

УДК 534.784

**К ВОПРОСУ О СТРУКТУРНЫХ ИЗМЕНЕНИЯХ РЕЧИ,
ПРОИЗНОСИМОЙ В ГЕЛИЕВО-КИСЛОРОДНОЙ СРЕДЕ**

Л. С. Бутырский, Б. И. Петленко

Оцениваются нарушения основных параметров речи, произносимой в гелиево-кислородной атмосфере при различных давлениях среды: формантной структуры, соотношений мощностей отдельных звуков, частоты основного тона. Эксперименты показали, что такая речь характеризуется значительным, приблизительно пропорциональным, увеличением формантных частот, уменьшением отношения мощностей согласный/гласный, незначительным изменением частоты и мелодии основного тона по сравнению с речевым сигналом, произносимым в обычной атмосфере.

Как известно [1, 2], разборчивость речи, произносимой в гелиево-кислородных смесях, особенно при повышенных давлениях, весьма низка и, если не принимать специальных мер, направленных на восстановление разборчивости, в этих условиях имеет место практически срыв устной связи. Сказанное иллюстрируется табл. 1 [2], где приведены результаты измерения звуковой разборчивости речи, произносимой в гелиево-кислородной смеси при различных давлениях.

Таблица 1

Дыхательная смесь и ее давление, <i>атм</i>	Воздух 1	HeO ₂				
		1	7	12	17	21
Звуковая разборчивость, <i>D</i> , %	90	87	61	53	38	21

Нарушений характера артикуляции дикторов в этих условиях практически замечено не было, за исключением тенденции к замедлению темпа речи и растягиванию гласных звуков. Поэтому прогрессирующее падение разборчивости с увеличением давления и концентрации гелия в дыхательной смеси можно связать главным образом с изменениями структуры речевых сигналов, вызванными сменой газовых условий. Ввиду этого определение характеристик речи в гелиево-кислородных средах является необходимым для построения эффективных систем речевой коммуникации при использовании гелиевых дыхательных смесей.

Однако исследование характеристик и свойств речевых сигналов в газовых условиях, отличных от обычных, диктуется не только утилитарными соображениями создания эффективных систем устной связи в подобных ситуациях. Изучение такой речи полезно и потому, что на ее примере возможна проверка и уточнение основных положений как теории речеобразования, так и теории слухового восприятия речевых сигналов. Совершенствование этих теорий в свою очередь позволит получить лучшие решения проблем компрессии, повышения разборчивости и автоматического распознавания речевых сигналов в любых условиях.

Исследования структуры речевого сигнала, произносимого в газовой среде, отличной от воздушной, преимущественно в гелиево-кислородной, производились рядом исследователей [1], причем было обнаружено, что в условиях измененной газовой атмосферы, речь претерпевает заметные изменения. Однако результаты были получены для сравнительно небольших диапазонов изменения давления и концентрации компонентов дыхательной смеси, а по ряду данных относительно структуры речевого сигнала в гелиевой среде имеются противоречия. Последнее, в частности, относится к изменению частоты основного тона голосовых звуков при повышенном давлении.

Целью настоящей работы и явилась оценка нарушений изменений основных параметров речи в гелиевой атмосфере при более широком диапазоне изменения давления. Количество гелия и кислорода в используемых дыхательных смесях поддерживалось таким, что в большинстве случаев парциальное давление кислорода составляло примерно 1,3 атм. Акустические условия речеобразования и регистрации сигналов были всегда идентичными.

В формировании формантной структуры речи ведущая роль отводится речеобразующему тракту, представляющему собой трубку переменного сечения, начинающуюся у голосовых связок и заканчивающуюся у отверстия рта, конфигурация которой изменяется в зависимости от произносимого звука. С некоторыми допущениями трубка переменного сечения может быть аппроксимирована последовательно соединенными цилиндрическими звеньями однородного сечения, у которых эквивалентные распределенные акустические величины на единицу длины определяются выражениями [3]:

$$L_a = \frac{\rho}{A}; \quad C_a = \frac{A}{\rho c^2}; \quad R_a = \frac{S}{A^2} \sqrt{\frac{\omega \rho \mu}{2}}; \quad G_a = S \frac{\eta - 1}{\rho c^2} \sqrt{\frac{\lambda \omega}{2 C_p \rho}}, \quad (1)$$

где A — площадь сечения единицы длины тракта, S — периметр сечения, ρ — плотность газовой среды, c — скорость распространения звука в газовой среде, μ — коэффициент вязкости смеси, λ — коэффициент ее теплопроводности, η — адиабатическая постоянная среды, C_p — удельная теплоемкость при постоянном давлении. Для случая малых потерь ($R_a \ll \omega L_a$ и $G_a \ll \omega C_a$) фазовая постоянная β и постоянная затухания α определяются таким образом:

$$\beta \approx \omega \sqrt{L_a C_a} = \frac{\omega}{c}; \quad (2)$$

$$\alpha = \frac{R_a}{2} \sqrt{\frac{C_a}{L_a}} + \frac{G_a}{2} \sqrt{\frac{L_a}{C_a}} = \frac{S}{2\sqrt{2}A} \left[\sqrt{\frac{\omega \mu}{\rho c^2}} + (\eta - 1) \sqrt{\frac{\omega \lambda}{C_p \rho c^2}} \right], \quad (3)$$

а характеристическое сопротивление $Z_0 = \sqrt{L_a / C_a} = \rho c / A$.

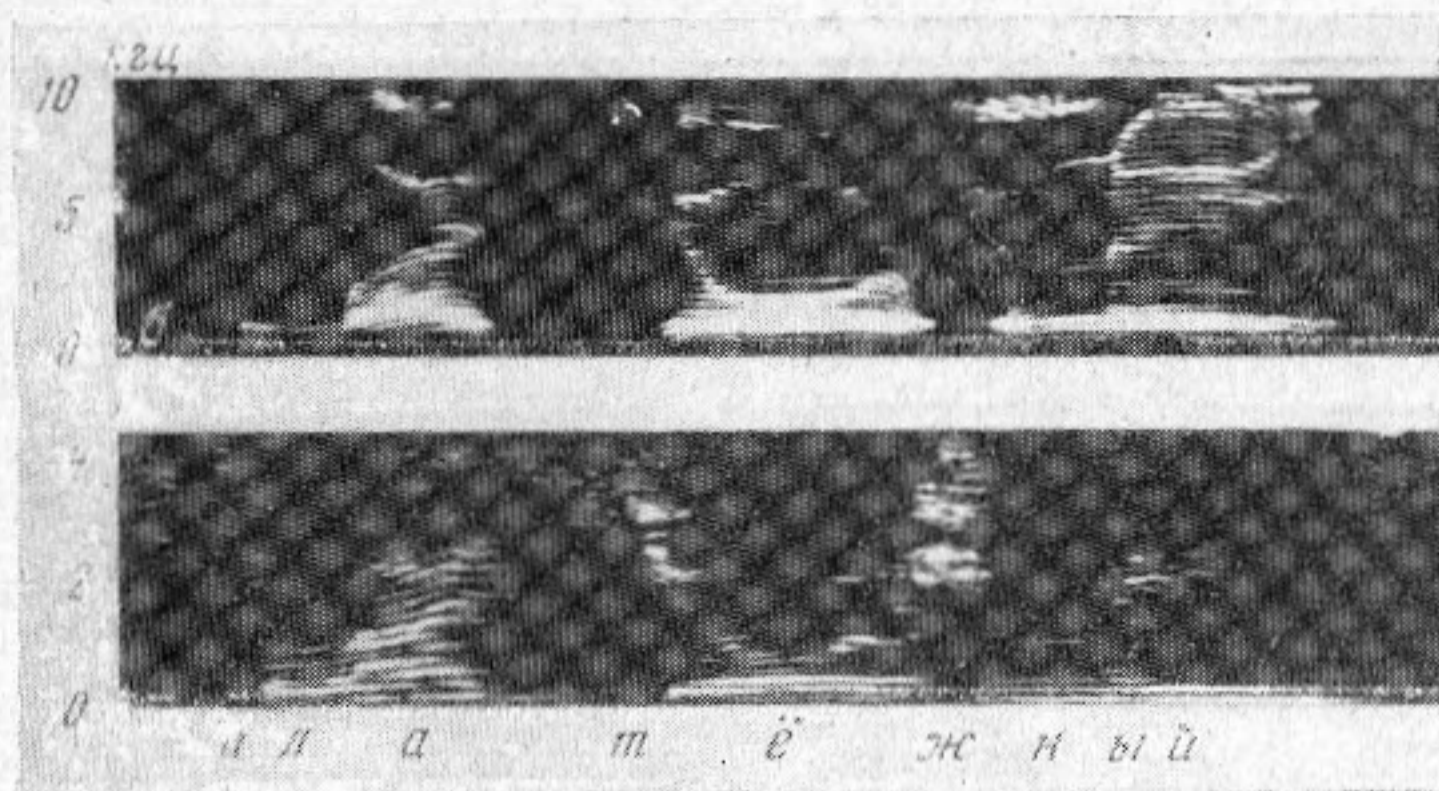
Некоторые сравнительные характеристики воздуха и гелия при давлении 1 атм и температуре 37° даны в табл. 2.

Учитывая данные табл. 2, можно отметить, что наиболее изменчивыми характеристиками среды, которыми определяются и изменения акустических параметров речеобразующего тракта, являются ρ , c , λ . Так как в пер-

Таблица 2

Среда	ρ , кг/м ³	c , м/сек	$\mu \cdot 10^{-7}$, нсек/м ²	$\lambda \cdot 10^{-3}$, вт/мград	C_p , м ² /сек ² град	η
Воздух	1,14	350	1,85	23	1000	1,4
Гелий	0,16	1000	2,02	135	800	1,66

вом приближении мы считаем, что смена состава дыхательной смеси не изменяет характера движений органов речи*, то геометрические размеры речеобразующего тракта в соотносимые моменты времени одинаковы в различных дыхательных средах. Обозначим относительные значения плотности и скорости звука в гелиевой смеси по сравнению с воздухом как $\rho_* = \rho_c / \rho_v$, $c_* = C_c / c_v$, где ρ_c , c_c , ρ_v , c_v — плотности среды и скорости звука



Фиг. 1

в ней, соответствующие гелиевой смеси и воздуху. Тогда частоты резонансов единичных отрезков тракта, определяемые как $(L_a C_a)^{-1/2}$, изменятся в гелиево-кислородной среде в c_* раз. Резонансы речевого тракта однозначно определяют формантную структуру речевых сигналов, поэтому должно иметь место пропорциональное смещение всех деталей формантной огибающей спектра по шкале частот также в c_* раз. Этот вывод, однако, является лишь приблизительным, так как мы пренебрегли влиянием на частоты формант изменения импеданса излучения, сопротивления голосового источника и характера колебаний стенок речеобразующего тракта.

Как известно [3], в воздушной среде влияние импеданса излучения на резонансы речеобразующего тракта заключается в незначительном (единицы процентов) их понижении по частоте, а наличие импеданса голосовой щели приводит, наоборот, к увеличению формантных частот, еще меньшему в процентном отношении, причем это изменение относится главным образом к низкой первой форманте.

Влияние на резонансы голосового тракта упругости его стенок учитывается введением в эквивалентную схему дополнительной шунтирующей проводимости $G_w = 1/R_w$ и индуктивности L_w . Собственные механические резонансы, обусловленные упругостью тканей, лежат ниже обычного диапазона частот речевых сигналов и поэтому $\omega C_a \gg 1/\omega L_w$. Как следует из приближенного уравнения для фазовой постоянной

$$\beta' \approx \omega \sqrt{L_a (C_a - 1/\omega^2 L_w)} = \frac{\omega}{c'} \quad (c' > c), \quad (4)$$

влияние вибраций стенок полостей состоит в некотором уменьшении β' или эквивалентном увеличении скорости c' распространения звука в тракте с мягкими стенками по сравнению со скоростью распространения звука c в свободном пространстве. Из формулы (4) следует, что эффект максимален на низких частотах и резонансные частоты смещаются вверх. Оцен-

* Об этом можно косвенно судить, например, на основании оценок качества звучания и разборчивости фонограммы гелиевой речи, скорректированной путем замедленного воспроизведения.

ка этого смещения затруднительна; вероятно оно не превышает при воздушной среде нескольких процентов даже для области низкой первой форманты.

Если предположить, что при замене воздуха гелиевой смесью относительное влияние импеданса голосовой щели и сопротивления излучения на

Таблица 3

Давление дыхательной смеси, атм	Форманты	Частота, гц	Коэффициент сдвига формант $c'_+ = F'n_c / F'n_b$
1 (воздух)	F ₁	420	1,00
	F ₂	2230	1,00
	F ₃	2860	1,00
	F ₄	3815	1,00
1 (HeO ₂)	F ₁	620	1,48
	F ₂	3160	1,41
	F ₃	3980	1,39
	F ₄	5350	1,40
	Среднее		1,42
21 (HeO ₂)	F ₁	1030	2,45
	F ₂	5350	2,39
	F ₃	6600	2,31
	F ₄	8800	2,31
	Среднее		2,36

резонансные частоты останется прежним, то все же можно считать, что в гелиевой среде произойдет пропорциональное смещение деталей спектральной огибающей по шкале частот. В противном случае можно ожидать и нелинейного сдвига различных формант по частотному диапазону.

Форма огибающей спектра речи определяется не только резонансными частотами, но и потерями энергии в речеобразующем тракте, которые определяют ширину формант. Эти потери обусловлены вязкостью среды и теплопроводностью стенок речевого тракта; кроме того, определенная часть энергии

излучается и теряется на сопротивлении голосовой щели. Теоретический расчет потерь для нейтрального гласного Э [4] показывает, что в гелиевой атмосфере потери энергии растут. Это приводит к трех — пятикратному расчетному увеличению ширины форманты при давлении гелиево-кислородной смеси 21 атм. Следовательно, есть основания ожидать увеличения ширины формант и в реальной «гелиевой» речи.

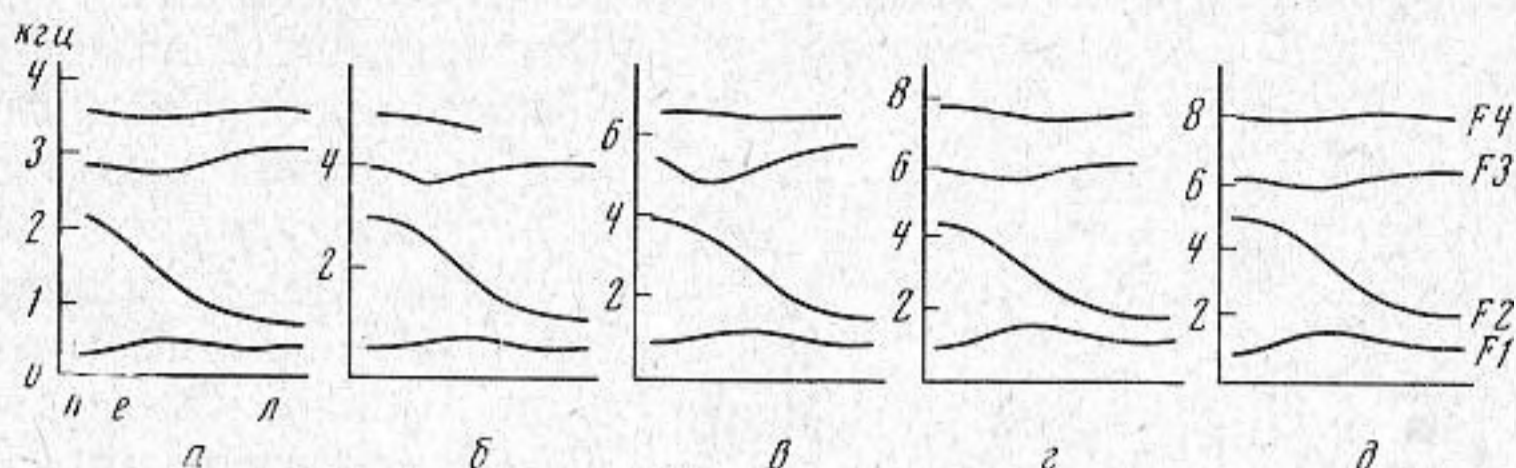
Для проверки соответствия предполагаемых искажений речи в гелиево-кислородной среде изменениям, имеющим место в действительности, речевой материал четырех дикторов, работавших в гелиево-кислородных смесях, был подвергнут детальному спектрографическому исследованию. Длительность совокупности речевых отрезков составила несколько десятков минут. Изоморфизм изображений «видимой речи» одних и тех же слов, произносимых в различных газовых условиях, позволяет достаточно уверенно сопоставлять спектральные сечения и частоты формант, относящиеся к сходным фонетическим сегментам. Сказанное иллюстрируется фиг. 1, где представлена видеодиаграмма слова «платежный», произнесенного в воздухе при 1 атм и в гелиевой атмосфере при 21 атм (а и б соответственно).

Измерение частот одноименных формант при различных условиях речеобразования позволило вычислить коэффициент их сдвига по частотной шкале. Пример определения коэффициента смещения формант по частоте для четырех формант спектрального сечения сегмента, соотносимого со звуком «Е» в слове «невода», дан в табл. 3. При этом точность вычисления коэффициентов сдвига ограничена степенью неизменности артикуляционной деятельности диктора во всех случаях, погрешностями выбора сечений в сходных звуковых сегментах и погрешностями определения положения форманты.

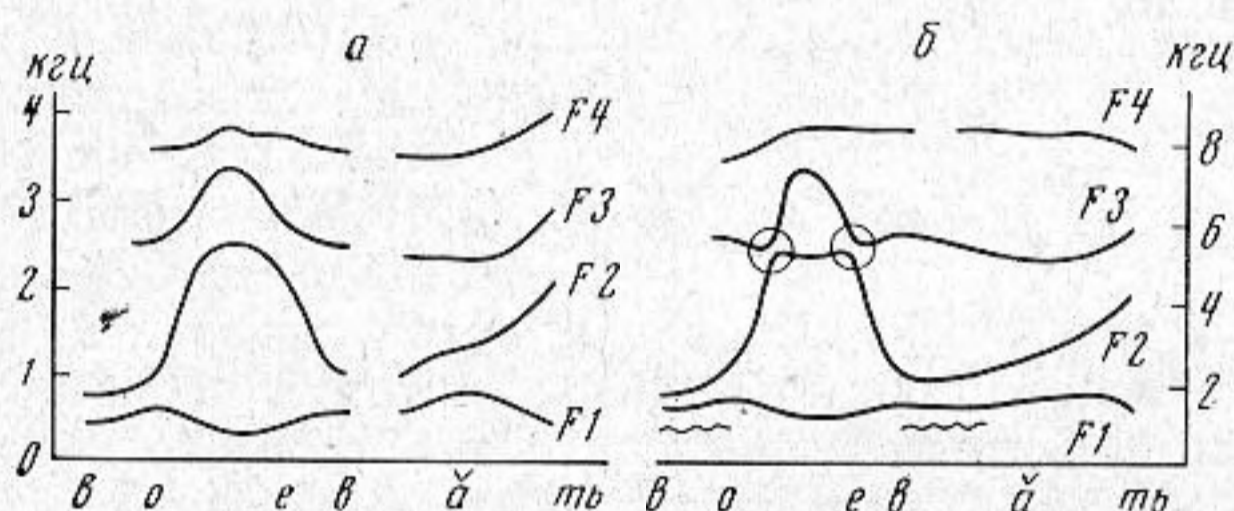
Средний коэффициент сдвига формант определялся и по другой методике, согласно которой его значение c'' выбиралось таким, чтобы при пропорциональном сжатии шкалы частот на видеодиаграмме суммарное отклонение траекторий всех формант целого слова, произнесенного в гелиево-кислородной смеси, от соответственных траекторий в воздухе достигало минимума. Пример такой пропорциональной частотной нормализации схемати-

зированных видеogramм слога «пел» в различных условиях гелиевой атмосферы дан на фиг. 2.

На этой фигуре изображение *a* соответствует воздуху при давлении 1 атм, изображения *b—d* гелиево-кислородным смесям с параметрами $P = 1 \text{ атм}$, $c'' = 1,4$; $P = 7 \text{ атм}$, $c'' = 1,7$; $P = 12 \text{ атм}$, $c'' = 2,1$ и $P = 21 \text{ атм}$,



Фиг. 2



Фиг. 3

$c'' = 2,3$. Так же как и из табл. 3, из фиг. 2 видно, что с повышением давления гелиевой смеси коэффициент сдвига формант по частоте для низкочастотной первой форманты несколько превышает аналогичные коэффициенты для 2,3 и 4-й формант.

В табл. 4 дана сравнительная оценка расчетного пропорционального коэффициента сдвига c_* и экспериментально измеренного по второй методике c''_* .

Экспериментальные исследования речи в гелиевой среде показали, что у всех дикторов во всех случаях c''_* было меньше c_* . Лучшее совпадение теоретического и экспериментального сдвигов формант наблюдалось для низкочастотных резонансов речи.

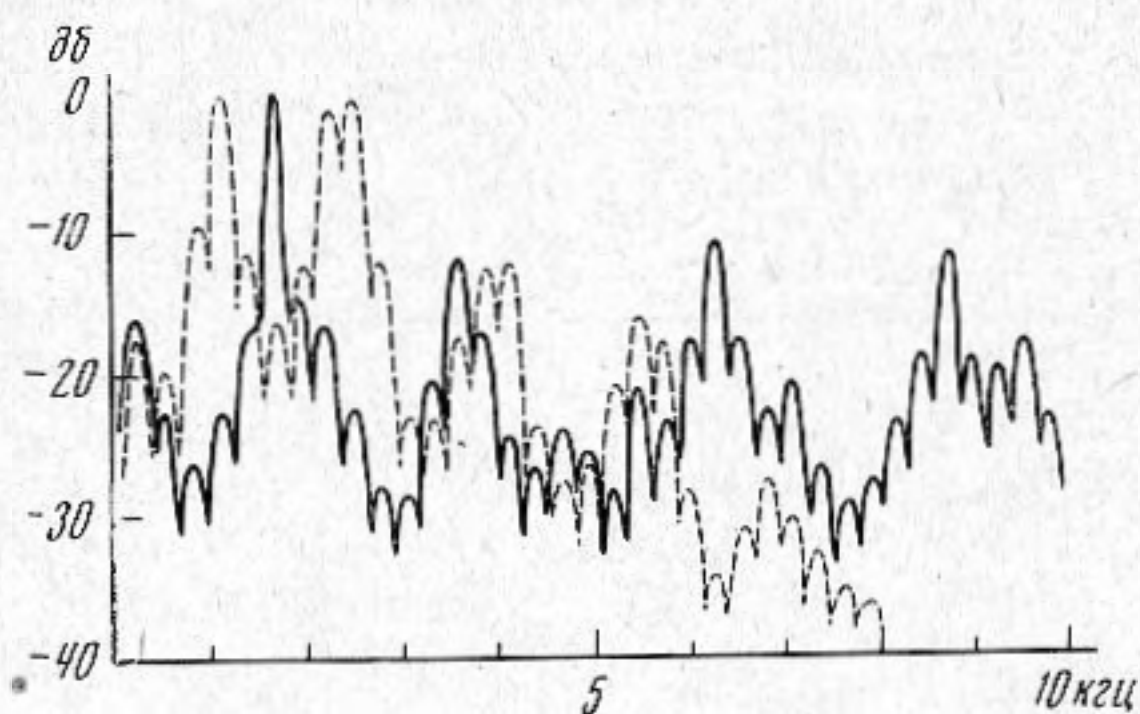
Таблица 4

Давление дыхательной смеси, атм	Воздух 1	HeO ₂			
		1	7	12	21
Скорость звука c , м/сек	343	600	652	768	851
$c_* = c/c_B$	1,00	1,75	1,90	2,24	2,48
c''_*	1,00	1,40	1,70	2,1	2,3

Непропорциональность смещения отдельных формант особенно заметна на переходных участках гелиевой речи, проявляющаяся в эффекте «притяжения» формантных траекторий, расположенных относительно близко по частоте. Это относится к любым из четырех формант и иллюстрируется схематизированной видеogramмой с такими эффектами для 1—2-й

и 2—3-й формантных траекторий в слове «воевать» (фиг. 3, а — воздух при 1 атм, б — гелиево-кислородная смесь при 21 атм). «Притяжение» отчетливо наблюдается уже при давлениях 7 атм гелиево-кислородной атмосферы и при дальнейшем росте давления не усиливается.

Тем не менее, как видно из табл. 3 и фиг. 2, непропорциональность смещения формант речи в гелиевой среде составляет всего несколько



Фиг. 4

процентов и поэтому линейная аппроксимация увеличения резонансных частот является удовлетворительным для практики приближением. Исследование спектральных разрезов речи в гелиевой атмосфере показало наряду с возрастанием формантных частот и увеличение ширины формант. При этом спектральная картина речи в гелиево-кислородной смеси обладает более четкой структурой формант, меньшим числом дополнительных нулей и полюсов, вследствие чего в ней легче проводить аппроксимацию огибающей спектра формантного типа. Фигура 4, на которой представлены спектральные разрезы сегмента, соотносимого со звуком «а» в слове «таблица», на воздухе (пунктир) и при 21 атм гелиево-кислородной смеси (сплошная кривая) для одного и того же диктора иллюстрирует сказанное.

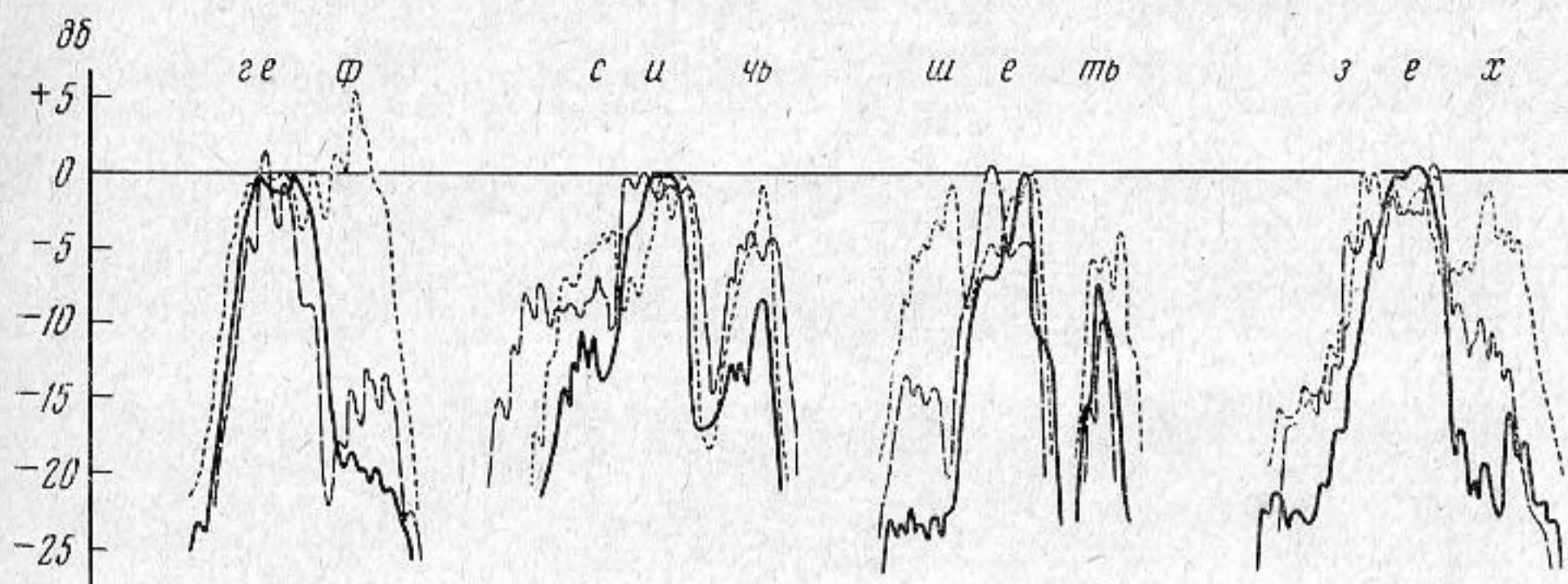
Большая отчетливость формантной структуры речи в гелиевой смеси по сравнению с речевыми сигналами в воздухе, из-за значительного расширения шкалы частот при сравнительно небольших изменениях добротности резонансов, высокая относительная точность определения положения и поведения формант позволяет предположить, что речь в гелиевой атмосфере может оказаться весьма полезным средством при детальном изучении переходных участков нормального речевого потока, а также при решении задач фонетической классификации.

При анализе амплитудных соотношений в гелиевой речи следует иметь в виду два обстоятельства. Во-первых, при использовании гелиевых дыхательных сред частотный диапазон звуков речи расширяется пропорционально изменению скорости звука c . Поэтому адекватные сопоставления мощностей звуков в различных по составу средах требуют искусственного ограничения полосы частот речи таким образом, чтобы измерения мощностей проводились для звуков с одинаковой по составу формантной структурой.

Во-вторых, на измерения абсолютной и относительной мощностей звуков речи значительно влияют индивидуальные особенности голосов различных дикторов и эмоциональное состояние одного и того же диктора. Ввиду этого корректная количественная оценка мощностных соотношений в речи весьма затруднительна и сравнения результатов различных измерений дает скорее качественную картину явлений.

Интегральная излучаемая мощность пропорциональна характеристическому сопротивлению среды $Z_c = \rho_c c_c$, которое может быть больше или меньше соответственного значения для воздуха $Z_b = \rho_b c_b$. Используемые для наших исследований гелиево-кислородные смеси имели $Z_c < Z_b$ при давлении примерно до 3 атм, после чего с ростом давления сред наблюдалось превышение характеристического сопротивления над соответствующим значением для воздуха соответственно увеличению плотности смеси. Поэтому оказывается, что при нормальном давлении в 1 атм гелиевой сре-

ды средняя мощность речевых сигналов падает по сравнению с таковой на воздухе. Различные исследователи указывают цифру в 5—14 дБ [1]. Отметим, что снижение среднего уровня речи в средах с $Z_c < Z_v$ не представляет особого интереса, так как оно существенно с точки зрения эффективности устной связи, ибо изменение средней мощности речевых сигналов легко компенсируется установкой желаемого режима усиления при использовании электроакустических трактов, а в случае непосредственного



Фиг. 5

общения людей в искусственной атмосфере для этой цели оказывается достаточно резерва голосовой мощности говорящих.

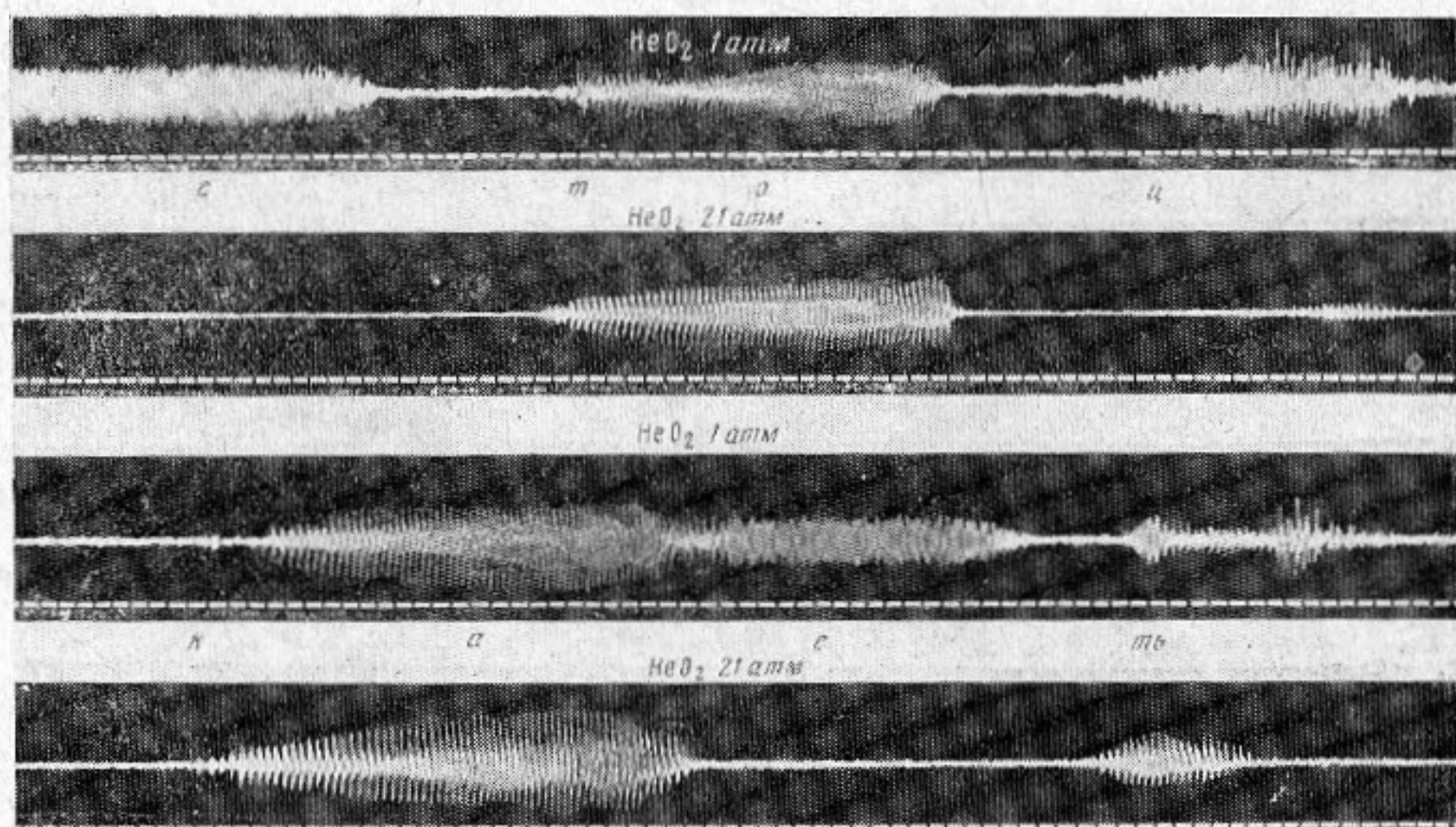
Значительно более важным для коммуникативных целей параметром речи является соотношение мощностей различных звуков, поэтому этот вопрос мы рассмотрим более подробно. Согласно теоретическим предположениям [1], из-за изменения характеристик газовой среды наибольшим нарушениям подвержена работа шумового источника; поэтому есть основания ожидать изменения соотношения мощностей различных звуков. Сравнительное исследование огибающих мощностей звуков речи в воздухе и гелиево-кислородной смеси при давлении 1 атм показало систематический рост соотношения мощностей согласный/гласный для фрикативных звуков, несмотря на то, что общий уровень речи в этих условиях падает, как это было отмечено выше. Сказанное иллюстрируется фиг. 5, где приведены нормализованные по средним уровням гласных огибающие функции звукового давления для нескольких слогов типа СГС, произнесенных одним и тем же диктором в гелиевых атмосферах при различных давлениях*. Сплошные кривые соответствуют гелиево-кислородной среде при давлении 21 атм, пунктирные — то же при 1 атм и штриховые — воздушной среде при давлении 1 атм. Повышение же давления, ведущее к увеличению плотности дыхательной смеси, приводит, наоборот, к уменьшению соотношения мощностей рассматриваемых звуков. Это следует из фиг. 5 и 6, где даны осциллограммы слогов, произнесенных одним диктором в гелиево-кислородных средах при давлениях 1 и 21 атм**.

Увеличение уровня фрикативных звуков типа «С», «Ш», «Ф», «Х», а также фрикативных сегментов взрывных звуков и аффрикат относительно уровня гласных при уменьшении плотности дыхательной среды в достаточной степени согласуются с данными работы [5], согласно которым уменьшение давления воздуха и последующая замена его чистым кислородом на высотах до 10670 м ($P = 0,17$ атм) сопровождается примерно двукратным увеличением соотношения уровней согласный/гласный для фрикативных и особенно взрывных звуков.

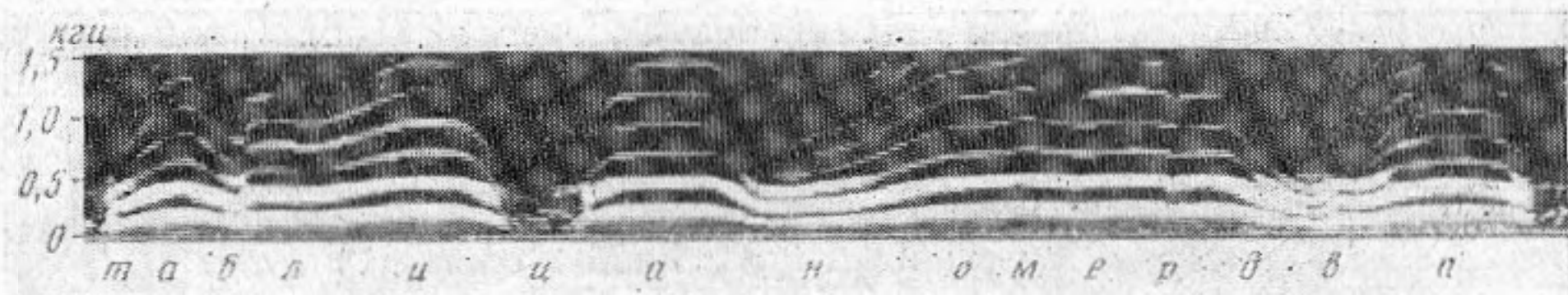
* Согласно сказанному ранее, частотный диапазон сигналов речи выбирался в каждой среде пропорционально скорости звука c .

** Сигналы получены с применением микрофона ДЭМШ-1а и неизменном частотном диапазоне регистрации.

Отмеченные изменения отношения мощностей согласный/гласный и уменьшение этого отношения при произнесении речи в условиях высокого давления гелиевой атмосферы являются весьма существенным фактором, влияющим на восприятие речи, поэтому его необходимо особо учитывать при конструировании систем связи в искусственных атмосферах.



Фиг. 6



Фиг. 7

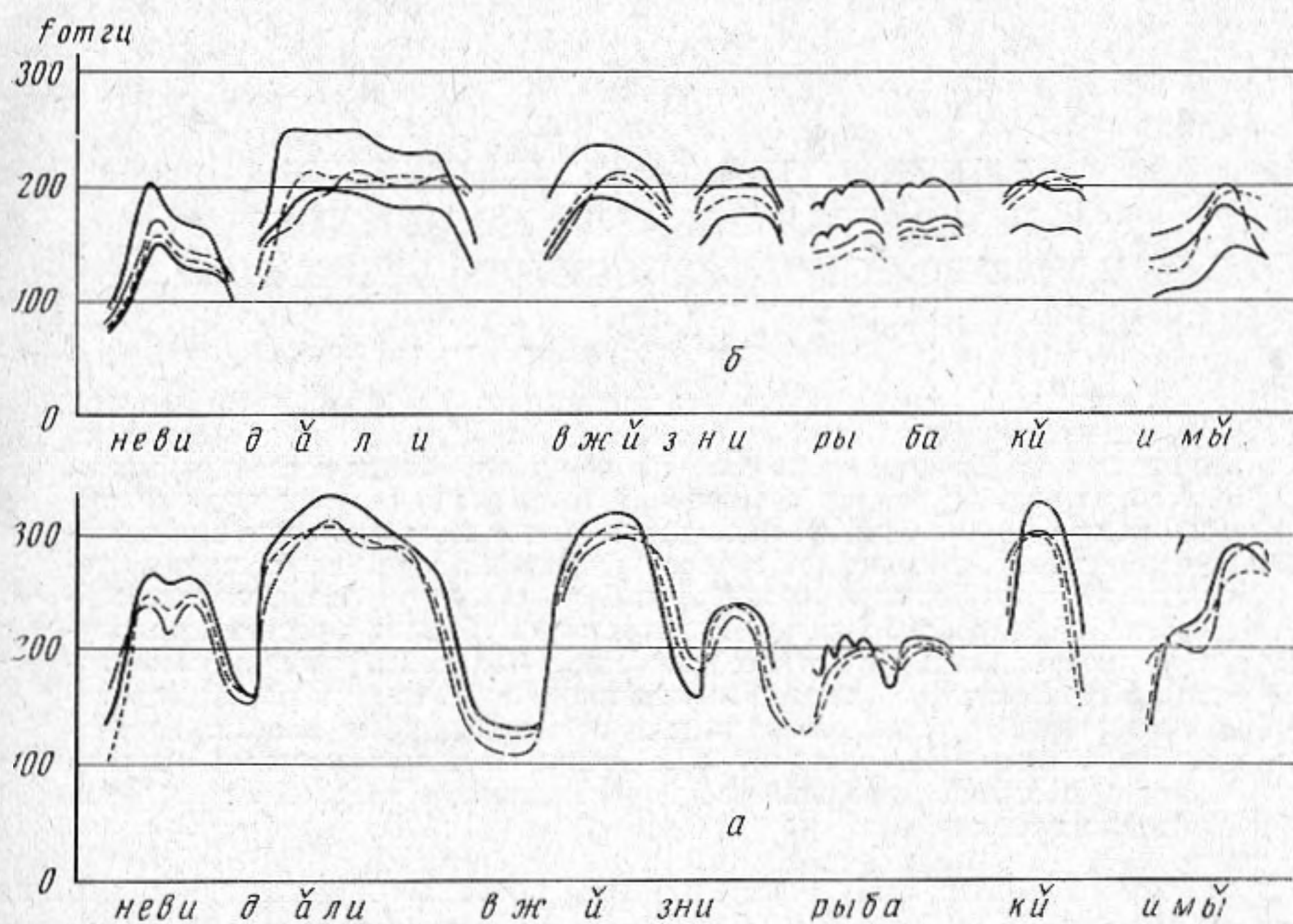
Ввиду заметного различия в данных, относящихся к изменению частоты основного тона речи в гелиевой атмосфере, приведенных рядом авторов [1], а также с целью выяснения особенностей поведения основного тона в условиях высоких давлений среды, мы исследовали изменения мелодий основного тона в гелиево-кислородной среде при давлениях от 1 до 21 атм.

Материалом для сравнений служили видеогаммы речи, снятые узкополосным фильтром типа видеогаммы фразы «таблица номер два», представленной на фиг. 7. Одновременное измерение группы гармонических составляющих голосовых колебаний позволяет оценить мелодию основного тона с точностью не хуже 1–2% по частоте.

Сопоставления мелодий основного тона на фразовом, словесном и слоговом материале в диапазоне давлений гелиевой среды 1–21 атм и воздуха при давлении 1 атм для четырех дикторов показали, что индивидуальные особенности мелодии, свойственные конкретным диктором, устойчиво сохраняются. Кроме того, измерения показали, что при повышении давления гелиевой среды у части тренированных дикторов, имеющих опыт в артикуляционных испытаниях, обнаружилась тенденция к незначительному (10–15%) увеличению частоты основного тона. Сказанное иллюстрируется фиг. 8, а, на которой приведены мелодии основного тона фразы, прочитанной диктором Б в воздушной среде при давлении 1 атм (штри-

ховая кривая), в гелиево-кислородной атмосфере при давлении 1 атм (пунктир) и при давлении 21 атм (сплошная кривая).

Однако отмеченная тенденция наблюдалась не всегда. Так для диктора О. выявить эффект устойчивого повышения частоты основного тона с ростом давления среды не удалось, что видно из рис. 8, б, где даны две реали-



Фиг. 8

зации одной и той же фразы, прочитанной при давлении гелиевой среды 17 атм. Две сплошные кривые отвечают этим двум реализациям; смысл остальных обозначений тот же, что и на фиг. 8, а. В совокупности с фактом неизменности индивидуальных особенностей мелодии основного тона в широком диапазоне давлений среды (1—21 атм) это показывает, что человек в достаточной степени контролирует процесс генерации голосовых импульсов, и этот процесс мало зависит от свойств газовой среды.

ЛИТЕРАТУРА

1. Л. С. Бутырский, Б. И. Петленко. Состояние исследований и способы повышения разборчивости речи в гелиево-кислородной среде. *Зарубежная радиоэлектроника*, 1970, 5, 71—92.
2. Б. И. Петленко. Разборчивость речи, произносимой в гелиево-кислородной среде. *Акуст. ж.*, 1972, 18, 1, 82—89.
3. Дж. Фланаган. Анализ, синтез и восприятие речи. Перев. с англ. под ред. А. А. Пирогова. М., «Связь», 1968.
4. Б. И. Петленко. К вопросу нарушений спектра речи, вызываемых изменением условий речеобразования. *Тр. Московск. энергетич. ин-та*. Вып. 82, 1971, 156—167.
5. К. С. Clark. The effects of high altitude on speech. *J. Acoust. Soc. Amer.*, 1948, 20, 776.

Московский автомобильно-дорожный институт

Поступила
9 июня 1972 г.