

«ТО ЕСТЬ, НЕ УБИЛИ, А ЗАРЕЗАЛИ САБЛЕЙ»: САМОИСПРАВЛЕНИЯ ГОВОРЯЩЕГО В УСТНЫХ РАССКАЗАХ

Подлеская В. И. (podlesskaya@ocrus.ru)

Российский государственный гуманитарный
университет, Москва, Россия

На материале электронного корпуса устных рассказов предпринят качественный и количественный анализ самоисправлений говорящего, затрагивающих лексику, морфологию, синтаксическую структуру, сегментную фонетику и просодию. Показано, что в наиболее частотных паттернах самоисправлений, фрагмент, подлежащий коррекции, и его откорректированный коррелят являются формально и функционально изоморфными. Именно этот изоморфизм позволяет преодолеть последствия речевого затруднения.

Ключевые слова: устный дискурс; самоисправление; русский язык; корпус

“THEY SHOT HIM DEAD, OH, NO, THEY KNIFED HIM DEAD WITH A SABER”: SELF-REPAIRS IN ORAL STORIES

Podlesskaya V. I. (podlesskaya@ocrus.ru)

Russian State University for the Humanities, Moscow, Russia

The paper introduces a discourse oriented classification of repair types in Russian by addressing, inter alia, the following questions: (i) whether or not self-repairing entails speech disfluency; (ii) whether or not the fragment under repair and its repaired correlate are structurally isomorphic; (iii) does the speaker revise a lexical, a morpho-syntactic, or a phonologic shape of the reparandum. Basing on the data from the Prosodically Annotated Corpus of Spoken Russian, established classes of repairs were analyzed qualitatively and quantitatively. Fluent isomorphic repairs appeared to be the most frequent in the corpus, although fluent non-isomorphic repairs, as well as disfluent isomorphic and disfluent non-isomorphic repairs are also attested.

Keywords: spoken discourse; self-repair; Russian; corpus

1. Постановка вопроса

Задача работы — выявить языковые средства, которые задействует говорящий в тех случаях, когда обнаруживается несоответствие порожденного речевого фрагмента исходному замыслу¹. Материалом послужили самоисправления, или (авто)коррекции, говорящего в четырех корпусах электронной коллекции «Рассказы о сновидениях и другие корпуса звучащей речи» (Prosodically Annotated Corpus of Spoken Russian, PrACS-Russ), содержащей аудиофайлы с синхронизированными просодически размеченными транскриптами (SpokenCorpora 2013):

- NDS «Рассказы о сновидениях» (129 монологов, респонденты от 7 до 17 лет);
- SLS «Рассказы сибиряков о жизни» (17 монологов, респонденты от 19 до 70 лет);
- FLS «Веселые истории из жизни» (40 монологов, респонденты от 18 до 60 лет);
- SPS «Истории о подарках и катании на лыжах» (20 рассказов по картинкам и 20 пересказов тех же сюжетов по памяти, респонденты от 20 до 30 лет).

Для качественного и общего сравнительного количественного анализа использовались 817 эпизодов самоисправления, полученных методом сплошной выборки из всех четырех корпусов коллекции. Для детального количественного анализа использовались 194 эпизода самоисправления, зарегистрированных в FLS и 79 эпизодов самоисправления, зарегистрированных в SPS.

Основные принципы разметки дискурсивной информации в экспериментальных корпусах этой коллекции были сформулированы в работе [Кибрик, Подлеская (ред.) 2009]. В этой же работе были сформулированы базовые подходы к систематизации речевых сбоев и к их аннотированию. Для последующего изложения существенны следующие сведения о типах дискурсивной информации и способах их аннотирования:

- В транскриптах произведено деление на элементарные дискурсивные единицы (ЭДЕ). ЭДЕ — минимальный фрагмент дискурса, который способен вступать в смысловые отношения с другими фрагментами в структуре дискурса. Просодически, ЭДЕ — это автономная единица, интонационно организованная вокруг одного рематического (фразового) акцента; с синтаксической точки зрения ЭДЕ в сильной степени коррелирует с клаузой.
- Транскрипты снабжены просодической разметкой — в частности, указаны локализация акцентов и тональный тип акцента (нотируется перед словом иконически с помощью косых черт), размечены абсолютные и заполненные паузы. Ударный слог в слове — носитель фразового акцента подчеркивается.
- Точка прерывания (речевого сбоя) маркируется знаком «= \Rightarrow » на границе ЭДЕ, и знаком «||» внутри ЭДЕ

¹ Работа выполнена при поддержке РГНФ, грант 12-04-00258

Дальнейшее изложение строится следующим образом. В разделе 2 вводятся два параметра классификации самоисправлений, которые позволяют описать, каким образом самоисправление встраивается в текущий дискурс; в разделе 3 один из этих параметров обсуждается более подробно; в разделе 4 рассматривается количественное распределение отдельных классов коррекций; в заключительном разделе 5 извлекаются уроки из полученных результатов.

2. Онлайн vs. оффлайн исправления; изоморфные vs. неизоморфные исправления

2.1. Трехчастная схема самоисправления

Прототипическими самоисправлениями, начиная с классических работ Levelt 1993, Shriberg 1994 и др., считаются самоисправления, строящиеся по трехчастной схеме «репарандум (фрагмент, подлежащий исправлению) / точка прерывания / репаранс (откорректированный коррелят)». В точке прерывания возможны заполненные паузы, лексические маркеры хезитации и прочие сигналы сбоя, ср.:

(1) *SLS*

... ээ ""всякь= || .. различные /-зме-еи,,,

(2) *FLS*

.. Поскольку-у \дверь ээ к тому моменту-у || "" \замок к тому моменту не /
отмёрз,

2.2. Два свойства канонического трехчастного самоисправления

Прототипические трехчастные самоисправления обладают следующими двумя важнейшими свойствами. Во-первых, они строятся с использованием онлайн стратегии, при которой проблема устраняется сразу, как только она обнаружена, ср. (Levelt 1993: 478), или в уточненной формулировке (Finlayson et al.2011: 79–80) — «как только говорящий завершил перепланирование и готов продолжить артикуляцию». Плавное развертывание дискурса прерывается, наличие прерывания показывает говорящему, что репарандум следует устранить из текста, чтобы получить правильный с точки зрения говорящего фрагмент. Онлайн стратегия сопряжена с речевым сбоем, т.е. с нарушением лексико-грамматической и/или просодической когерентности дискурса, в том числе, с обрывом текущего фрагмента. По нашим корпусным данным в 40% коррекций в точке прерывания имеется обрыв слова.

Во-вторых, для прототипических трехчастных самоисправлений характерен формальный и семантический изоморфизм фрагмента, подлежащего

исправлению, и его откорректированного коррелята. Ср. почти полную синонимию, тождество грамматического оформления и синтаксической функции репарандума *всяк[ие]* и репаранса *различные* в (1) или репарандума *дверь к тому моменту* и репаранса *замок к тому моменту* в (2).

2.3. Оба свойства канонического самоисправления не являются обязательными

Во-первых, не является единственно возможной онлайн стратегия. Говорящий может предпочесть оффлайн стратегию: исправление ошибки откладывается до завершения текущего речевого отрезка, и лишь после того, как этот отрезок транслирован слушающему, следуют эксплицитные признания ошибки и извинения по типу «О, простите, я высказался неверно/неточно». Оффлайн стратегия не предполагает нарушений плавного развертывания речи. Так, например, самоисправление в (3), вынесенное в заголовок статьи, и самоисправление в (4) полностью встроены в когерентный дискурс:

(3) *NDS*

*/Потом он один раз вышел на /балкон,
.. и /его /-пристрели-или...
.. То есть не /пристрелили,
а-а \за= =\резали.
\Саблей.*

(4) *SPS*

*.. /-Утро^h,,,
.. /мужик \-просыпается^w.
.. Даже не /мужик наверное,
/-парень,,,*

Поскольку самоисправление при оффлайн стратегии встроено в плавно развертываемый дискурс, точка прерывания выпадает из трехчастной схемы. Ввиду отсутствия самоперебива и таких внешних сигналов сбоя, как например, обрывы слов, оффлайн самоисправления плохо поддаются автоматической аннотации; поэтому они редко исследуются на корпусных массивах (среди редких исключений — Maguama, Sano 2006; Plug 2006).

Во-вторых, не является обязательным изоморфизм между фрагментом, подлежащим исправлению и его откорректированным коррелятом. Так, онлайн коррекции могут быть неизоморфными, если говорящий в принципе отказывается от забракованного фрагмента, изменяет исходный замысел и переходит к новому, когерентно построенному эпизоду. Фактически, в неизоморфных онлайн коррекциях репарандум есть, а репаранса — нет:

(5) *NDS*

… И-и /\мы-ы с Саньком .. тогда /пошли,
… в \магазин,
чего-нибудь \покупать,
.. а-а ==
…(0.5) и Са= ==
и я \проснул↑ся,
на /\самом там-м .. интересном –месте.

(6) *FLS*

У нас был .. чёрный потом огромный /потолок ==
\А,
ещё \больше того,
он же стал это \тушить.

2.4. Независимость двух параметров

Два параметра самоисправлений — онлайн vs. оффлайн и изоморфные vs. неизоморфные — являются независимыми, и в реальном дискурсе наблюдаются все четыре возможных комбинации их значений.

В каноническом трехчастном самоисправлении, примеры (1)–(2), представлена комбинация «+ онлайн, + изоморфизм». В примерах (5)–(6) представлены неизоморфные онлайн коррекции, т. е. комбинация «+ онлайн, — изоморфизм». Дальше всего от канонической трехчастной трехчастной схемы отстоят неизоморфные оффлайн коррекции («– онлайн, — изоморфизм») — в них самоисправление встроено в полностью когерентный дискурс, нет точки прерывания, откорректированный коррелят (а иногда и забракованный фрагмент) представлен диффузно в протяженном фрагменте дискурса. Так, в примере (7) первая ЭДЕ произносится с падающим тоном во фразовом акценте, что сигнализирует завершенность; исправление добавлено пост-фактум и сигнализируется дискурсивным маркером *в смысле (что)*:

(7) *FLS*

.. Но-о /он кстати был очень \приличн^{ый} молодой человек.
Ну в смысле что он был /молодой,
и-и .. /\ну-у .. как бы /внешне не \выглядел как фрик.
Единственное что был \голый.

Самоисправления этого рода стоят в одном ряду с такими дискурсивно когерентными построениями, как постпозитивные уточнения, дополнения и проч. Внутри этого класса самоисправления могут выделяться лишь по семантическому критерию, т. е. лишь в том случае, когда говорящий объявляет «недействительным» какой-то фрагмент транслированного значения. В примере (7) говорящий понял, что употребил слово *приличный* неточно: он хотел

сказать, что молодой человек, если не считать отсутствия одежды, ВЫГЛЯДЕЛ прилично, но не имел в виду квалифицировать его поведение как поведение приличного (т. е. соблюдающего нормы) человека. Часто грань между исправлением и уточнением провести довольно трудно. Так, в (8) и (9) фрагмент, следующий после дискурсивного маркера *то есть*, следует, скорее, квалифицировать как уточнение/дополнение, поскольку говорящий не отменяет ничего из сказанного:

(8) *FLS*

Ну мы пошли купили билет на тот же самый /самолёт,
… стоило это раза в три /дороже,
то есть за эту сумму можно было туда-обратно и ещё раз /туда слетать,

(9) *FLS*

И он должен был поехать в \Ташкент.
… {ЧМОКАНЬЕ} И там … торчать какое-то \время.
… {ЧМОКАНЬЕ} \Вот.
То есть у него был /доклад,
а потом он ещё должен был … какое-то время там \оставаться.

Наконец, последняя из четырех возможных комбинаций — оффлайн изоморфные коррекции («– онлайн, + изоморфизм») — представлена в примерах (10)–(11), где откорректированный коррелят появляется после когерентного и полностью завершённого фрагмента с падающим тоном во фразовом акценте. При этом морфологическое оформление коррелята полностью изоморфно забракованному фрагменту. Например, в (10) форма винительного падежа единственного падежа слова *коробку* идентична форме забракованного фрагмента *посылку* и лицензируется глаголом *возьму*. Точно также в (11) откорректированный коррелят *кошки* дублирует подлежащее *коты*, хотя и вынесено за пределы интонационно и грамматически цельнооформленной ЭДЕ:

(10) *FLS*

<НРЗБ> «В-вам как \посылку нужно оформить.»
Я говорю
«Какую /–посылку?,
где я вам возьму \посылку?!
Вот это || … –к-коробку?»

(11) *FLS*

ну у нас там у \многих жили кстати в общаге-е ээ \коты.
… \Кoшки.

Изоморфизм в оффлайн коррекциях может проявляться и опосредованно — например, через встраивание откорректированного коррелята в конструкцию, изоморфную конструкции с забракованным фрагментом. Так,

в (3)–(4) откорректированный коррелят вводится через изоморфную конструкцию с эксплицитным отрицанием; а в примере (12) происходит замена обстоятельного слота в структурно идентичных ЭДЕ:

(12) *NDS*

… Я вечером /\уснул,
… просыпаюсь \вечером.
… Днём /\уснул,
а /–вечером … \просыпаюсь.

Изоморфизм репарандума и репаранса — одно из наиболее эффективных средств, позволяющих говорящему поддержать структурную целостность текста, нарушенную сбоем, а слушающему — помочь восстановить исходный замысел говорящего. Этот параметр будет более подробно исследован в следующем разделе.

3. Типы изоморфизма между забракованным фрагментом и откорректированным коррелятом

3.1. Повторы и модификации

Внутри изоморфных самоисправлений выделяется два основных типа — повторы и модификации.

Максимально полно изоморфизм репарандума и репаранса проявляется при повторах. Повторы могут быть вызваны следующими функциональными причинами:

- говорящий пытается выиграть время для планирования (хезитация);
- говорящий не удовлетворен тем, что начал произносить, пытается подобрать более удачный вариант, не справляется, решает за неимением лучшего довершить начатое;
- говорящий понимает, что приступил к фрагменту слишком рано, добавляет недостающую информацию и возобновляет фрагмент

В (13) приведен пример онлайн повтора, а в (14) — пример частичного оффлайн повтора:

(13) *FLS*

… Начал ==
… Просёк эту /фишку,
… и начал ==
они ж коты \чего делают,
они … –писают,
… начал описывать им … –тапочки например какие-нибудь,,,

(14) *FLS*

- И один раз когда мы с сестрой /спали,
- (Кстати конечно я тоже этого не /помню,
- нам рассказала /\бабушка с мамой.
- \Да.)
- когда мы /спали,
- наша **бабушка** нав= || · намыла **столько** /–винограда,

Частичный изоморфизм репарандума и репаранса представлен в модификациях — репарандум и репаранс относятся к одному семантическому и\или грамматическому классу и претендуют на один и тот же слот в грамматической или, шире, дискурсивной структуре, но противопоставлены по некоторому параметру, ради которого и осуществляется коррекция. Наиболее распространенный тип модификации — лексическая. Лексическая онлайн модификация демонстрируется в (15); лексическая офлайн модификация, маркируемая контрастной просодией, представлена выше в (12), а также в (16) и (17):

(15) *NDS*

- Назвали его И^w= ” || · \–Ва-аня...

(16) *FLS*

- И вот папа ··· шестого числа ·· значит /доложил,
- {ЦОКАНЬЕ} ·· \пятого числа папа /доложил,
- и шестого пошёл бегать ·· ээ по всяким /авиакассам,

(17) *SPS*

- Сел на /–лыжи,,,
- {ЦОКАНЬЕ} ·· \встал на лыжи,
- /и \↑поехал!

Лексические модификации в общем случае удовлетворяют требованию Правильной Структуры (“Well-formedness Rule”, Levelt 1993: 485–489): забракованный фрагмент и его откорректированный коррелят входят в отношение близкое к альтернативному сочинению, и в частности, они правильным образом могут быть встроены в соответствующий альтернативный вопрос, ср. «Назвали Иван или Ваня?»; «Доложил пятого или шестого?»; «Сел на лыжи или встал на лыжи?». Однако даже для лексических модификаций подобного рода вопросы не всегда адекватны. Трудности возникают, например, при коррекции референциальных выражений, см. (18)–(19) («*кто-то или какой-то мужчина бежит?»; «*это какой-то или некий знак?»):

(18) *NDS*

- И я /смотрю,
- кто-то' ==
- какой-то \мужчина за нами бежит.

(19) *FLS*

… и тут Брентон …пробегаёт мимо \/меня-а,
… с таким .. \конусом.
.. Ну вот которые \ставят там-м ” для чего-то на /дорогу,
то есть это какой-то= || некий \знак,

Еще менее применимым оказывается требование Правильной структуры к модификациям, затрагивающим грамматику и фонологию, которым посвящен следующий раздел.

3.2. Изоморфные самоисправления, затрагивающие грамматику и фонологию

3.2.1. Морфологические модификации

Изоморфные модификации внутрисловно выражаемых категорий обычно требуют повторного воспроизведения слова:

(20) *SLS*

…на правом берегу реки Воронеж …… находилась || находилась маленькая станция \Отрожка.

(21) *SPS*

У его /жены … /скоро должно было случиться день \рождения.
.. –\Был случиться.

В редких случаях однако обнаруживаются морфологические модификации без повторного воспроизведения слова, что необычно для флективного фузионного языка (ср.обсуждение в Wouk 2005), ср. замену основы без необходимой замены алломорфа приставки:

(22) *NDS*

… /Потом^h …’ …… подо’= .. =/бежала ко мне \розоваяw .. мышка^w,

3.2.2. Модификации с использованием согласуемых препаративных подстановок

Особый класс модификаций в русском языке составляют самоисправления, с использованием прономинальных выражений, проецирующих грамматическую форму отложенной составляющей (английский термин placeholders, см. Podlesskaya 2010), ср. мужской род репарандума и репаранса в (23) и женский в (24):

(23) *NDS*

А /вы ещё нас /поведёте в \этот ||.. в тренажёрный \зал?

(24) *FLS*

… и-и … \Шороха мы держали .. на-а этой ||.. на /панели,,,

3.2.3. Модификации сегментного фонологического состава

В этот класс отнесены оговорки *per se* — модификации, при которых забракованный вариант представляет собой звуковую цепочку, не являющуюся фактическим словом языка:

(25) *FLS*

и они значит вмесело || весело маршируют \наверх.

3.2.4. Модификации просодии

В редких случаях говорящий может исправлять просодическую конфигурацию репаранса. Так, в (26) предпоследняя строка была первоначально замыслена как незаключительная и произнесена, соответственно, с подъемом тона, однако в результате автокоррекции строка повторена в качестве заключительной с падением тона во фразовом акценте:

(26) *NDS*

… а с= || … /двойняшки,
… /они' .. /\`огораживали,
… там где \посажено,
ставили /колышки ==
… ставили \колышки^h,

3.2.5. Конструкционные модификации

При самоисправлениях этого класса репарандум подвергается структурным изменениям, которые делают невозможным встраивание откорректированного коррелята в исходную структуру (изменения порядка слов, замена именной группы на глагольную и др.). Вместе с тем, «остатки» исходной структуры сохраняются, что позволяет считать эти коррекции изоморфными. Так в (27) представлено эксплицитное отрицание забракованного фрагмента, а в (28) — замена первоначально задуманного именного сказуемого (*больной?/хромой?/несчастный?*) на глагольную клаузу:

(27) *FLS*

.. обслуживали быто= || ” ну всякую ==
не /бытовую,
а-а .. \копировальную технику.

(28) SPS

и /он с перелом-м= =мом \руки-/ноги,

… теперь такой-й ==

” короче на /лыжах он теперь больше не \будет кататься.

4. Количественное распределение самоисправлений

Таблица 1 демонстрирует частотность самоисправлений в четырех корпусах коллекции «Корпуса звучащей речи» (PrACS-Russ). Зарегистрированная частотность самоисправлений — 2,3–3,4 в минуту, 1,8–2,9 на 100 слов — хорошо согласуется с данными по другим языкам. В целом, частотность самоисправлений обычно фиксируется в примерном интервале от одного до семи в минуту; в диалогах — выше, чем в монологах, в бытовой речи — выше, чем в официальной. Так, для китайских диалогов приводятся сведения о 5,4 случаях самоисправлений в минуту (Tseng 2006); для английских устных пересказов — 1,9–3,7 на 100 слов (Fraundorf and Watson 2008); для венгерских диалогов — 3,8 в минуту (Németh 2012); для японских монологов — 1,2 на 100 слов (Maruyama and Sano 2006). Таким образом, полученные нами данные позволяют предположить, что межъязыковое, межжанровое и межвозрастное варьирование частотности самоисправлений находится в пределах интервала, отражающего универсальные тенденции речепроизводства. При этом тексты корпуса SPS, отличающиеся от трех остальных корпусов по жанру — это рассказы по картинкам, а не «истории из жизни» — по-видимому, демонстрируют свойства меньшей спонтанности, так как в них частота самоисправлений ниже, чем в трех других корпусах:

Таблица 1. Общее число и частотность самоисправлений в исследованных корпусах

Корпус	Число слов	Время звучания (мин)	Общее число испр	Испр/мин	Испр/100 слов
SLS	5 000	40	132	3,3	2,6
NDS	14 000	120	412	3,4	2,9
FLS	7 000	70	194	2,8	2,8
SPS	4 500	35	79	2,3	1,8

По-видимому, именно меньшая спонтанность, приводит к тому, в корпусе SPS выше, чем в трех других корпусах, доля онлайн коррекций, хотя, в целом, по всем корпусам доля онлайн коррекций систематически значительно превосходит долю офлайн коррекций, см. Таблицу 2:

Таблица 2. Распределение онлайн и оффлайн самоисправлений в исследованных корпусах

Корпус	Общее число испр	Онлайн (# / %)	Оффлайн (# / %)
SLS	132	108 / 82%	24 / 18%
NDS	412	343 / 83%	69 / 17%
FLS	194	164 / 85%	30 / 15%
SPS	79	71 / 90%	8 / 10%

Более частные типы самоисправлений исследовались на материале двух корпусов «Веселые истории из жизни», FLS, и «Истории о подарках и катании на лыжах», SPS. Таблицы 3а и 3б показывают, что «золотым стандартом» являются онлайн изоморфные самоисправления. В корпусе FLS 146 онлайн изоморфных коррекций составляют 75% от общего числа (194) и 89% от числа онлайн коррекций (164); в корпусе SPS 64 онлайн изоморфные коррекции составляют 81% от общего числа коррекций (79) и 90% от числа онлайн коррекций (71):

Таблица 3а. Распределение основных типов самоисправлений в корпусе FLS

	Онлайн	Оффлайн	Всего
Изоморфные	146	16	162
Неизоморфные	18	14	32
Всего	164	30	194

Таблица 3б. Распределение основных типов самоисправлений в корпусе SPS

	Онлайн (дисфлу)	Оффлайн (флу)	Всего
Изоморфные	64	8	72
Неизоморфные	7	–	7
Всего	71	8	79

Таблицы 4а и 4б показывают, что среди изоморфных самоисправлений число модификаций незначительно превышает число повторов — в FLS соотношение 45,9% : 54,1%, в SPS соотношение 40,6% : 59,4%. Среди модификаций наиболее частотными являются лексические и конструкционные; самыми редкими — просодические, хотя частные классы изоморфных модификаций демонстрируют больший разброс между двумя корпусами, чем более общие типы самоисправлений. Можно предположить, что и здесь оказывает влияние жанр текстов, но это нуждается в более тщательной проверке.

Показательно распределение частот оговорок в исследованных корпусах. В корпусе FLS 7000 слов, следовательно, 15 оговорок в корпусе дают долю 1,7 на 1000 слов. В корпусе SPS 4500 слов, следовательно, 5 оговорок в корпусе дают долю 1,1 на 1000 слов. Снижение доли оговорок в рассказах по картинкам — еще одно проявление меньшей спонтанности таких текстов по сравнению в рассказами из жизни. Частотность оговорок оказалась сопоставима с данными, которые имеются в литературе по другим языкам (1–2 на 1000 слов), ср. Garnhem et al. (1981) по данным Лондонско-Лундского корпуса — 1 оговорка на 1000 слов в английских диалогах; Eklund (2004: 258ff.) по данным корпуса шведских диалогов (включая диалоги человека с человеком и человека с компьютером) — от 1,5 до 2,4 на 1000 слов. Это позволяет предположить, что частотность оговорки отражает действие общих механизмов речепорождения, и в меньшей степени связана с лингвоспецифическими и жанроспецифическими факторами. Однако — это предположение, как и все другие приведенные соображения о количественном распределении речевых сбоев, безусловно, нуждаются в проверке на более объемных выборках, сбалансированных по языкам, возрасту говорящих, жанрам и другим параметрам.

Таблица 4а. Распределение изоморфных самоисправлений в корпусе FLS

Тип коррекции		Онлайн	Оффлайн	Всего
Повторы		67	8	75
Модификации	лексические	31	6	39
	морфологические	4	–	3
	с использованием согласуемых препаративных подстановок	11	–	11
	сегментного фонологического состава (оговорки)	15	–	15
	просодические	2	–	2
	конструкционные	16	2	17
	Итого:	79	8	87
Всего		146	16	162

Таблица 46. Распределение изоморфных самоисправлений в корпусе SPS

Тип коррекции		Онлайн	Оффлайн	Всего
Повторы		26	1	27
Модификации	лексические	10	3	12
	морфологические	7	2	8
	с использованием согласуемых препаративных подстановок	–	–	–
	сегментного фонологического состава (оговорки)	5	–	7
	просодические	–	–	–2
	конструкционные	16	2	18
	Итого:	38	7	45
Всего		64	8	72

5. Выводы

- Были рассмотрены самоисправления говорящего в электронной коллекции «Рассказы о сновидениях и другие корпуса звучащей речи» — первом открытом ресурсе с систематически размеченной фразовой сегментацией, просодией, речевыми сбоями, паузацией и другими явлениями спонтанной устной речи.
- Была предложена классификация самоисправлений, позволяющая оценить способ встраивания самоисправления в текущий дискурс.
- Параметры предложенной классификации ориентированы на следующие исследовательские вопросы: (а) связано ли самоисправление с нарушением структурной целостности текущего дискурса; (б) являются ли забракованный фрагмент и его откорректированный коррелят структурно изоморфными; (в) затрагивает ли самоисправление лексику, морфологию, синтаксис, сегментную или супraseгментную фонологию забракованного фрагмента.
- Было показано, что говорящие отдают существенное предпочтение онлайн изоморфным самоисправлениям по сравнению со всеми остальными классами, зарегистрированными в исследованных корпусах. Можно предположить, что данный способ для говорящего оказывается наименее трудозатратным, а слушающему позволяет наиболее точно реконструировать исходный замысел говорящего. Однако это предположение нуждается в проверке с привлечением не только корпусных, но и экспериментальных методов.

Литература

1. Кибрик А. А., Подлеская В. И. (Ред.). Рассказы о сновидениях: Корпусное исследование устного русского дискурса. М.: ЯСК, 2009.

References

1. *Kibrik, A. A. and V. I. Podlesskaya* (eds.): 2009, *Rasskazy o snovidenijax: korpusnoe issledovanie usntogo russkogo diskursa* [Night Dream Stories: A corpus study of spoken Russian discourse]. Moskva: Jazyki Slavjanskix Kul'tur.
2. *Eklund, Robert*: 2004, *Disfluency in Swedish human-human and human-machine travel booking dialogues*. PhD thesis, Linköping Studies in Science and Technology, Dissertation No. 882, Department of Computer and Information Science, Linköping University, Sweden.
3. *Finlayson et al.*: 2011, 'Fluency or accuracy: What matters when correcting errors in spoken dialogue?', in Ian Finlayson, Robin Lickley, and Martin Corley (eds.), *Architectures and mechanisms of language Processing (AMLaP 2011)*, Centre Universitaire Paris Descartes, 79–80.
4. *Fraundorf, S. H. and D. G. Watson*: 2008, 'Dimensions of variation in disfluency production in discourse', in J. Ginzburg, P. Healey, and Y. Sato (eds.), *Proceedings of LONDIAL 2008, the 12th Workshop on the Semantics and Pragmatics of Dialogue*, London: King's College London, 131–138.
5. *Garnham, Alan, Richard C. Shillcock, Gordon D. A. Brown, Andrew I. D. Mill and Anne Cutler*: 1981. 'Slips of the tongue in the london-lund corpus of spontaneous conversation', *Linguistics* 19(7–8), 805–818.
6. *Levelt, Willem J. M.*: 1983, 'Monitoring and Self-Repair in Speech', *Cognition* 14, 41–104.
7. *Levelt, Willem J. M.*: 1993, *Speaking: From Intention to Articulation* [ACL-MIT Series in Natural Language Processing].
8. *Maruyama, Takehiko and Sano, Shin'ichiro*: 2006, 'Classification and Annotation of Self-Repairs in Japanese Spontaneous Monologues', in *LPSS — Linguistic Patterns in Spontaneous Speech*, Taipei, November 2006, 283–298.
9. *Németh, Zsuzsanna*: 2012, 'Recycling and replacement self-repairs in spontaneous Hungarian conversations', in *Proceedings of the First Central European Conference in Linguistics for postgraduate Students*, 211–224.
10. *Plug, L.*: 2006, 'Speed and reduction in postpositioned self-initiated self-repair', *York Papers in Linguistics* 2(6), 143–162.
11. *Podlesskaya, Vera*: 2010, 'Parameters for typological variation of placeholders', in N. Amiridze, Boid H. Davis and Margaret Maclagan (eds), *Fillers, Pauses and Placeholders*. [Typological Studies in language (TSL), vol. 93]. Amsterdam/Philadelphia: John Benjamins, 11–32.
12. *Shriberg, Elizabeth Ellen*: 1994, *Preliminaries to a Theory of Speech Disfluencies*. PhD thesis, University of California, Berkeley.
13. *Spokencorpora*: 2013, «Рассказы о сновидениях и другие корпуса звучащей речи». Prosodically Annotated Corpus of Spoken Russian. Pilot version. Online: <http://spokencorpora.ru>
14. *Tseng, S.-C.*: 2006, 'Repairs in Mandarin conversation', *Journal of Chinese Linguistics* 34(1), 80–120.
15. *Wouk, Fay*: 2005, 'The syntax of repair in Indonesian', *Discourse Studies* 7 (2), 237–258.