

Разметка кореференции в brat

[tl;dr к документации](#)

Что выделять?

- **мёншны** (англ. *mentions* - упоминания)
Называют конкретных людей, организации, предметы, события, процессы, даты и др.
Типы меншнов:
 - X, Y, X-ADJ, Y-ADJ, FACT, X-HEAD, X-ATTR
- **корреферентные связи**
между меншнами
 - LINK (IDENT, ANAPH, CATAPH)
 - APPOS
- **уточнения** неполных меншнов
(DEF, ELLIPSIS)

Корреференция | Анафора

Корреференция (англ. *co-reference*) – вид текстовой или синтаксической связности, при которой два или более упоминания-меншна называют один и тот же объект (**референт**)

Ана́фора (англ. *anaphora*) – явление, при котором интерпретация одного упоминания зависит от другого, выше или ниже в тексте

Корреферентными и анафорическими связями могут объединяться только **равнозначные** по количеству/объему/охвату меншны

Как выделять меншны?



Примечание: Выделяем по аналогии с именованными сущностями

Нужно выделить цепочки из двух и более меншнов, которые указывают на один и тот же конкретный, реальный объект.

Меншны, которые не с чем связать, выделять не нужно.

Каждый меншн должны быть обязательно связан с каким-то другим.

В brat - выделить меншн, выбрать один из доступных лейблов

X - максимально конкретные меншны во всей цепочке

Y - меншны-анафоры, которые ссылаются на более конкретные X

X-ADJ, **Y-ADJ** - те же X и Y, но в форме прилагательных

FACT - глагольные фразы, обозначают факты

X-HEAD, **X-ATTR** - основная и атрибутивная части аппозитивной связи

DEF - фраза или слово-уточнитель для неполного меншна

Какие бывают меншны?

- **Именные группы**

существительное + зависимые слова

- **Глагольные группы или фразы**

глагол + зависимые слова или фраза целиком

- **Предикативы**

слова и фразы - статическое состояние

- **Местоимения**

указывают на другие меншны

- **Некоторые наречия**

напр., наречия времени, места и др.

- **Прилагательные**

Относительные прилагательные

Какие бывают меншны?

- **Именные группы**

существительное + зависимые слова

- **Глагольные группы или фразы**

глагол + зависимые слова или фраза целиком

- **Предикативы**

слова и фразы - статическое состояние

- **Местоимения**

указывают на другие меншны

- **Некоторые наречия**

напр., наречия времени, места и др.

- **Прилагательные**

Относительные прилагательные

- [доклад]
- [три года]
- [первый доклад о [готовности [страны] к [введению в действие [новой реформы]]]]*
- [Книги [Хэмингуэя]]

* обратите внимание - меншны могут быть вложенными

Типы меншнов: X, Y, X-HEAD, X-ATTR

Какие бывают меншны?

- **Именные группы**
существительное + зависимые слова
- **Глагольные группы или фразы**
глагол + зависимые слова или фраза целиком
- **Предикативы**
слова и фразы - статическое состояние
- **Местоимения**
указывают на другие меншны
- **Некоторые наречия**
напр., наречия времени, места и др.
- **Прилагательные**
Относительные прилагательные

- Я решил начать [бегать трусцой по утрам], потому что [это] поможет мне улучшить свое здоровье.
- Акции [выросли на 790%]. Такой [рост] обусловлен...
- [Министерство образования перевело студентов Москвы на дистанционное обучение]. Об [этом] сообщил Интерфакс.

Типы меншнов: **FACT**

Какие бывают меншны?

- **Именные группы**

существительное + зависимые слова

- **Глагольные группы или фразы**

глагол + зависимые слова или фразы целиком

- **Предикативы**

слова и фразы - статическое состояние

- **Местоимения**

указывают на другие меншны

- **Некоторые наречия**

напр., наречия времени, места и др.

- **Прилагательные**

Относительные прилагательные

- Вчера было [очень холодно], из-за [этого] на дорогах гололед.
- Она была [самая красивая и умная из всех] и благодаря [этим качествам] победила в Мисс Мира.

Типы меншнов: **FACT**

Какие бывают меншны?

- **Именные группы**
существительное + зависимые слова
- **Глагольные группы или фразы**
глагол + зависимые слова или фразы целиком
- **Предикативы**
слова и фразы - статическое состояние
- **Местоимения**
указывают на другие меншны
- **Некоторые наречия**
напр., наречия времени, места и др.
- **Прилагательные**
Относительные прилагательные

- я, ты, он, она, они, мой, твой, свой, себя, который...
- За несколько недель до [его] смерти, [отец] завещал мне [свой] дом.

Типы меншнов: Y, X-HEAD, X-ATTR

Какие бывают меншны?

- **Именные группы**

существительное + зависимые слова

- **Глагольные группы или фразы**

глагол + зависимые слова или фразы целиком

- **Предикативы**

слова и фразы - статическое состояние

- **Местоимения**

указывают на другие меншны

- **Некоторые наречия**

напр., наречия времени, места и др.

- **Прилагательные**

Относительные прилагательные

- Сегодня, завтра, тогда, там, здесь, сюда, туда, отсюда, куда...
- Пойдешь гулять [послезавтра]? - Нет, мне [в этот день] надо к стоматологу.
- Я завтра еду в [Урюпинск]. - А мы только вчера [там] были.

Типы меншнов: X, Y, X-HEAD, X-ATTR

Какие бывают меншны?

- **Именные группы**
существительное + зависимые слова
- **Глагольные группы или фразы**
глагол + зависимые слова или фразы целиком
- **Предикативы**
слова и фразы - статическое состояние
- **Местоимения**
указывают на другие меншны
- **Некоторые наречия**
напр., наречия времени, места и др.
- **Прилагательные**
Относительные прилагательные

- Люблю творчество [Пушкина]. Считаю, что [пушкинская] проза замечательная.
- В [Австралии] обитает десятки тысяч видов животных. [Австралийская] фауна, без сомнения, одна из самых разнообразных.
- В [Краснодарской области] вторую неделю идут ливни. [Региональные] власти объявили эвакуацию населения.

Типы меншнов: X-ADJ, Y-ADJ

Как выделять **связи**?

Перед связыванием убедитесь, что меншны размечены верно

1. кликаем на лейбл первого меншна и вытаскиваем из него стрелку
2. перетаскиваем стрелку на второй меншн
3. выбираем подходящий тип связи



Если связь надо протянуть на меншн, который не помещается в экран -

используем прокрутку мыши
или стрелки ↑ и ↓

Какие бывают связи между меншнами?

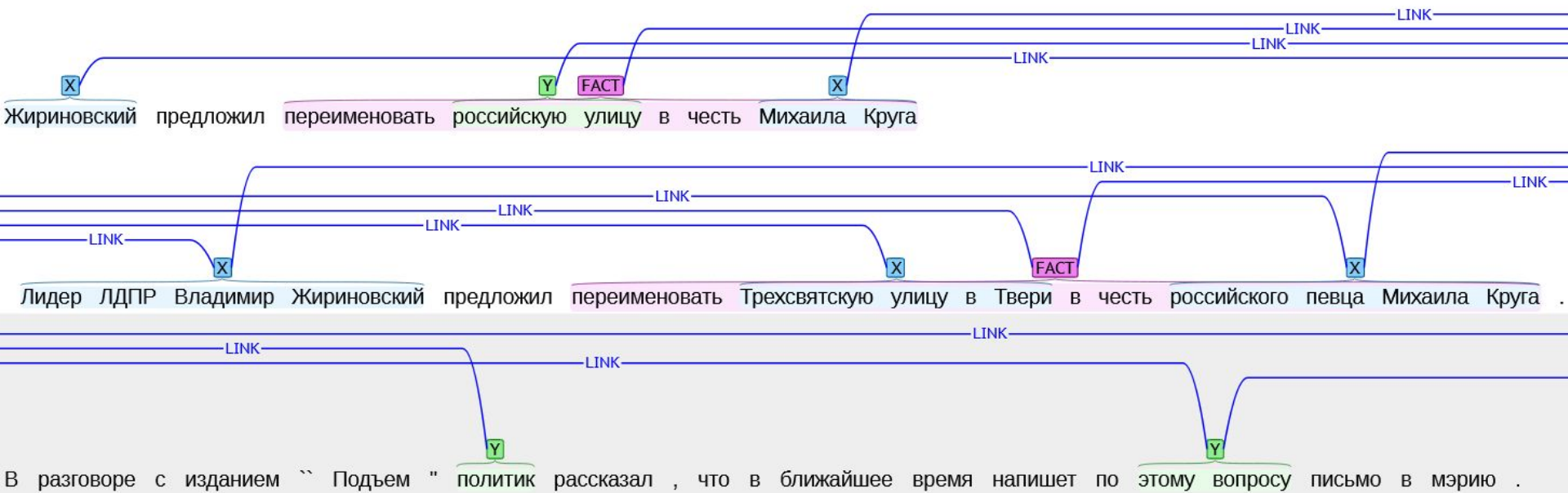
- **LINK** - связь между элементами кореферентной цепочки, в которой все обозначают один и тот же референт*.
→ Связывает меншны X, Y, X-ADJ, Y-ADJ, FACT по цепочке
- **APPOS** (аппозитив) - конструкция-приложение с доп. информацией, выделенная запятыми, тире или скобками
→ Попарно связывает меншны X-ATTR → X-HEAD
- **ELLIPSIS** (эллипсис) - дополнение наименования меншна с помощью близлежащих слов или фраз-уточнителей
→ Попарно связывает меншн DEF с другими меншнами

* **референт** - конкретный объект, на который указывает меншн: человек, организация, событие и т. д.

Связи между меншнами: **LINK**

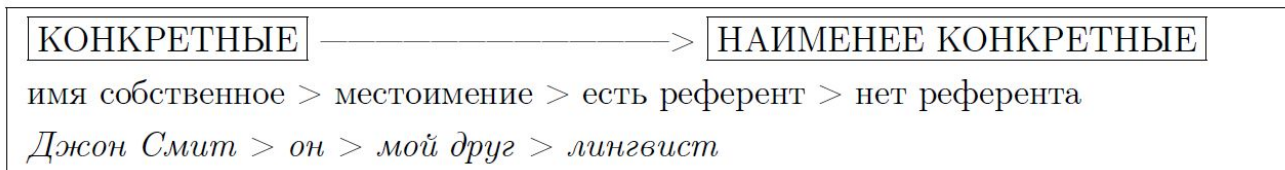
- Все элементы **LINK**:
 - обозначают **один и тот же референт**
 - друг с другом **связаны по цепочке**
→ ненаправленная связь

Связью **LINK** можно соединять
X, **Y**, **X-ADJ**, **Y-ADJ**, **FACT**



Связи между меншнами: APPOS

- **Состоит из двух частей-меншнов: X-HEAD и X-ATTR**
 - X-HEAD - основной, более конкретный меншн из двух
 - X-ATTR - дополнительный меншн с добавочной информацией об X-HEAD
 - X-HEAD и X-ATTR стоят непосредственно друг за другом, одна из частей обособлена знаками препинания - запятые, тире или скобки
 - связь направленная: X-ATTR → X-HEAD
- Выбрать X-HEAD можно с помощью шкалы конкретности. Если оба элемента равнозначны, X-ATTR - обособленный.



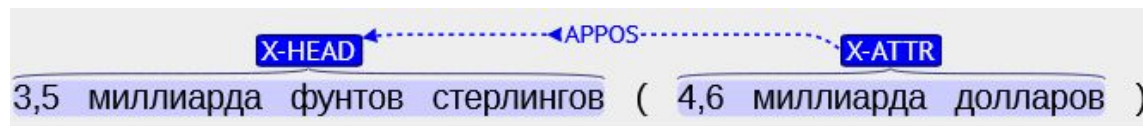
Частые случаи APPOS:

- названия к статьям законы
- денежные эквиваленты
- переводы фраз на другие языки

Не APPOS:

- Деепричастные и причастные обороты
- Придаточные части сложноподчиненных предложений
- Копулярные конструкции типа "А – это В"

Примеры: APPOS



- ✓ обособление знаками препинания
- ✓ **X-ATTR** - доп. инфа об **X-HEAD**

APPOS-конструкция может целиком объединяться в X и входить в цепочку LINK


DEF, ELLIPSIS и дополненные меншны

DEF - “присыпка-уточнитель” - уточняющее слово или фраза, которым можно дополнить наименование другого меншна

ELLIPSIS - это связь **DEF** → [любой меншн], когда наименование другого меншна дополняется с помощью **DEF**



Разное

- У меншнов есть **шорткаты** (в окне выбора подчеркнута нужная буква)
- **LINK:**
 - Если есть цепочка 1-3, то при добавлении связи 1-2 цепочка будет выглядеть как 1-2-3 а не 
 - Если в 1-2-3-4 удалить связь 2-3, оба элемента удаляются целиком из цепочки, останется только цепочка 1-4.
Чтобы, например, оставить цепочку 1-2-4, нужно удалить меншн 3. Тогда связь оборвется только с этим меншном.
- Атрибут **NoRef**
 - Если все элементы цепочки LINK выражены местоимениями или нереферентными меншнами (непонятно, о ком или о чем идет речь - возможно, текст обрезан) - к первому элементу нужно добавить атрибут **NoRef**.

Предположительно, такие цепочки будут встречаться очень редко.