

А. М. АКМАРОВ (Ижевск)

НЕКОТОРЫЕ ФИЗИЧЕСКИЕ ОСОБЕННОСТИ ГЛАСНЫХ УДМУРТСКОГО ЯЗЫКА

Методика исследования

Для более полного и глубокого знания звукового состава языка необходимо изучить как физиологическую сторону звуков речи, место и способы их образования, так и их физические особенности, высотные, словесые, темпоральные и спектральные характеристики.

В данной работе рассматриваются два физических признака удмуртских гласных — спектральная характеристика и долгота. При спектральном анализе важно было определить формантную структуру гласных и выяснить зависимость ее от позиции гласного. Выделено пять позиций: 1) гласный в изолированном положении (Г): *a, e, i, o, e, u, i*; 2) в позиции открытого ударного слога после согласного *p* (СГ): *apa* 'тетя', *tupe* 'мой мяч', *ṭṣežpi* 'утенок', *gopo* 'неровный', *pipi* 'осина', *tipi* 'дуб'; 3) в закрытом ударном слоге между двумя *p* (СГС): *rap, per, pip, por, per, pir, pip*; 4) в закрытом ударном слоге между некоторыми палатальными согласными (С'ГС'): *ḥaḥ* 'хлеб', *džetš* 'добрый', *šiš* 'гнилой', *šeš* 'хищник', *šuš* 'воск', *ḥil* 'четыре' и *fat, tet, tit, teṭ, tot, tut, tiṭ*; 5) в позиции предударного слога (СГ): *pipi* 'осина', *gopo, tupe, tipi*; (ГС): *apa*; (ГС'): *eṭe* 'брат'; (СГС): *ṭṣežpi*.

Данный материал был записан на магнитную ленту в произнесении двух дикторов: Семена Константиновича Бушмакина (Б.) и Юлии Васильевны Симаковой (С.). Оба диктора — носители средних говоров удмуртского языка — якшур-бодьинского (дер. Суровай) и нылгинского (дер. Мельниково). Произношение их отвечает орфоэпическим нормам литературного языка.

Спектральный анализ гласных произведен на динамическом спектрографе (сонаграфе) типа Кенига, Данна и Лейси в одном из научно-исследовательских институтов Москвы. Диапазон спектрографа — от 85 до 8000 гц. Наличие двух фильтров — широкополосного на 300 гц и узкополосного на 45 гц — дает возможность получить два вида сонаграмм для анализа.¹

В работе использованы три вида визуализации: широкополосные и узкополосные сонаграммы и разрезы, высший уровень частотной шкалы которых 4500 гц. В этом диапазоне обычно лежат те форманты, частота которых служит ключом для характеристики спектров гласных (и некоторых согласных).

¹ Принцип работы сонаграфа подробно описан в работе: М. А. Сапожков, Речевой сигнал в кибернетике и связи, Москва 1963.

Под формантой при этом понимается (вслед за Комиссией по акустике АН СССР) область концентрации энергии в спектре звука речи. Для обозначения формант употребляется символ F , к которому прибавляется номер форманты. Так, первые четыре форманты обозначаются F_1, F_2, F_3, F_4 , а формантные частоты — соответственно F_1, F_2, F_3, F_4 . Частота основного тона передается через F_0 .

Говоря о частоте форманты, имеется в виду частота центральной части, середины форманты по широкополосной сонаграмме.

При вычислении частоты формант удмуртских гласных были использованы сонаграммы и разрезы. По широкополосным рисункам предварительно находились формантные области и определялся характер их проявления; по узкополосным вычислялась частота основного тона; по разрезам определялись вершины спектра (полюсы) и расположение их по частоте.

Частоту полюсов можно установить по шкале, которая приводится с левой стороны или внизу каждого рисунка. Градуировка шкалы произведена из того расчета, что высота рисунков, равная 67 мм, отражает диапазон в 4500 гц. Следовательно, каждый миллиметр рисунка по вертикали равен 67,164 гц. Нечетное число 67,164 гц неудобно для пользования, поэтому на нашей шкале нанесено 15 делений по 300 гц. Надписи сделаны 900 гц, 1800 гц, 2700 гц, 3600 гц, 4500 гц.

Следует отметить один недостаток сонаграфа — трудность регулировки нулевой линии, что может привести к неточностям при вычислениях. Чтобы избежать этого, частоту полюсов лучше определять по частоте основного тона, считая, что F_0 — это первая составляющая, за которой через равные интервалы следуют другие. Если F_0 какого-либо гласного равна 160 гц, то вторая составляющая появится на частоте 320 гц, третья — на частоте 480 гц и далее через каждые 160 гц.

Для примера покажем, как определялась частота формант гласного i в позиции Г. Сначала рассмотрим широкополосную сонаграмму этого i ; на ней видны четыре области почернения, из которых три (первая, третья и четвертая) выделяются наибольшей степенью концентрации энергии: это формантные области данного гласного. Согласно разрезу в области первой форманты по интенсивности выделяется вторая гармоника, в области второй — тринадцатая и в области третьей — семнадцатая и восемнадцатая гармоники. Указанные гармоники принимаем за вершины спектра, частота их является частотой форманты. Если высота основного тона голоса при произнесении этого i равна 210 гц, то частота второй гармоники будет $210 \times 2 = 420$ гц, тринадцатой — $210 \times 13 = 2730$ гц, семнадцатой — $210 \times 17 = 3570$ гц. Таким образом, форманты гласного i в изолированном положении в произнесении диктора С. проявляются на частотах: 420 гц, 2730 гц, 3600 гц (середина третьей форманты лежит между 17-й и 18-й гармониками — выше, чем 3570 гц).²

Для определения частоты основного тона есть два способа измерения F_0 по сонаграммам. Первый способ — измерить частотный интервал между гармониками узкополосной сонаграммы или разреза. Для этого с помощью частотной шкалы определяется количество герц между какими-либо двумя гармониками. В целях получения более точного результата берут не соседние гармоники (это слишком маленький промежуток для измерения), а гармоники, отстоящие друг от друга на определенном расстоянии — через десять и более гармоник. Полученное число делится на количество частотных интервалов между точками измерения. Так, если между первой и шестнадцатой гармониками 2400 гц,

² В тех случаях, когда невозможно было определить порядковый номер гармоники, частота форманты устанавливалась просто по шкале.

это число делится на 15, т. е. на количество интервалов, и определяется искомая величина F_0 , равная 160 гц.

По второму способу F_0 определяется как величина, соответствующая числу импульсов, создаваемых голосовыми связками в единицу времени:

$$F_0 = \frac{1}{T_0}$$
 где T_0 — расстояние между периодами основного тона на шкале времени. Начало периода, как было изложено выше, легко определить по вертикальным штрихам на широкополосной сонаграмме. Если известно, что каждый миллиметр шкалы времени равняется 7 мсек, то нетрудно определить время каждого периода, но в этом нет необходимости, ибо можно оперировать более крупными единицами времени, скажем, 100 мсек, что соответствует 14 мм рисунка вдоль шкалы времени. При произнесении *a* в слове *ñai* 'хлеб' за 100 мсек отмечается 16 импульсов голосовых связок. За одну секунду таких импульсов будет 160. Следовательно, $F_0 = 160$ гц.

В данном эксперименте основной тон голоса измерялся по первому способу, второй способ применялся лишь в тех случаях, когда возникла необходимость проверить результаты, полученные первым методом.

Формантная структура удмуртских гласных

Результаты анализа сонаграмм, приведенные в табл. 1, дают возможность сделать некоторые выводы по формантной структуре удмуртских гласных.

Так, гласный между палатальными согласными характеризуется увеличением частоты второй форманты и уменьшением частоты первой. В артикуляционных терминах это означает, что гласные в позиции $C'G'S'$ являются более закрытыми и более переднего образования.

Несколько менее устойчивая закономерность выявляется при сравнении формантных структур гласных в позициях G и $C'G'$. Мы наблюдали тенденцию произносить гласные в открытом ударном слоге несколько более открыто, чем, скажем, изолированные гласные. Об этом говорит увеличение частоты F_1 в позиции $C'G'$ для *e*, *a*, *o*, *u*. Это замечание не относится к гласным *i* и *i*. Здесь следует еще проверить, не являются ли последние в позиции $C'G'$, наоборот, более закрытыми.

Относительно формантной структуры гласных в закрытом ударном слоге между двумя *p* можно сказать, что в одних случаях она ближе к формантной структуре гласных в позиции G , а в других — $C'G'$. А при знакомстве с формантной структурой гласных предударного слога бросается в глаза изменение частоты основного тона голоса: у диктора Б. F_0 предударных гласных всегда меньше, чем F_0 ударных гласных (например, в позиции $C'G'$), а у диктора С. — наоборот, выше. Естественно отсюда несколько высокое местонахождение второй, самой чувствительной форманты предударных гласных в произнесении диктора С.

1. Гласный *e*. Формантные области данного гласного распределяются относительно равномерно по всему спектру в заданном диапазоне частот. На широкополосной сонаграмме полосы затемнения располагаются одна над другой через приблизительно равные промежутки: для мужского голоса первая на уровне 450—500 гц, вторая — 1450—1500 гц, третья — 2500—2550 гц, четвертая, если она проявляется, — на уровне 3500—4000 гц. Из четырех областей концентрации энергии наиболее интенсивны, как правило, первые две, а то и три (для женского голоса), интенсивность других формант значительно меньше.

Таблица 1

	ε		i		a		o		u		i		e	
	Б.	С.	Б.	С.	Б.	С.	Б.	С.	Б.	С.	Б.	С.	Б.	С.
Г														
F ₀	150	207	150	200	150	187	150	200	150	207	160	210	145	200
F ₁	450	620	350	400	750	1112	450	600	450	414	320	420	580	600
F ₂	1450	1635	1500	1600	1350	1496	950	1000	750	1245	2500	2730	2030	2400
F ₃	2550	3150	2400	3040	2150	2805	2100	2200	—	—	3100	3600	2610	3040
F ₄	4100	—	4400	4050	3700	4000	—	—	—	—	—	—	—	—
СГ														
F ₀	160	207	171	190	160	178	163	200	150	180	161	187	160	190
F ₁	500	600	342	380	1280	1112	525	600	510	400	320	374	580	570
F ₂	1450	1656	1539	1900	2150	1683	1110	1000	1400	1260	2400	2800	1920	2100
F ₃	2500	2900	2400	3040	—	2431	2200	2100	2400	2150	3100	3550	2450	3040
F ₄	3500	—	4300	4150	—	—	—	—	—	—	4320	—	—	4200
СГС														
F ₀	160	190	166	207	160	190	168	190	180	218	160	207	166	192
F ₁	500	570	332	414	1280	760	500	570	500	436	320	400	498	576
F ₂	1440	1450	1500	1656	2240	1520	900	950	1260	1090	2400	2890	1950	2300
F ₃	2600	2850	2500	2900	3000	3100	2100	2100	—	—	3150	3700	2500	3070
F ₄	3250	4400	4400	4200	4000	—	—	—	—	—	—	—	—	—
СГС														
F ₀	150	190	163	210	160	187	170	190	160	218	160	192	166	190
F ₁	500	570	400	420	640	748	510	570	320	436	320	380	400	400
F ₂	2000	1900	1300	2500	1600	2057	1200	1140	1280	1308	2400	2880	2100	2600
F ₃	2400	2850	2445	3900	2100	3000	2100	—	2050	2180	3000	3840	2950	3550
F ₄	4300	—	3300	—	3000	—	—	—	—	—	4500	—	—	—
СГ (или ГС, ГС и СГС)														
F ₀	125	218	140	225	140	200	140	218	140	180	140	225	—	175
F ₁	500	600	350	450	1260	1000	500	500	450	540	300	400	—	575
F ₂	1500	1980	1400	2000	2100	1600	1000	1100	1250	1080	2400	2700	—	2600
F ₃	2450	—	2500	2950	—	2400	—	2200	—	1980	3000	—	—	3600
F ₄	—	—	4300	4100	—	—	—	2900	—	—	—	—	—	—
Усредненные частоты														
F ₁	490	592	355	413	1042	946	497	568	446	445	316	394	514	544
F ₂	1568	1720	1448	1931	1888	1671	1032	1038	1188	1196	2400	2800	2000	2400
F ₃	2500	2937	2449	3076	2416	2747	2125	2160	2283	2103	3083	3672	2627	3260

Ввиду того, что формантные области ϵ расположены на шкале частот относительно равномерно, данная структура напоминает резонансы простейшей модели речевого тракта в форме однородной акустической трубы. При длине 17,5 см (средняя длина надставной трубы от голосовых связок до губ для мужчин) такая труба дает резонансы примерно на 500, 1500, 2500, 3500 и т. д. через каждые 1000 гц. А при длине 15 см, что соответствует длине речевого тракта для диктора С., первый, основной, резонанс появляется приблизительно на частоте 550 гц, второй — 1650, третий — 2750 и т. д. через каждые 1100 гц. Такое сходство дает возможность принять формантную структуру ϵ за исходную при анализе сонаграмм остальных гласных звуков удмуртского языка.

2. Гласный i . Характерной особенностью этого гласного является увеличение частотного интервала между первой и второй формантами по сравнению с исходной структурой. Первая форманта i во всех позициях находится ниже на шкале частот, чем для ϵ , в то время как вторая форманта практически не меняет своего положения, если взять среднее значение. Далее, в отличие от ϵ из четырех формантных областей наиболее интенсивна первая, степень интенсивности второй и третьей почти одинакова.

3. Гласный a . Основной спектральный признак гласного a — очень компактное расположение первых двух формант. Причем первая форманта обладает наибольшей интенсивностью (для мужского голоса). В произнесении диктора-женщины спектральная огибающая гласного в ее нижней части имеет весьма неопределенное очертание: вместо пиков на ней появляются плоские вершины, затрудняющие определение местоположения первых двух формант. Таким образом, полностью подтверждается утверждение Б. Линдблума, что «чем выше частота основного тона голоса, тем меньше информации на огибающей спектра».³

Следует также заметить, что a дает очень большое колебание в расположении первой форманты: ср. F1, равную 750 и 640 гц в позициях Г и С'ГС', с 1280 и 1260 гц в позициях СГ' и ГС.

4. Гласный o характеризуется понижением второй и третьей формант по частоте, первая форманта находится практически на том же уровне, что и в спектре исходного гласного ϵ .

На широкополосной сонаграмме очень четко проявляются первые две форманты в виде двух почти сливающихся полосок. По компактности расположения F1 и F2 гласный o напоминает a , однако у последних эти форманты расположены по частоте несколько выше.

5. Гласный u . В спектре этого гласного произносительная энергия распределяется не равномерно, ее основная часть находится на низких частотах. Особенно показательны в этом отношении сонаграммы изолированного u , где вся энергия сконцентрирована между частотами 450 и 750 гц у диктора Б. и 414 и 1242 гц у диктора С. Естественно, наибольшей амплитудой интенсивности обладает первая форманта. Сравнение с исходной структурой показывает, что все три форманты u сдвинуты по частоте вниз.

6. Гласный i отличается от ϵ и других гласных тем, что произносительная энергия в его спектре сконцентрирована больше по краям. На широкополосной сонаграмме четко выделяются три-четыре полосы почернения, одна внизу на частоте 320 гц, две рядом в верхней половине

³ B. Lindblom, Accuracy and Limitations of Sonagraph Measurements. — Proceedings of the Fourth International Congress of Phonetic Sciences, The Hague 1962, стр. 188—202.

на частотах 2400 и 3100 гц, а четвертая захватывает верхний край сонаграммы в области 4320 гц. Для женского голоса три форманты *i* в этой позиции СГ лежат на уровнях 374, 2800 и 3550 гц. При этом мы не рассматриваем область концентрации энергии на частоте 1200 гц как форманту ввиду ее малой амплитуды интенсивности и ввиду того, что она отмечается только в двух случаях: на сонаграмме изолированного *i* у диктора С. и на сонаграмме *i* в слове *tšežpi* у диктора Б. Кроме того, подобное явление встречается на сонаграммах других гласных, например, *e* в *pe* и *tšežpi* и *i* слога *fiť* в произнесении диктора С.

Анализ сонаграмм и разрезов показывает далее, что *i* имеет ярко выраженную топологию спектральных вершин, причем вершины, расположенные в области высоких частот, имеют большую амплитуду интенсивности, лишь немногим меньше амплитуды первой форманты.

7. Гласный *e*. Ему присуще наличие одной форманты на низких частотах, двух или трех — на высоких. От исходного *e* этот гласный отличается несколько большим расстоянием между первой и второй формантами, с одной стороны, и уменьшением расстояния между второй и третьей, с другой. Иными словами тем, что формантные области *e* расположены почти как у гласного *i*. Так, широкополосные сонаграммы *e* в ряде позиций очень напоминают соответствующие сонаграммы гласного *i*, разница лишь в уровне расположения F1, F2, F3. В спектре *i* F1, с одной стороны, и F2 и F3, с другой, как бы отталкиваются друг от друга, тогда как у второго гласного они приближаются друг к другу.

Длительность удмуртских гласных

Долгота или краткость удмуртских гласных не является релевантным признаком соответствующих фонем, но, как выясняется, каждый аллофон характеризуется определенной длительностью произнесения и для правильного распознавания его в слове в потоке речи совсем не безразлично, с какой длительностью он произносится.

Из данных литературы известно, что в ударном положении гласные обладают большей длительностью, чем в безударном⁴, и независимо от ударения длительность гласных в начальном и конечном положениях больше, чем в середине слова между согласными, во фразах длительность гласных еще более уменьшается.⁵ Согласно исследованиям И. В. Тараканова⁶, В. И. Лыткина и Т. И. Тепляшиной⁷, первое из вышеупомянутых положений относится также к гласным удмуртского языка. Что же касается остальных положений, то они никем не исследовались. В настоящей работе не ставилась задача специально исследовать длительность удмуртских гласных, вопрос этот имеет самостоятельную научную ценность и может служить темой отдельного исследования. Однако имевшийся экспериментальный материал дал возможность произвести анализ длительности гласных в позициях Г, СГ, СГС; С'ГС', СГ (или ГС, ГС', ГС').

⁴ См. Г. П. Торсуев, Экспериментальное исследование английского словесного ударения. — Ученые записки 1-го Московского государственного педагогического института иностранных языков, т. XVIII. Экспериментальная фонетика и психология речи, Москва 1960.

⁵ Л. А. Варшавский, И. М. Литвак, Исследование формантного состава и некоторых других физических характеристик звуков русской речи. — Проблемы физиологической акустики III, Москва—Ленинград 1956.

⁶ И. В. Тараканов, Фонетические особенности бавлинского диалекта удмуртского языка. Канд. дисс., Тарту 1958.

⁷ Грамматика современного удмуртского языка, Ижевск 1962, стр. 7—50.

Таблица 2

Позиции	Дикторы	Длительность гласного, мсек							Среднее
		е	і	а	о	и	і	е	
Г	Б.	350	350	325	325	350	325	325	298
	С.	225	250	275	250	250	300	275	
СГ	Б.	300	300	300	300	300	250	275	265
	С.	300	150	250	250	250	200	275	
СГС	Б.	200	200	250	200	200	200	200	205
	С.	200	225	200	200	200	200	200	
СГС'	Б.	225	150	200	—	200	200	225	209
	С.	225	225	200	250	225	200	200	
СГ, ГС, ГС' и СГС	Б.	125	125	125	150	112	125	—	129
	С.	125	125	150	150	125	125	125	
Средняя длительность		212	187	209	214	201	187	214	204

На основании полученных (усредненных) данных (см. табл. 2) можно сделать следующие выводы:

1. В отдельно взятых словах длительность удмуртских гласных колеблется примерно от 129 до 265 мсек.

2. В открытом ударном слоге длительность гласных в два с лишним раза больше, чем в безударном (предударном) слоге (ср. 265 и 129 мсек).

3. В конечном положении в ударном слоге гласные обладают большей длительностью, чем в середине слова между двумя согласными (ср. 265 мсек и 205—209 мсек).

Как видно, длительность удмуртских гласных зависит от ударения (находится ли гласный в ударном или безударном слоге) и от типа слога.

4. Средняя длительность удмуртских гласных в изолированных словах равна приблизительно 204 мсек. Нарушение этой количественной характеристики ведет к тому, что удмуртская речь звучит неестественно и затруднена для понимания.

А. М. АКМАРОВ (Iževsk)

SOME PHYSICAL CHARACTERISTICS OF THE VOWELS OF THE UDMURT LANGUAGE

The article deals with two physical characteristics of Udmurt vowels: formant structure and length. These two characteristics are studied in five vowel positions: V, CV, CVC, C'VC' and CV (in three cases the pretonic syllable is represented by different sound combinations such as VC, VC' and CVC).

The experimental material was read by two natives (one male and one female), recorded and analyzed by means of a sound spectrograph (sonograph). The sonagrams were interpreted according to the method described in this paper.

The results obtained are shown in Table 1 (formant structure) and Table 2 (length of the vowels).