - Составляющая семантически доминирует, если она – не пресуппозиция и не имеет ссылки в контексте
- Условие на вынос
  - Выносить элементы можно только из тех составляющих, которые могут рассматриваться как доминирующие в некотором контексте

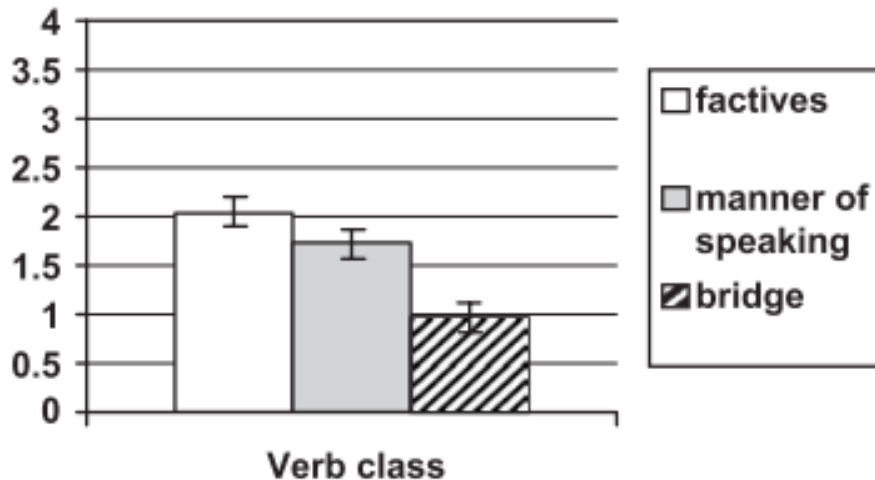
# Тест «это ложь»

- Тест на доминирование
- *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.*

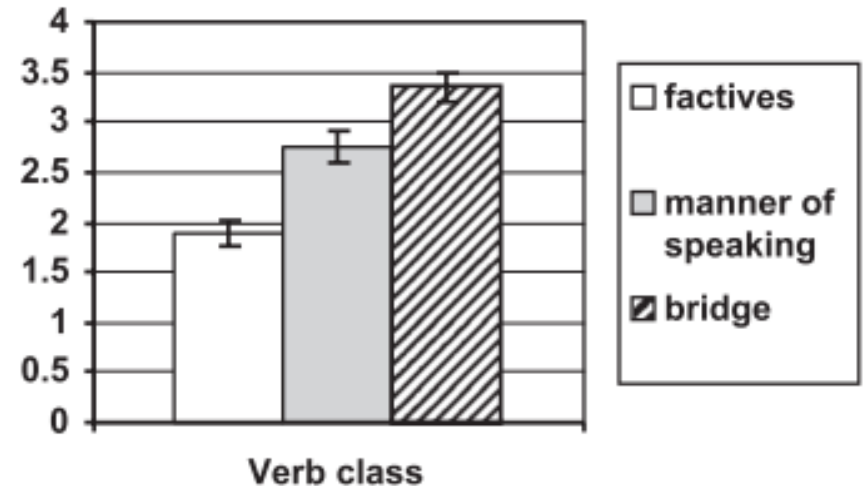
# Goldberg 2006

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

Неграмматичность выноса



Фокусность



$p < 0.001$

correlation = 0.83

# Объяснение

- Почему элементы заднего плана – острова?
- Ответ
  - Каждый элемент заднего плана – отдельная инструкция
  - Она выполняется отдельно
  - Дискурсивные референты, от которых она зависит, уже должны быть найдены/созданы к этому моменту
  - Если это не так (например, из-за порочных кругов в зависимостях), набор инструкций невыполним
  - А значит предложение не интерпретируемо

# Пример не острова

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

update *b*:

<i>p</i>
see ( <i>b</i> , <i>p</i> ) picture ( <i>p</i> ) of ( <i>p</i> , <i>j</i> )
<b>find <i>b</i>:</b> <span style="border: 1px dashed black; padding: 2px;"><i>b</i></span>
<span style="border: 1px dashed black; padding: 2px;">bill (<i>b</i>)</span>
<b>find <i>j</i>:</b> <span style="border: 1px dashed black; padding: 2px;"><i>j</i></span>
<span style="border: 1px dashed black; padding: 2px;">john (<i>j</i>)</span>

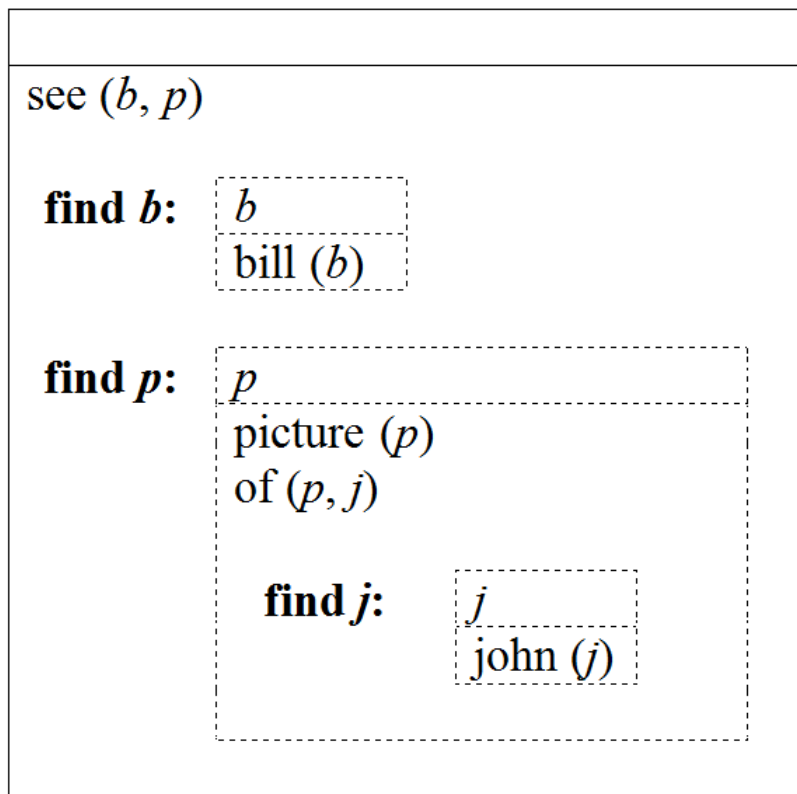
retrieve *w*:

<i>w</i> , <i>p</i>
see ( <i>b</i> , <i>p</i> ) picture ( <i>p</i> ) of ( <i>p</i> , <i>w</i> )
<b>find <i>b</i>:</b> <span style="border: 1px dashed black; padding: 2px;"><i>b</i></span>
<span style="border: 1px dashed black; padding: 2px;">bill (<i>b</i>)</span>

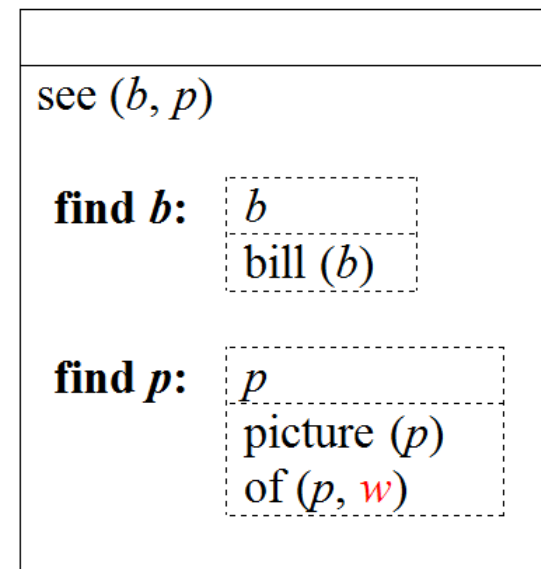
# Пример острова

- *Bill saw [the picture of John]*
- \**Who did Bill see [the picture of \_\_]?*

update  $b$ :



retrieve  $w$ :





# Нарушение релевантности

- *Bill [raised a son and planted a tree]*
- \***Who** did Bill [raise \_\_\_ and planted a tree]?

update *b*:

<i>s, t</i>
raise ( <i>b, s</i> )
son ( <i>s, b</i> )
plant ( <i>b, t</i> )
tree ( <i>t</i> )
find <i>b</i> : <span style="border: 1px dashed black; padding: 2px;"><i>b</i></span>
<span style="border: 1px dashed black; padding: 2px;">bill (<i>b</i>)</span>

retrieve *w*:

<i>w</i>
raise ( <i>b, w</i> )
plant ( <i>b, t</i> )
tree ( <i>t</i> )
find <i>b</i> : <span style="border: 1px dashed black; padding: 2px;"><i>b</i></span>
<span style="border: 1px dashed black; padding: 2px;">bill (<i>b</i>)</span>

# Заключение

- Предложен единый подход к представлению разных типов составляющих заднего плана в DRT
  - Пресуппозиции
  - Specific indefinites
  - «Скобочные» конструкции
- Предложено объяснение, как синтаксическая структура предложения возникает из семантического представления в нашей голове
- Предложено объяснение, почему составляющие заднего плана – это синтаксические острова

# Литература

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

# Литература

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.

Спасибо за внимание!  
Вопросы?