

Звуковая реальность словоизменительных аффиксов (по данным звукового корпуса русского языка)¹

Phonetic realization of Russian inflections in the oral speech corpus of everyday communication

Степанова С. Б. (stsvet_2002@mail.ru),
Асиновский А. С. (a.s.asinovsky@gmail.com),
Рыко А. И. (aryko@mail.ru),
Шерстинова Т. Ю. (sherstinova@gmail.com)

Факультет филологии СПбГУ

На материале записи спонтанных монологов из Звукового корпуса русского языка проведен анализ словоизменительных аффиксов. Выявлены наиболее часто встречающиеся грамматические значения, передаваемые флексиями, их возможные фонетические варианты и случаи недореализации флексий.

Общеизвестным является факт, что русский язык относится к языкам флективного строя, то есть грамматическое значение словоформы (род, число, падеж, лицо) выражается в русском языке через окончание (аффикс), которое служит для указания на синтаксические отношения, связи данного слова с другими словами во фразе. При этом практически все русские флексии являются омонимичным: одно и то же звучание может выражать различные грамматические значения: так, флексия [oj] может иметь значение «прилагательное мужского рода, единственного числа, именительный падеж — *большой*» или «существительное женского рода, единственного числа, творительный падеж — *с добротой*» и др. Думается, при создании систем автоматического распознавания речи небесполезными могли бы быть сведения, какие грамматические значения одной и той же звуковой манифестации являются наиболее частотными.

С другой стороны, флексии, находясь в артикуляторно слабой позиции (заударный слог, конец слова, конец фразы), в наибольшей степени подвержены качественной редукции и отличаются максимальной вариативностью реализаций. Анализ реального звучания словоформ и, в частности, словоизменительных аффиксов на материале представительного корпуса спонтанной речи — одно из направлений исследований, выполняемых на матери-

але ЗКРЯ, первые результаты такого исследования представлены в данном докладе.

Для проведения **корпусного** исследования реализаций словоизменительных аффиксов в спонтанной русской речи, в качестве первого шага, для нескольких файлов из Звукового корпуса русского языка (ЗКРЯ) было осуществлено аннотирование по методу сплошной выборки: вычленились и транскрибировались **все** встречающиеся в материале флексии существительных, прилагательных, финитных форм глаголов, причастий, порядковых числительных, изменяемых разрядов местоимений (притяжательных, указательных и т. д.).

Необходимо сразу оговориться, что в некоторых случаях, мы отходили от классического понимания флексии и выделяли «финаль слова» — так называемый *морфный комплекс*, завершающий словоформу. Это относится к существительным на *-ие* (напр., *поколен{uj-e}*, *настроен{uj-e}* и т. п.): здесь мы нарушаем традиционный принцип выделения флексии и рассматриваем финальную последовательность *-ие* как единый элемент. Мы сочли возможным «отдать» флексии гласную основы, так как:

- во-первых, финаль в подавляющем большинстве случаев реализовалась как достаточно краткий, однородный гласный, делить который на основу и флексию мы сочли нецелесообразным;

¹ Исследование выполнено при поддержке гранта РГНФ «Разработка информационной среды для мониторинга устной русской речи» (09-04-12115в).

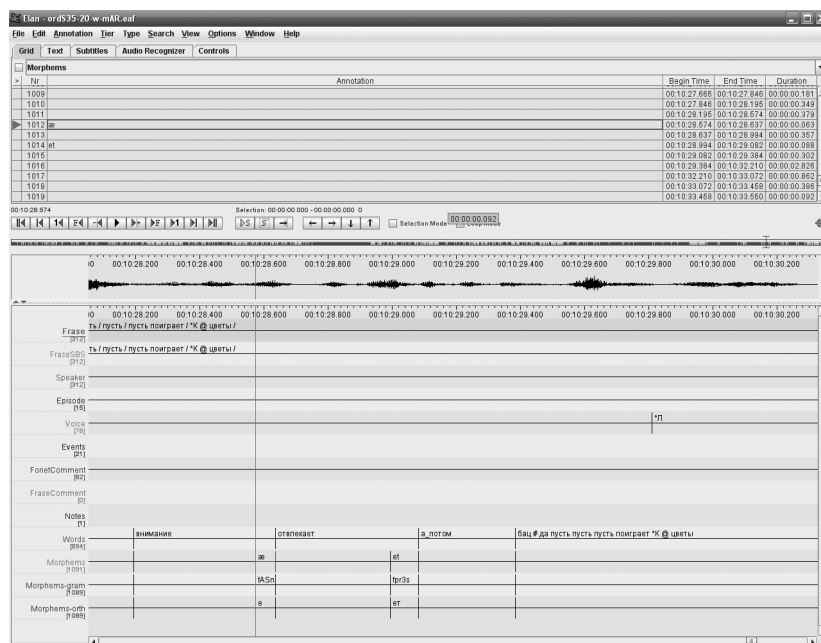


Рис. 1. Аннотирование на уровне Morphems-gram и Morphems-orth в программе ELAN (часть фразы из файла ordS35-20: ага / это он сейчас внимание отвлекает / а потом бац / # да пусть / пусть / пусть поиграет / *K @ цветы /)

- во-вторых, считаем маловероятным, что эта финаль может на фонетическом уровне реализоваться как-то иначе, чем, например, флексия в прилагательных (*синие*).

Процедура выделения флексии методом слухового анализа, поддержанного возможностью наблюдать реализацию звучания на спектрограмме и осциллограмме, а также транскрибирование слухового впечатления от флексии проводились в программе Praat. Транскрипция осуществлялась в символах МФА. В случае если морфема не была реализована или «скрылась» в предыдущем согласном, выделялся короткий символический отрезок в конце словоформы, который обозначался знаком [-] (отсутствие реализации, полная редукция).

Затем файл Praat конвертировался в файл программы ELAN и проводилось аннотирование флексий на уровнях Morphems-gram (грамматическое значение морфемы) и Morphems-orth (орфографическое представление морфемы) — см. рис. 1.²

Для аннотирования грамматического значения словоизменяемых флексий мы использовали следующую систему обозначений³. Первым символом в каждом обозначении является «f» — указание на принадлежность элемента к классу флексий.

1. Для имен — указание на падеж и число: fNS — именительный ед.ч., fGS — родительный ед.ч., fG2S — «второй родительный» (партитивный)

ед.ч., fDS — дательный ед.ч., fAS — винительный ед.ч., fASi — винительный ед.ч. (для неодушевленных существительных мужского и среднего рода и согласованных с ними прилагательных), fIS — творительный ед.ч., fLS — предложный ед.ч., fL2S — местный ед.ч. (*в лесу*), fNP — именительный мн.ч., fGP — родительный мн.ч., fDP — дательный мн.ч., fAP — винительный мн.ч., fAPi — винительный мн.ч. (для неодушевленных существительных и согласованных с ними прилагательных), fIP — творительный мн.ч., fLP — предложный мн.ч.

1.2. Для существительных после указания на падеж и число отмечен тип склонения:

fNS1 — 1-е склонение (*сестра, земля*), fNS2 — 2-е склонение (*кот, конь; село, поле*), fNS3 — 3-е склонение (*лошадь*), fNS4 — существительные 2-го склонения с исходном основы на -j, которые, по нашим данным, образуют особый тип, в котором невозможно отделить окончание от финали основы (*санаторий*).

1.3. Для прилагательных (а также местоимений-прилагательных, порядковых числительных) после указания на падеж и число отмечаем род или принадлежность к подпарадигме мн.ч.:

fNSm — мужской, fNSf — женский, fNSn — средний, fNPp — подпарадигма мн.ч.

2. Глаголы

2.1. Подсистема настоящего/будущего времени.

Отмечаем принадлежность к этой подсистеме (pr), лицо (1, 2, 3) и число (s, p):

fpr1s (*хожу*), fpr2s (*ходишь*), fpr3s (*ходит*), fpr1p (*ходим*), fpr2p (*ходите*), fpr3p (*ходят*)

² О системе аннотирования на других уровнях см. (Шерстинова, Рыко, Степанова 2009)

³ В силу специфики данного конкретного исследования нулевая флексия не выделялась и не аннотировалась.

2.2. Подсистема прошедшего времени.

Отмечаем принадлежность к этой подсистеме (р) и к роду (множественному числу):
fprf (*ходила*), fprn (*ходило*), fprp (*ходили*)

2.3. Краткие пассивные причастия прошедшего времени

fpartf (*написана*), fpartn (*написано*), fpartp (*написаны*)

Всего к настоящему времени таким образом проаннотировано 880 морфем из 4 файлов корпуса ОРД.

В табл.1 представлены результаты автоматической обработки морфемных уровней аннотирования. Приведены флексии, встретившиеся в проанализированном материале не меньше 10 раз.

Таблица 1. Частота встречаемости морфем с различными грамматическими значениями

КОД грамматического значения	Абсолютное количество	%
fNS1	72	8,18
fNSm	71	8,07
fprf	52	5,91
fpr2s	46	5,23
fpr3s	43	4,89
fAS1	41	4,66
fNSf	35	3,98
fpp	35	3,98
fNP2	33	3,75
fGS1	26	2,95
fpr1s	26	2,95
fGS2	25	2,84
fNPp	24	2,73
fpr3p	23	2,61
fNSn	22	2,50
fpr1p	20	2,27
fISm	19	2,16
fASf	18	2,05
fASim	18	2,05
fLS2	16	1,82
fLS1	15	1,70
fNS4	14	1,59
fAPi2	11	1,25
fNS2	11	1,25
fNP1	10	1,14

Таблица 2. Частота встречаемости различного типа морфем, представленных в одинаковой орфографической форме

Орфографическое представление	Абсолютное количество	%
а	172	19,55
и	92	10,45
ой	63	7,16

Орфографическое представление	Абсолютное количество	%
у	62	7,05
е	37	4,20
ешь	36	4,09
ый	34	3,86
о	30	3,41
ая	29	3,30
ы	29	3,30
ие	22	2,50
ое	22	2,50
ю	22	2,50
ет	21	2,39
ит	18	2,05
им	17	1,93
ом	16	1,82
ые	13	1,48
ей	12	1,36
ем	12	1,36
ий	12	1,36
ого	9	1,02
юг	9	1,02
ишь	7	0,80
ут	7	0,80
ую	7	0,80
ия	6	0,68
ым	6	0,68
я	6	0,68
ят	6	0,68
ете	5	0,57
ём	4	0,45
ёт	4	0,45
ов	4	0,45
ых	4	0,45
ами	3	0,34
ёшь	3	0,34
ими	3	0,34
ому	3	0,34
ах	2	0,23
ев	2	0,23
ии	2	0,23
ях	2	0,23
ам	1	0,11
ат	1	0,11
его æ	1	0,11
ием	1	0,11
ию	1	0,11
Всего:	880	

Сопоставление этих таблиц позволяет сделать некоторые предварительные наблюдения над функционированием словоизменяемых аффиксов в русской спонтанной речи и наметить пути их дальнейшего исследования.

1. Типы и реализации флексий, представленных орфографическим {а}

Как видно из табл. 2, самой частотной орфографической манифестацией флексии на нашем материале является {а} — почти 20 % от всех словоизменительных финалей. Рассмотрим, какие грамматические значения она чаще всего выражает и каким образом реализуется.

В табл.3 видно, что самое частотное значение {а} — Им. п. имен существительных ед. ч. 1-ого склонения (*мама, папа*). Чуть реже встречается глагольное окончание (ж. род, прош. вр.: *была, делала*), значительно реже — Род. п. ед. ч. существительных 2-ого склонения (*друга*). На долю остальных 9 значений {а} приходится около 13 % её употреблений в проанализированном материале.

Таблица 3. Количество различных грамматических значений морфемы, представленной {а}

Грамматическое значение	Абсолютное количество	%
fNS1	68	39,535
fpf	52	30,233
fGS2	25	14,535
fNP2	8	4,6512
fAS2	6	3,4884
fNSf	6	3,4884
fASi	2	1,1628
fAP2	1	0,5814
fAPi7	1	0,5814
fASf	1	0,5814
fFS2	1	0,5814
fpartf	1	0,5814
	172	100

Реализация {а} характеризуется максимальным разнообразием представленных аллофонов. Впрочем, об этом, как об особой изменчивости именно фонемы /а/ в зависимости от позиционных и комбинаторных условий, неоднократно писали фонетисты (см., напр.: Бондарко 1998; Кузнецов 1997; и др.). Обратим внимание на самые частотные аллофоны /а/.

Необходимо сразу оговориться, что в нашей транскрипции имеется в виду под тем или иным знаком МФА.

При транскрибировании самыми частыми обозначениями являлись следующие знаки для аллофонов /а/:

[а] — гласный среднего ряда нижнего подъема, более или менее соответствующий изолированному произнесению /а/ в русском языке. Этот

знак в МФА используется для обозначения переднего гласного, однако мы выбрали именно его для гласного среднего ряда (а не[e]) из соображений максимального удобства при транскрибировании самого частотного русского аллофона фонемы /а/.

[а] — гласный заднего ряда, нижнего подъема,

[Λ] — гласный заднего ряда средне-нижнего подъема, отличающийся от [а] не столько подъёмом, сколько краткостью, и именно из-за краткости реализующийся как менее продвинутый назад гласный,

[æ] — гласный переднего ряда средне-нижнего подъема.

Остальные знаки совпадают по значениям с теми описаниями, которые имеются в литературе и в Интернете относительно МФА.

В табл. 4 представлено процентное соотношение различных аллофонов, реализующих финальное {а}.

Таблица 4. Аллофоны, реализующие финальное {а}

а	45	ө	2
э	33	aj	1
-	17	е	1
Λ	16	u	1
ɑ	13	ɑ	1
æ	11	е	1
з	6	ã	1
ɑ̣	5	əj	1
в	4	ı	1
i	3	õ	1
ə̣	3	oe	1
ɔ	3	æa	1
			172

Видно, что чаще всего морфема {а} реализуется в нашем материале как [а] или как [э] (45 и 33 употребления соответственно).

Таблица 5 показывает, каким образом конечное [а] реализуется в разных флексиях {а}.

Таблица 5. От реализации — к грамматическому значению

а	Λ	æ
fpf 8	fpf 11	fNS1 7
fGS2 2	fNS1 5	fNSf 6
fNS1 2	fNS2 1	fGS2 4
fAS2 1	fNSm 1	fGS4 2
fNS2 1	fpartn 1	fAPi2 1
		fAS4 1
		fLS2 1
		fNP1 1
		fNPp 1
		fNSn 1
		fpf 1

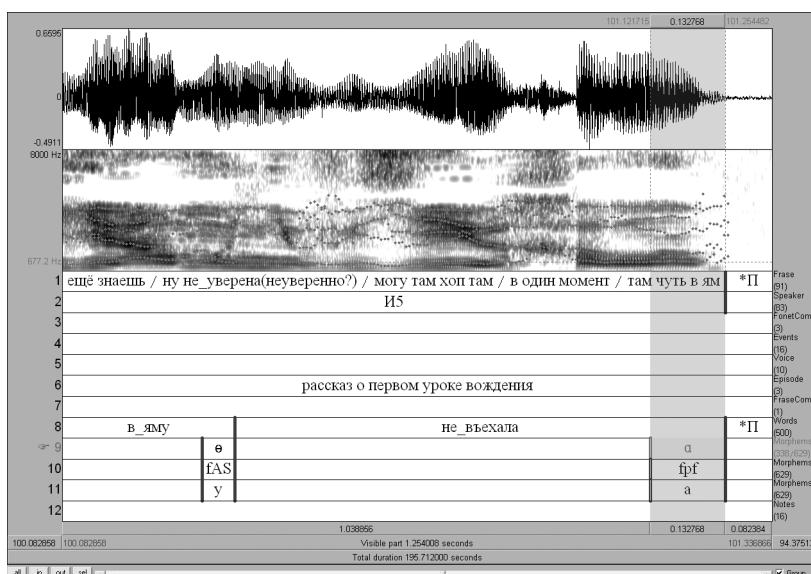


Рис. 2. Осциллограмма, спектрограмма, многоуровневая разметка словоформы *не въехала* из фразы *...там чуть в яму не въехала // *П*

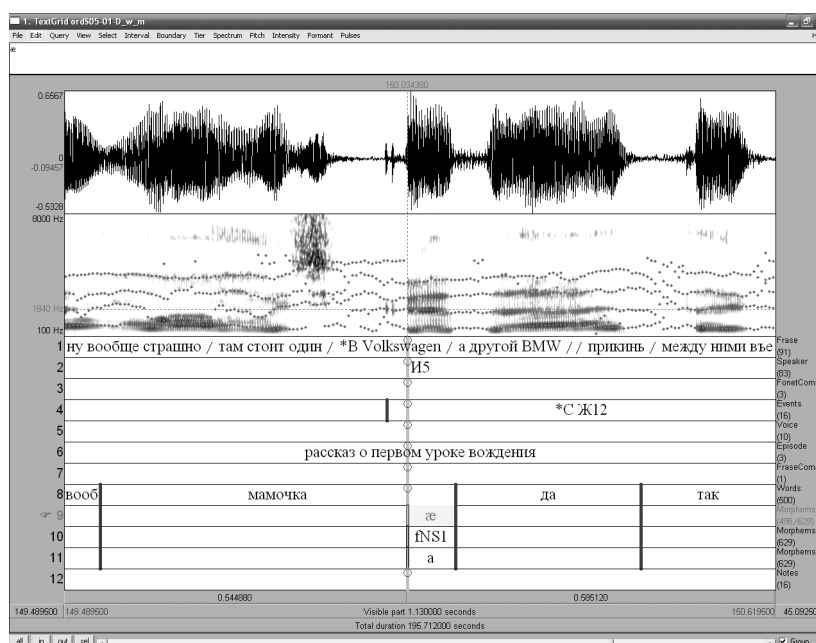


Рис. 3. Осциллограмма, спектрограмма, многоуровневая разметка словоформы *мамочка* из фразы *вообще мамочка! да / так*

Обращает внимание на себя тот факт, что как задний гласный [a] или [ʌ] реализуется чаще всего морфема со значением «прошедшее время глагола, ж. р.», а аллофон [æ] практически не встречается в морфеме с этим значением. Причем акустическое и перцептивное различие между [a] и [æ] оказалось весьма существенным. На рис. 2 и 3 можно видеть осциллографическое и спектральное представления аллофонов гласного /a/ из флексий в словах *въехала* и *мамочка*.

Различие в спектрах очевидно: средние значения формант4 [a] из *въехала* F1 — 700 гц, F2 —

1200 гц, [æ] из *мамочка* — F1 — 600 гц, F2 — 1700 гц. В результате при изолированном прослушивании этих гласных они воспринимаются как два абсолютно непохожих звука. Однако это различие не связано с противопоставлением разных морфем {a}, оно имеет чисто фонетическую природу. Задний характер гласного во флексиях глаголов прошедшего времени вызван исключительно соседством с сильно веляризованным согласным /l/. Неожиданный после заднеязычного согласного передний характер гласного [æ] во флексии слова *мамочка* можно объяснить мягким [ç] перед [k] и последующим, произнесенным без паузы, *да* с переднеязычным согласным.

4 Точечные линии на рисунках — значения формант.

О других аллофонах гласного /а/ никак нельзя сказать, что они могут быть привязаны к какой-то определенной грамматической форме — см. табл. 6 (еще раз подчеркнем, что и со значением прошедшего времени аллофоны [a] и [ʌ] соотносятся лишь через суффикс [л]).

Таблица 6. От реализации — к грамматическому значению

[a]	æ	ə	полная редукция				
fNS1	13	fNS1	7	fNS1	14	fNS1	9
fpf	13	fNSf	6	fNSm	9	fpf	5
fNSf	8	fGS2	4	fpf	9	fAS1	4
fGS2	7	fGS4	2	fGS2	6	fNSm	4
fNP2	3	fAPi2	1	fpp	5	fpr1s	4
fNS2	3	fAS4	1	fGS1	4	fGS1	3
fNSn	3	fLS2	1	fASn	3	fNP2	2
fpn	3	fNP1	1	fGSf	3	fpn	2
fAS2	2	fNPp	1	fNSf	3	fpp	2
fAP2	1	fNSn	1	fNSn	3	fpr2s	2
fASf	1	fpf	1	fpn	3	fAPi1	1
fASi2	1			fASf	2	fASf	1
fASim	1			fASi2	2	fASi2	1
fDS2	1			fLSf	2	fGS3	1
fpartn	1			fNP2	2	fLS1	1
				fpartn	2	fLS2	1
				fAPi2	1	fNS2	1
				fAS1	1	fNSn	1
				fAS2	1	fpartf	1
				fGSm	1	fpr3s	1
				fIS1	1		
				fISf	1		
				fLS1	1		
				fNPp	1		
				fNS2	1		
				fpr3p	1		

Во всех случаях максимально частотное появление аллофонов /а/ — в словоформах существительных ж. рода Им. п. ед. ч. (fNS1), что связано с общей частотностью этой формы в речи наших информантов.

2. Типы и реализации флексий, представленных орфографическими {и} и {е}

Таблица 7 показывает, что чаще всего одиночная гласная {е} представляет окончание пр. п. существительных ед. ч. 1-ого или 2-ого склонения (30 из 36 встретившихся в материале {е}). Одиночное {и} чаще всего встречается в морфемах, имеющих значение «прошедшее время глагола, мн. ч.», «им. п. существительных, мн. ч.» и «род. п. существительных, ед. ч., 1-ое склонение».

Таблица 7. От орфографии — к грамматическому значению

е		и	
fLS2	16	fpp	35
fLS1	14	fNP2	17
fDS1	3	fGS1	15
fGS1	1	fNP1	6
fNS2	1	fNPp	5
fASi	1	fAPi7	3
		fGS3	3
		fLS3	3
		fAPi	1
		fAPi3	1
		fAPip	1
		fASi	1
		fLS1	1

Вопрос о реализации этих морфем в безударном положении тесно связан с проблемой ээкающего или иэкающего произношения в современной русской речи, которая активно обсуждалась в русистике последних десятилетий. Можно сказать, что точку в дискуссии поставили исследования, проведенные под руководством Л. А. Вербицкой, в которых была проанализирована речь 150 ленинградцев. Экспериментальные данные, полученные в ходе инструментального анализа убедительно показали, что ведущим в конце XX века стал иэкающий вариант нормы (Вербицкая 1976).

Нисколько не оспаривая этого факта (в нашем распоряжении пока лишь описание 880 флексивных морфем из речи 4-х дикторов), мы вынуждены признать, что наш материал предоставляет и другие варианты реализации гласных на месте как орфографического {е}, так и орфографического {и} (см. табл. 8).

Таблица 8. Реализации гласных на месте орфографических {е} и {и}

Реализация Е	Абс. кол-во	%	Реализация И	Абс. кол-во	%
e	13	35,1	i	40	43,5
i	6	16,2	e	15	16,3
ε	4	10,8	ɪ	10	10,8
–	3	8,1	ə	7	7,6
y	2	5,4	ɨ	5	5,4
ɪ	1	2,7	—	4	4,3
ɨ	1	2,7	ɨ	2	2,2
ĩ	1	2,7	ĩ	2	2,2
iaɛ	1	2,7	ea	1	1,1
jo	1	2,7	i:	1	1,1
ə	1	2,7	ш	1	1,1
з	1	2,7	y	1	1,1
ø	1	2,7	e	1	1,1
æ	1	2,7	oe	1	1,1
	37	100	æ	1	1,1
				92	100

Причем {е} как [e] наблюдалась не только в ударном слове (что естественно), но и в безударных окончаниях.

Реализация {и} как [i] преобладает во флекси-ях (43,5 %), однако и здесь возможны реализации более открытого гласного — [e] (16 %) и др. Это наводит на мысль, что закрытые передние гласные вследствие вялой артикуляции в конечной безударной позиции отодвигаются чуть назад и чуть вниз, что и дает восприятие гласного как *e*-образного. Несомненно, подготовленный к настоящему времени материал не дает возможности сказать, насколько это явление распространено в современной спонтанной речи. Возможно, это характерно лишь для некоторых наших информантов.

3. К постановке вопроса о реализации финалей, представленных некоторыми орфографическими двухбуквенными сочетаниями

Речь в данном случае идет главным образом о тех явлениях, о которых писала Л. В. Бондарко: «в словах *тихий, серый* нет никаких следов конечного /j/, вообще неустойчивость среднеязычного сонанта проявляется в заударных комплексах, где он оказывается между гласными. В этих случаях единственным следом сонанта оказывается передний характер гласного, который следует за ним» (Бондарко 1998: 266). Именно этот тезис можно подтвердить, наблюдая материал, обработанный к настоящему времени.

Из 46 финалей {ый}/{ий} в различных формах /j/ был реализован лишь в одном случае (см. табл. 9)!

Иначе обстоит дело с окончанием {ой}. В табл. 10 можно видеть, что чаще всего эта финаль встречается в окончаниях имен прилагательных (Им. п., м. род, ед. ч.). Как известно, это окончание всегда ударно. Думается, именно поэтому почти в 30 % случаев конечное буквосочетание {ой} реализовано как [oj].

Что касается заударных морфных комплексов, где среднеязычный сонант оказывается между гласными, то, как показывает наш материал (табл.11), исчезает не только этот сонант, но и само сочетание гласных в подавляющем большинстве случаев стягивается до одного однородного аллофона (ср.: Овчаренко 1988:9).

Таблица 9. От орфографии — к реализации

ий		ый		ой	
i	8	i	11	oj	20
i̇	3	ə	7	ə	8
ɪ	1	—	4	ø	7
Всего:	12	ɨ̇	3	o	6
		i̇	2	ɔ	4
		a	1	i	3
		i	1	əj	3
		ɪ	1	ij	2
		ij	1	oe	2
		y	1	e	1
		ʌ	1	ɪ	1
		z	1	oi	1
		Всего:	34	õ	1
				u	1
				uj	1
				ɜ	1
				ø	1
				Всего:	63

Таблица 10. От орфографии — к грамматическому значению

ой	
fNSm	36
fASim	7
fGSf	6
fLSf	5
fIS1	4
fISf	3
fDSf	2

Из 99 финалей только в 28 реализовано бифонемное сочетание. Думается, они приходится на ударные окончания, что предстоит еще проверить.

В целом можно сказать, что корпусный подход к накоплению, систематизации и — что особенно важно — к анализу материала русской звучащей речи позволяет проверить многие сделанные ранее наблюдения фонетистов, подтвердить или опровергнуть многие научные гипотезы, ответить на многие поставленные ранее вопросы и, может быть, поставить новые, на которые предстоит искать ответам новым поколениям лингвистов.

Таблица 11. От орфографии — к реализации

ая	ие	ия	ое	ую	ые
æ 6	е 10	æ 2	оə 5	и 2	і 3
а 3	і 5	і 1	ə 4	иј 1	іə 3
аə 3	г 3	іə 1	а 2	иу 1	і-ə 3
аз 2	æ 2	іæ 1	з 2	и 1	іе 1
ə 2	ѣ 1	і̃ 1	– 1	и 1	із 1
æе 2	іə 1	Всего 6	о 1	иə 1	ə 1
ае 1	Всего 22		оə 1	Всего 7	ε 1
е 1			оі 1		Всего 13
іæ 1			ој 1		
иј 1			оу 1		
а-ə 1			оз 1		
а̣ 1			оæ 1		
əə 1			ε 1		
ѣ 1			Всего 22		
з 1					
ө 1					
æј 1					
Всего 29					

Литература

1. Бондарко Л. В. Фонетика современного русского языка. СПб., 1998.
2. Вербицкая Л. А. Русская орфоэпия: К проблеме экспериментально-фонетического исследования особенностей современной произносительной нормы. Л., 1976.
3. Кузнецов В. И. Вокализм связной речи. СПб., 1997.
4. Овчаренко Е. Б. Реализация и восприятие заударных морфных комплексов и функциональная нагрузка морфем. // Автореф. канд. дисс. Л., 1988.
5. Шерстинова Т. Ю., Рыко А. И., Степанова С. Б. Система аннотирования в звуковом корпусе русского языка «Один речевой день». // Формальные методы анализа русской речи. Материалы секции 18-ой Международной конференции. СПбГУ, 2009. С. 66–75.