

# Перцептивно значимые элементы редуцированных словоформ



*Риехакайнен Е. И.*

В статье описываются результаты эксперимента по восприятию изолированных редуцированных словоформ, извлечённых из записей русской спонтанной речи. Предметом исследования является выявление тех элементов фонологического уровня, на которые опирается слушающий при распознавании редуцированных словоформ. На материале русского языка согласные оказываются более перцептивно значимыми элементами, чем гласные. Наибольшей перцептивной значимостью обладают: порядок следования согласных, начальный одиночный смычный согласный и ряд устойчивых для восприятия согласных в любой позиции.

The results of the experiment in which subjects were listening to isolated reduced word forms, extracted from spontaneous dialogues in Russian, show that consonants are more reliable sources of information for a listener than vowels in the process of Russian speech recognition. The most important elements are consonant order, single initial stop consonant and some perceptually strong consonants in any position. These results are proved by the recognition of the same word forms in context.

## Теоретическая база исследования

В процессе восприятия речи, т.е. при необходимости «приписывания языковой структуры речевому сигналу» [1: 53], слушающий может пользоваться различными стратегиями. Исследования на материале изолированных фонем, слогов и слов часто предполагают пофонемное распознавание<sup>1</sup> звучащего сигнала (см., например: [2], [3]). В моделях пофонемного распознавания осуществляется последовательный анализ всех фонем словоформы.

Однако в процессе распознавания естественной звучащей речи использование только фонологической информации не является преобладающей стратегией. В условиях ограниченности во времени и фонетического несовершенства и/или неполноты сигнала<sup>2</sup> слушающий предпочитает стратегии, сочетающие «грубый, т.е. неполный фонологический

<sup>1</sup> В данной работе термины «восприятие» и «распознавание» считаются синонимами.

<sup>2</sup> «Чем большей спонтанностью обладает текст, тем больше в нём участков неполного типа» [4: 6].

анализ с одновременным лексико-семантическим и грамматическим» [5: 185].<sup>3</sup> В большинстве современных моделей восприятия речи осуществление неполного фонологического анализа описывается через понятие активации. Предполагается, что некоторые элементы словоформы являются более перцептивно значимыми, чем другие. В процессе распознавания словоформы в словаре слушающего активируется всё множество единиц, которое соответствует последовательности её перцептивно значимых элементов. Впоследствии из этого множества активированных словоформ выбирается единственно возможная с учётом частотности, лексико-семантических и грамматических параметров.

При такой трактовке выделение перцептивно значимых элементов словоформы является необходимым этапом описания процесса восприятия речи.

Решение данной проблемы предполагает рассмотрение, по крайней мере, двух вопросов:

- какие части слова (начало, конец, ударный слог, консонантный «скелет» и т.п.) наиболее надёжно воспринимаются слушающими;
- существуют ли фонемы, которые распознаются лучше, чем остальные, и какие дифференциальные признаки фонем являются наиболее перцептивно устойчивыми.<sup>4</sup>

Первый вопрос получил достаточно подробное освещение в литературе по восприятию речи и решается по-разному в различных моделях. Так, в ранней версии модели когорты [8] первоначальная активация лексических единиц осуществляется исключительно по началам слов. При этом не существует единого мнения относительно того, каким должен быть размер начальной подцепочки: называются варианты от 1–2 до 3–5 фонем или первого слога.<sup>5</sup> Если идентификация словоформы по начальным элементам оказывается невозможной, то необходимо продолжать анализ и привлекать следующие элементы, пока цепочка не станет уникальной.

Конец слова, напротив, считается ненадёжной для восприятия частью: в процессе речевосприятия слушающий не распознаёт конечный согласный слова гораздо чаще, чем начальный. Однако ослышки, вызванные неверной интерпретацией одного из звуков слова, связаны, как правило, именно с заменой начального, а не конечного согласного [9].

Тот факт, что ударный гласный практически никогда не редуцируется и произносится в слове наиболее отчётливо,<sup>6</sup> позволяет ряду исследователей считать ударный гласный (или ударный слог) наиболее перцептивно значимым элементом [11]. Значимость ударного слога в целом при этом объясняется его большей продолжительностью и большей интенсивностью по сравнению с другими слогами в слове. Противники данной гипотезы

<sup>3</sup> Обращение к пофонемному распознаванию оправдано только при необходимости восприятия новых, незнакомых, слов.

<sup>4</sup> Термин «перцептивная устойчивость» заимствован из [6]. Встречающийся в более ранних работах (см., например, [7]) термин «перцептивная помехоустойчивость» решено было не использовать, поскольку в данной статье речь идёт о распознавании словоформ в условиях отсутствия каких-либо значимых помех (шума, акцента и т.п.).

<sup>5</sup> См. ссылки на соответствующие работы в [1: 40].

<sup>6</sup> См. обзор типичных случаев редукции гласных в разговорной речи в [10: 43–48].

указывают на то, что выводы о перцептивной значимости ударного гласного и начала слова делаются преимущественно на материале английского языка, в котором часто именно первый слог является ударным. При этом невозможно определить, какой именно признак — ударность гласного или его положение в начале слова — оказывается решающим [12: 4—47]. О большей значимости именно начального положения слога свидетельствуют результаты экспериментов, описанных в [13]: выполнение заданий на восприятие со словами, в которых ударным являлся не первый слог, занимало у испытуемых больше времени и требовало большего объёма памяти, чем работа со словами, в которых ударным был первый слог.

Проблема значимости ударного гласного связана с определением значимости гласных в целом для надёжности распознавания словоформы. На материале английского и немецкого языков было показано, что при ослышках ошибки в восприятии согласных происходят намного чаще, чем ошибки в восприятии гласных, а ударные гласные вообще практически не подвергаются изменениям. На основе данного наблюдения делается вывод о том, что по крайней мере в рассмотренных языках гласные в слове (особенно ударные) воспринимаются лучше и являются более надёжным источником информации, чем согласные [14].

Авторы монографии «Русская разговорная речь» отмечают, что в русской спонтанной речи слово узнаётся главным образом по согласным, гласные же служат преимущественно для передачи ритмики слова [10: 40-41]. Следовательно, в данной работе согласные признаются более значимыми элементами для восприятия речи, чем гласные. А исследования распознавания слов русского языка при наличии разного рода помех (шум, тугоухость, восприятие синтезированных звуков в различных условиях и т.п.) показали, что наиболее значимым фонетическим фактором для восприятия слов оказывается «ударная гласная», в большинстве экспериментальных условий значим также фактор «длина слова в слогах» [15].

В [4] описан эксперимент, в котором испытуемым предлагались для восприятия наиболее частотные редуцированные словоформы, извлечённые из записей русской речи. Была проанализирована роль 5 групп признаков: ударный гласный, ритмическая структура, начальная фонема словоформы, конечная фонема словоформы, количество согласных в словоформе. Результаты исследования показали, что количество согласных является самым существенным для опознания словоформы вне контекста: «чем большее число согласных выпадает из словоформы при её реализации, тем хуже она распознаётся» [4: 129]. При опущении одного из согласных количество правильно опознанных словоформ снижается на 37%. Кроме того, на надёжность распознавания влияют также сохранность ударного гласного и количества слогов: при искажении одного из этих двух признаков количество правильно опознанных словоформ снижается на 32%.

Современные модели распознавания звучащей речи, принимая во внимание вышеописанные наблюдения, в той или иной степени основываются на понятии повсеместной активации (*radical activation*). Предполагается, что входной речевой сигнал, соответствующий слову *dog*, может активировать и слово *bog* на основании совпадающих гласного и конечного согласного, несмотря на то что начальные согласные этих слов различны, при этом слову *dog* будет отдано предпочтение в процессе распознавания [16: 595]. В.Б. Касевич называет все те признаки словоформы, на которые слушающий может опираться в процессе восприятия, «просодиями» слова и относит к ним, в частности, ритмическую структуру слова (число слогов и распределение ударений и тонов), интонационный контур, тип распределения в слове согласных, все виды гармонии и сингармонизма [5: 201—202, 353—354].

Рассмотрение воспринимаемых фонем с точки зрения их дифференциальных признаков связано в первую очередь с созданием признаковых моделей восприятия речи (см. обзор в [1: 15—29]). Подобные модели не нашли широкого применения, однако необходимость анализа восприятия дифференциальных признаков до сих пор признаётся многими исследователями. Так, достаточно широко распространены исследования, направленные на выявление наиболее устойчивых дифференциальных признаков. Как правило, в подобных исследованиях рассматриваются прежде всего дифференциальные признаки согласных.

Анализ ослышек в английском языке показывает, что самым перцептивно устойчивым признаком является признак «шумность vs. сонорность»: при неверной интерпретации согласного в слове шумный в большинстве случаев заменяется на другой шумный, а сонант — на другой сонант, тогда как изменения по остальным признакам происходят намного чаще (например, смычный шумный переднеязычный /t/ в слове *great* может заменяться на /p/, /k/, /d/, /f/, /ç/) [9].

Существуют работы, посвящённые анализу восприятия дифференциальных признаков согласных фонем, являющихся начальными компонентами открытых слогов. Ряд исследований на материале языков, отличных от русского, свидетельствует о том, что «в составе стимулов типа CV наиболее устойчивыми, вероятно, являются признаки участия голоса и назальности, в наибольшей степени подверженным изменению оказывается признак места образования» [6: 29]. Эксперименты же по восприятию открытых слогов, являющихся элементами связного текста, на русском материале показали, что наиболее устойчивыми являются признаки глухости-звонкости, шумности-сонорности и мягкости-твёрдости, тогда как признаки способа образования и активного артикулирующего органа являются менее устойчивыми для восприятия. Кроме того, в результате этих экспериментов было обнаружено, что самые большие сложности для опознания представляют согласные в составе сложных консонантных кластеров [6: 32—155].

При наличии разного рода помех существенными факторами оказываются звонкость-глухость (и в слогах, и в словах), шумность-сонорность (в основном, для слогов), место образования (для слогов), способ образования (для слов и слогов в ряде экспериментальных условий). Твёрдость-мягкость оказалась существенной лишь в небольшом числе экспериментов [15].

В настоящей статье рассматриваются результаты экспериментального исследования восприятия редуцированных словоформ, извлечённых из записей спонтанной русской речи. Одной из задач являлось выделение перцептивно значимых элементов словоформ.

### Материал эксперимента

Под редуцированными словоформами в данной статье понимаются формы слов, которые на сегментном уровне представлены меньшим количеством элементов, чем в полном варианте, предусмотренном нормами кодифицированного литературного языка. Подобные словоформы являются неотъемлемой

составляющей спонтанной речи. В 9 спонтанных бытовых диалогах на русском языке<sup>7</sup> общей продолжительностью звучания 16 минут 33 секунды, которые были выбраны в качестве исходного материала для проведения экспериментов, встретилось 459 редуцированных словоформ (что составляет примерно 18% от общего числа словоформ экспериментальных текстов).<sup>8</sup> Многочисленность редуцированных словоформ делает их распознавание важным элементом восприятия спонтанной речи в целом, а их использование в экспериментах по восприятию речи — перспективным для описания данного процесса и, в частности, для выявления перцептивно значимых элементов словоформ.

Поскольку основной задачей описываемого эксперимента было сопоставление надёжности распознавания редуцированных словоформ при изолированном и контекстном предъявлении, тестовая последовательность состояла из двух частей: в первой части испытуемым предъявлялись расположенные в произвольном порядке изолированные словоформы, во второй — они же в контексте.

Для проведения эксперимента были отобраны 24 редуцированные словоформы: 22 словоформы в женском произнесении (4 диктора) и 2 — в мужском (1 диктор).<sup>9</sup>

Контекст, в котором предъявлялись словоформы во второй части тестовой последовательности, мог быть как правым, так и левым — в зависимости от того, что позволял исходный текст диалогов.

Каждый стимул повторялся три раза: в первой части с интервалом в 0,5 секунды, а во второй — в 1 секунду. Межстимульный интервал составил для первой и второй частей соответственно 3,5 и 10 секунд. Порядок предъявления словоформ в первой и второй частях тестовой последовательности не совпадал. Всего в эксперименте было представлено 48 стимулов (по 24 в каждой части), общая продолжительность эксперимента составила 9 минут 20 секунд.

В таблице 1 представлен список всех словоформ-стимулов, их транскрипция, составленная на основании инструментального анализа сигналов, а также контекст, в котором они предъявлялись во второй части эксперимента.

Таблица 1

### Материал эксперимента

Стимул	Данные инструментального анализа*	Контекст
Ходит	kóit	Она <b>ходит</b> к психологу
Наверно(е)	nám <sup>i</sup>	<b>Наверно</b> , тоже чтобы как-то

<sup>7</sup> Участниками диалогов были 7 женщин и 1 мужчина в возрасте от 23 до 26 лет, для которых русский язык являлся родным. Параметры записи: 44100 Гц; 16 бит. Записи диалогов были предоставлены нашими французскими коллегами, которые проводили исследования русской и французской спонтанной речи.

<sup>8</sup> Подробное описание редуцированных словоформ, встретившихся в данных спонтанных диалогах, представлено в: [17: 6—16].

<sup>9</sup> Поскольку в качестве материала эксперимента выступал континуум спонтанной речи, при составлении тестовой последовательности из всех словоформ, представленных в исходных диалогах, были выбраны только те, при выделении которых из текста не возникало технических трудностей: не было искажений начального и конечного звуков слова под воздействием звуков соседних слов (например, озвончения конечного согласного под влиянием начального звонкого следующего слова), смеха, посторонних шумов.



кажется	kaʒs	мне <b>кажется</b> , таких людей
открыть	atkrʹ	не может <b>открыть</b> фотографии
представляешь себе	ptěšyt	и он об этом не знает, <b>представляешь себе</b>
нравятся	nrec	мне очень <b>нравятся</b> ситуации
какую-то	kógyt	продала <b>какую-то</b> картину
очень	eč	не <b>очень</b> хорошее
последний	pasýnʹ	сказала в <b>последний</b> раз
тебе	tʹe	она <b>тебе</b> рассказывает
следующий	ʹsl'eš':	на <b>следующий</b> день
километров	kájna	двести <b>километров</b> может ехать
целый	cée	<b>целый</b> лист белой бумаги
по-моему	rom	он нам, <b>по-моему</b> , принёс
слушай / слышь	slyš	<b>слушай</b> , ты про него что-нибудь слышала?
понимаешь	рыывáš	<b>понимаешь</b> , она на каждом шагу останавливается
а она говорит	aongýt	<b>а она говорит</b> : «Да, но это не в первый раз!»
фильмов	ʹf'ımyf	из серии <b>фильмов</b> ужасов
достаточно	tstáčna	<b>достаточно</b> светло
потому что	ptyšt	<b>потому что</b> налоговая инспекция может
буквально	bʹál'n	<b>буквально</b> за час
появился	pʹlc	<b>появился</b> какой-то дополнительный коридор
в пятницу	p'an'c	нет, это было не <b>в пятницу</b> , это было в среду
просят	p'os't	<b>просят</b> прислать фотографии

\* Во втором столбце данные представлены с помощью фонематической транскрипции, созданной на основе фонетической транскрипции, предложенной Л.В. Щербой (см. [18: 272—273], [19: 38]); звуки, записанные как верхний индекс, обозначают короткие звуки или их призвуки (/ʰ/).

Анализ транскрипции показывает, что в рассматриваемых словоформах наряду с выпадением звуков встречается и другой тип редукции: некоторые звуки (преимущественно гласные) в составе словоформ подвергаются количественным и качественным изменениям.

Далее в статье будут подробно рассмотрены результаты восприятия изолированных редуцированных словоформ, т.е. стимулов, вошедших в первую часть тестовой последовательности. В конце статьи будут приведены

некоторые результаты распознавания исследуемых словоформ в контексте (во второй части тестовой последовательности и в другом эксперименте), позволяющие судить о перцептивной значимости элементов редуцированных словоформ.

## Методика эксперимента и испытуемые

Тестовая последовательность, записанная на магнитную ленту, предъявлялась испытуемым через громкоговорители.

Участникам эксперимента предлагалось прослушать предъявляемые стимулы и записать то, что они услышали, буквами русского алфавита. Разрешалось пропускать сигналы (ставить прочерк в анкете), если испытуемый не мог определить, что он слышит, или не успевал идентифицировать стимул.

В инструкции было указано только то, что испытуемым будут предъявлены фрагменты русской речи, поэтому испытуемые могли интерпретировать услышанное как любую последовательность звуков русской речи (необязательно осмысленную) и записывать в ответах не только слова и словосочетания, но и отдельные слоги и даже звуки.

Испытуемыми были студенты и преподаватели филологического и восточного факультетов Санкт-Петербургского государственного университета.

В эксперименте приняли участие 40 человек.

## Результаты эксперимента

### Восприятие изолированных словоформ

В ходе обработки полученных данных были подсчитаны процентные соотношения по следующим параметрам: количество распознанных словоформ, количество отказов, число осмысленных буквосочетаний в ответах испытуемых. Считалось, что словоформа распознана верно, если испытуемый записывал в ответе то слово, редуцированный вариант которого был представлен в эксперименте (например, *тебе* для /t'e/ или *очень* для /eč/). Кроме того, поскольку в эксперименте проверялась надёжность восприятия редуцированных словоформ с искажениями в основе, то при обработке полученных данных считалось, что словоформа распознана верно, даже если испытуемый записывал в бланк ответа флексию, отличную от исходной. Таким образом, в данном эксперименте правильными ответами считались: *нравится* (для стимула *нравятся*), *ходят* (для *ходит*), *просит* (для *просят*), *пятница* (для *(в) пятницу*). Для стимула *слушай* /slyš/ верным признавался и ответ *слышь*, поскольку редуцированные варианты этих вводных слов очень близки между собой и невозможно определить, какое же слово было произнесено на самом деле в диалоге, на материале которого составлялась тестовая последовательность.

При анализе перцептивной значимости элементов словоформ подсчитывался процент верных отражений звуков на соответствующих позициях в ответах испытуемых. В данном случае верными считались те ответы, в которых было записано то, что соответствовало транскрипции, представленной в таблице 1, а не то, что должно было звучать в редуцированном варианте соответствующей словоформы (например, *к*, а не *х* в начале

словоформы *ходит /kóit/* или *a*, а не *e* на месте ударного гласного в *наверно /nám/*). Для сопоставления результатов вычислялись доверительные интервалы на 5%-ном уровне значимости. Статистически были выделены 2 параметра: надёжность распознавания того или иного элемента словоформы и наличие преобладающего варианта при отражении этого элемента испытуемыми. Распознавание считалось надёжным, если испытуемые воспринимали реально звучащую фонему достоверно чаще, чем все остальные фонемы: проводилось сравнение доверительных интервалов для верного варианта и суммы всех неверных. Вариант считался преобладающим, если процент испытуемых, зафиксировавших его в своих ответах, был достоверно выше любого другого варианта для данного стимула: доверительный интервал для наиболее частотного варианта сопоставлялся с каждым из менее частотных вариантов для данного стимула.

Прежде всего, необходимо отметить, что процент распознавания редуцированных словоформ при изолированном предъявлении оказался очень низким. Так, 58,3% всех предъявляемых словоформ не были правильно опознаны ни одним испытуемым. В таблице 2 представлены результаты распознавания словоформ, которые были опознаны хотя бы одним испытуемым.

Таблица 2

#### Надёжность распознавания изолированных редуцированных словоформ

Стимул	Количество верных распознаваний с учётом доверительных интервалов, %
достаточно	4,1<12,5<26,8
просят	2,7<10,0<23,7
слушай	1,5<7,5<20,4
понимаешь	1,5<7,5<20,4
в пятницу	1,5<7,5<20,4
нравятся	0,5<5,0<17,0
очень	0,2<2,5<13,2
тебе	0,2<2,5<13,2
а она говорит	0,2<2,5<13,2
буквально	0,2<2,5<13,2

Только для двух словоформ распознавание превышает или равно 10%. Четыре словоформы были правильно опознаны только 1 испытуемым из 40. Таким образом, при анализе восприятия редуцированных словоформ, предъявляемых изолированно, нельзя говорить о надёжном распознавании ни одной из них.

Стимулы, вошедшие в тестовую последовательность, не были специально отобраны для проведения эксперимента по изучению перцептивной значимости элементов словоформ: по таблице 1 видно, что фонетический облик исследуемых словоформ достаточно разнообразен. Тем не менее, полученный в ходе эксперимента материал позволяет сделать ряд предвари-



тельных выводов о перцептивной значимости элементов редуцированных словоформ и устойчивости некоторых признаков фонем.

Опираясь на описанные выше гипотезы о перцептивной значимости различных элементов словоформы, рассмотрим, насколько надёжно были восприняты испытуемыми ритмическая структура и консонантный «скелет» (в частности, начальный согласный). В инструкции, предъявлявшейся испытуемым, не требовалось расставлять ударения, поэтому при оценке перцептивной значимости ритмической структуры будет проанализирована только надёжность распознавания количества слогов и того гласного, который является ударным; позиция же ударения не может быть оценена по данным проведённого эксперимента.

Инструментальный и слуховой анализ показали, что количество слогов практически во всех исследуемых редуцированных словоформах меньше, чем в их нередуцированных вариантах: реально звучат как трёхсложные только 3 стимула (*последний* /pasʲiɲi/, *понимаешь* /ryɪvʲaʃ/, *а она говорит* /aongʲit/), как двусложные — 9 словоформ, большинство же словоформ (12) являются односложными. Причём в двусложных и трёхсложных словоформах безударные гласные часто являются очень короткими (иногда это характерно и для ударных гласных).<sup>10</sup>

Только для 13 стимулов (54,2%) можно выделить преобладающий вариант при отражении количества слогов в услышанной словоформе. Интересно отметить, что в стимуле *открыть* /atkrʲ/ большинство испытуемых (10 из 14 давших ответ) вообще не услышали гласного звука. Результаты для 13 стимулов представлены в таблице 3.

Таблица 3

## Отражение количества слогов в ответах испытуемых

Стимул	Отражение кол-ва слогов в ответах испытуемых, чел.					Преобладающий вариант, %
	0 слогов	1 слог	2 слога	3 слога	4 слога	
ходит /kóit/	0	34	2	0	0	87,2
наверно /námi/	0	27	6	1	0	79,4
кажется /kaʒs/	4	33	0	0	0	89,2
нравятся /nrec/	0	34	2	2	0	89,5
какую-то /kógyt/	0	23	10	0	0	63,8
следующий /sl'eʃ':/		30	3			83,3
цельй /cée/	5	25	4	0	0	73,5
по-моему /pom/	1	35	2	0	0	92,1
слушай /slyʃ/	6	29	2	1	0	76,3
а она говорит /aongʲit/	0	4	21	8	1	61,8
фильмов /f'i'myɪf/	0	31	4	0	0	88,6
достаточно /tstáčna/	0	3	24	4	6	63,2
появился /p'ɪc/	2	32	1	0	0	91,4

<sup>10</sup> Столь значительная редукция ритмической структуры исследуемых словоформ, по-видимому, и является причиной их плохой распознаваемости при изолированном предъявлении (ср.: [4: 131—132]).

Серым цветом в таблице выделено число испытуемых для каждого из стимулов, которые верно отразили количество слогов; жирным шрифтом выделен тот вариант, который преобладал в ответах испытуемых. В последнем столбце отмечен процент, который составляет преобладающий вариант от общего числа полученных ответов.

В 7 из 13 стимулов большинство испытуемых услышали столько слогов, сколько в них звучало на самом деле, в шести — меньше, чем в предъявлявшемся стимуле. Для 6 из 7 стимулов, количество слогов, в которых было верно отражено испытуемыми, распознавание можно считать статистически надёжным. Все эти стимулы представляют собой односложные словоформы (*/kažs/*, */nrec/*, */s' eš' /*, */pom/*, */slyš/*, */p'lc/*). Таким образом, только в 25% из всех предъявлявшихся в эксперименте стимулов количество слогов было отражено испытуемыми верно. Ни для одной из двух- и трёхсложных словоформ распознавание количества слогов не является надёжным.

Анализируя ударные гласные в предъявлявшихся стимулах, можно заметить, что редукция затрагивает не только безударные гласные, но и ударный. Данные инструментального и слухового анализа свидетельствуют о том, что ударный гласный сохраняется только в 15 словах тестовой последовательности (см. табл. 1). Следовательно, в ходе оценки перцептивной значимости ударного гласного будет рассмотрена надёжность распознавания того гласного, который находился в предъявляемом стимуле на позиции ударного, а не того, который должен быть на данной позиции при полном произнесении.

Гласный в ударной позиции был надёжно распознан только в 3 стимулах: *ходит* */kóit/*, *просят* */p'os't/* и *по-моему* */pom/*<sup>11</sup> (79,5%, 84,2% и 84,2% правильных опознаваний ударного гласного соответственно). Ещё для 5 стимулов можно выделить преобладающие варианты ответов испытуемых. Для словоформы *понимаешь* */ryiváš/* преобладающий вариант (69,7% ответов) совпадает с той фонемой, которая звучала в предъявляемом стимуле (*/a/*). Для стимулов *слушай* */slyš/*, *появился* */p'lc/* и *фильмов* */f'ityf/* в ответах испытуемых преобладает фонема */u/* на месте ударного гласного (55,2%, 60% и 68,6% ответов соответственно). В стимуле *а она говорит* */aongýt/* преобладающим вариантом интерпретации ударного гласного является фонема */e/* (38,2%).

Можно предположить, что перцептивно важным должен быть единственный гласный в односложном стимуле, поскольку гласный традиционно считается вершиной слога. Однако, кроме уже описанных выше результатов по стимулам *просит* и *по-моему*, вряд ли можно говорить об однозначном распознавании гласных в односложных стимулах.

Таким образом, ни количество слогов, ни ударный гласный не являются на проанализированном материале перцептивно значимыми признаками словоформы: испытуемые склонны распознавать в предъявляемом стимуле

<sup>11</sup> В стимуле *по-моему*, который звучит в эксперименте как */pom/*, не ясно, является гласный */o/* исходным ударным гласным или безударным гласным первого слога, но для изучения надёжности восприятия гласных этот вопрос не является столь принципиальным.

меньше гласных, чем в нём содержится на самом деле, а ударный гласный был надёжно распознан только в 3 стимулах.

В рамках описываемого в настоящей статье исследования не проводилось последовательного анализа перцептивной значимости всех признаков гласных, однако необходимо отметить, что, по-видимому, особенную важность в процессе речевосприятия приобретает признак огубленности. Этот вывод подкрепляется следующими наблюдениями: — хотя в целом ударный гласный воспринимается испытуемыми не очень надёжно, в тех словах, где он является огубленным (/o/ или /u/), распознавание ударного гласного выше, чем в большинстве других слов (при этом иногда испытуемые могут воспринимать /o/ вместо /u/ и наоборот, но признак огубленности сохраняется); — в ряде слов большинство испытуемых восприняло ударный гласный как огубленный под влиянием соседних губных согласных: например, *фильмов* /f'ɪmɨf/ (гласный был интерпретирован как **у** или **ю** 54,3% испытуемых: например, *чуф*, *тюф* и др.), *буквально* /bʊ'ál'n/ (буквы **у** или **о** на позиции гласного встретились в ответах 75,7% испытуемых: *бун*, *больн* и др.), *появился* /p'ɪls/ (**у** или **о** на месте единственного гласного в словоформе — 62,9%: *пуск*, *возг* и др.).

Данные наблюдения являются, несомненно, интересными, но при их интерпретации необходимо учитывать, что огубленный гласный может появляться в некоторых случаях по законам фонотактики.

В целом, сделанные выше наблюдения свидетельствуют, по-видимому, о том, что на материале русского языка гласные не являются наиболее перцептивно значимыми элементами слова (по крайней мере, при восприятии редуцированных словоформ).

При анализе восприятия консонантного «скелета» необходимо рассмотреть несколько аспектов. Во-первых, стоит отметить, что, в целом, испытуемые достаточно точно отражают порядок следования согласных в исследуемых словоформах. Этот факт приобретает особую значимость, если принять во внимание то обстоятельство, что в ходе эксперимента словоформы предъявлялись триадами с небольшими временными интервалами внутри триады, что увеличивало вероятность неверного осмысления границ словоформы, а также влияния фонотактики.

Экспериментальный материал позволяет сделать некоторые выводы относительно того, насколько адекватно воспринимается начальный согласный в редуцированных словоформах. Для этого были рассмотрены все стимулы из тестовой последовательности, начинающиеся с одиночного согласного звука, а также с последовательности одиночного согласного и призвука другого согласного (всего 14 словоформ). Нужно отметить, что в 13 из анализируемых стимулов начальный согласный был смычным, в одном — аффрикатой (/сéе/). В шести стимулах данный согласный был распознан надёжно (выделен серым цветом в таблице 4), ещё для трёх верный вариант является преобладающим вариантом ответа.

Таблица 4

#### Словоформы, начальный согласный в которых был распознан хорошо

Словоформа	% верных распознаваний начального согласного	Примечания
ходит /kóit/	97,4	
какую-то /kóгыt/	91,7	

тебе /t'e/	70,3	Для данных словоформ получено примерно равное количество ответов для твёрдого и мягкого начального согласных. Оба варианта считались в данном случае верным распознаванием
километров /kájna/	88,6	
по-моему /pom/	60,5	
понимаешь /pъывáš/	60,6	
буквально /b'ál'n/	56,8	
(в) пятницу /p'an'c/	84,6	
просят /p'os't/	81,6	

Кроме того, стоит отметить надёжность распознавания согласного /g/ в стимуле *а она говорит /aongýt/* (94,1%).

Плохо распознавался начальный согласный в словоформах *наверно /námí/*, *кажется /kažs/*, *последний /pasýni/* и *появился /p'íc/*, а также аффриката /c/ в словоформе *целый /céé/*, на месте которой большинство испытуемых (52,9%) услышали щелевой /f/.

Сочетания согласных в начале слова распознавались плохо, что соотносится с данными о плохом распознавании консонантных кластеров, полученными в ходе экспериментов других исследователей [6]. Начальные консонантные кластеры были восприняты большинством испытуемых как одиночные согласные (часто щелевые).

Отдельно необходимо рассмотреть результаты распознавания начального согласного /n/ в стимулах *наверно* (звучал как /námí/) и *нравятся /nrec/*. В этих стимулах, вопреки распространённому мнению о том, что признак шумности-сонорности является одним из наиболее устойчивых перцептивных признаков, начальный сонорный /n/ воспринимается большинством испытуемых как шумный смычный /p/. И если в словоформе *нравятся* этот факт можно объяснить более высокой частотностью начального сочетания /pr/ по сравнению с сочетанием /nr/ (т.е. испытуемые, даже не зная, что им предлагаются для прослушивания слова, подсознательно ориентируются на законы фонотактики и правила построения слов русского языка), то для стимула *наверно* это объяснение не подходит.

Анализ перцептивно устойчивых признаков начального согласного на материале остальных стимулов подтверждает тезис о наибольшей устойчивости признаков шумности-сонорности и глухости-звонкости [6], хотя в ответах испытуемых встречались практически все возможные варианты замен согласных. Признак мягкости-твёрдости начальных согласных является менее надёжным для восприятия (в словоформах *тебе* и *километров* было получено примерно равное количество ответов с мягким и твёрдым начальным согласным), что согласуется с результатами по восприятию слогов и слов при наличии разного рода помех [15].

В позициях, отличных от начальной, замены согласных по различным признакам встречались чаще, чем в начале слова. Относительно этих позиций можно делать выводы скорее об устойчивости восприятия отдельных согласных,

чем об устойчивости определённых позиций в слове, поскольку консонантные «скелеты» исследуемых словоформ достаточно разнообразны. Наиболее последовательно воспроизводятся смычные шумные согласные (особенно в конце словоформы), щелевые /š/, /š':/, аффриката /č/, а также сонант /n/, являющийся суффиксом наречий (в словоформе *буквально* /b'ál'n/ в 86,5%, в словоформе *достаточно* /tstáčna/ — в 100% ответов). Хуже всего, как и в начале слова, распознаются консонантные кластеры, чаще всего они заменяются щелевыми согласными.

В качестве особо интересных случаев стоит отметить появление согласного /r/ в ответах на те стимулы, в составе которых он на самом деле отсутствует (данный согласный не представлен ни в нередуцированном, ни в редуцированном варианте соответствующих словоформ): *понимаешь* /ryiváš/ (63,6%, например, ответы *пэраш*, *пурваф*, *дурак*), *кажется* /kažs/ (51,4%: *кырс*, *кэрс*, *трэст* и др.), *целый* /céel/ (35,3%: *фёр*, *фер* и др.), *какую-то* /kógyit/ (30,6%: *курт*, *кург* и др.), *потому что* /ptyšt/ (27,6%: *трст*, *фирст* и др.), *(в) пятницу* /p'an'c/ (25,6%: *пер*, *перед* и др.), — а также достаточно последовательное отражение этого звука при восприятии тех стимулов, в которых он изначально присутствует.

Появление на месте ударного гласного в стимуле *кажется* /kažs/ гласного *о* в окружении *к\_сть* или *г\_рст* может также объясняться тем, что слушающий даже при работе с изолированными редуцированными словоформами, которые он не может соотнести с каким-либо конкретным словом, опирается на знание законов фонотактики и частотности звуко сочетаний в русском языке. Такая трактовка является дополнительным подтверждением того, что при восприятии речи слушающий опирается на консонантный «скелет» словоформ.

Таким образом, результаты проведённого эксперимента позволяют предположить, что при необходимости идентификации редуцированной словоформы в процессе восприятия речи слушающий опирается в первую очередь на консонантный «скелет» словоформы (последовательность согласных, начальный смычный согласный, некоторые устойчивые для восприятия согласные).

### Восприятие редуцированных словоформ в контексте. Некоторые наблюдения

Выдвинутая гипотеза подтверждается результатами восприятия описанных словоформ в контексте. Прежде всего, уже при появлении небольшого контекста редуцированные словоформы распознаются испытуемыми достаточно надёжно (14 словоформ были распознаны более чем половиной испытуемых).

Сопоставление результатов распознавания словоформ при изолированном предъявлении и в контексте представлено на диаграмме 1.

При необходимости интерпретации стимула в контексте испытуемые достаточно легко восстанавливают количество слогов до того, которое должно быть в нередуцированном варианте. Изменение же согласных для придания осмысленности фразе встречается намного реже.

Отдельного внимания заслуживают результаты другого эксперимента, направленного на определение роли контекста в процессе восприятия. В этом эксперименте описанные выше редуцированные словоформы были помещены в контекст, который должен был спровоцировать их неверную интерпретацию. Например, словоформа *просит* /p'os't/

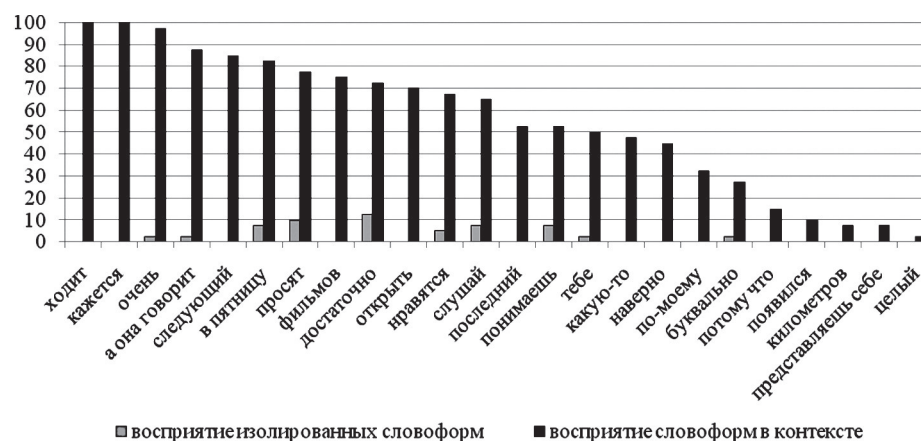


Диаграмма 1. Количество верных распознаваний редуцированных словоформ при изолированном предъявлении и в контексте (в процентах от общего числа испытуемых)

помещалась в контекст **просят раздувает** и **просят живёт** (ожидалось появление словоформы *пусть* в ответах испытуемых), словоформа *говорит /gɨt/* помещалась в контексты **говорит свечку**, **свечка говорит** (ожидалось появление ответов *горит*, *гасит* и т.п.) и т.д. Методика проведения данного эксперимента совпадала с методикой эксперимента по распознаванию редуцированных словоформ в естественном контексте (см. выше).

Результаты восприятия большинства стимулов в данном эксперименте подтверждают важность начального согласного как перцептивно значимого элемента словоформы, а также значимость консонантного «скелета». Так, словоформа *говорит /gɨt/* в контексте **свечка говорит** была воспринята 33,3% испытуемых как *горит*, для словоформы *просит /pʲosʲtʲ/* в обоих контекстах самым частотным вариантом ответа был *пусть*, при интерпретации словоформы *фильмов /ʲfʲimɨfʲ/* в контекстах и **фильмов машина** и и **фильмов приснился** испытуемые предпочитали вариант *Ефим(ов)(а)у* (12,8% и 23,1% соответственно), словоформа *тебе /tʲe/* в контексте **тебе рядом с этим** была воспринята 36% испытуемых как *ты*, 5,1% испытуемых как *то* и 3% испытуемых как *те* (правильный ответ *тебе* встретился у 6,1% испытуемых). Кроме того, интересен пример со словоформой *наверное /návɨ/* в контекстах **то же самое делается**, **наверно** и **порошки лучше покупаются**, **наверно**, где начальный согласный в контексте в большинстве случаев воспринимается как /n/ (тогда как при изолированном предъявлении наиболее частотным был согласный /p/), однако из-за влияния фонотактики в ответах испытуемых появляются также слова *сами* и *сам*.

На количество слогов и качество гласных (в том числе ударного) испытуемые опираются, по-видимому, меньше, чем на согласные; в процессе интерпретации редуцированной словоформы они склонны заменять гласные и увеличивать количество слогов (например: замены гласных в словоформе *тебе /tʲe/* (/e/->/ы/), *просят /pʲosʲtʲ/* (/o/->/u/), «восстановление» слога с ударным гласным в словоформе *говорит /gɨtʲ/* в контексте **говорит свечку** (*гáсит*, *гláдит*), «восстановление» гласных в окончаниях: *фильмов /ʲfʲimɨfʲ/* (*Ефимову*, *Ефимова*), *следующий /sʲlʲeʂʲː/*).

## Выводы

Результаты описанных выше экспериментов показывают, что редуцированные словоформы распознаются в естественном (неизменённом) контексте намного лучше, чем при изолированном предъявлении.

Количество слогов и ударный гласный отображаются испытуемыми менее надёжно, чем консонантный «скелет» словоформы, что, по-видимому, свидетельствует о том, что при распознавании редуцированных словоформ согласные являются более перцептивно значимыми элементами словоформы, чем гласные. Наибольшей перцептивной значимостью обладают: порядок следования согласных, начальный одиночный смычный согласный, смычные шумные согласные, щелевые /š/, /š':/, аффриката /č/ в любой позиции.

Наиболее перцептивно устойчивыми признаками начальных согласных являются шумность-сонорность и глухость-звонкость. Кроме того, результаты эксперимента указывают на перцептивную значимость признака огубленности гласных.

Таким образом, результаты, полученные в описанном исследовании, расходятся с результатами, полученными при анализе ослышек на материале английского и немецкого языков, которые указывают на перцептивную значимость гласных в слове.

В то же время данные эксперимента не указывают на то, что начальный согласный является наиболее перцептивно значимым элементом словоформы: существует ряд согласных, которые являются перцептивно значимыми вне зависимости от позиции; с другой стороны, в начале слова надёжно распознавались только одиночные смычные согласные, консонантные же кластеры воспринимались плохо как в начале словоформы, так и в любой другой позиции.

Следовательно, наиболее правдоподобным объяснением полученных результатов является модель повсеместной активации с опорой на консонантный «скелет».

Полученный материал позволяет предположить, что при необходимости распознавания редуцированной словоформы в спонтанной речи слушающий выбирает из перцептивного словаря ту словоформу, которая подходит по консонантному «скелету» и по контексту.

Безусловно, данная стратегия не является единственно возможной при восприятии речи. По всей видимости, в распоряжении слушающего находится целый набор стратегий, из которых он может выбрать наиболее рациональную в той или иной ситуации. Возможно также, что данный набор стратегий обусловлен конкретным языком и именно этим объясняются различия, полученные на материале английского и немецкого языков, с одной стороны, и русского языка — с другой. Представленная же в статье стратегия с опорой на консонантный «скелет» позволяет, по крайней мере, описать механизм распознавания редуцированных словоформ в русской спонтанной речи. Возможность применения данной стратегии в других ситуациях является предметом дальнейших исследований.

## Литература

1. Венцов А.В., Касевич В.Б. Проблемы восприятия речи. М.: Едиториал УРСС, 2003.
2. Чистович Л.А., Кожевников В.А., Алякринский В.В. и др. Речь. Артикуляция и восприятие. М., Л.: Наука, 1965.

3. Чистович Л.А., Венцов А.В., Гранстрем М.П. и др. Физиология речи. Восприятие речи человеком. Л.: Наука, 1976.
4. Бондарко Л.В., Вербицкая Л.А., Гейльман Н.И. и др. Фонетика спонтанной речи / Под ред. Н.Д. Светозаровой. Л.: Изд-во Ленинградского университета, 1988.
5. Касевич В.Б. Труды по языкознанию: В 2-х т. Т.1 / Под ред. Ю.А. Клейнера. СПб: Филологический факультет СПбГУ, 2006.
6. Ягунова Е.В. Восприятие согласных фонем и их дифференциальных признаков (экспериментально-фонетическое исследование на материале русского языка) / Диссертация на соискание учёной степени кандидата филологических наук. Рукопись. СПб, 1994. 261 с.
7. Штерн А.С. Влияние лингвистических факторов на восприятие речи. / Диссертация на соискание ученой степени кандидата филологических наук. Рукопись. Л., 1981. 251 с.
8. Marslen-Wilson, W.D., Welsh, A. Processing interactions during word recognition in continuous speech // *Cognitive Psychology*. 1978. Vol. 1. N1. P. 29-63.
9. Bond Z.S. Slips of the Ear // *The Handbook of Speech Perception* / Ed. by D.B. Pisoni, R.E. Remez. Oxford: Blackwell, 2005. P. 290-310.
10. Русская разговорная речь / Под ред. Е.А. Земской. М.: Наука, 1973.
11. Cutler A., Norris D. The Role of Strong Syllables in Segmentation for Lexical Access // *Journal of Experimental Psychology: Human perception and performance*. 1988. Vol. 14, N1. P. 113-121.
12. Protopapas A. Perspectives of Syllables, Stress, and Interactions in Speech Perception: Experimental and Connectionist Approaches / Theses for the Degree of Doctor of Philosophy. Brown University, 1997. 183 p.
13. Mattys S., Samuel A. Implications of stress pattern differences in spoken word recognition // *Journal of Memory & Language*. Vol. 42, №4, 2000. P. 571—596.
14. Aitchison J. *Words in the Mind: An Introduction to the Mental lexicon*. Oxford: Blackwell, 2003.
15. Штерн А.С. Перцептивный аспект речевой деятельности: (Экспериментальное исследование). СПб: Издательство Санкт-Петербургского университета, 1992.
16. Luce P.A., McLennan C.T. Spoken Word Recognition: The Challenge of Variation // *The Handbook of Speech Perception*. / Ed. by D.B. Pisoni, R.E. Remez. Blackwell Publishing Ltd, Berlin, Oxford, 2005. P. 592—609.
17. Горбова Е.В., Слепокурова Н.А., Комовкина Е.П., Макарова А.Б., Малов Е.М., Риехакайнен Е.И. Современная русская спонтанная телевизионная речь: некоторые итоги исследования // *Современная русская речь: состояние и функционирование: Сборник аналитических материалов* / Под ред. О.И. Глазуновой, Л.В. Московкина, Е.Е. Юркова. СПб: Издательский дом «МИРС», 2008. С. 5—82.
18. Бондарко Л.В. Фонетика современного русского языка: Учебное пособие. СПб: Изд-во Санкт-Петербургского университета, 1998.
19. Бондарко Л.В., Вербицкая Л.А., Гордина М.В. Основы общей фонетики. СПб: Филологический факультет СПбГУ; М.: Издательский центр «Академия», 2004.

### **Риехакайнен Елена Игоревна**

Аспирантка кафедры общего языкознания Факультета филологии и искусств Санкт-Петербургского государственного университета, научный сотрудник Лаборатории экспериментальной фонетики Института филологических исследований СПбГУ. Сфера научных интересов: восприятие речи, изучение спонтанной речи, фонетика, психолингвистика. Автор 20 научных работ.