



Ивтушок Е.И.
Шульгина К.М.
Толдова С.Ю.
Акутина Н.С.

Выбор референциальных средств разграничения эпизодов на примере текстов корпуса Russian CLiPS

Структура доклада

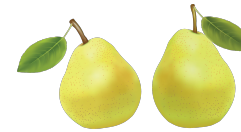
1. Корпус Russian CliPS
2. Экспериментальный подкорпус
3. Цели
4. Разметка кореференции
5. Разметка эпизодов
6. Результаты

1.

Корпус Russian CliPS

= Russian Clinical Pears Stories

Корпус Russian CLiPS



Базовая информация:

- ▶ Состоит из пересказов «Фильма о грушах» (Pear film) [Chafe 1980], собранных у здоровых носителей русского языка и людей с повреждениями мозга (афазией и ППП)
- ▶ Предобработка текстов сделана при помощи программы ELAN [[https:// tla.mpi.nl/tools/tla-tools/elan/](https://tla.mpi.nl/tools/tla-tools/elan/)]

Достоинства корпуса:

Структура подобного нарратива (четкая сюжетная линия, ограниченное количество появляющихся сущностей, отсутствие речи) предоставляет возможность исследования:

- ▶ грамматических и лексических особенностей на уровне клауз
- ▶ макроструктуры: общей структуры рассказа, понятности и связности, нарративной стратегии, системы интеракции

Экспериментальный подкорпус

Афазия

Акустико-мнестическая афазия

- 5 текстов
- средний возраст 48,8 (40-59)

Эфферентно-моторная афазия

- 5 текстов
- средний возраст 47,8 (30-55)

Контрольная группа

- 10 текстов
- средний возраст 58 (40-84)

	НР	Ас-мн	Eff
Количество текстов	10	5	5
Максимальная длина	391	419	408
Минимальная длина	106	195	266
Средняя длина	277,8	292,2	312,8

Цели исследования

1. Применить схему разметки на эпизоды к материалу корпуса Russian Clips
2. Совместить две схемы разметки
3. Оценить влияние границ эпизода на выбор референциального средства у трех групп говорящих

2.

Разметка

1. Разметка кореференции
2. Разметка на эпизоды

1. Разметка кореференции

Разметка кореференции — это определение всех языковых средств, использованных для референции к определенной сущности.

...мальчик остановился и взял грушу из корзины. Но потом он понял, что одной ему недостаточно.

Рассматриваемые языковые средства:

- ▷ Существительное
- ▷ Полные ИГ (существительное и модификатор)
- ▷ Анафорические местоимения
- ▷ Нулевые местоимения

1. Разметка кореференции - пример

HP-v03

которая ему нравилась.⊘
он увлёкся⊘
и наехал на камень.⊘
○ рассыпались груши ,⊘
и тут вышли друзья его по этому посёлку.⊘
○ которые не агрессивны , а наоборот⊘
помогли ему собрать всё ○ вот⊘
и даже ○ [вдох] ○ э ○ потом он уронил шляпу ,⊘
○ ему тоже дружок поднял ему шляпу⊘
○ а в знак что благодарности он ему ○ несколько груш дал
○ вот⊘

2. Разметка на эпизоды

Эпизод — единица макроструктуры нарратива.

Граница эпизода определяется по **смене перспективы**.

Смена перспективы может выражаться:

- ▷ в смене действующего лица или темы
- ▷ в переходе от одного типа компонента к другому («устойчивые» компоненты: *Зачин, Кода*, «неустойчивые»: *Основная линия, Интеракция, Описание*)
- ▷ в переключении между *Миром Рассказа* и *Миром Рассказывания*

Эпизод 1: Основная линия

...

[1] мальчик [n;boy;first]

[1] для него [ana;boy;rep] это была конечно неожиданность

[1] это как-то его [ana;boy;rep] дестабилизировало

[1] и он [ana;boy;rep] отвлекся от дороги [n;road;first]

[1] и наехал [zero;boy;rep] на камень [n;stone;first]

[1] ну от удара о камень [n;stone;rep] он [ana;boy;rep] упал

[1] груши [n;pears;first] соответственно тоже все рассыпались

смена эпизода

Эпизод 2: Основная линия

[2] тут навстречу идут ещё три подростка [NP;3boys;first] навстречу мальчику [ana;boy;first]

[2] увидев [PRO;3boys;rep]

[2] что на дороге [n;road;first] валяется велосипед [n;bike;first]

[2] мальчик [n;boy;rep] у него [ana;boy;rep] лежит пустая корзина [NP;basket_boy;first] на боку [n;basket_boy_part;first]

[2] рассыпаны груши [n;pears;first]

[2] они [ana;3boys;first] подошли

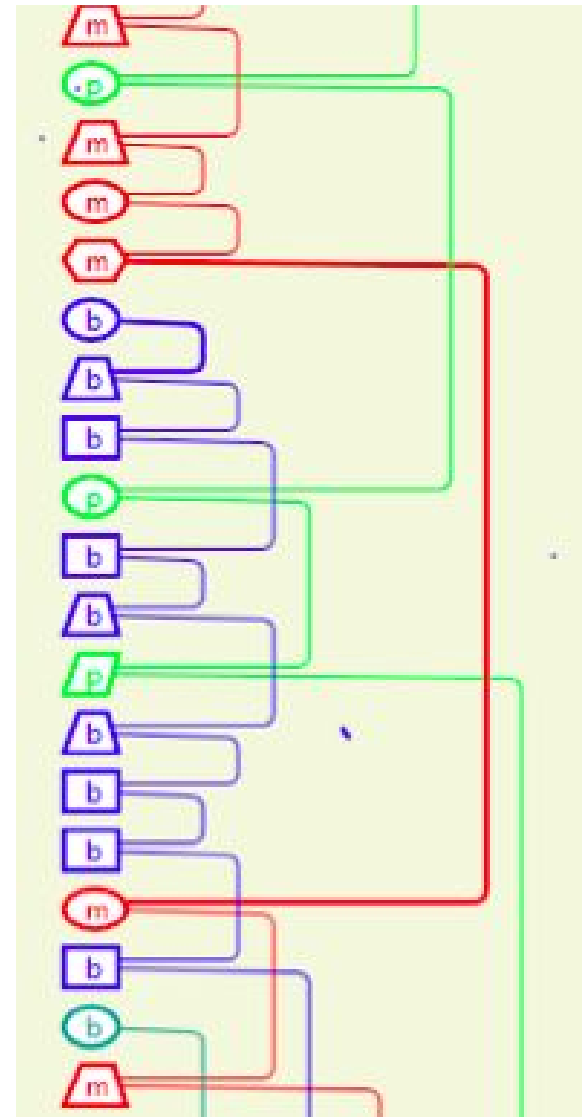
[2] помогли [pro_synt;3boys;rep] собрать [PRO;3boys;rep]

...

Способ визуализации разметки - PearVisual

Entity	Short	Color	Type	Shape
boy	b	Blue circle	n	Blue circle
man	m	Red circle	synt	Blue square
pears	p	Green circle	refl	Blue hexagon
3boys	3	Dark red circle	zero	Blue rounded rectangle
basket_boy	b	Teal circle	ana	Blue trapezoid

n	Blue circle
synt	Blue square
refl	Blue hexagon
zero	Blue rounded rectangle
ana	Blue trapezoid
NP	Blue parallelogram
dem	Blue circle
rel	Blue square



Разметка на эпизоды - гипотеза

Иерархия доступности [Ariel 1988; Ariel 1990]:

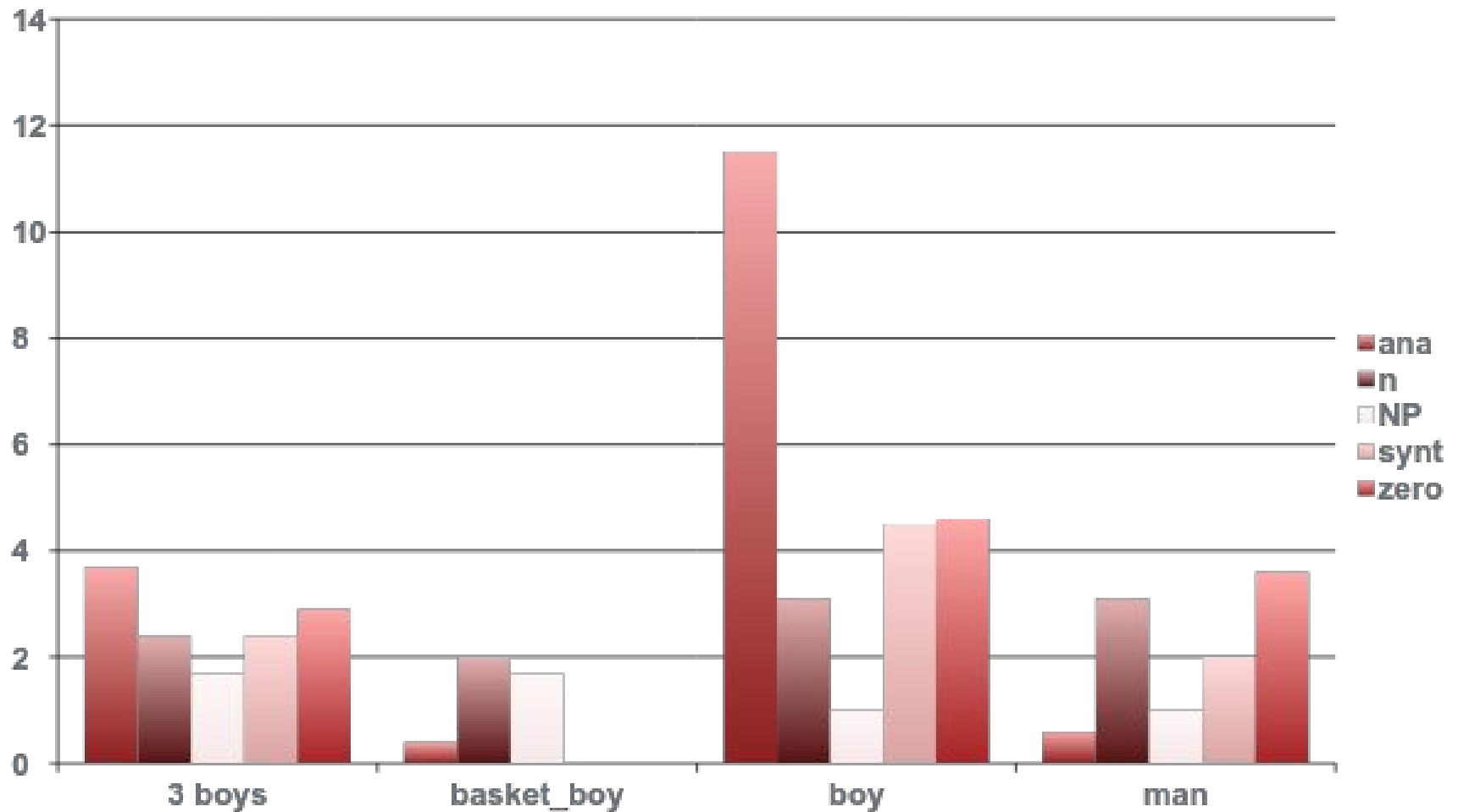
- ▶ маркеры легкодоступности (*нулевая анафора и анафорические местоимения*, обычно местоимения 3 лица)
- ▶ маркеры среднедоступности (*указательные местоимения*)
- ▶ маркеры труднодоступности (*полные ИГ*)

Гипотеза: при смене эпизода говорящий переходит от легкодоступных средств к труднодоступным

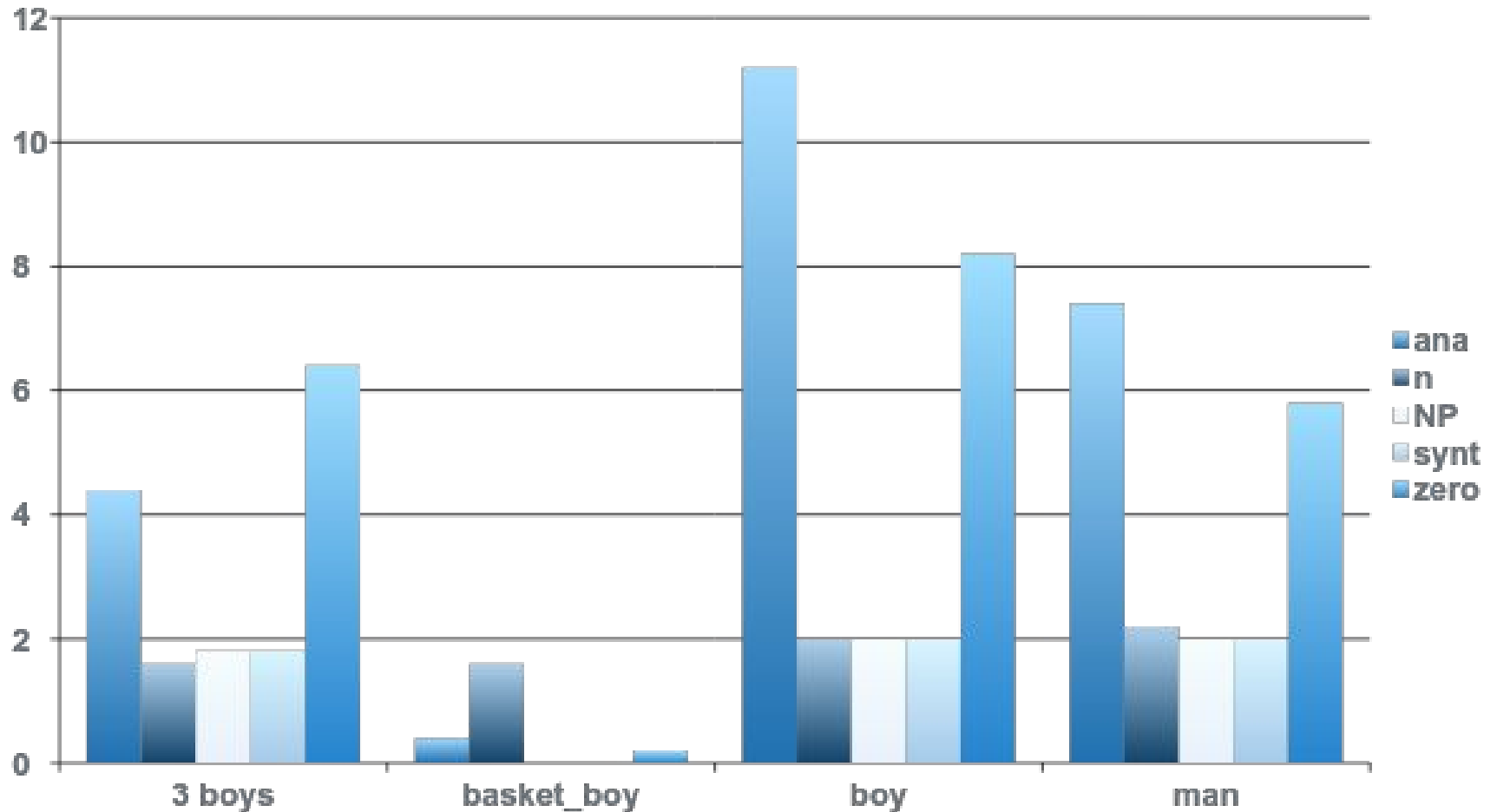
3.

Результаты

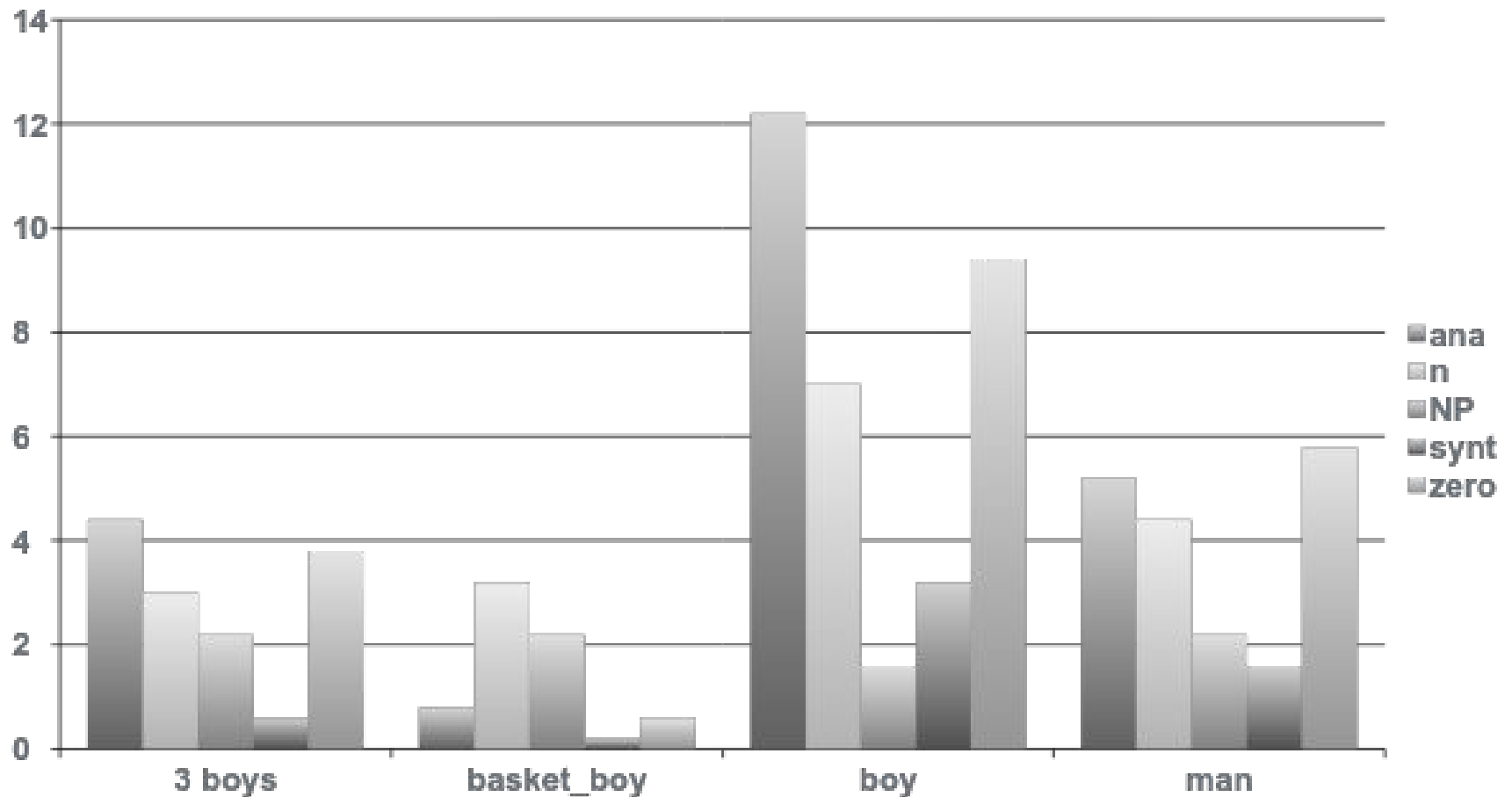
Распределение референциальных средств при упоминании четырех основных сущностей у здоровых людей



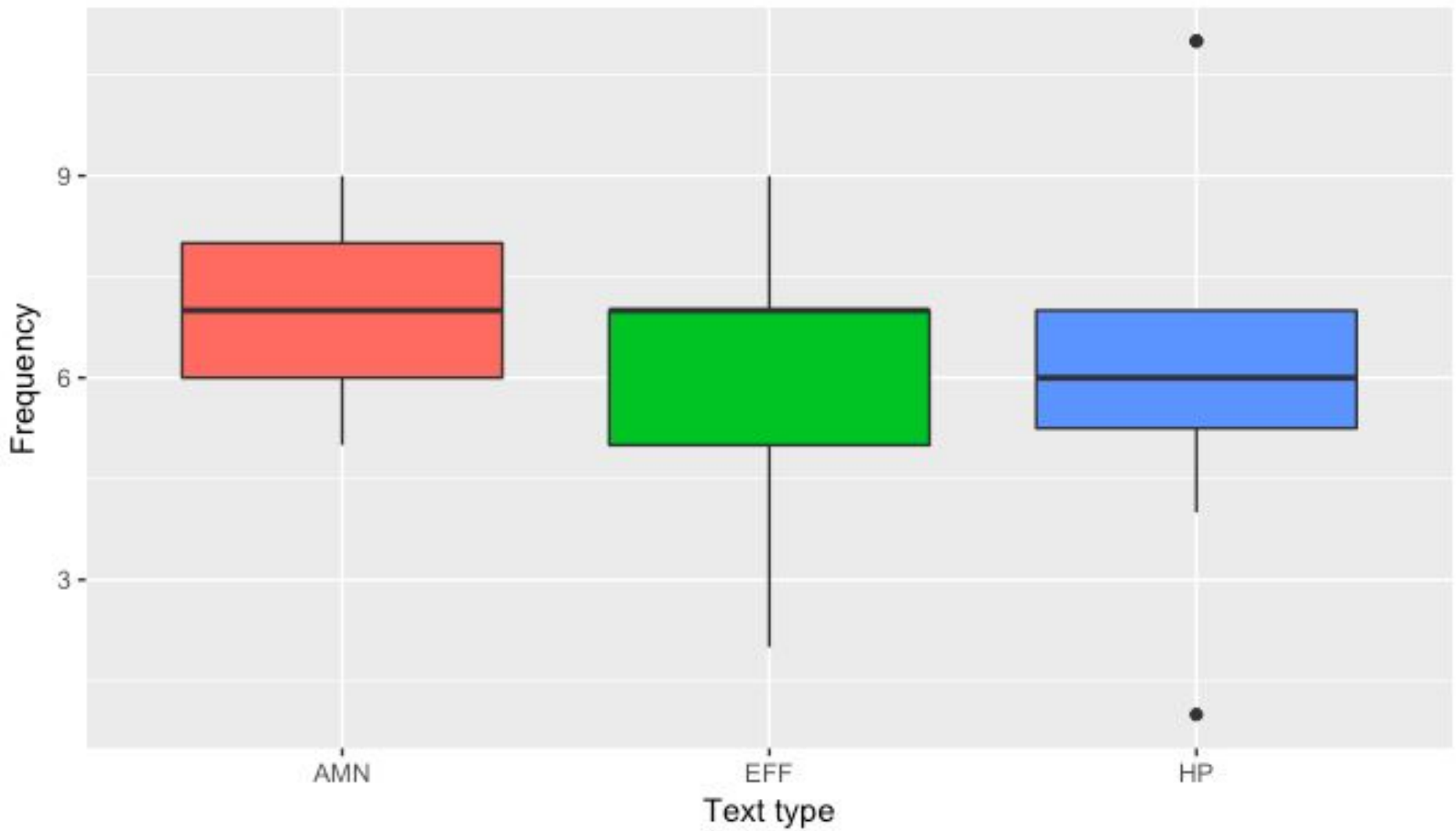
Распределение референциальных средств при упоминании четырех основных существей у людей с акустико-мнестической афазией



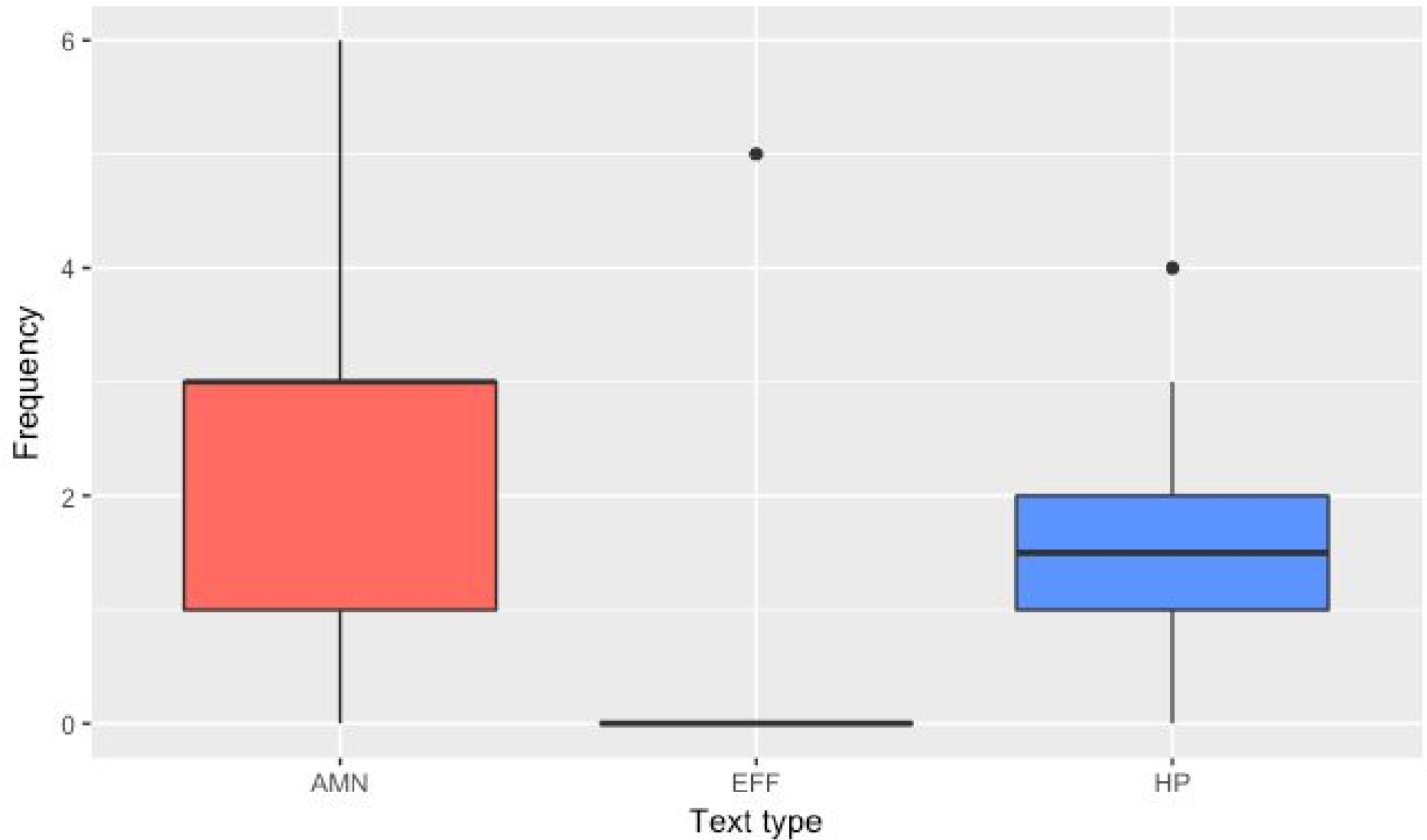
Распределение референциальных средств при упоминании четырех основных существей у людей с эфферентной моторной афазией



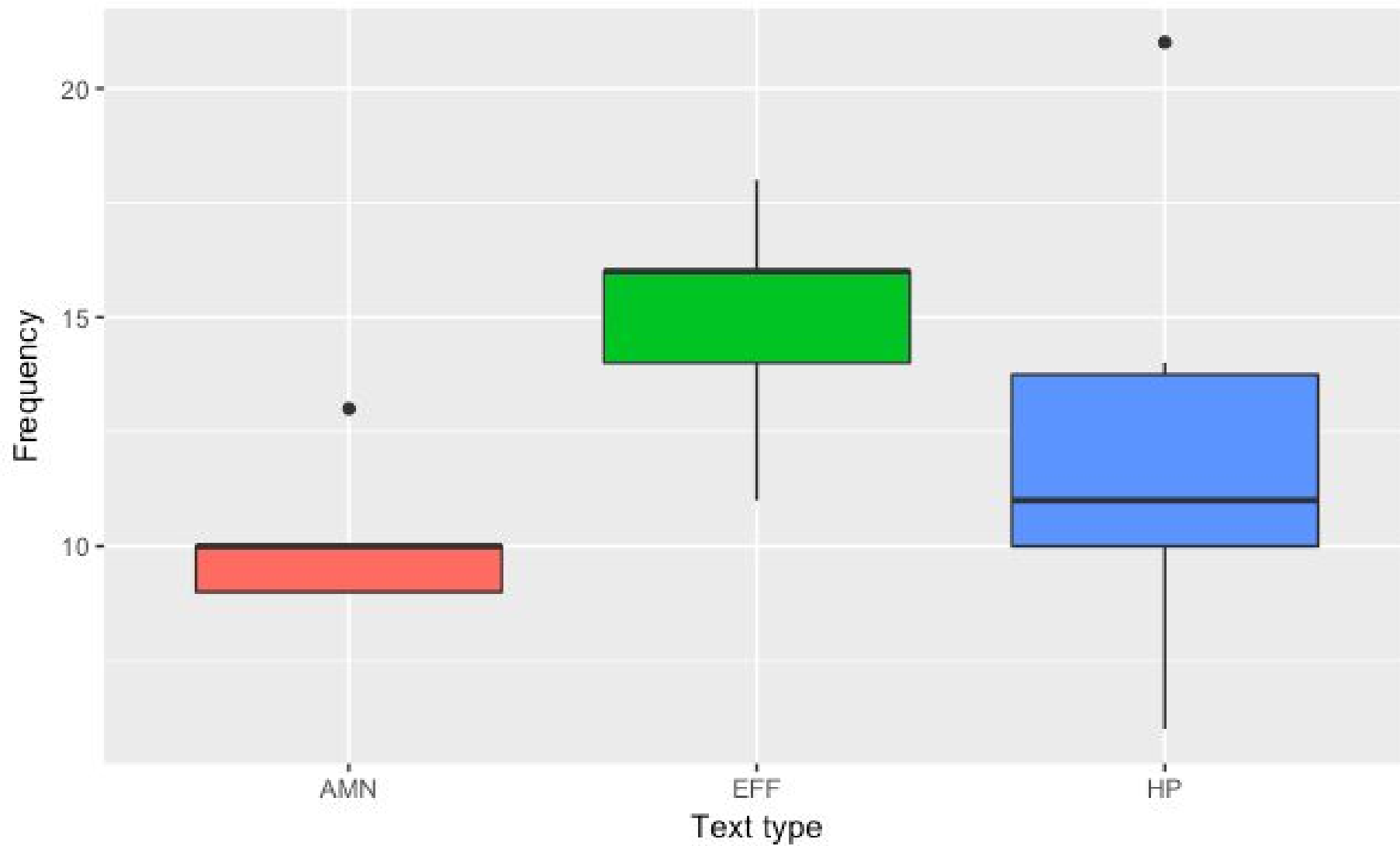
Anaphoric pronoun



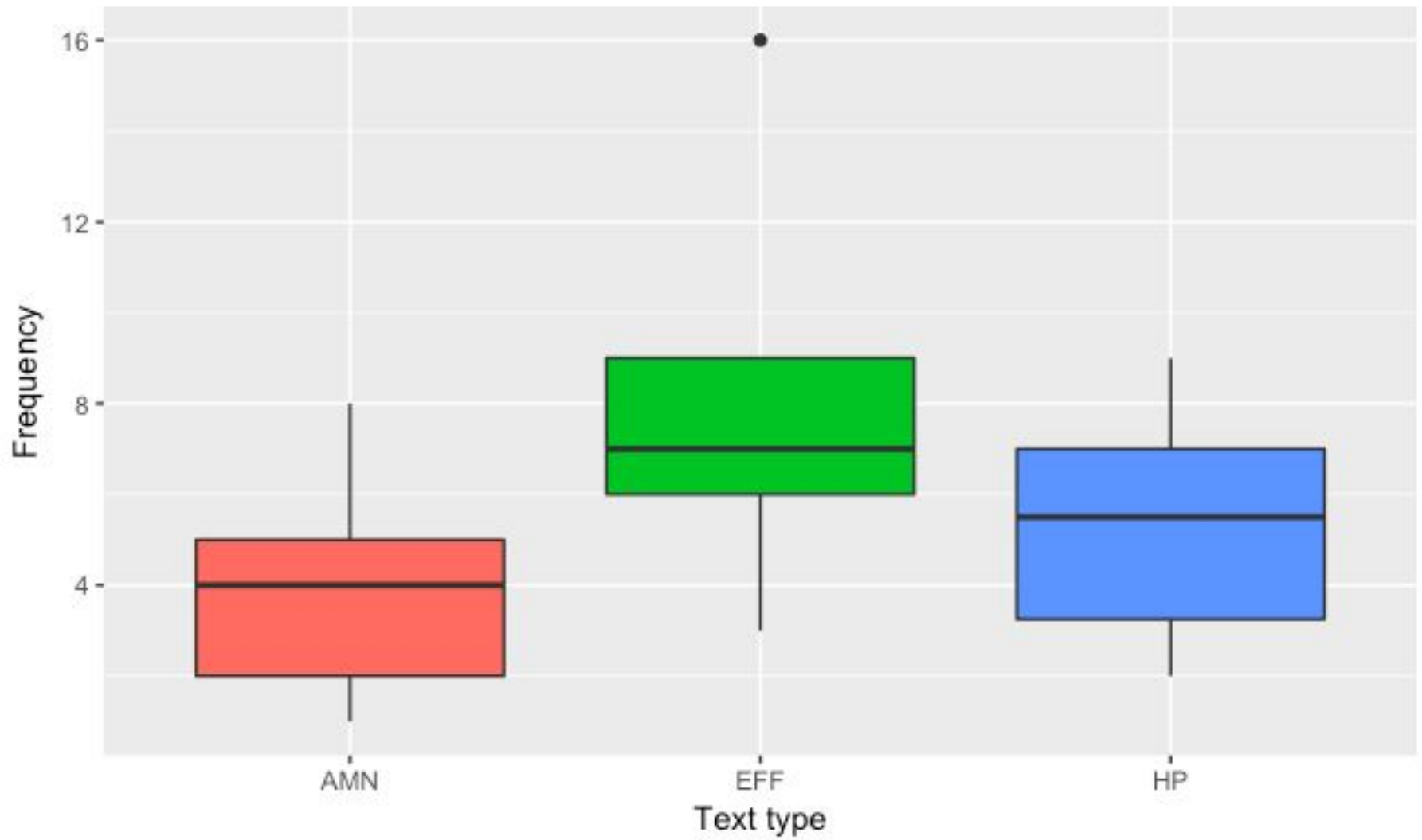
Demonstrative pronoun



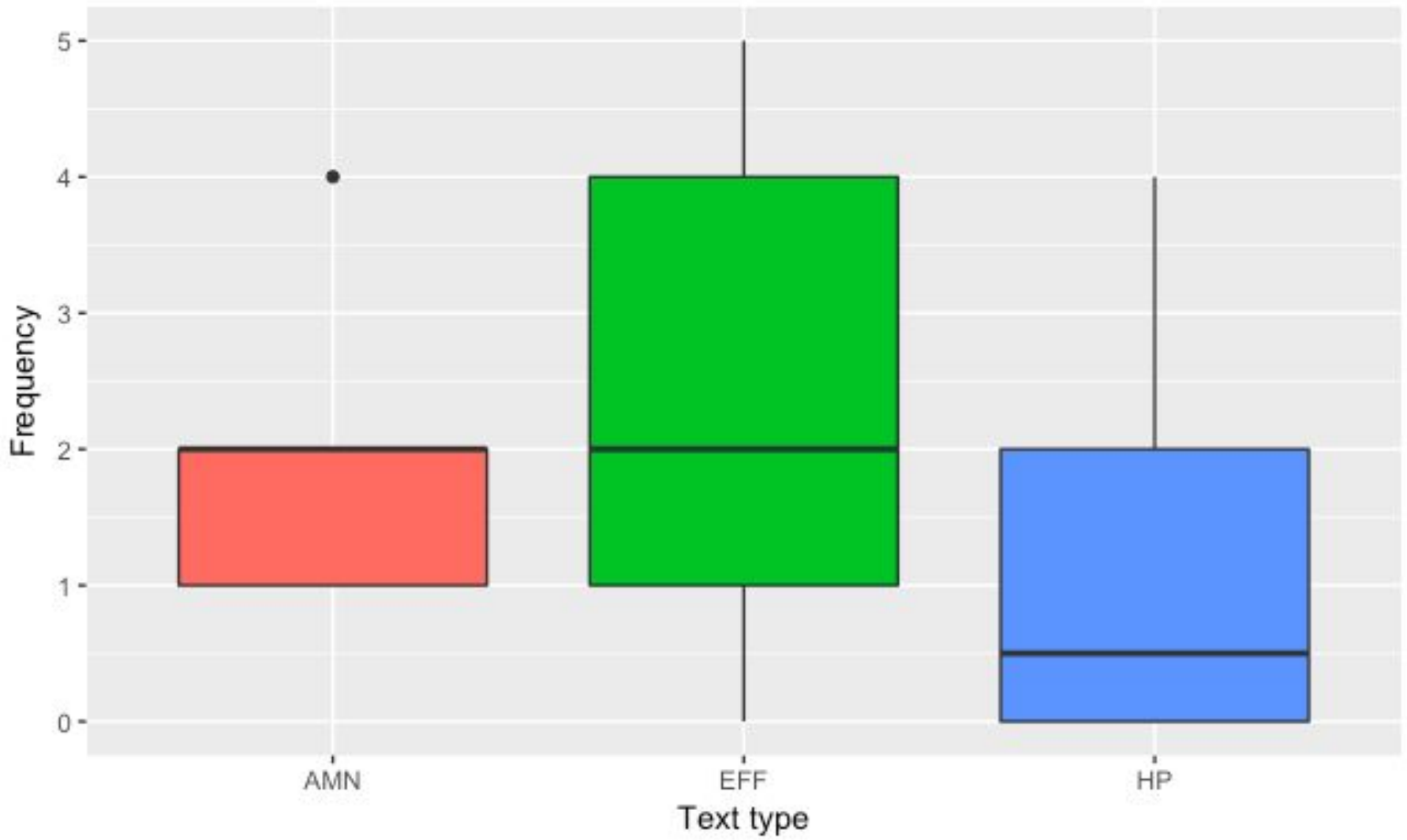
Bare noun



Noun phrase



Zero

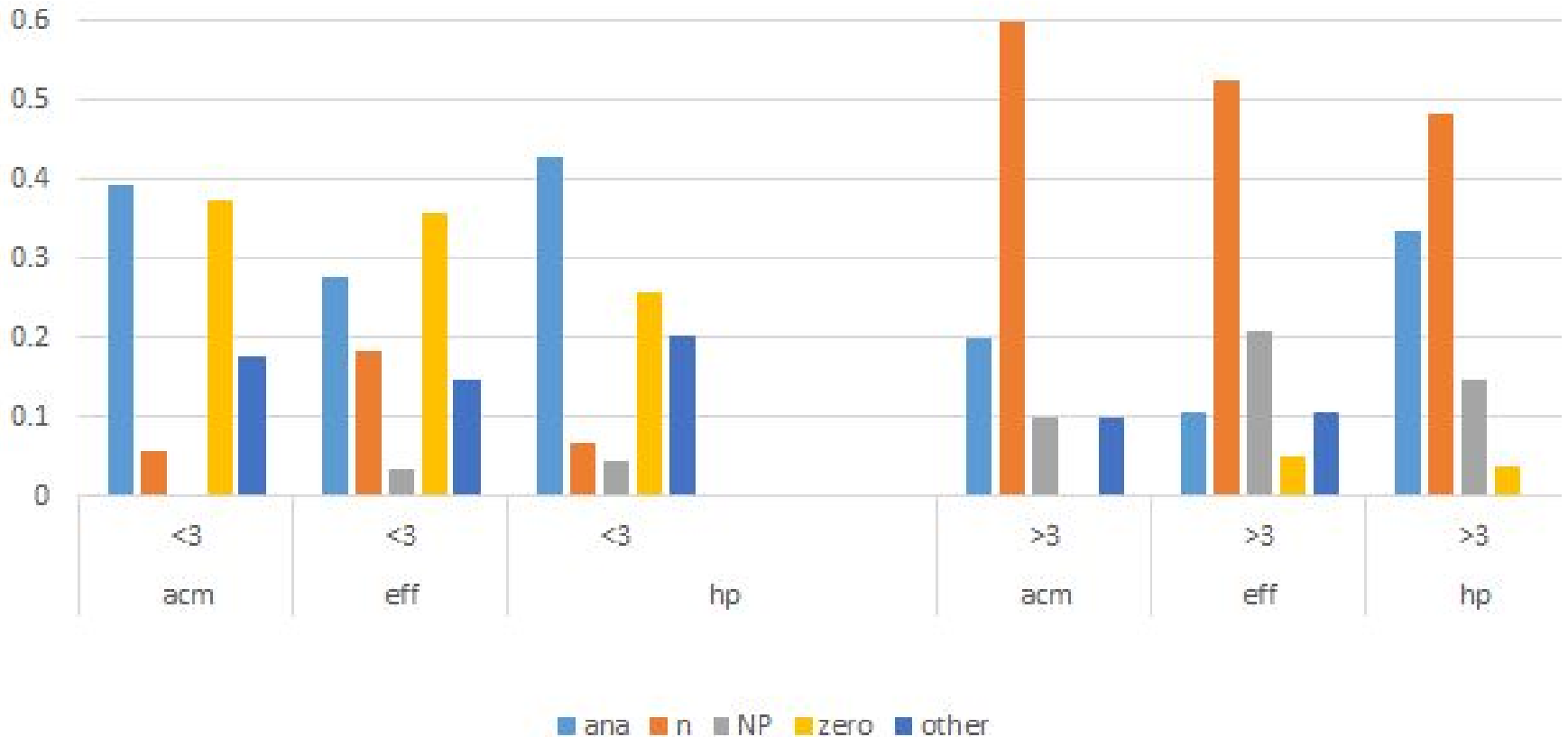


Результаты

- ★ Статистически значимых различий в предпочтении определенных референциальных средств для введения в эпизод новой сущности не обнаружено:
 - между двумя группами афазий
 - между группой афазии и группой нормы
- ★ Однако прослеживаются различия в употреблении нулевых местоимений:
- ★ Также не обнаружено статистически значимых различий в средней длине следующих кореферентных цепочек:
 - цепочка употреблений между двумя n/NP
 - цепочка нулевых местоимений
 - цепочка анафорических местоимений

Альтернативы...

Референциальное расстояние



Спасибо за внимание!