



DOI: 10.22363/2618-8163-2018-16-3-322-343

УДК 811.161.1'34

КАЧЕСТВО И ДЛИТЕЛЬНОСТЬ БЕЗУДАРНЫХ ГЛАСНЫХ РУССКОГО ЯЗЫКА КАК АКУСТИЧЕСКИЕ КЛЮЧИ ДЛЯ ОПРЕДЕЛЕНИЯ СЛОВЕСНОЙ ГРАНИЦЫ: ПЕРЦЕПТИВНЫЙ ЭКСПЕРИМЕНТ НА МАТЕРИАЛЕ ПСЕВДОСЛОВ

П. В. Дурягин

Национальный исследовательский университет «Высшая школа экономики»
Российская Федерация, 101000, Москва, ул. Мясницкая, 20

DURATION AND FORMANT VALUES OF UNSTRESSED VOWELS IN RUSSIAN AS ACOUSTIC CUES FOR SEGMENTATION: A PERCEPTIVE EXPERIMENT BASED ON NONCE WORDS

P. V. Duryagin

National Research University Higher School of Economics
20, Myasnitskaya str., Moscow, 101000, Russian Federation

Настоящее исследование посвящено изучению механизма естественного восприятия словесных границ носителями русского языка. Характерной особенностью ритмики русского слова является наличие просодического ядра — качественная и количественная выделенность гласных ударного и первого предударного слогов по сравнению с гласными остальных слогов. Целью работы стала проверка возможностей носителей русского языка воспринимать различия между гласными разных степеней редукции (а именно, гласного [e] первого предударного слога и гласного [ъ] прочих слогов в позиции после твердого согласного) для определения места границы слова. В качестве материала эксперимента были выбраны последовательности из пяти слогов с двумя ударениями, представленные респондентам в виде вымышленных имен. Участникам эксперимента предлагался выбор из двух различных вариантов сегментации этих пятисложных последовательностей на слова (вымышленные имя и фамилия). Результаты проведенного перцептивного эксперимента позволяют предположить, что респонденты в некоторых случаях использовали характеристики гласных в качестве акустических ключей, последовательно верно определяя место прохождения словесной границы. Однако для половины стимулов доля верных ответов оказалась близкой к 50%, что свидетельствует о непоследовательности использования акустических ключей респондентами. Искусственное изменение длительности гласного первого предударного слога показало, что этот акустический параметр оказывает влияние на то, как носители русского языка воспринимают словесные границы.

Ключевые слова: русский язык, фонетика, сегментация, редукция, внешнее сандохи, восприятие звучащей речи

The research deals with natural perception of word boundaries by native speakers of Standard Russian. A specific feature of Russian word rhythmic structure is a so-called “prosodic core”: not only stressed, but also first pre-stressed vowels differ in duration and quality from vowels that occur in other positions, a phenomenon that is also commonly described as two degrees of reduction. The purpose of this study is to find out whether native Russian speakers are able to use acoustic differences between vowels [ɐ] (Degree 1 reduction) and [ə] (Degree 2 reduction) in order to recognize word boundaries correctly. The stimuli for the experiment were nonce words, five-syllable sequences including two stressed vowels; they were presented to the participants of the experiment in a form of fictional foreign names. The listeners were asked to choose between two possible ways of segmentation of these five-syllable sequences into a first name and a second name of a person. The results of the experiment show that native Russian speakers used the acoustic differences between vowels for segmentation, but the results were statistically significant only for some of the stimuli. However, for half of stimuli the listeners performed correct segmentation at chance level. In addition, artificial modification of first pre-stressed vowel duration was performed for some of the stimuli; the participants’ responses show that vowel duration influences the degree of success in the segmentation task.

Key words: Standard Russian, phonetics, segmentation, reduction, external sandhi, speech perception

Введение

Успешное восприятие звучащей речи на родном или иностранном языке включает в себя несколько этапов, один из которых определение места словесных границ в высказывании. Необходимость и важность этого этапа объясняется тем, что «основной единицей словаря слушающего применительно к русскому языку признается словоформа» (Риехакайнен, 2016: 56), или фонетическое слово: «... мы не обладаем бесконечным хранилищем, необходимым для того, чтобы хранить репрезентации всех возможных высказываний. Следовательно, сегментация является необходимой операцией» (Cutler & Butterfield, 1992: 232). Основные источники информации о месте словесных границ в родном языке (при отсутствии паузы, которая, однако, также не является надежным маркером стыка фонетических слов) — «ритмическая структура слов, фонотактика, знание лексики и фонетические признаки» (Weber & Broersma, 2012: 6–7).

Роль лексического знания в определении границ слов многообразна. Простейший пример привлечения словаря — случай, когда последовательность звуков (обычно большой длительности, не способная быть частью более длинного фонетического слова) однозначно интерпретируется слушающим и позволяет установить границы этого фонетического слова, являющиеся также границами двух соседних слов. Несколько более сложные механизмы используются в том случае, когда слушающий использует стратегию, опирающуюся на контекстную предсказуемость и частотность слов.

Знание правил фонотактики может привлекаться в процессе сегментации в том случае, если носитель языка вычленяет в звуковом потоке последовательность из двух и более сегментов, сочетание которых невозможно в данном языке в позиции внутри фонетического слова. Так, в русском языке в пределах фонетического слова (за исключением некоторых сочетаний ударного слова с энклитикой) запрещены сочетания двух согласных, различающихся только по признаку твердости/мягкости (Панов, 1979: 170). Таким образом, в последовательности звуков [ты лл’у...] носитель языка может определить наличие границы слова, не прибегая

к поиску слова *тыл* в своем ментальном лексиконе. При этом открытым остается вопрос о том, могут ли такую же роль в восприятии играть консонантные кластеры, которые «фонетически возможны для русского языка», но «не представлены в середине общераспространенных слов» (там же), например, предложенное М.В. Пановым в качестве примера сочетание [цф], возможное на стыках слов (*отец фотографировал*), но внутри фонетического слова встречающееся только в редких именах собственных (*Цфасман*).

Под «фонетическими признаками» в процитированной работе подразумеваются пограничные сигналы, основанные на употреблении разных аллофонов одной и той же фонемы. Традиционным примером подобного рода служит различие между английскими сочетаниями слов типа *keeps parking* и *keep sparking*, где наличие/отсутствие придыхания у глухого взрывного служит маркером словесной границы. Проведенные экспериментальные исследования демонстрируют, что изучающие английский язык как иностранный носители испанского (Altenberg, 2005) и японского (Ito & Strange, 2009) используют этот признак для верной сегментации значительно хуже, чем те информанты, для которых английский язык — родной. В качестве примера из русского языка можно привести позицию гласного в абсолютном начале слова после мягкого согласного предыдущего слова при отсутствии паузы (*рассказать Алене*). Экспериментальное исследование спектральных характеристик гласных в подобных позициях показало, что «минимального отличия тембра безударного начального гласного слова после мягкого согласного предшествующего слова в словосочетании, произнесенном без паузы между словами, от безударного гласного внутри слова достаточно для маркирования межсловной границы» (Моисеева, 2015: 6).

Наиболее надежной опорой для слушающего, в особенности в условиях затрудненного восприятия (например, при наличии помех в виде шума), является ритмическая структура слова (Cutler & Butterfield, 1992: 226). Ее делимитативная роль сохраняется в языках не только с фиксированным, но и со свободным (лексическим) ударением. Так, результаты экспериментов (Cutler & Norris, 1988; Cutler & Butterfield, 1992) показывают, что носители английского языка склонны значительно чаще считать начальными в слове «сильные слоги», т. е. слоги, содержащие гласные полного образования. Напротив, слоги с редуцированными гласными участники эксперимента принимали за начальные только в качестве служебных слов. Такую стратегию авторы исследования объясняют неодинаковой частотностью разных ритмических моделей слов английского языка: в исследованном ими корпусе спонтанной речи на английском языке «от 85 до 90% знаменательных слов содержали в первом слоге гласный полного образования» (Cutler & Carter, 1987: 133).

В русском языке ударение подвижно и разноместно, при этом статистические закономерности, подобные приведенной ранее для английского языка, предположительно, отсутствуют. Таким образом, делимитативная функция ударения в русском языке несколько ослаблена. При этом основной типологической особенностью ритмической структуры слова в современном русском литературном языке (далее — СРЛЯ) является «наличие двусложного просодического ядра (т.е., противопоставление ударного и первого предударного слогов всем остальным)»

(Князев, Пожарицкая, 2005: 123). Эти ритмические правила накладывают определенные ограничения на сочетаемость гласных внутри фонетического слова. Вероятно, эти запреты способны функционировать в роли «пограничных сигналов»¹: подобно последовательности из двух ударных слогов, последовательность из слогов с редуцированным гласным (гласным второй степени редукции) и ударным гласным [т'ьтá] также может свидетельствовать о наличии между слогами словесной границы. Особый случай составляет подсистема двусложных составных единиц типа *стон-кран*, в которой «в слабой фразовой позиции при отсутствии ударения на первом слоге гласные среднего подъема [o] и [e] в первом предударном слоге после твердых согласных чередуются с [ъ], а не с гласным полного образования»; такие единицы могут рассматриваться как сочетания двух фонетических слов (Князев, 2015: 277).

Цель

Настоящее исследование было проведено с целью выяснить: могут ли носители СРЛЯ при отсутствии прочих источников информации о месте словесной границы внутри синтагмы (в том числе лексических) пользоваться пограничными сигналами, содержащимися в ритмической структуре слова. В частности, предполагалось определить, в какой степени акустические различия между последовательностями «слог с редуцированным гласным» + «слог с ударным гласным» типа [т'ьтá] (запрещенное в пределах фонетического слова сочетание) и «слог с безударным гласным полного образования» + «слог с ударным гласным» типа [т'ьтá]² могут восприниматься и использоваться носителями русского языка, поставленными перед задачей сегментации.

Материалы и методы

Исследование было проведено с применением экспериментально-фонетических методов и включало в себя два этапа: продуктивный и перцептивный эксперименты. Материалом исследования на первом его этапе послужили записи фраз, включающих сочетания псевдослов, записанные от одного носителя литературного произношения. Акустические характеристики гласных в полученных записях были проанализированы с использованием компьютерной программы *Praat* (Voersma & Weenink, 2017). На втором этапе был проведен перцептивный эксперимент с участием 30 информантов, которые были поставлены перед задачей сегментирования стимулов — псевдослов, записанных на первом этапе эксперимента. Качественные и количественные характеристики гласных, а так-

¹ Трубецкой Н.С., в «Основах фонологии», представивший первое системное описание делимитативной функции звуковых единиц, отмечал, что «...ритмическая структура слова поддерживается реализацией всех фонем, и любое нарушение этой ритмической инерции, свидетельствующее во всех случаях о конце одного слова и о начале другого, приобретает тем самым особую отчетливость» (Трубецкой, 2000: 298).

² Символ [v] используется в настоящей работе вслед за Р.Ф. Касаткиной (Касаткина, 2005) для обозначения фонетической реализации <a> и <o> в позиции первого предударного гласного после твердого согласного

же данные перцептивного эксперимента были проанализированы с использованием методов математической статистики.

Результаты

Проведенное исследование показало, что в речи диктора, принимавшего участие в эксперименте, гласные первой и второй степеней редукации в позиции после твердого согласного последовательно противопоставлены по своим акустическим характеристикам, а именно, по длительности и частоте первой форманты. В то же время перцептивный эксперимент продемонстрировал, что информанты непоследовательно использовали эти акустические ключи, будучи поставленными перед задачей сегментации псевдослов. Для половины тестовых примеров (4 из 8) процент ответов с верной сегментацией оказался статистически незначимым. В то же время еще в 3 из 8 примеров, а также в 2 из 3 стимулов с искусственно модифицированной длительностью гласного количество верных ответов значительно превысило 50%.

Обсуждение

В связи с необходимостью исключить влияние лексического знания на информантов в процессе сегментации в качестве материала для эксперимента были выбраны псевдослова¹. Критерий для правомерного использования не существующих в исследуемом языке слов в подобного рода экспериментах выделял С.С. Высотский: «...облик звуковой структуры слова как обобщение реальных словоформ данной языковой системы могут представлять и квазислова, если их звуковая структура подчинена тем же правилам» (Высотский, 1973: 21). В настоящем эксперименте псевдослова представляли собой последовательности из пяти открытых слогов. Каждая последовательность имела в своем составе по два ударных слога, таким образом, все они представляли собой пары фонетических слов. Каждый слог состоял из твердого согласного [т], [к] или [с] и гласного [á], [e] или [ъ]. Выбор структуры слога и качества согласных был обусловлен удобством сегментации (Кузнецов, Отт, 1989: 39). В результате применения всех указанных критериев были созданы четыре пары сочетаний псевдослов, различающиеся только характером третьего гласного: [тáкъ#свкáтъ] и [тáкъсь#кáтъ]; [кáсь#твсáкъ] и [кáсьтъ#сáкъ]; [твкá#свкáтъ] и [твкáсь#кáтъ]; [квсá#твкáтъ] и [квсáтъ#кáтъ]. Для упрощения задачи все тестовые сочетания были представлены участникам эксперимента в виде имен собственных, а именно — в виде имен и фамилий жителей вымышленной страны: Тáка Сакáто и Тáкаса Кáто, Касá Такáто и Касáта Кáто, и т.п.

Далее к экспериментальному материалу были добавлены филлеры двух типов. Во-первых, для контроля внимательности информантов было добавлено восемь

¹ Другим возможным источником могли бы послужить омофоны (точнее, сочетания слов, различающиеся характером одного гласного) типа *это ж[ъ] над[ъ] было* и *это ж[ы]на д[в]б[ы]ла*, однако при составлении экспериментального материала не было обнаружено достаточного количества подобных примеров.

примеров с единственно возможным делением на слова (например, содержащих два ударных слога подряд): *Сата́ Ка́тако*. Во-вторых, в эксперимент были включены пары «неоднозначных» примеров — тех сочетаний, в которых должны полностью отсутствовать сигналы о месте словесной границы, следовательно, должны быть в равной мере возможны оба варианта членения на слова: *Ка́та Касака́* и *Ка́така Сакá*, *Такáса Катá* и *Такá Сакатá*.

Всего было выбрано 20 пятисложных последовательностей, далее для краткости изложения будем называть их «псевдоименами». Для каждого «псевдоимени» была придумана «рамочная» фраза так, чтобы тестовое сочетание слов было употреблено в конце этой фразы. Контроль позиции слова во фразе объясняется необходимостью учитывать существующее в русском языке влияние фразовой позиции на акустические характеристики гласных (Князев, 2005). Фразы с филлерами и тестовыми «псевдоименами» были перемешаны в случайном порядке, в результате был получен текст, начинающийся так:

Здравствуйте, меня зовут Та́са Ка́сата.

Моего друга зовут Ка́са Тасáка.

Вас ждет господин Сатака́ Сатá, и т.п.

Полученный текст был прочитан вслух диктором Д.Б. — носителем русского языка в возрасте 25 лет, москвичом, имеющим высшее филологическое образование и владеющим литературной нормой. В общей сложности было сделано 4 записи текста в произношении Д. Б. — дважды по 2 раза с промежутком в неделю (в итоге в материале эксперимента были использованы по два первых корректных прочтения каждой фразы). Текст предъявлялся диктору в виде слайдов в орфографической записи с проставленным ударением на «псевдоименах». Чтение вымышленных слов не вызвало трудностей у диктора.

Гласные в 40 записанных фразах были проанализированы с помощью компьютерной программы *Praat* (Voersma & Weenink, 2017). Была поставлена цель изучить акустические характеристики гласных в «псевдоименах» в произношении диктора, а именно: частоту первой и второй формант на стабильном участке в центральной части гласного и длительность гласного. Акустические характеристики последнего, пятого, гласного в «псевдоименах» было решено не измерять, поскольку позиция в конце синтагмы является особой в связи с действием универсального правила усиления артикуляции гласных и согласных рядом с просодическими границами, описанного, в частности, в (Fougeron & Keating, 1997). В частности, в русском языке «в словах, расположенных на конце синтагмы или фразы, удлиняется ударный гласный, а также заударный, если он находится в абсолютном конце слова» (Качковская, 2014: 68). Результаты измерения средних значений частот двух первых формант и средних длительностей 160 гласных приведены в таблице 1.

Таблица 1

Средние значения длительности и частоты формант гласных в произношении диктора Д.Б. (в скобках указаны стандартные отклонения)

Гласный	Кол-во	Длительность, мс	F1, Гц	F2, Гц
ударный [á]	68	106 (18)	704 (56)	1316 (38)
первый предударный [e]	44	60 (7)	573 (60)	1421 (62)
редуцированный [ь]	48	37 (8)	454 (39)	1452 (88)

Количественные и качественные различия между ударными, предударными и редуцированными гласными СРЛЯ неоднократно становились объектом экспериментальных исследований. В частности, изучалось акустическое сходство ударного и первого предударного гласных в СРЛЯ. Так, эксперимент, описанный в работе «Редукция гласного как показатель его ударности в современном русском литературном языке» показал, что «акустически контраст между гласными просодического ядра в СРЛЯ надежно выражен только в сильной фразовой позиции и только при помощи длительности» (Князев, 2005: 50). В работе (Barnes, 2006) также были обнаружены малые различия в средней длительности гласных просодического ядра, при этом частота первой форманты первого предударного гласного [e], вопреки ожиданиям, оказалась даже выше, чем F1 ударного [á]. В то же время автором исследования была обнаружена положительная корреляция между длительностью безударных гласных ([e] и [ь]) и значением частоты первой форманты: частота F1 уменьшалась с уменьшением их длительности. Таким образом, подтверждается объяснение механизма второй степени редукции фонетическим «недострелом» (*undershoot*), в то время как первую степень редукции автор предлагает описывать как «категориальный фонологический процесс» (Barnes, 2006: 65).

Несколько иные данные были получены в работе (Padgett & Tabain, 2005). Так, в позиции после твердого согласного во всех контекстах первый предударный гласный был значительно короче ударного гласного, при этом второй предударный в тех же позициях обладал меньшей длительностью, чем [e], в речи только части (7 из 9) дикторов (противоположный результат был получен для позиции после мягких согласных). Что касается качественных характеристик гласных, то статистически значимые различия были обнаружены для всех гласных, причем не только для F1, но и для F2.

Материал настоящего исследования был ограничен гласными после твердых согласных в открытых слогах в произношении единственного диктора, в результате его анализа были получены результаты, схожие с данными работы (Padgett & Tabain, 2005). Однофакторный дисперсионный анализ позволил обнаружить статистически значимые различия между всеми парами переменных (длительность, частота F1 и частота F2 гласных [á], [e] и [ь]), исключение составила разница между частотами вторых формант [e] и [ь], оказавшаяся статистически незначимой ($p = 0,056$). Таким образом, в речи диктора Д.Б. в описываемых позициях гласные первого предударного слога короче гласных ударного слога, но длиннее гласные прочих слогов (второй предударный, заударные); все гласные различаются частотой F1. Формантные различия между [á], [e] и [ь] в схематическом виде продемонстрированы на рисунке.

Анализ произношения диктора показал, что в его речи первый предударный гласный отличается от гласного других безударных слогов длительностью и значением частоты F1. Проведенный перцептивный эксперимент имел целью выяснить, могут ли носители русского языка использовать это различие как ключ к сегментации псевдослов.

Материалом для перцептивного эксперимента послужили описанные ранее записи 20 фраз, по два раза корректно произнесенных диктором Д.Б. Из каждой

пары произнесений необходимо было выбрать одно — то, которое будет предъявлено информантам. Выбор филлеров был произвольным (обычно использовалось первое произнесение). Напротив, из двух произнесений, содержащих исследуемое различие между [v] и [ʏ], в качестве стимула для перцептивного эксперимента выбиралось то, в котором гласный, содержащий «ключ» для сегментации, имел длительность и формантные характеристики, наиболее близкие к средним для этого гласного (см. табл. 1).

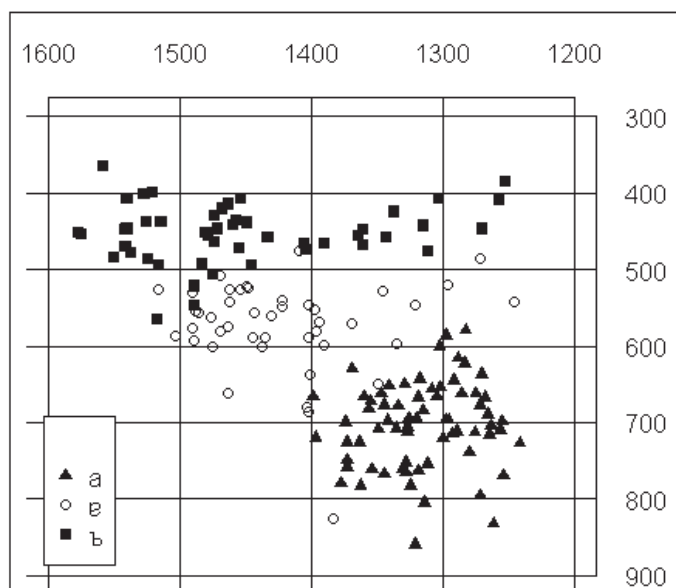


Рисунок. Значения формант гласных [á], [v] и [ʏ] в центральной части, Гц.

Горизонтальная ось — значение второй форманты, вертикальная — значение первой форманты

Кроме этого, к 20 выбранным стимулам были добавлены три стимула с гласными, длительность которых была искусственно изменена в программе *Praat*, а именно: 1) [v] с длительностью, увеличенной до 86 мс; 2) [ʏ] с длительностью, увеличенной до 74 мс; 3) [v] с длительностью, сокращенной до 40 мс. Манипуляции с длительностью проводились вручную путем копирования (или удаления) целых периодов колебаний в центральной части гласных, естественность звучания модифицированных стимулов была проверена и не вызвала нареканий у информантов. Таким образом, материал эксперимента составили 23 фразы, каждая из них содержала один стимул (12 филлеров и 11 тестовых примеров, в том числе 3 — с модифицированными гласными).

Информантам предлагалось прослушать каждую фразу дважды (с интервалом в 5 с) и выбрать один из двух вариантов, ответив на вопрос: «Какие имя и фамилию вы услышали?» (вариантами ответа были два возможных способа сегментации услышанного псевдоимени, например, *Така Саката* или *Такаса Ката*; ударения в бланках ответов проставлены не были). В эксперименте приняли участие 30 информантов (16 женщин и 14 мужчин) в возрасте от 21 до 31 года (средний возраст — 26,5 лет), носителей литературного произношения, либо родившихся в Москве, либо получивших в Москве высшее образование. Каждый информант

проходил эксперимент индивидуально; 12 человек вносили ответы в бумажный бланк, 18 — проходили опрос в электронном виде.

В 240 филлерах первого типа (8 псевдоимен с единственно возможным делением на слова) участниками было допущено лишь 4 ошибки в сегментации (по 2 в двух разных стимулах 4 разными информантами). Указанные ошибки составляют менее 2% от общего количества ответов. Анализ реакций на 4 филлера второго типа (с двумя в равной степени возможными способами сегментации) показал, что 6 из 30 информантов, сталкиваясь с указанной неоднозначностью, выбирали всегда либо первый, либо второй вариант. Остальные 24 информанта, по всей видимости, отыскивали акустические ключи для выбора «верного», по их мнению, варианта. Однако доля таких «верных» ответов (т.е., случаев, когда информанты выбирали именно тот вариант, который был прочитан диктором) составила 50% (48 из 96; при этом ни один из информантов не дал 4 «верных» ответа из 4), что может служить косвенным подтверждением отсутствия каких-либо скрытых акустических ключей (пауз, интонационных средств) при прочтении диктором псевдослов.

Результаты перцептивного эксперимента для тестовых примеров (не филлеров) приведены в таблице 2; стимулы представлены в том порядке, в котором они предлагались участникам эксперимента. Статистическая значимость результатов эксперимента для каждого стимула и для всех стимулов в совокупности определялась с применением биномиального критерия ($\alpha = 0,05$). Таким образом, при указанном уровне значимости нулевая гипотеза о том, что участники эксперимента верно делили последовательность слогов на слова с вероятностью 50% (т.е., их ответы были случайными), может быть отвергнута для 4 из 8 примеров с немодифицированными гласными и для 2 из 3 примеров с модифицированными гласными.

Таблица 2

**Результаты перцептивного эксперимента для тестовых примеров
(серым выделены стимулы, результат для которых признан статистически значимым)**

№	Гласный	Длительность, мс	F1, Гц	Кол-во верных ответов (из 30)
1	[e]	64	542	14
2	[ь]	35	430	24
3	[ь]	33	454	26
4	[e]	62	580	29
5	[e]	62	661	8
6	[ь]	31	458	10
7	[e]	52	593	17
8	[ь]	36	476	12
Примеры с измененной длительностью гласного				
1*	[e]*	64→86	542	21
6*	[ь]*	31→74	458	12
4*	[e]*	62→40	580	21

В то время как реакция участников эксперимента на филлеры имеет тривиальное объяснение (единственно возможная сегментация в филлерах первого типа, невозможность выбора единственного варианта ответа в филлерах второго типа), эксперимент с тестовыми примерами, различающимися качеством одно-

го гласного, дал противоречивые результаты. В общей сложности информантами была проведена верная сегментация 140 из 240 тестовых примеров с немодифицированными гласными (58,3%). Информанты проводили сегментацию с разной степенью успешности (от 3 до 6 верных ответов из 8). Ни одному из информантов не удалось правильно сегментировать все 8 тестовых примеров.

Результаты (табл. 2) показывают, что количество верных ответов существенно колеблется от примера к примеру. На первый взгляд, крайне высокая доля ответов с верной сегментацией в примерах № 2, 3 и 4 может свидетельствовать либо о том, что информанты действительно используют различие между гласными разных степеней редукции для сегментации, либо же о том, что в этих примерах диктор использовал некие иные акустические ключи (помимо частоты первой форманты и длительности гласного) для маркирования словесной границы. Однако предположение о наличии скрытых акустических ключей, способствующих верной сегментации, во-первых, противоречит результатам эксперимента для филлеров второго типа, во-вторых, не объясняет результаты для примеров 1 и 6—8, где процент верных ответов оказался близок к 50, а в-третьих, не объясняет примера 5, в котором более двух третей информантов восприняли [e] не как первый из двух слогов просодического ядра, а как конечный гласный первого слова.

Отдельно следует рассмотреть результаты для стимулов с модифицированными гласными. Изменение длительности гласных в 2 из 3 стимулов дало предсказуемый результат. Увеличение длительности [e] с 64 до 86 мс значительно повысило количество верных ответов (с 14 до 21), в то время как искусственное сокращение длительности первого предупредительного гласного с 62 до 40 мс привело к обратному результату (количество верных сегментаций сократилось с 29 до 21). При этом увеличение длительности редуцированного [ъ] с 31 до 74 не дало статистически значимого результата.

Таким образом, проведенный эксперимент показал, что его участники не смогли последовательно успешно использовать количественные и качественные различия между [e] и [ъ] для сегментации псевдослов. Однако отдельные стимулы (3 из 8) были верно сегментированы значительным большинством респондентов, что свидетельствует о существовании тенденции к использованию указанных акустических ключей. Другим признаком того, что носители языка в некоторой степени опирались на характеристики гласных при сегментации, является результат эксперимента для стимулов с модифицированными гласными. Искусственное сокращение и удлинение гласного первого предупредительного слога повлияло на способность участников эксперимента верно сегментировать услышанные псевдослова.

Заключение

Проведенное исследование является первой попыткой экспериментальной проверки гипотезы о том, что особенность ритмической структуры русского слова, заключающаяся в различии между качеством и длительностью гласных первой и второй степеней редукции, может использоваться носителями русского языка в качестве акустического ключа при делении звучащей речи на слова. В дальнейшем методика проведения эксперимента может быть скорректирована в целях

получения менее противоречивого результата. В частности, материал настоящего эксперимента был ограничен записью одного диктора, в дальнейшем планируется изучить характеристики безударных гласных в речи носителей других групп, например, в женском произношении (гендерные особенности ритмики слова в СРЛЯ, связанные с длительностью первого предупредного гласного, описаны Р.Ф. Касаткиной (Касаткина, 2005)). Кроме того, расширение материала эксперимента возможно за счет использования иных гласных (не только <a> и <o>, но и <i>, <e>, <y>) и иных позиций (не только после твердого, но и после мягкого согласного). Наконец, несмотря на то, что использование псевдослов представляется надежным методом исследования произношения и восприятия, методика проведения эксперимента может быть значительно упрощена в том случае, если удастся найти достаточное количество минимальных пар, состоящих из реальных словосочетаний русского языка.

В заключение следует добавить, что описанный эксперимент изначально задумывался для изучения восприятия словесных границ носителями других языков, изучающими русский как иностранный. Сегментация звучащей речи на неродном языке и на материале псевдослов неоднократно становилась объектом экспериментальных исследований (в частности, использование ключей, связанных с ритмической структурой слова, описано в Cutler et al., 1986; Sanders & Neville, 2002; Hay & Diehl, 2007; Kabak, Maniwa & Kazanina, 2010 и др.). Специфическая ритмика русского слова в этом аспекте ранее не рассматривалась. На материале настоящего исследования был проведен пилотный эксперимент, участниками которого были носители английского языка, изучающие русский как иностранный. Полученные результаты свидетельствуют о том, что англоговорящие участники эксперимента при сегментации стимулов, произнесенных русскоговорящим диктором и обрамленных высказываниями на русском языке, склонны опираться на особенности ритмики родного языка (высокая частотность слов с ударением на первом слоге, допустимость редуцированного гласного в первом предупредном слоге). Таким образом, изучение переноса ритмических ключей к сегментации из родного языка в изучаемый иностранный (в частности, из английского в русский) представляется перспективным направлением для дальнейшего исследования.

СПИСОК ЛИТЕРАТУРЫ

- Высотский С.С.* О звуковой структуре слова в русских говорах / Исследования по русской диалектологии. М.: Наука, 1973. С. 17—41.
- Касаткина Р.Ф.* Московское аканье в свете некоторых диалектных данных // Вопросы языкознания. М.: Наука, 2005. № 2. С. 29—45.
- Качковская Т.В.* Использование темпоральных характеристик для сегментации речевого потока на крупные смысловые единицы (на материале русского языка) // Тр. СПИИРАН. СПб., 2014. № 32. С. 68—81.
- Князев С.В.* Редукция гласного как показатель его ударности в современном русском литературном языке // Компьютерная лингвистика и интеллектуальные технологии. По материалам ежегод. Междунар. конф. «Диалог». М.: Изд-во РГГУ, 2015. С. 277—285.
- Князев С.В., Пожарицкая С.К.* Современный русский литературный язык: фонетика, графика, орфография, орфоэпия. М.: Академический проект, 2005. 320 с.

- Кузнецов В.Б., Отт А.В.* Автоматический синтез речи. Алгоритмы преобразования «буква-звук» и управление длительностью речевых сегментов. Таллин: Валгус, 1989. 136 с.
- Моисеева Е.В.* Реализация гласных после мягких согласных на стыках слов в современном русском языке: дисс. ...канд. филол. наук. М., 2015. 217 с.
- Панов М.В.* Современный русский язык. Фонетика: учебник для ун-тов. М.: Высш. школа, 1979. 256 с.
- Риехакайнен Е.И.* Восприятие русской устной речи: контекст + частотность. СПб.: С.-Петербург. гос. ун-т, 2016. 270 с.
- Трубецкой Н.С.* Основы фонологии. М.: Аспект-пресс, 2000. 352 с.
- Altenberg E.P.* The perception of word boundaries in a second language. *Second Language Research*. 2005. 21 (4), P. 325—358.
- Barnes J.* Strength and weakness at the interface: positional neutralization in phonetics and phonology. Berlin & New York: Mouton de Gruyter, 2006.
- Boersma P., Weenink D.* Praat: doing phonetics by computer [Computer program]. Version 6.0.33, retrieved 26 September 2017 from <http://www.praat.org/>
- Cutler A., Mehler J., Norris D. & Segui J.* The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language*. 1986. № 25. P. 385—400.
- Cutler A., Butterfield S.* Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language*. 1992. № 31. P. 218—236.
- Cutler A., Carter D.* The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language*. 1987. № 2. P. 133—142.
- Cutler A., Norris D.* The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*. 1988. № 14. P. 113—121.
- Fougeron C., Keating P.A.* Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America*. 1997. № 101 (6). P. 3728—3740.
- Hay J.S.F., Diehl R.L.* Perception of rhythmic grouping: Testing the iambic/trochaic law. *Perception and Psychophysics*. 2007. № 69 (1). P. 113—122.
- Ito K., Strange W.* Perception of allophonic cues to English word boundaries by Japanese second language learners of English. *Journal of the Acoustical Society of America*. 2009. № 125 (4). P. 2348—2360.
- Kabak B., Maniwa K. & Kazanina N.* Listeners Use Vowel Harmony and Word-Final Stress to Spot Nonsense Words: A Study of Turkish and French. *Laboratory Phonology*. 2010. № 11. P. 207—224.
- Padgett J., Tabain M.* Adaptive dispersion theory and phonological vowel reduction in Russian. *Phonetica*. 2005. № 62 (1). P. 14—54.
- Sanders L.D., Neville H.J. & Woldorff M.G.* Speech segmentation by native and non-native speakers: The use of lexical, syntactic, and stress-pattern cues. *Journal of Speech, Language, and Hearing Research*. 2002. № 45 (3). P. 519—530.
- Weber A., Broersma M.* Spoken Word Recognition in Second Language Acquisition. In C.A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics*. Bognor Regis: Wiley-Blackwell, 2012.

© Дурягин П.В., 2018

История статьи:

Дата поступления в редакцию: 22.11.2017

Дата принятия к печати: 15.03.2018

Для цитирования:

Дурагин П.В. Качество и длительность безударных гласных русского языка как акустические ключи для определения словесной границы: перцептивный эксперимент на материале псевдослов // Русистика. 2018. Т. 16. № 3. С. 322—343. DOI: 10.22363/2618-8163-2018-16-3-322-343

Сведения об авторе:

Дурагин Павел Васильевич, кандидат филологических наук, преподаватель школы лингвистики факультета гуманитарных наук НИУ «Высшая школа экономики». *Сфера научных интересов:* экспериментальная фонетика, теоретическая фонетика, русская диалектология, методика преподавания русского языка как иностранного. *Контактная информация:* pduyagin@hse.ru

ENG

Introduction

The successful perception of speech in native and foreign language consists of several stages, one of them is the ability to define word boundaries within a phrase. The necessity and importance of this ability is due to the hypothesis that “phonetic word is the basic unit of the listener’s vocabulary, in particular with regard to Russian speakers” (Riekhakainen, 2016: 56); “we do not have the infinite storage space which would be required to contain a representation of every utterance with which we might possibly be presented. Therefore segmentation is a necessary operation...” (Cutler & Butterfield, 1992: 232). The sources of information about word boundaries in native language (in the absence of pauses which are, however, relatively unreliable sources, too), are rhythmic structure, phonotactic constraints, lexical knowledge and phonetic detail» (Weber & Broersma, 2012: 6—7).

One of the basic examples that show the role of lexical knowledge in segmentation is the case when a listener that is presented with a long sequence of sounds unambiguously recognizes an embedded word of his mother tongue which in turn helps him define the boundaries of two adjacent words. More complex mechanisms based on mental lexicon are contextual predictability and word frequency.

Phonotactic information can be involved when listener identifies sequences of two or more phonetic segments that cannot occur in a position within a phonetic word in his language. For example, in Standard Russian the sequences of two consonants that differ only in the feature hardness/softness are only allowed at word boundaries and at the boundary of a clitic and a host (Panov, 1979: 170). Thus, a native speaker of Russian presented with the sound sequence ['tillʊ...] might be able to detect the boundary without first accessing rather infrequent word / 'til/ ‘backing, support’ in his mental lexicon. Yet question remains whether this relates to consonant sequences in Russian that are not prohibited within word boundaries but are only present in rare words, such as the suggested

in (Panov, 1979) sequence [tsf] that can occur at word boundaries but only is present in proper nouns such as the surname *Tsfasman*.

The phonetic details mentioned in (Weber & Broersma, 2012) are the differences based on allophones of the same phoneme. The traditional example of this kind is English minimal pair *keeps parking* and *keep sparking* where presence/absence of aspiration in the bilabial plosive marks the place of the word boundary. The experiments reported in (Altenberg, 2006) and (Ito & Strange, 2009) have shown that English speakers use this phonetic detail for segmentation with a significantly higher rate of success than Spanish and Japanese learners of English. Russian language provides comparable cases, for example, the phonetic realization of word-initial open vowel /a/ after a preceding word ending with soft consonant (in the absence of a pause). The formant measurements of vowels in this position showed that “minimal differences in vowel quality... are sufficient for marking word boundaries” (Moiseeva, 2015: 6).

Rhythmic structure is the most reliable source of information for segmentation, in particular in case of noise-masked speech (Cutler & Butterfield, 1992: 226). Its delimitative role is preserved even in languages with lexical stress. The experiments conducted in (Cutler & Norris, 1988; Cutler & Butterfield, 1992) show that for the native speakers of English “strong syllables (containing full vowels) are most likely to be the initial syllables of lexical words, whereas weak syllables (containing central, or reduced, vowels) are nonword-initial, or, if word-initial, are grammatical words”. This segmentation strategy is explained by the frequency effect of different rhythmic models of English words: an analysis of spontaneous British English corpus in (Cutler & Carter, 1987: 133) showed that “in this corpus, 90% of lexical words were found to begin with strong syllables”.

Russian language, like English, has variable stress, but no indications of comparable predictability level can be found. Thus, the delimitative function of stress is not as strong as in English. On the other hand, with regard to word rhythmic structure, Modern Standard Russian has a specific feature: “disyllabic prosodic core, where stressed and first pre-stressed syllables are contrasted with all other syllables” (Knyazev & Pozharitskaya, 2005: 123). This rhythmic structure imposes restrictions on the repertory of vowels that can be present in certain syllables. We suggested that these constrictions can have a delimitative function in Russian: if the sequence of two stressed vowels indicates the presence of word boundary, then hypothetically the sequence of a syllable with Degree 2 reduction and a stressed syllable can indicate the boundary as well because such disyllabic sequences are prohibited by Russian phonotactics. It is necessary to mark here that a subsystem of Russian complex words such as *стоп-кран* ‘emergency brake’ is left aside here. In these lexemes, according to (Knyazev, 2015: 277), “in position without phrasal accent and in pronunciation without the stress on the first syllable, phonemes /o/ and /e/ in first pre-stressed syllable have phonetic realization [ə]”; these units, however, should be regarded as two different phonetic words.

Purpose

The purpose of the experiment was to find out whether speakers of Standard Russian can utilize the information about word boundaries that is contained in Russian word rhythmic structure when all other sources of information for segmentation (including

lexical) are absent from the signal. In particular, whether the acoustic differences between the sequences “a syllable with [ə] + a syllable with a stressed vowel” (tə'ta, prohibited by Russian phonotactics within a word) and “a syllable with [e]¹ + a syllable with a stressed vowel” (te'ta, allowed by Russian phonotactics within a word) can be perceived by Russian speakers and used for segmentation purposes. In order to achieve the purpose a perceptive experiment was conducted.

Methods and materials

The methods of experimental phonetics were used in the present research. The experiment included two parts: productive and perceptive stages. During the first stage the phrases including stimuli were recorded from one native speaker. Acoustic features of vowels in his pronunciation were analyzed by means of Praat software (Boersma & Weenink, 2017). During the second stage a perceptive experiment was conducted: 30 participants were instructed to segment the stimuli. Acoustic differences between different vowels and listeners' responses were analyzed with statistical methods.

Results

The research confirmed that in the recorded pronunciation of the speaker Degree 1 and Degree 2 reduction vowels significantly differ in duration and F1 frequency. The perceptive experiment has shown that listeners used these acoustic keys inconsistently when they were asked to segment nonce-words. For half of natural experimental stimuli (4 out of 8) the number of responses containing correct segmentation didn't differ significantly from chance level. On the other hand, for 3 other natural stimuli (and also for 2 out of 3 stimuli containing artificially modified vowels) the result was statistically significant.

Discussion

Due to the need to exclude the influence of lexical knowledge on the listeners doing the segmentation task, nonce-words were chosen as stimuli². The stimuli in question were the sequences of five open syllables first presented in orthography with stress marked by acute. Every sequence included two stressed syllables, therefore, all of them could be identified as two phonetic words of different length. Every syllable included a hard consonant *m* [t], *κ* [k] or *c* [s] and a vowel *a* or *o* (*o* was used in order to make nonce-words less monotonous and was only present in final unstressed syllables). The choice of syllable structure and obstruent consonants was intended to facilitate further sound segmentation and duration measurements (Kuznetsov & Ott, 1989). Four minimal pairs of nonce-words based on these criteria were formed, in Standard Russian they are supposed to be pronounced as ['takə#sək'atə] and ['takəsə#'katə]; ['kasə#tə'sakə] and

¹ The symbol [e] that refers to the open vowel in the first pre-stressed syllable after a hard consonant is used in the present paper following (Kasatkina, 2005).

² Homophonic phrases, such as *это ж[э] над[э] было* (possible translation ‘it was necessary’) and *это ж[і]на д[в]была* ‘wife has gained it’, could have served as another possible source of stimuli but the number of such phrases was considered not sufficient for the present experiment.

[ˈkasətə#ˈsakə]; [tɐˈka#sɐˈkatə] и [tɐkaˈsə#ˈkatə]; [kɐˈsa#tɐˈkatə] и [kɐsˈatə#ˈkatə]. To make both production and perception tasks easier for participants of the experiments all these nonce-words were presented to them in Russian orthography as the names and surnames of people from some fictional country: *Та́ка Сака́то* and *Та́каса Ка́то* (*Taka Sakato vs. Takasa Kato*), *Каца́ Така́то* and *Каца́ма Ка́то* (*Kasa Takato vs. Kasata Kato*), etc.

The experimental material was further expanded by adding two groups of fillers. Firstly, in order to control listeners' involvement, eight stimuli with the only acceptable segmentation were added (for example, sequences containing two stressed syllables in a row): [sɐˈta#ˈkatəkə] (*Сама́ Кама́ко*). Secondly, several pairs of “ambiguous” stimuli were added: the sequences where no delimitative information is present, consequently, both segmentation answers had equal chances to be chosen: [ˈkatə#kəsɐˈka] and [ˈkatəkə#sɐˈka] (*Ка́ма Касака́* and *Ка́мака Сака́*).

In total, 20 stimuli were selected; every “fictional name” was embedded in the final part of a carrier phrase. The need to control phrasal position of stimuli is due to the influence of phrasal position on the acoustic characteristics of vowels (Knyazev 2006). The phrases were presented in a pseudo-random order; thus a text was produced. The beginning of this text is presented below:

Здравствуйте, меня зовут Та́са Ка́сата (*Hello, my name is Tasa Kasata*)

Моего дру́га зовут Ка́са Таца́ка (*My friend's name is Kasa Tasaka*)

Вас жде́т господин Са́така Са́та (*You are expected by Mr. Sataka Sata*), etc.

The text was recorded from speaker D.B., a 25-year-old native Russian speaker, a Muscovite with a degree in Russian philology but no professional interest in phonetics or linguistics. Four recordings were made in total: two sessions of two recordings with an interval of two weeks. The first two correct pronunciations of every phrase were further used for the analyses. The text was presented in a form of PowerPoint slides, every slide contained one phrase in orthography with marked stress. The informant reported no difficulties in reading the nonce-words.

The acoustic characteristics of vowels in 40 phrases were analyzed in Praat software (Boersma & Weenink, 2017). Vowel durations and frequencies of F1 and F2 at the center of vowels were measured. Only the first four vowels in the stimuli were analyzed; the last vowels in all phrases were not included in the experiment due to the possible effects of prosodic domain strengthening (Fougeron & Keating, 1997). For example, it was previously shown that in Russian «stressed and post-stressed word-final vowels lengthen in words at the end of phrases (Kachkovskaya, 2014: 68). The results of these measurements for all 160 vowels are presented in Table 1.

Table 1

Mean duration, F1 and F2 frequencies for vowels in nonce-words pronounced by speaker D.B. (with standard deviations)

Vowel	Number of measurements	Duration, ms	F1, Hz	F2, Hz
stressed [á]	68	106 (18)	704 (56)	1316 (38)
first pre-stressed [ɐ]	44	60 (7)	573 (60)	1421 (62)
other non-stressed [ə]	48	37 (8)	454 (39)	1452 (88)

The qualitative and quantitative differences between stressed, first pre-stressed and other unstressed vowels have been previously studied experimentally. In particular,

common acoustic features of stressed [a] and Degree 1 [ɐ] were shown. An experiment in (Knyazev, 2006) showed that the acoustic contrast between the prosodic core vowels in Standard Russian is significant only under phrasal accent and only in terms of duration. Significant durational differences (and no differences in F1 frequency) between the prosodic core vowels were reported in (Barnes, 2006: 65). Also, in this study strong correlation between duration and F1 of unstressed vowels lead to the conclusion that Russian has “one phonological process (Degree 1 reduction) and one phonetic reduction process (Degree 2 reduction)”.

Different results are reported in (Padgett & Tabain, 2005), where first pre-stressed vowels were shown to be significantly shorter than stressed vowels in all positions after hard consonants, while other unstressed vowels (Degree 2 reduction [ə]) in the same contexts were significantly shorter than [ɐ] for 7 out of 9 speakers. As for the qualitative characteristics, significant differences in F1 and F2 frequency were shown for all degrees of reduction.

The purpose of the present experiment was mostly to study perception, so the experimental material was restricted to vowels in open syllables pronounced by only one speaker of Standard Russian. The results of acoustic analysis resemble the findings of (Padgett & Tabain, 2005). Single-factor analysis of variance has shown significant differences for all pairs of variables (duration, F1 and F2 frequencies for all three vowels [á], [ɐ] and [ə]), with the exception of the difference in F2 frequency between Degree 1 and Degree 2 reduction vowels [ɐ] and [ə] that was slightly above p-value ($p = .056$). Therefore, for speaker D.B. the three studied groups of vowels differ in duration and F1 frequency; stressed vowels differ in F2 frequency from unstressed vowels. Differences in F1 and F2 measurements are illustrated schematically at Figure.

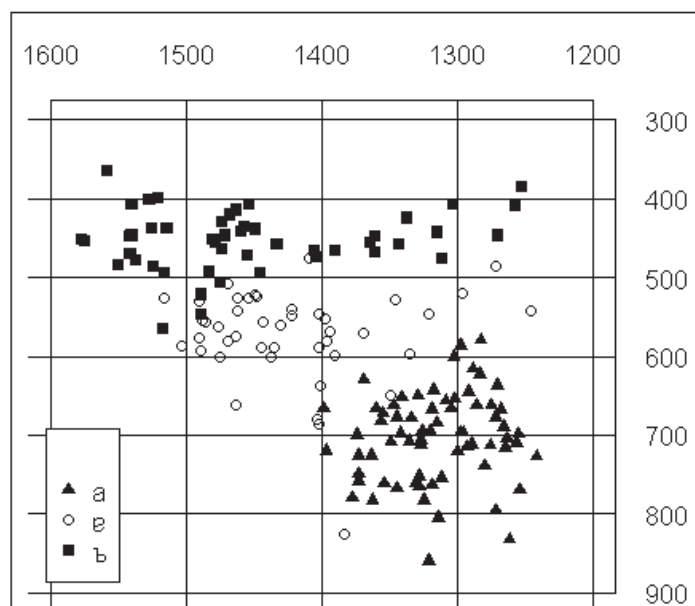


Figure. Formant frequencies of studied vowels measured in central stable part, Hz.
Horizontal axis — F2 frequency, vertical axis — F1 frequency

The analysis of vowels pronounced by the speaker has shown that in his pronunciation Degree 1 and Degree 2 reduction vowels differ in duration and F1 frequency. A perceptive experiment was held in order to find out whether native Russian speakers can utilize these acoustic differences while performing a task of segmentation.

For each of 20 experimental phrases one pronunciation was chosen as a stimulus for perceptive experiment. While the choice of filler phrases was random, the choice of stimuli that included the difference between [ɐ] and [ə] was based on the acoustic characteristics of the studied vowels. The tokens in which vowels that presumably contained keys for segmentation had “typical” (nearest to the mean values shown in Table 1) duration and F1 frequency for [ɐ] and [ə] were chosen for this part of the experiment.

In addition, three stimuli with artificially modified vowels were included in the perceptive experiment: 1) [ɐ] with the duration increased to 86 ms; 2) [ə] with the duration increased to 74 ms; 3) [ɐ] with the duration reduced to 40 ms. Manipulations with duration were made in Praat by copying whole periods of oscillations in the central part of the vowel. The modified stimuli sounded naturally; tested informants were unable to distinguish them from non-modified stimuli. In total, the material of the perceptive experiment included 23 phrases, each of them included one stimulus (“name of a foreigner”): 12 filler phrases, 8 non-modified target phrases and 3 target phrases with modified vowel duration.

The phrases were presented to the participants of the experiment twice with five-second intervals. The informants were asked to answer the question “What were the name and the surname of the person that you have heard?”. The informants had to choose one of two possible variants of segmentation for each phrase, e.g. *Така Саката* (Taka Sakata) or *Такаса Ката* (Takasa Kata). The stress was not marked in the forms that the participants were asked to fill in.

30 native Russian speakers took part in the experiment (16 females and 14 males), aged 21–31 (mean age 26,5 years old). All of them either were born in Moscow or obtained higher education in Moscow. Every informant had an individual experimental session in headphones; 12 of them filled in the forms with answers in paper, 18 — in electronic format.

In 240 presentations of fillers with the only possible segmentation the participants made only 4 “mistakes” (2 “mistakes” in 2 different stimuli made by 4 different informants). The reaction to 120 presentations of 4 fillers with two equally possible correct responses showed two strategies. 6 out of 30 informants when they faced this uncertainty always chose the first or the second answer in the form. Other 24 informants were apparently trying to find acoustic keys for segmentation but nevertheless their success rate (the number of answers when they correctly chose the segmentation that was initially presented to speaker D.B.) was at chance level (48 out of 96; none of the participants gave 4 “correct” answers out of 4). These results support the assumption that no acoustic keys for segmentation (including pauses and pitch contour features) were present in this group of fillers.

The results of the perceptive experiment for test stimuli are presented in Table 2, the stimuli are given here in the order of presentation to listeners. To measure the statistical significance binomial test was applied ($\alpha = 0,05$). The test shows that the null hypothesis

that the informants performed segmentation at chance level should be rejected for 4 out of 8 test stimuli with non-modified vowels and for 2 out of 3 stimuli with artificially modified vowels.

Table 2

Perceptive experiment results for test stimuli. Statistically significant results are marked with grey colour

№	Vowel	Duration, ms	F1, Hz	Number of correct responses (out of 30)
1	[ɐ]	64	542	14
2	[ə]	35	430	24
3	[ə]	33	454	26
4	[ɐ]	62	580	29
5	[ɐ]	62	661	8
6	[ə]	31	458	10
7	[ɐ]	52	593	17
8	[ə]	36	476	12
The stimuli with artificially modified vowels				
1*	[ɐ]*	64→86	542	21
6*	[ə]*	31→74	458	12
4*	[ɐ]*	62→40	580	21

While the response to fillers shown above has a trivial explanation, the experiment with test stimuli showed partly unexpected and inconsistent results. In general, 140 out of 240 test stimuli with non-modified vowels (58,3%) that were presented to listeners were segmented correctly. The informants' success rate varied (3 to 6 correct responses out of 8), but none of them segmented the nonce-words 100% correctly.

Table 2 shows that the success rate for individual stimuli varies significantly. At first glance, the extremely high success rate for the stimuli 2, 3 and 4 might either mean that the informants used the acoustic differences between the unstressed vowels or that they utilized some other acoustic keys, for example, prosodic features or pauses. However, the results for fillers tend to support the first hypothesis. None of the hypotheses explain the results for the stimuli 1 and 6-8 where the informants performed at chance level and stimulus 5 where the majority of listeners perceived [ɐ] as a word-final vowel.

The results for the stimuli with artificially modified vowels should be considered separately. Table 2 shows that the modification of vowel duration worked as predicted in 2 out of 3 cases. The lengthening of [ɐ] from 64 to 86 ms increased the number of correct answers (from 14 to 21), and the shortening of a similar vowel from 62 to 40 ms led to the opposite result (the number of correct answers changed from 29 to 21). However, artificial lengthening of [ə] from 31 to 74 ms had no result.

Thus, the perceptive experiment has shown that the listeners were unable to use the acoustic keys of unstressed vowels [ɐ] и [ə] for segmentation consistently. However, some individual stimuli were divided into words correctly at a high success rate. These findings might support the hypothesis that Russian speakers can use these acoustic keys. Other findings in favor of this assumption are the results for artificially modified vowels showing that lengthening and shortening of [ɐ] significantly changed the listeners' segmentation success rate.

Conclusion

The paper is an experimental verification of a hypothesis that the features of Russian word rhythmic structure, namely, the acoustic differences between Degree 1 and Degree 2 reduction vowel, can be used by native speakers for segmenting natural speech. The research can be continued by further modification of experimental design. Specifically, the production experiment included the recordings of only one speaker. Further experiments should include analyses of pronunciation by other groups of native Russian speakers, for example, female speakers (gender-based features of Modern Standard Russian phonetics have been mentioned in (Kasatkina, 2005)). Moreover, further extension of experimental material could include other Russian vowels in different positions (including the position after a palatalized consonant). Finally, despite the fact that using nonce-words can be considered a robust way to observe pronunciation and perception, using real Russian words (in case a sufficient number of minimal pairs will be found) would've made the experiment design less complicated for the informants.

Finally, the present experiment was initially planned for L2 research purposes. The language-specific strategies for segmentation have been observed in a number of experimental papers (e.g., Cutler et al., 1986; Sanders & Neville, 2000; Hay & Diehl, 2007; Kabak, Maniwa & Kazanina, 2010). However, there have been no studies to date about the rhythmic structure of Russian word in this aspect. A pilot experiment was held based on the material of the present paper. It showed that English-speaking students that studied Russian used their own specific strategy for segmenting Russian nonce-words. Presumably, this strategy could be based on the features of English word rhythmic structure. Therefore, it could be of interest to examine the transfer of segmentation strategies into L2 for learners of Russian.

REFERENCES

- Vysotskii, S.S. (1973). O zvukovoi strukture slova v russkikh govorakh. *Issledovaniya po russkoi dialektologii*. [On the sound structure of words in Russian dialects. *Studies on Russian dialectology*]. 17–41. Moscow: Nauka Publ. (In Russ).
- Kasatkina, R.F. (2005). Moskovskoe akan'e v svete nekotorykh dialektnykh dannykh. *Voprosy yazykoznaniiya*. [Moscow Akan'e in the light of some dialectal data. *Questions of Linguistics*]. 2, 29–45. Moscow: Nauka Publ. (In Russ).
- Kachkovskaya, T.V. (2014). *Ispol'zovanie temporal'nykh kharakteristik dlya segmentatsii rechevogo potoka na krupnye smyslovye edinitsy (na materiale russkogo yazyka)*. [The use of temporal characteristics for the segmentation of the speech stream into large semantic units (on the material of the Russian language)]. Tr. SPIIRAN Publ. 32, 68–81. Saint Petersburg. (In Russ).
- Knyazev, S.V. (2015). Reduktsiya glasnogo kak pokazatel' ego udarnosti v sovremennom russkom literaturnom yazyke. [Reduction of the vowel as an indicator of its impact in modern Russian literary language]. *Computer linguistics and intellectual technologies. Based on the materials of the annual International Conference "Dialogue"*. (pp. 277–285). Moscow: Izd-vo RGGU Publ. (In Russ).
- Knyazev, S.V., & Pozharitskaya, S.K. (2005). *Sovremennyi russkii literaturnyi yazyk: fonetika, grafika, orfografiya, orfoepiya*. [Modern Russian literary language: phonetics, graphics, orthography, orthoepia]. Moscow: Akademicheskii proekt Publ. (In Russ).

- Kuznetsov, V.B., & Ott, A.V. (1989). *Avtomaticeskii sintez rechi. Algoritmy preobrazovaniya «bukva-zvuk» i upravlenie dlitel'nost'yu rechevykh segmentov*. [Automatic speech synthesis. Algorithms for the conversion of “letter-sound” and control of the duration of speech segments]. Tallin: Valgus Publ. (In Russ).
- Moiseeva, E.V. (2015). *Realizatsiya glasnykh posle myagkikh soglasnykh na stykakh slov v sovremennom russkom yazyke*. [Realization of vowels after soft concordant words in modern Russian]. (Kandidat dissertation, Moscow). (In Russ).
- Panov, M.V. (1979). *Sovremennyi russkii yazyk. Fonetika: uchebnik dlya un-tov*. [The modern Russian language. Phonetics]. Moscow: Vyssh. shkola Publ. (In Russ).
- Riekhakainen, E.I. (2016). *Vospriyatie russkoi ustnoi rechi: kontekst + chastotnost'*. [Perception of Russian oral speech: context + frequency]. Saint-Petersburg: S.-Peterb. gos. un-t Publ. (In Russ).
- Trubetzkoy, N.S. (1969). *Principles of phonology*. [Fundamentals of phonology]. Berkeley: University of California Press. (In Eng).
- Altenberg, E.P. (2005). *The perception of word boundaries in a second language*. Second Language Research. 21 (4), 325—358. (In Eng).
- Barnes, J. (2006). *Strength and weakness at the interface: positional neutralization in phonetics and phonology*. Berlin & New York: Mouton de Gruyter. (In Eng).
- Boersma, P., Weenink, & D. Praat: doing phonetics by computer [Computer program]. Version 6.0.33, retrieved 26 September 2017 from <http://www.praat.org/>
- Cutler, A., Mehler, J., Norris, D. & Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language*. 25, 385—400. (In Eng).
- Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language*. 31, 218—236. (In Eng).
- Cutler, A., & Carter, D. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language*. 2, 133—142. (In Eng).
- Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*. 14, 113—121. (In Eng).
- Fougeron, C., & Keating, P.A. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America*. 101 (6), 3728—3740. (In Eng).
- Hay, J.S.F., & Diehl, R.L. (2007). Perception of rhythmic grouping: Testing the iambic/trochaic law. *Perception and Psychophysics*. 69 (1), 113—122. (In Eng).
- Ito, K., & Strange, W. (2009). Perception of allophonic cues to English word boundaries by Japanese second language learners of English. *Journal of the Acoustical Society of America*. 125 (4), 2348—2360. (In Eng).
- Kabak, B., Maniwa, K. & Kazanina N. (2010). *Listeners Use Vowel Harmony and Word-Final Stress to Spot Nonsense Words: A Study of Turkish and French*. *Laboratory Phonology*. 11, 207—224. (In Eng).
- Padgett, J., & Tabain, M. (2005). Adaptive dispersion theory and phonological vowel reduction in Russian. *Phonetica*. 62 (1), 14—54. (In Eng).
- Sanders, L.D., Neville, H.J. & Woldorff, M.G. (2002). Speech segmentation by native and non-native speakers: The use of lexical, syntactic, and stress-pattern cues. *Journal of Speech, Language, and Hearing Research*. 45 (3), 519—530. (In Eng).
- Weber, A., & Broersma, M. (2012). Spoken Word Recognition in Second Language Acquisition. In C.A. Chapelle (Ed.). *The Encyclopedia of Applied Linguistics*. Bognor Regis: Wiley-Blackwell. (In Eng).

Article history:

Received: 22.11.2017

Accepted: 15.03.2018

For citation:

Duryagin P.V. (2018) Duration and Formant Values of Unstressed Vowels in Russian as Acoustic Cues for Segmentation: a Perceptive Experiment Based on Nonce Words. *Russian language studies*, 16 (3), 322–343. DOI: 10.22363/2618-8163-2018-16-3-322-343

Bio note:

Duryagin Pavel Vasilievich, Candidate of Sciences in Philology, Lecturer at Faculty of Humanities, School of Linguistics, National Research University Higher School of Economics. *Research interests*: Russian dialectology, second language acquisition: Russian as L2, experimental phonetics. *Contact information*: pduryagin@hse.ru