

Splitting of complex prepositional phrases in Russian: an experimental approach

Daria Belova

Lomonosov Moscow State University

dd.belova@yandex.ru

This research has been supported by the Interdisciplinary Scientific and Educational School of Moscow University “Brain, Cognitive Systems, Artificial Intelligence”

Расщепление составляющих

- Расщепление (*separation construction / discontinuous constructions / XP-splits*)
 - такая конфигурация, при которой части составляющих реализуются в поверхностной структуре разделенными другими (ненулевыми) элементами предложения:

(1) Koje je Ivan zanimjive kupio knige?
which is Ivan interesting bought book
‘Which interesting books did Ivan buy?’
(Croatian, [Fanselow & Ćavar 2002])

Расщепление составляющих: русский язык

- [Sekerina 1997]: экспериментальное исследование расщепления при скрэмблинге на материале русского языка
- Чтение с саморегуляцией скорости

(2)

a. <i>O krasivoj</i> of beautiful-PREP	vspominal thought	moj dvojurodnjy brat my cousin-NOM	<i>studentke.</i> student- PREP	SPLIT SCRAMBLING, NO CATEGORICAL AMBIGUITY
'My cousin thought of the beautiful student.'				
b. <i>O studentke</i> of student-PREP	vspominal thought	moj dvojurodnjy brat my cousin-NOM	<i>postojanno.</i> constantly	XP-SCRAMBLING, NO CATEGORICAL AMBIGUITY
'My cousin thought of the student constantly.'				
c. <i>O znakomoj</i> of acquaintance-PREP	vspominal thought	moj dvojurodnjy brat my cousin-NOM	<i>studentke.</i> student- PREP	SPLIT SCRAMBLING, CATEGORICAL AMBIGUITY
'My cousin thought of the student acquaintance.' (Lit. 'the acquaintance student')				
d. <i>O znakomoj</i> of acquaintance-PREP	vspominal thought	moj dvojurodnjy brat my cousin-NOM	<i>postojanno.</i> constantly	XP-SCRAMBLING, CATEGORICAL AMBIGUITY
'My cousin thought of the acquaintance constantly.'				

Расщепление составляющих: русский язык

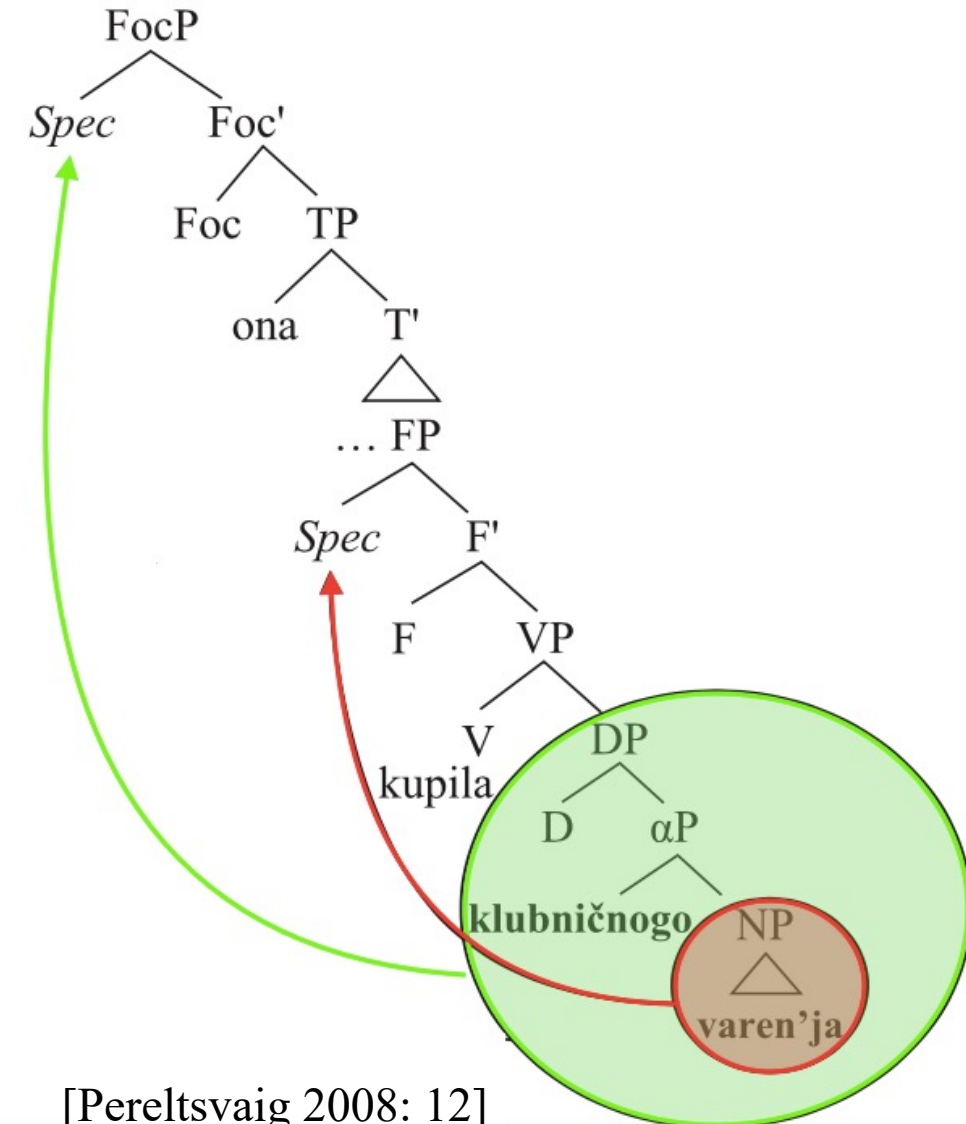
- [Sekerina 1997]: экспериментальное исследование расщепления при скрэмблинге на материале русского языка
- Чтение с саморегуляцией скорости:
 - глагол (второй элемент):
 - имя / адвербиал (последний элемент):
- Оба элемента при расщеплении читаются значимо дольше, чем при полном скрэмблинге
- Сплит-скрэмблинг требует больше ресурсов для обработки, чем полный скрэмблинг (а он – больше, чем конфигурация без скрэмблинга)

Расщепление составляющих: русский язык

- Семейство подходов с двухэтапной деривацией расщепления

- Double movement analysis ([Sekerina 1997])
- Remnant movement analysis (i.a. [Bašič 2004]):

- Клубничного она варенья купила.
- **NP** передвигается в Spec,FP
- **DP**, содержащая след от **NP**, передвигается выше, в Spec,FocP



Расщепление составляющих: русский язык

- [Pereltsvaig 2008]: корпусное исследование разговорного русского языка
- Контраргументы к двухэтапному подходу
 - Ни левая, ни правая части сплита не обязаны формировать составляющую:
(3) Незлобивый у него совершенно характер.
 - Возможны расщепления не на две, а на три части:
(4) Очень они хорошие были люди.
 - Анализ предсказывает немаркированный порядок слов, подходящий для хорватского, но не для русского:
(5) Курицу на большую положи тарелку.

Расщепление составляющих: русский язык

- Разрозненное удаление (scattered deletion, см. [Fanselow & Ćavar 2002])
- Copy and Deletion (см. [Chomsky 1993]): передвижение составляющей происходит путем создания копии этой составляющей в структурно более высокой позиции
 - В нормальном случае нижняя копия удаляется полностью
 - В случае разрозненного удаления составляющая озвучивается частично в верхней копии, частично – в нижней

Расщепление составляющих: русский язык

- Разрозненное удаление (scattered deletion): составляющая передвигается целиком, но озвучивается частично в верхней копии, частично – в нижней:

- (6)
- a. Против советской он выступал власти.
 - b. [Против [советской власти]] он выступал [против [советской власти]].
 - c. [Против [советской власти]] он выступал [~~против~~ [советской власти]].

- Считается более ресурсоемкой операцией, чем полное озвучивание (т.к. требует двойного применение операции удаления, а не одинарного, см. [Nunes 2004]).

Исследовательские предпосылки

- Расщепление ограничено A'-передвижениями (см. [Frey 2000]), однако в языках мира зафиксировано, что A'-передвижения внутри себя ведут себя негомогенно (в частности, в немецком, см. [Fanselow & Cavar 2002])

⇒ Есть ли в русском языке различия между расщеплением при разных типах A'-передвижений (вопросительном, относительном и скрэмблинге)?

Исследовательские предпосылки

⇒ Есть ли в русском языке различия между расщеплением при разных типах A'-передвижений (вопросительном, относительном и скрэмблинге)?

- Чувствительно ли расщепление к фонологическому весу и/или структурного статуса расщепляемых частей?

⇒ Есть ли различия в отщеплении левого элемента-посессора типа *чей* (Spec,DP) и адъектива типа *какой/этот* (D⁰)

Исследовательские предпосылки

- ⇒ Есть ли в русском языке различия между расщеплением при разных типах A'-передвижений (вопросительном, относительном и скрэмблинге)?
- ⇒ Есть ли различия в отщеплении левого элемента-посессора типа *чей* (Spec,DP) и адъектива типа *какой/этот* (D⁰)
- Remnant movement-анализ предполагает передвижение NP. Если у NP есть комплемент, то анализ предсказывает, что он будет двигаться вместе с ней.
- ⇒ Можно ли отщепить N⁰ от комплемента?

Экспериментальное исследование: дизайн

- Расщепление РР с именными группами, содержащими инфинитивный или РР-комплемент
- Три независимых переменных:
 - тип передвижения: вопросительное, относительное, скрэмблинг
 - тип левого элемента: *чей*/посессив, *какой*/детерминатор
 - позиция сплита: до именной вершины (ранний) / после вершины (поздний)

⇒ **10 условий**
- Шкала Ликерта 1–7 + чтение с саморегуляцией скорости
 - ⇒ **две зависимых переменных: оценка и время чтения**

Экспериментальное исследование: дизайн

- 30 экспериментальных блоков
- Шесть предлогов, каждый встречался пять раз: *о, за, против, от, на, в*
- Сбалансированы по типу компонента именной группы:
 - 15 блоков – инфинитивный, 15 блоков – зависимая РР
- В качестве детерминаторов выступали слова *этот* и *один*
- В конец стимулов с вопросительным и относительным передвижением были добавлены конечные вопросы, задающие пространство выбора, для облегчения интерпретации и избежания эффектов конца предложения при чтении
- По правилу латинского квадрата по трем независимым переменным распределены на 10 экспериментальных листов
- Каждый экспериментальный лист содержал 15 грамматичных и 15 неграмматичных филлеров

Экспериментальное исследование: блок а

- (7)
- a. О чьём Соня вспомнила совете **варить мясо на кости**, Томином или Женином?
 - b. О чьём совете Соня вспомнила **варить мясо на кости**, Томином или Женином?
 - c. Повар, о чьём Соня вспомнила совете **варить мясо на кости**, выпустил книгу рецептов.
 - d. Повар, о чьём совете Соня вспомнила **варить мясо на кости**, выпустил книгу рецептов.
 - e. О каком Соня вспомнила совете **варить мясо на кости**, полезном или нет?
 - f. О каком совете Соня вспомнила **варить мясо на кости**, полезном или нет?
 - g. Об Анином Соня вспомнила совете **варить мясо на кости**.
 - h. Об Анином совете Соня вспомнила **варить мясо на кости**.
 - i. Об одном Соня вспомнила совете **варить мясо на кости**.
 - j. Об одном совете Соня вспомнила **варить мясо на кости**.

Экспериментальное исследование: блок b

- (8)
- a. В чьей Женя ошиблась выкройке для **офисной юбки**, твоей или моей?
 - b. В чьей выкройке Женя ошиблась для **офисной юбки**, твоей или моей?
 - c. Женщина, в чьей Женя ошиблась выкройке для **офисной юбки**, отказалась платить за работу.
 - d. Женщина, в чьей выкройке Женя ошиблась для **офисной юбки**, отказалась платить за работу.
 - e. В какой Женя ошиблась выкройке для **офисной юбки**, заказанной или нет?
 - f. В какой выкройке Женя ошиблась для **офисной юбки**, заказанной или нет?
 - g. В твоей Женя ошиблась выкройке для **офисной юбки**.
 - h. В твоей выкройке Женя ошиблась для **офисной юбки**.
 - i. В одной Женя ошиблась выкройке для **офисной юбки**.
 - j. В одной выкройке Женя ошиблась для **офисной юбки**.

Экспериментальное исследование: проведение

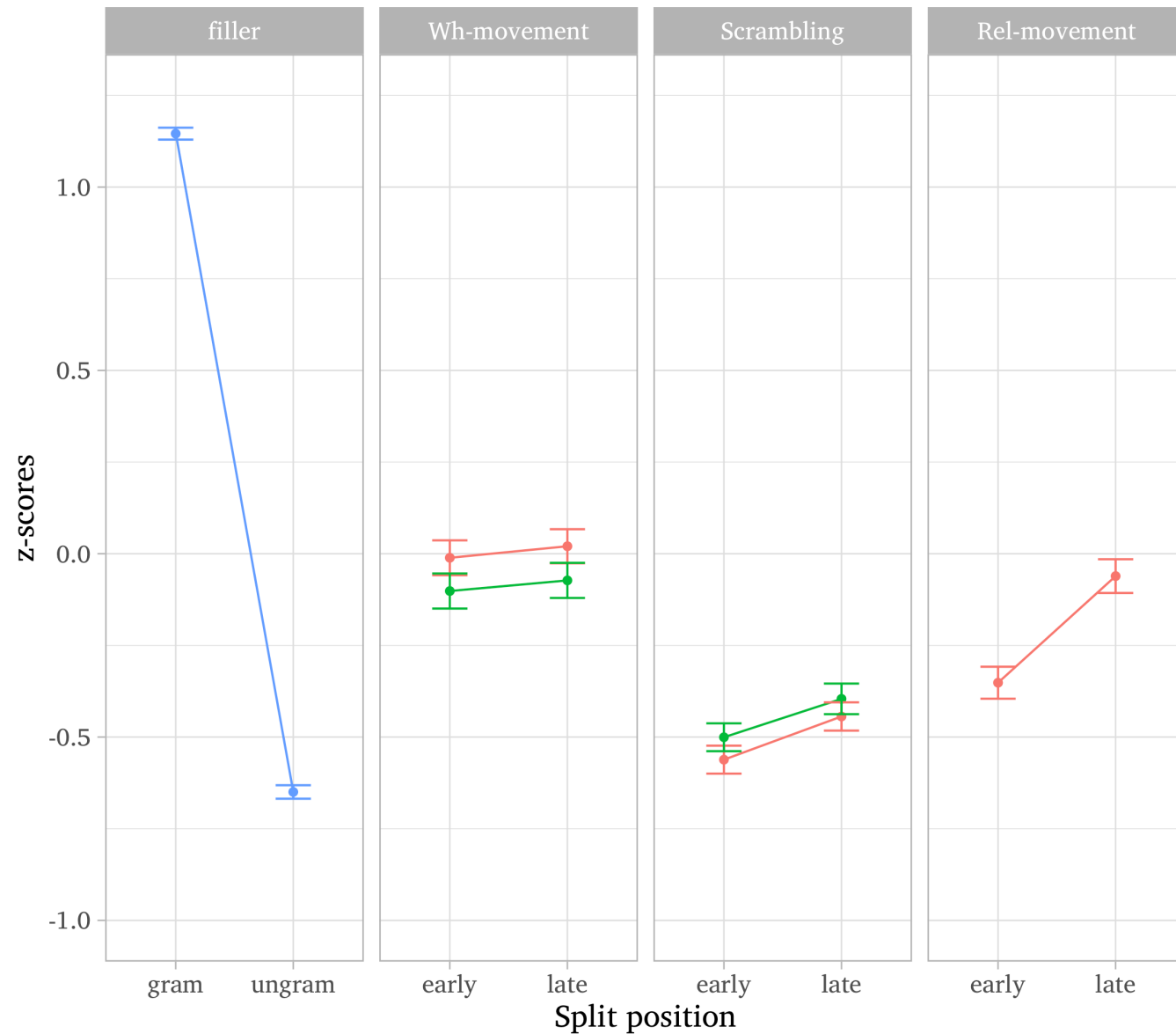
- В начале эксперимента респонденты заполняли социолингвистическую анкету и оценивали три филлерных предложения в качестве тренировки
- В середине эксперимента предлагалось сделать перерыв
- Респонденты привлекались в социальных сетях и на краудсорсинговой платформе «Яндекс.Толока»
- Всего эксперимент прошло 104 человека
 - 7 были отсеяны как аутлаеры: их оценки филлеров отклонялись от ожидаемых более чем на 2σ
 - Возраст 18–76 лет, $sr = 39$, $SD = 10.9$
 - 35 женщин
 - 6 респондентов указали наличие лингвистического образования

Результаты: обработка

- Перед статистической обработкой оценки по шкале Ликерта были нормализованы, чтобы нивелировать индивидуальные особенности использования шкалы:
- Формула нормализации: $Z_{ij} = (X_{ij} - \underline{X}_i) / \sigma_i$
где X_{ij} – исходная оценка, \underline{X}_i – среднее выборочное всех оценок респондента i ,
 σ_i – стандартное отклонение всех оценок респондента i .
- Полученные данные обрабатывались с помощью линейных смешанных моделей и апостериорного попарного сравнения по критерию Тьюки в R

Результаты: оценки

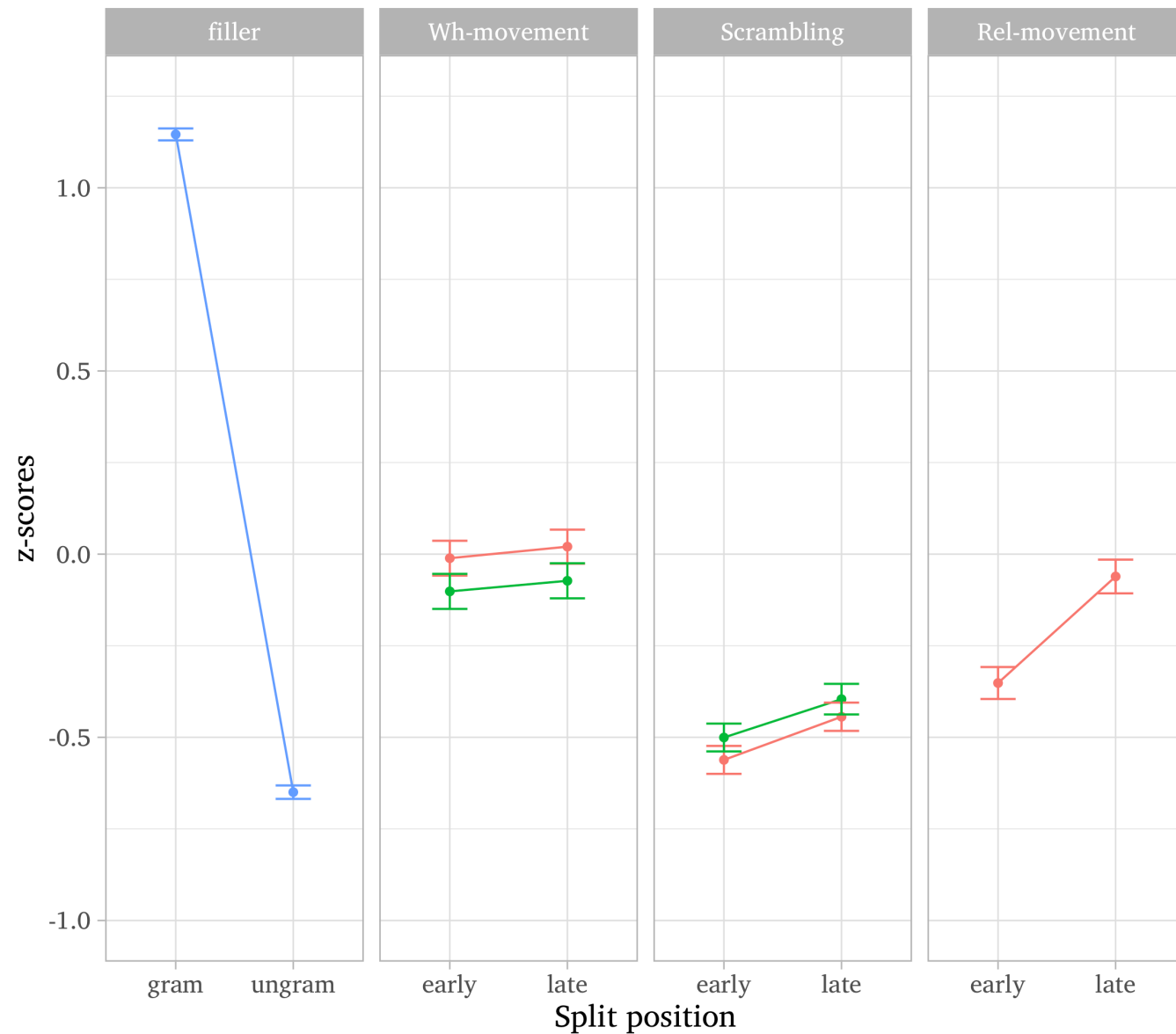
- В качестве случайных эффектов выступали:
 - ID респондента
 - номер предложения
 - тип конструкции (инфинитивный / предложный компонент ИГ)
- Финальная модель:
 $1 + \text{movement} * \text{split position}$
 $+ (1 | \text{sentence_id})$
 $+ (1 | \text{construction})$
 $+ (1 + \text{movement} + \text{split position} | \text{id})$



Left element — Chej / Possessive — Kakoj / Demonstrative — Filler

Результаты: оценки

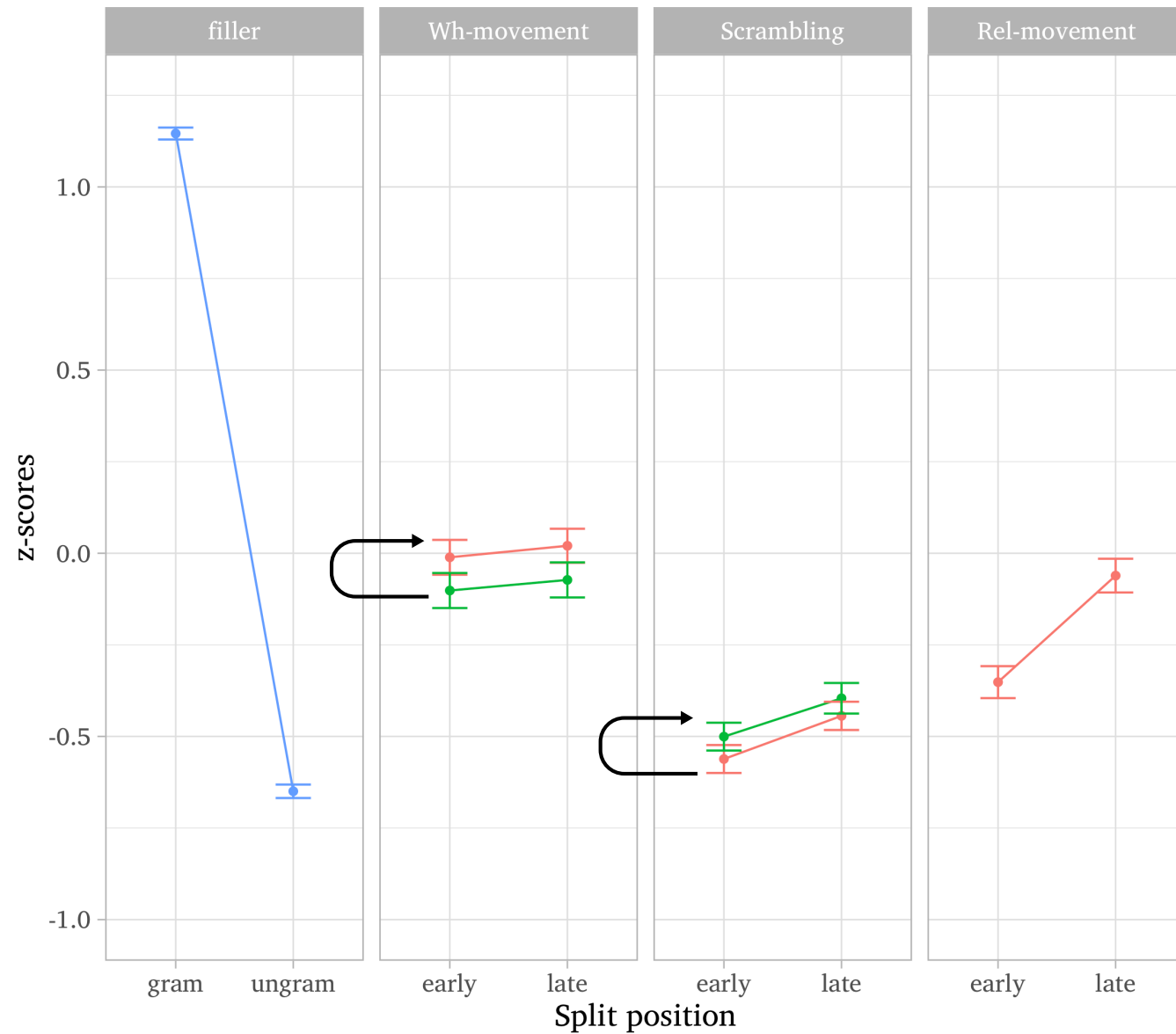
- Общий уровень оценок низкий по сравнению с грамматичными филлерами – из-за нетривиальной информационной структуры



Left element ● Chej / Possessive ● Kakoj / Demonstrative ● Filler

Результаты: оценки

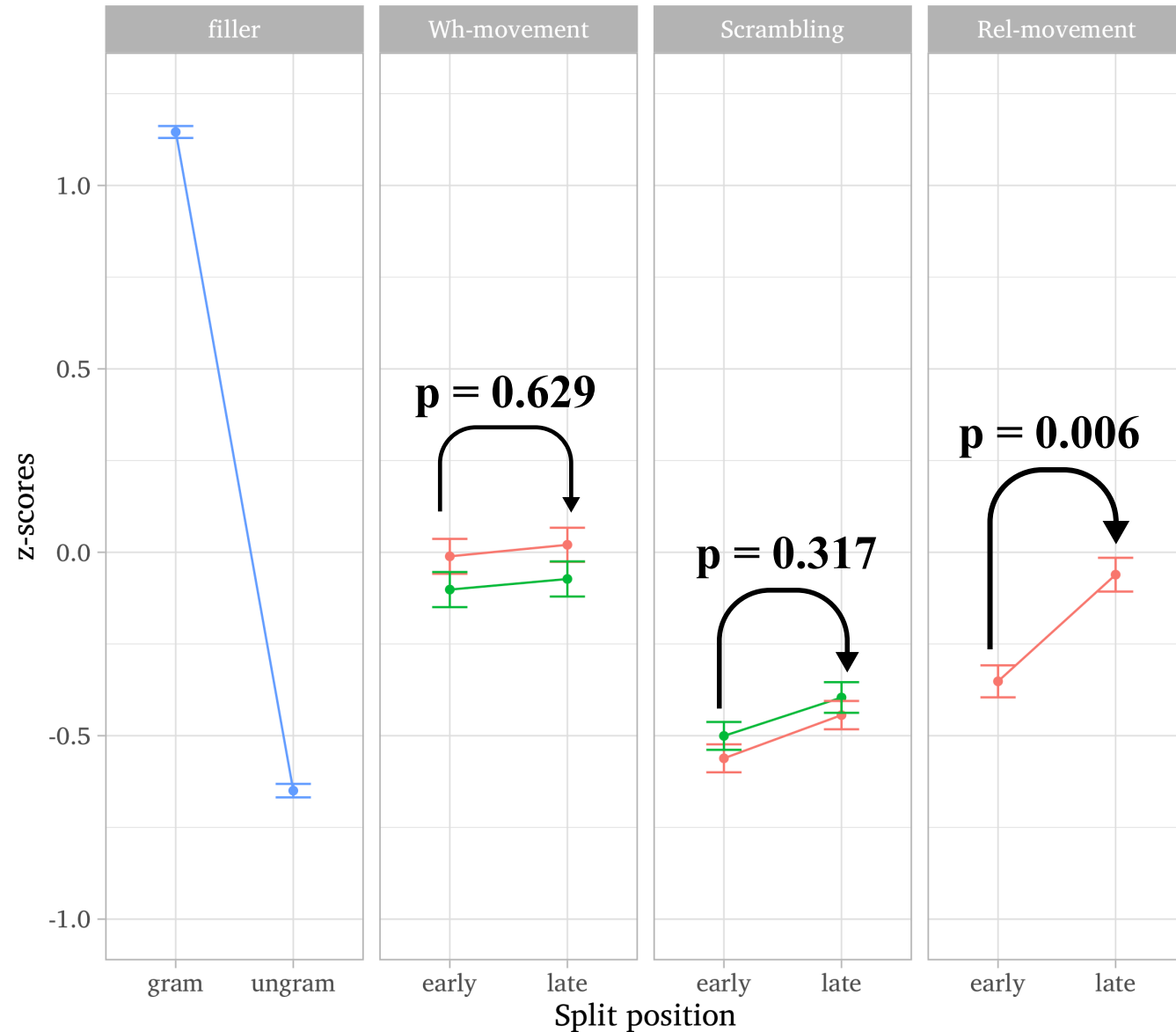
- Левый элемент – незначимый фактор



Left element ● Chej / Possessive ● Kakoj / Demonstrative ● Filler

Результаты: оценки

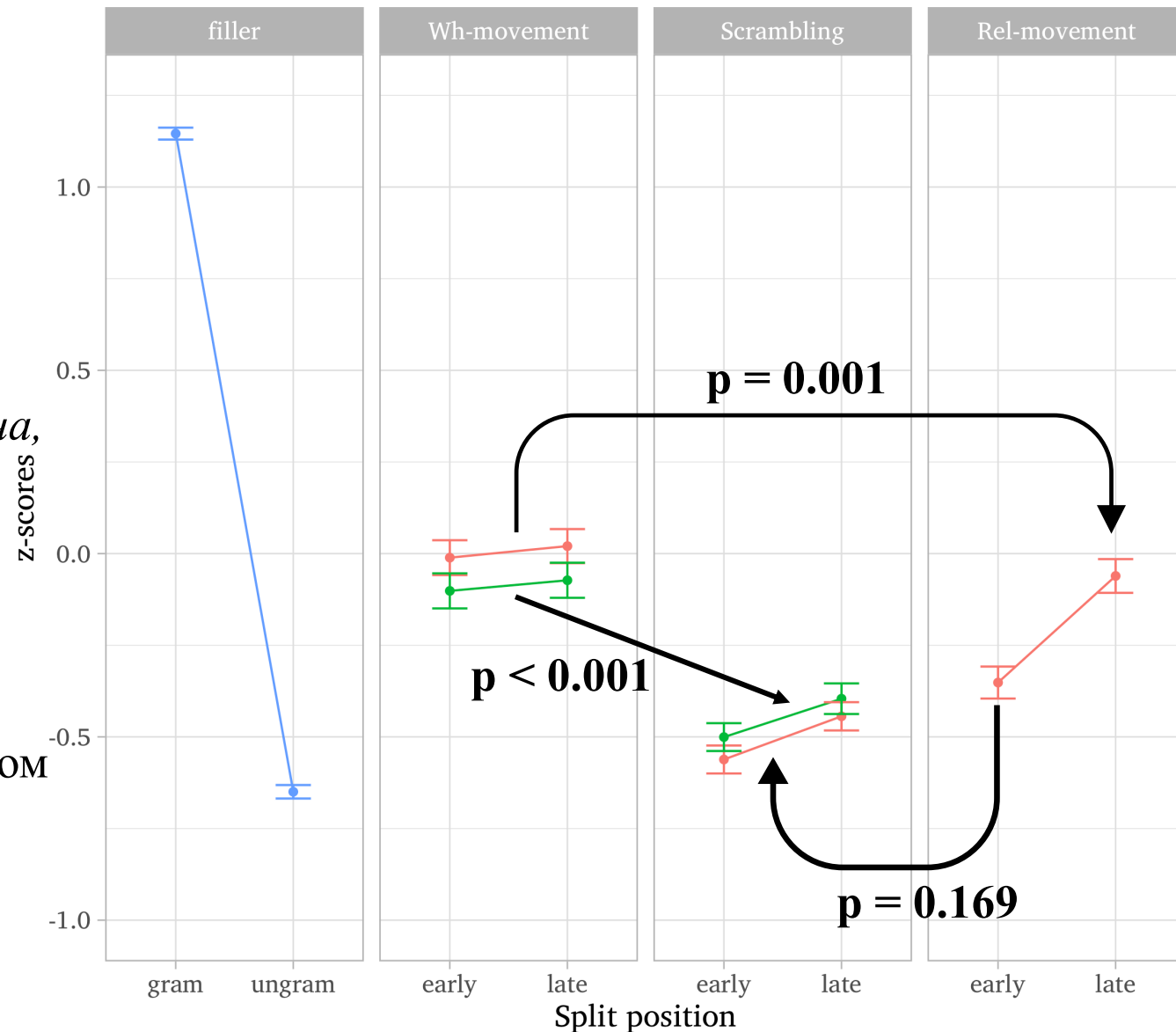
- **Левый элемент** – незначимый фактор
- Фактор **позиции сплита** статистически значим только для релятивизации: поздний сплит («женщина, в чьей выкройке Женя ошиблась для юбки») оценивается выше, чем ранний («женщина, в чьей Женя ошиблась выкройке для юбки»)



Left element — Chej / Possessive — Kakoj / Demonstrative — Filler

Результаты: оценки

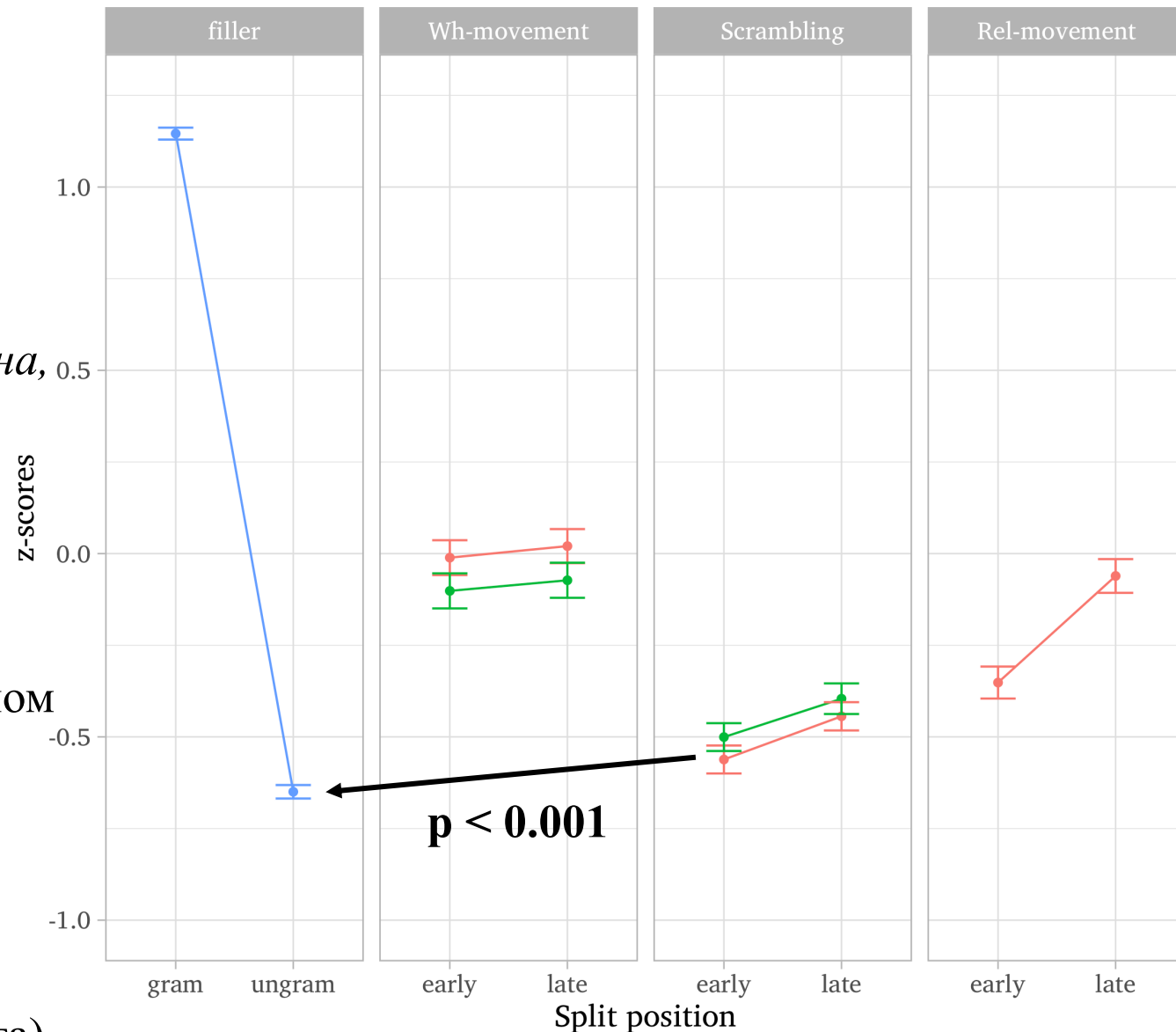
- **Левый элемент** – незначимый фактор
- **Фактор позиции сплита** статистически значим только для релятивизации: поздний сплит («женщина, в чьей выкройке Женя ошиблась для юбки») оценивается выше, чем ранний («женщина, в чьей Женя ошиблась выкройке для юбки»)
- Поздний сплит релятивизации оценивается на одном уровне с вопросительным передвижением
- Ранний сплит релятивизации оценивается на одном уровне со скрэмблингом
- Вопросительное передвижение оценивается значимо выше скрэмблинга



Left element — Chej / Possessive — Kakoj / Demonstrative — Filler

Результаты: оценки

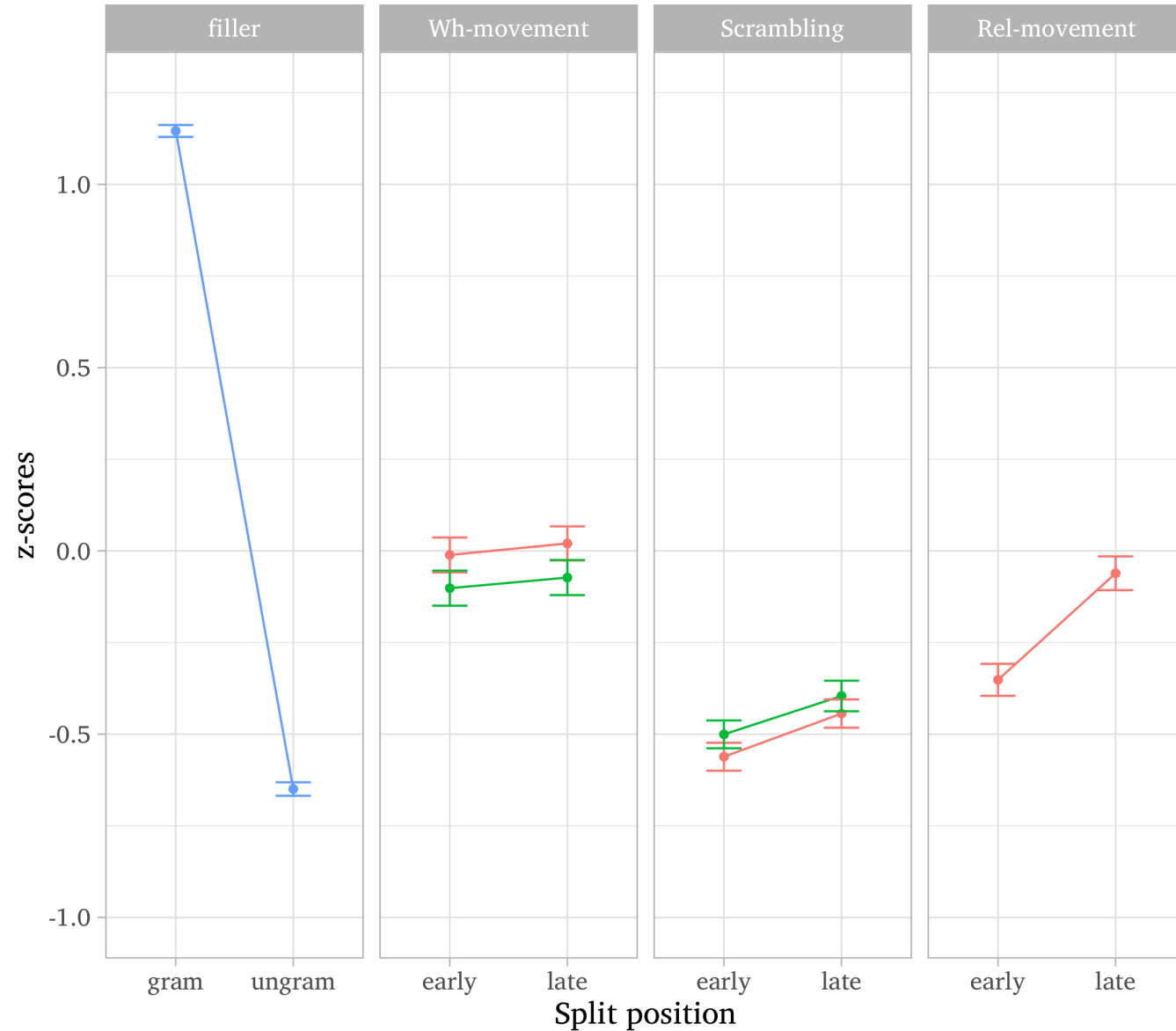
- **Левый элемент** – незначимый фактор
- Фактор **позиции сплита** статистически значим только для релятивизации: поздний сплит («женщина, в чьей выкройке Женя ошиблась для юбки») оценивается выше, чем ранний («женщина, в чьей Женя ошиблась выкройке для юбки»)
- Поздний сплит релятивизации оценивается на одном уровне с вопросительным передвижением
- Ранний сплит релятивизации оценивается на одном уровне со скрэмблингом
- Вопросительное передвижение оценивается значимо выше скрэмблинга
- Самое низкое условие (ранний сплит скрэмблинга) оценивается значимо выше неграмматичных филлеров (двусторонний t-тест Стьюдента)



Left element — Chej / Possessive — Kakoj / Demonstrative — Filler

Результаты: оценки

- Различия в паттернах взаимодействия с позицией сплита интересны
- При позднем сплите элементы на левом краю клаузы (предлог + wh + именная вершина) создают интерпретируемую составляющую, поэтому могут расцениваться как полный скрэмблинг
- Появление компонента требует реанализа этой структуры, в отличие от раннего сплита, где левые элементы не могут интерпретироваться, пока не появится именная вершина
- Можно было бы ожидать, что поздний скрэмблинг будет оцениваться ниже, чем ранний, из-за большей когнитивной ресурсоемкости

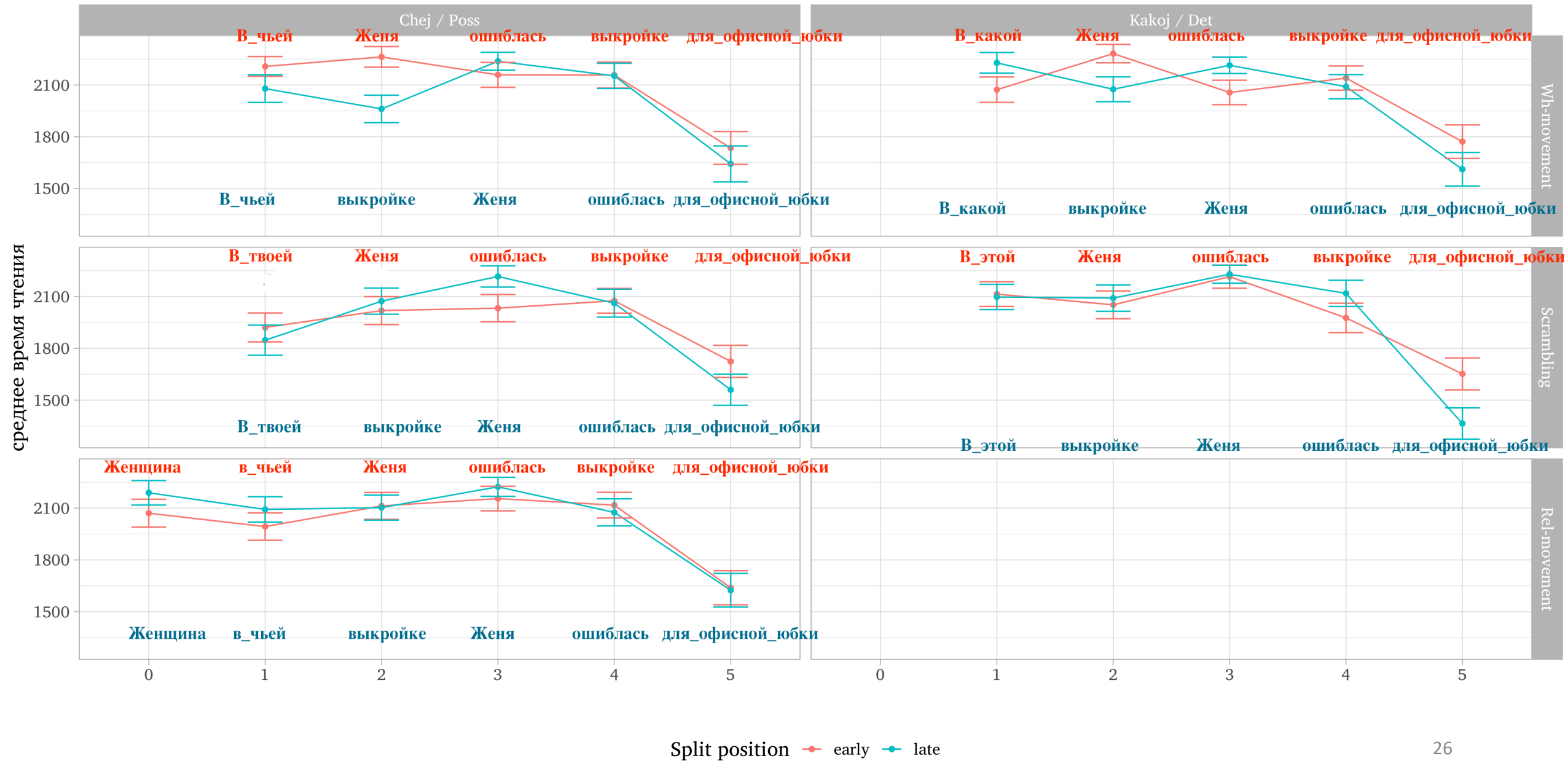


Left element — Chej / Possessive — Kakoj / Demonstrative — Filler

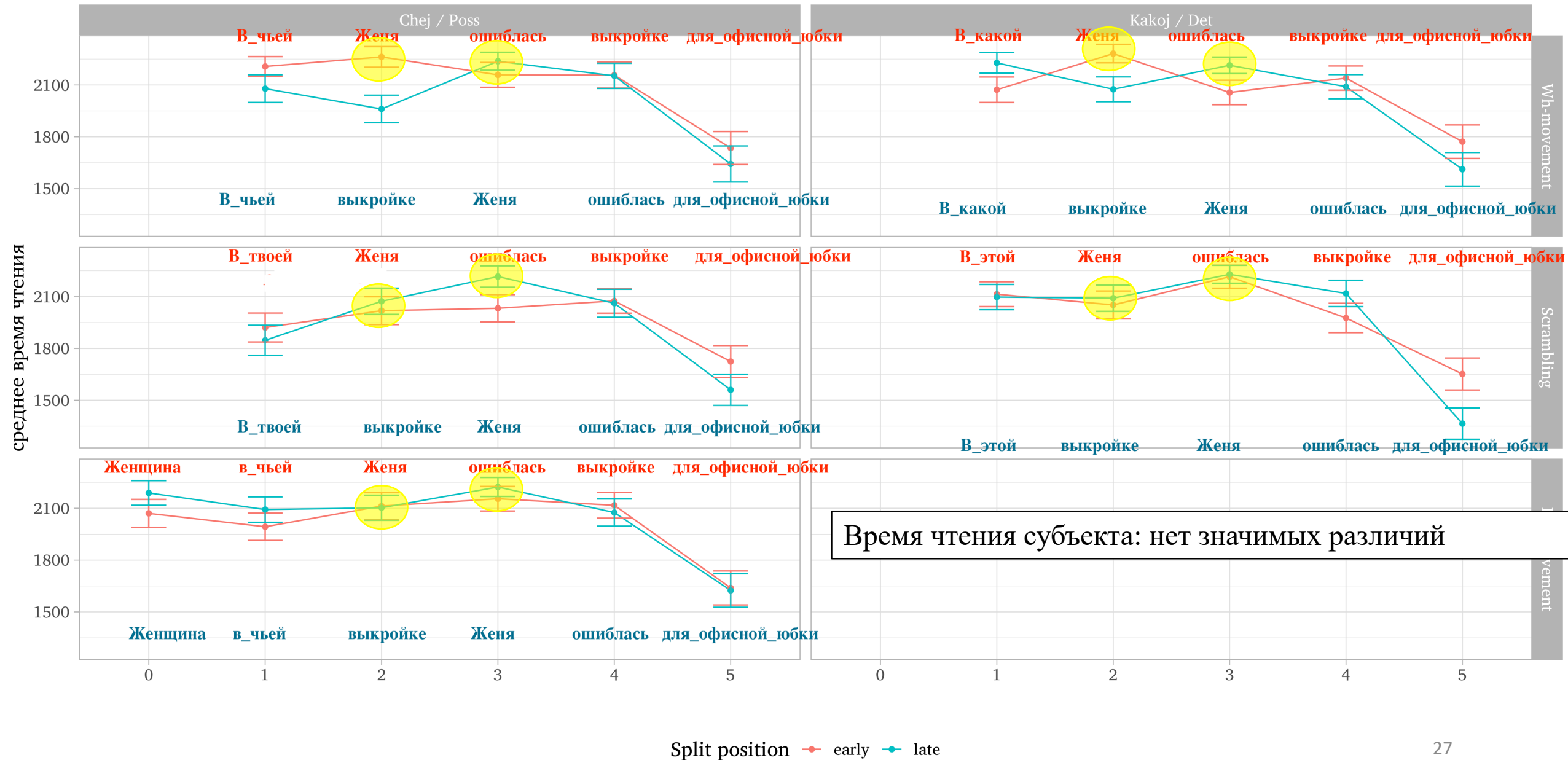
Результаты: время чтения

- Три пункта сравнения времени чтения: субъект (имя собственное), именная вершина и комплемент;
- Для каждого из пунктов создавалась своя ЛСМ, где в качестве фиксированных эффектов проверялись независимые переменные:
 - (10) a. Ранний сплит: Left Subj Verb Noun Comp
 - b. Поздний сплит: Left Noun Subj Verb Comp
- Рабочие гипотезы:
 - После чтения левого элемента респондент ожидает вершину, поэтому появление субъекта при раннем сплите требует перестройки структуры => субъект будет читаться дольше, чем при позднем;
 - Значимых различий во времени чтения именной вершины не ожидается;
 - Комплемент при позднем сплите не ожидается вообще, поскольку левый элемент + вершина формируют интерпретируемую потенциальную составляющую => будет читаться дольше, чем при раннем;

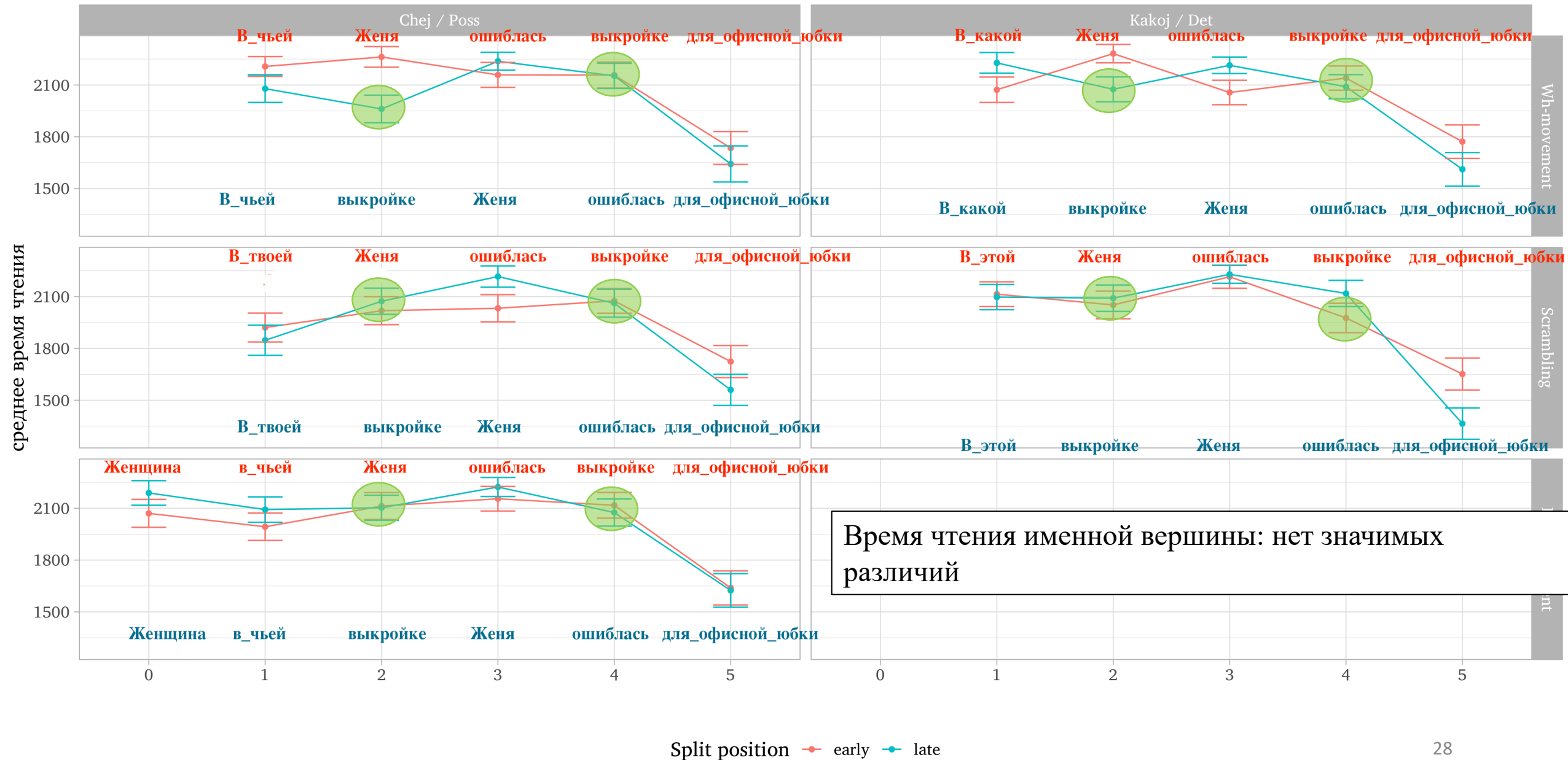
Результаты: время чтения, РР



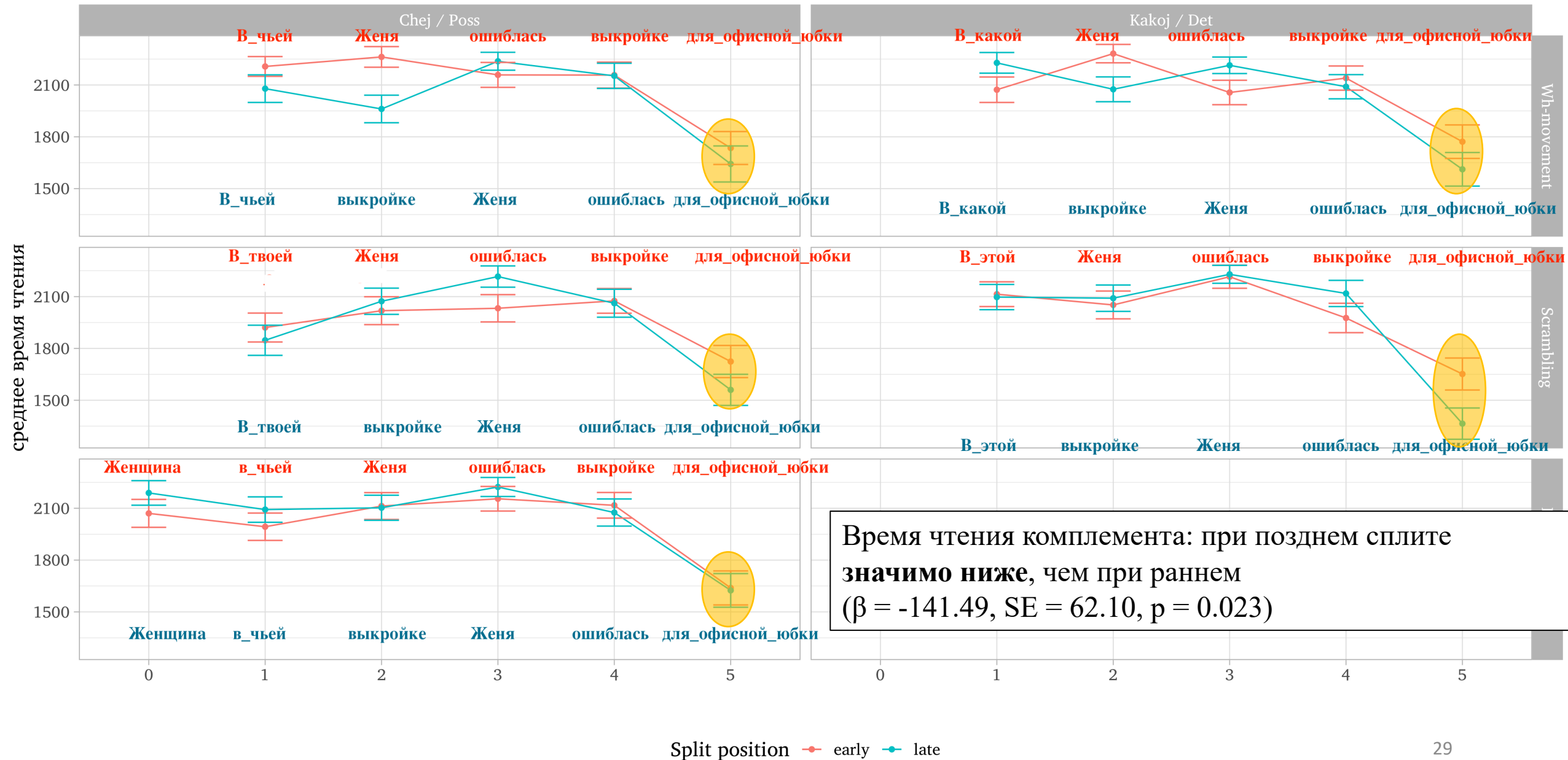
Результаты: время чтения, RR



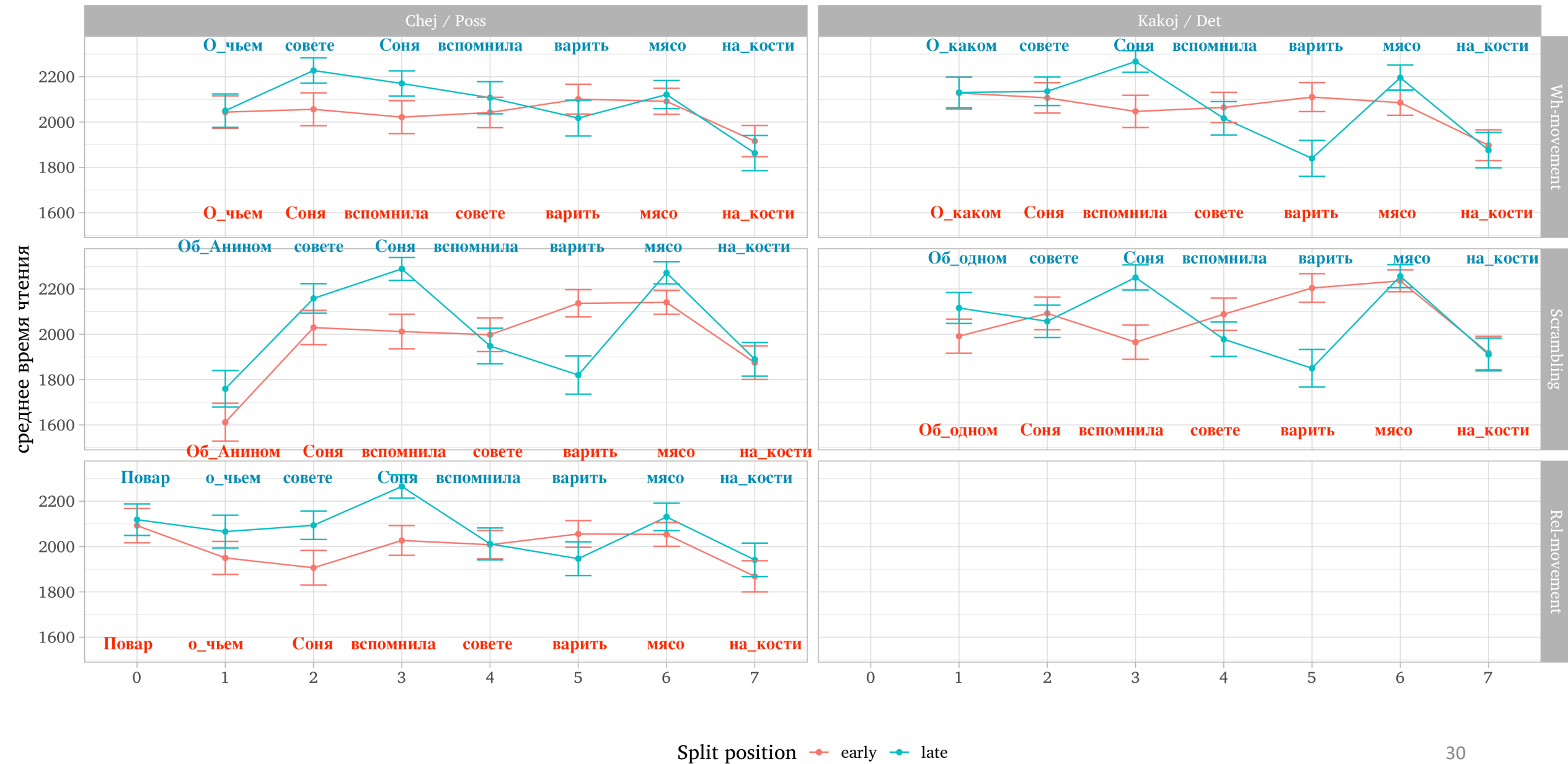
Результаты: время чтения, РР



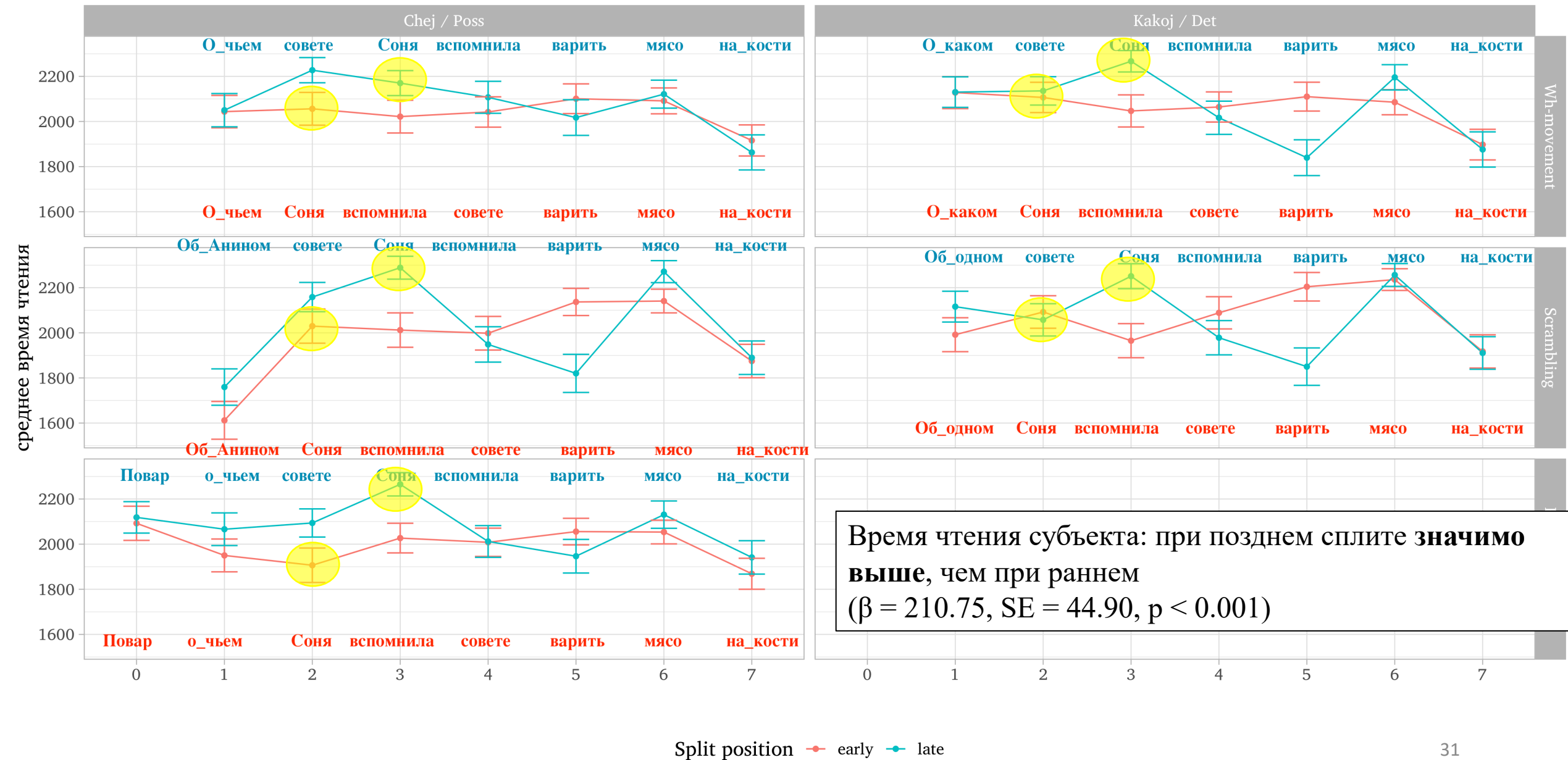
Результаты: время чтения, РР



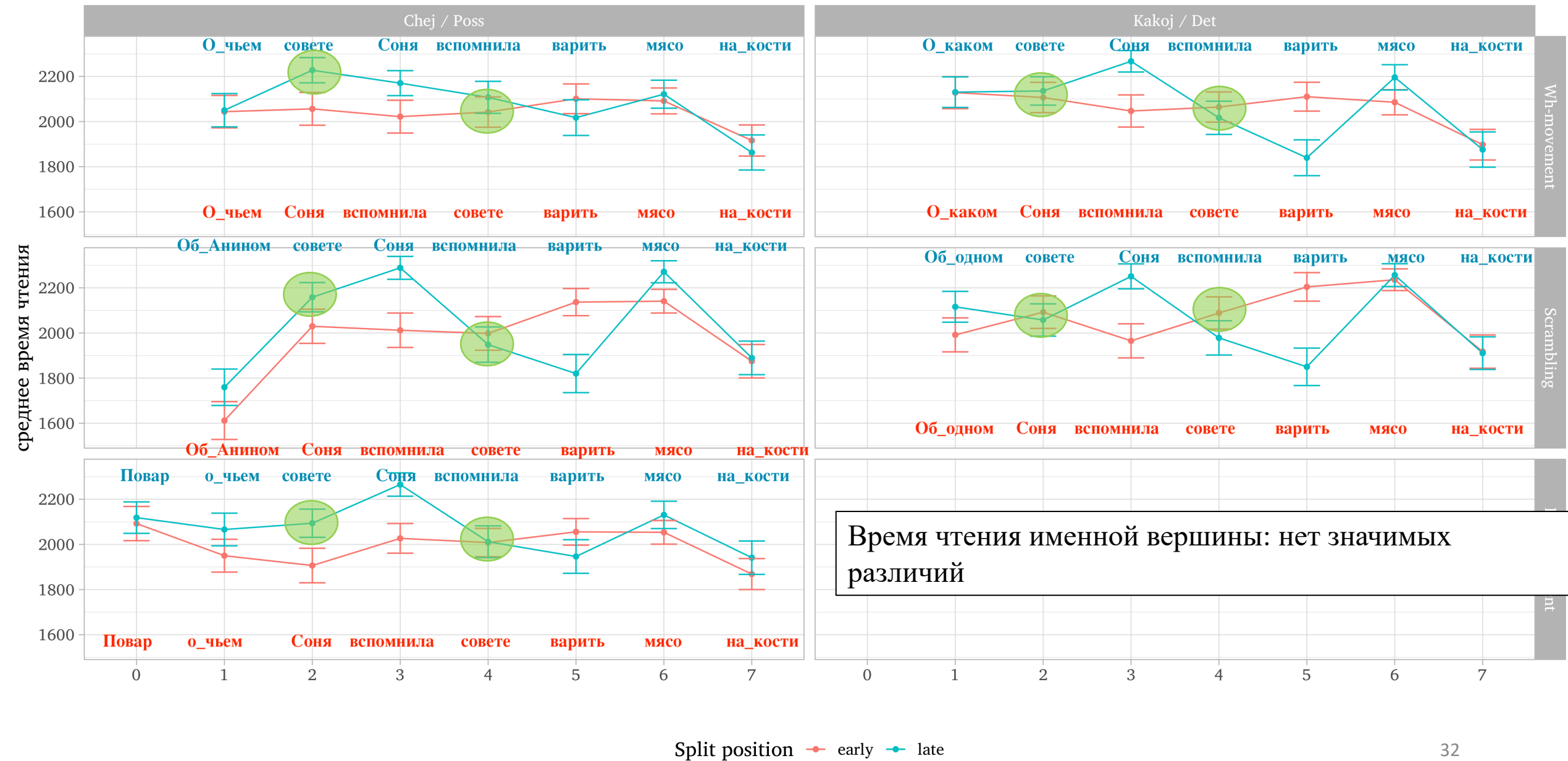
Результаты: время чтения, Inf



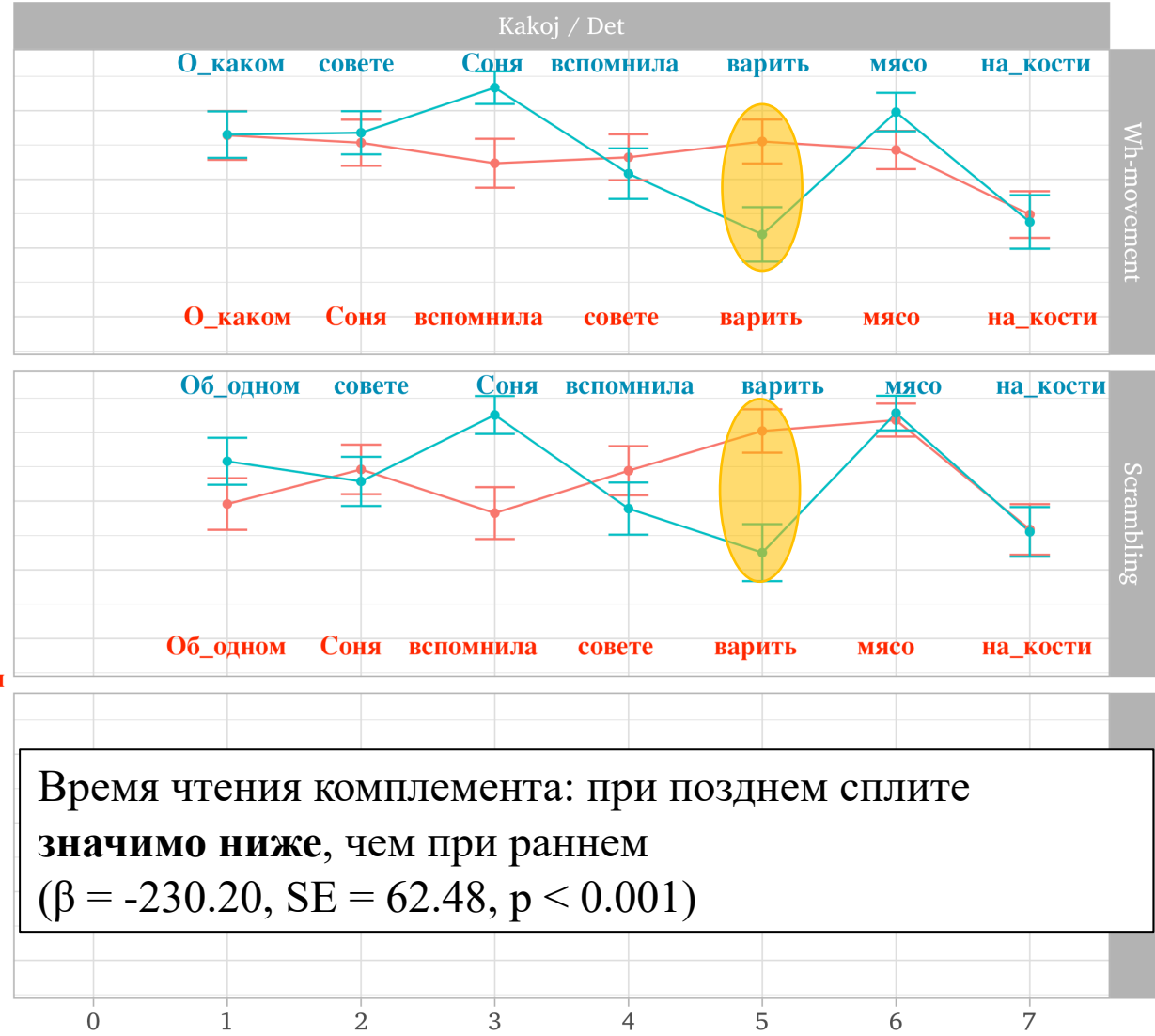
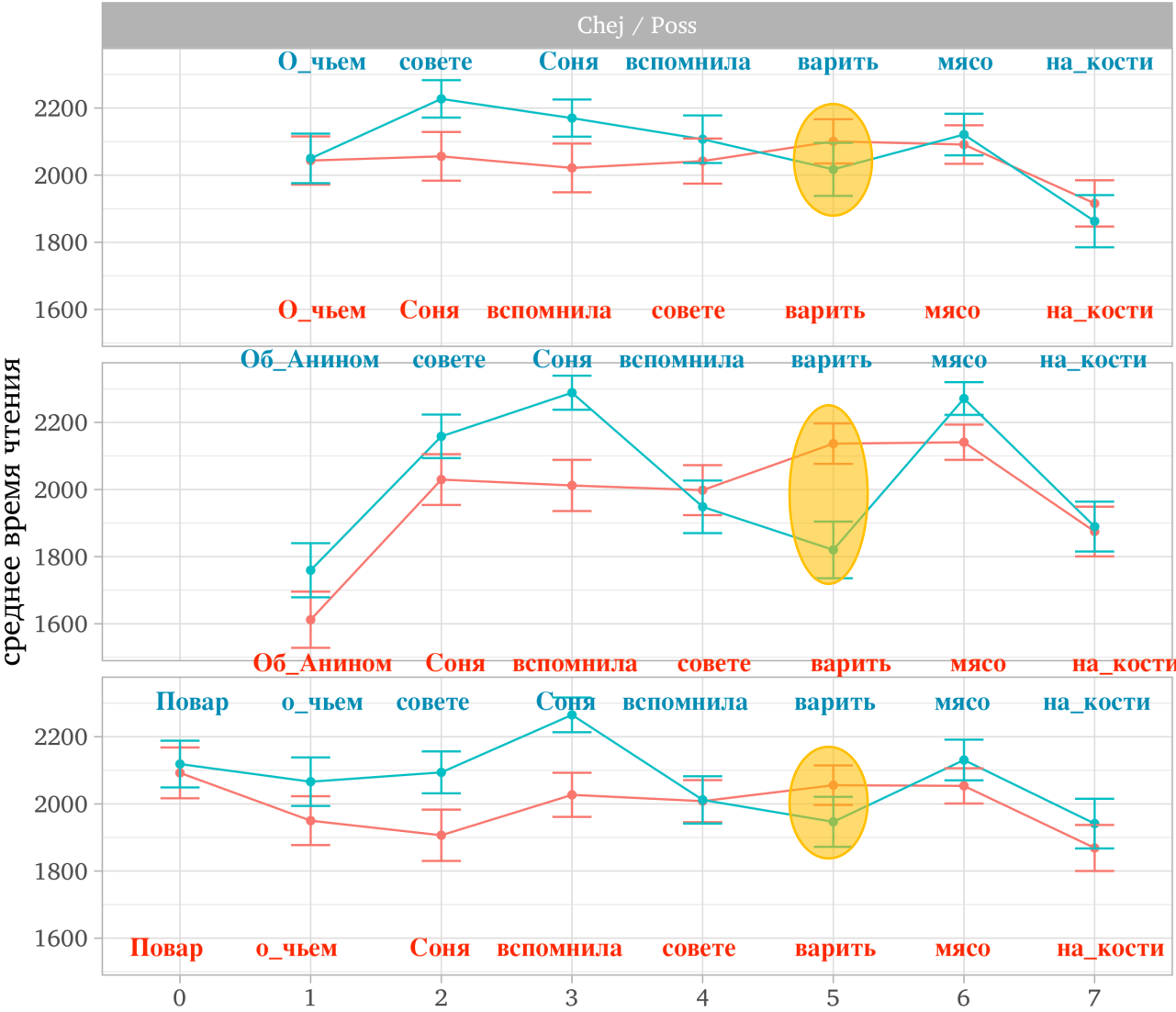
Результаты: время чтения, Inf



Результаты: время чтения, Inf



Результаты: время чтения, Inf



Время чтения компонента: при позднем сплите
значимо ниже, чем при раннем
 ($\beta = -230.20$, $SE = 62.48$, $p < 0.001$)

Split position — early — late

Результаты: время чтения

Элемент	Гипотеза	PP-комплемент	Inf-комплемент
Субъект	быстрее при позднем сплите	нет различий	быстрее при раннем сплите
Именная вершина	без различий	нет различий	нет различий
Комплемент	быстрее при раннем сплите	быстрее при раннем сплите	быстрее при раннем сплите

Обсуждение

Элемент	Гипотеза	PP-комплемент	Inf-комплемент
Субъект	быстрее при позднем сплите	нет различий	быстрее при раннем сплите
Именная вершина	без различий	нет различий	нет различий
Комплемент	быстрее при раннем сплите	быстрее при раннем сплите	быстрее при раннем сплите

- Различия между чтением субъекта в PP и Inf кажутся самыми неожиданными, потому что на этом моменте структуры стимулов никак не различаются;

Обсуждение

Элемент	Гипотеза	РР-комплемент	Inf-комплемент
Субъект	быстрее при позднем сплите	нет различий	быстрее при раннем сплите
Именная вершина	без различий	нет различий	нет различий
Комплемент	быстрее при раннем сплите	быстрее при раннем сплите	быстрее при раннем сплите

- Обработка комплемента при раннем сплите представляет из себя меньшую когнитивную нагрузку
- Тем не менее, прямой корреляции с оценками мы не видим: ранний сплит оценивается либо на одном уровне, либо ниже позднего
- Различия между ними объясняются не парсером, а грамматикой

Обсуждение

- Крайне низкие оценки скрэмлинга объясняются тем, что среди всех типов А'-передвижения он наиболее чувствителен к актуальному членению предложения;
- В нейтральном контексте у респондента нет оснований ожидать нестандартную информационную структуру => нет оснований для скрэмлинга вообще;
- В экспериментах И.Секериной не фигурировали оценки приемлемости, однако ее респонденты также отмечали «неестественность» или даже «невозможность» предъявляемых стимулов;

Обсуждение

- Различия между вопросительным и относительным передвижением также могут объясняться информационной структурой предложений;
- Вопросительное и относительное передвижение предпочитают различный коммуникативный статус частей расщепляемой составляющей

Обсуждение

- Различия между вопросительным и относительным передвижением также могут объясняться информационной структурой предложений;
- Вопросительное и относительное передвижение предпочитают различный коммуникативный статус частей расщепляемой составляющей
 - Вопросительное слово всегда фокусное (см. [Horvath 1985; Chomsky 1997]) и не накладывает обязательств на передвижение крупных составляющих

Обсуждение

- Различия между вопросительным и относительным передвижением также могут объясняться информационной структурой предложений;
- Вопросительное и относительное передвижение предпочитают различный коммуникативный статус частей расщепляемой составляющей

- Вопросительное слово всегда фокусное (см. [Horvath 1985; Chomsky 1997]) и не накладывает обязательств на передвижение крупных составляющих

Относительное местоимение, напротив, в своем предложении всегда топиальное и должно либо являться темой, либо входить в тему (см. [Лютикова 2019])

При релятивизации тематизации подвергается вся именная группа, поэтому чем бóльшая ее часть захватывается эффектом крысолова, тем «лучше» предложение.

Обсуждение

- Тот факт, что отщепление вершины от компонента оценивается не ниже отщепления левого элемента, не может объясняться remnant movement-семейством подходов, поскольку они предполагают передвижение NP;
- Вершины отдельно от компонентов не являются объектами A'-передвижений, потому что A'-передвижения оперируют составляющими;
- Таким образом, анализ с разрозненным удалением представляется более адекватным для описания результатов;