

ЛИТЕРАТУРА

1. Л. М. Бреховских. Волны в слоистых средах. М., Изд-во АН СССР, 1957, 57—63.
2. С. А. Рыбак, Б. Д. Тартаковский. Некоторые применения матрицы перехода к теории плоских волн в системе упругих слоев. Акуст. ж., 1962, 8, 1, 119—123.
3. W. T. Thomson. Transmission of elastic waves through a stratified solid medium. J. Appl. Phys., 1950, 21, 2, 89—93.

Москва

Поступило в редакцию
31 января 1968 г.

УДК 534.78

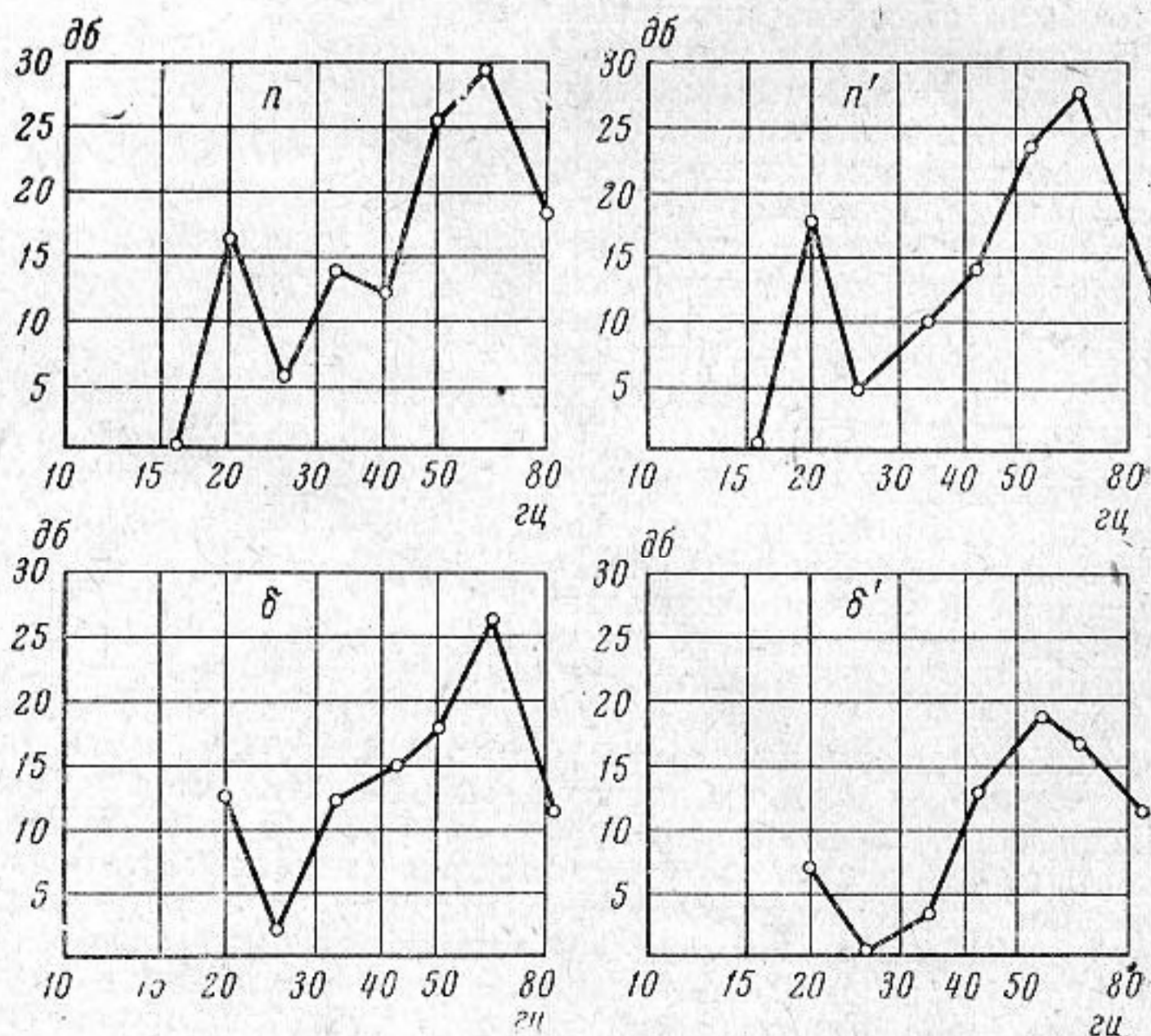
ИНФРАЗВУКОВЫЕ ПРИЗНАКИ ДЛЯ АВТОМАТИЧЕСКОГО РАСПОЗНАВАНИЯ ЗВУКОВ РЕЧИ

Л. Л. Мясников, Е. Н. Мясникова, М. Я. Пекельный,
А. И. Трилесник

Во многих задачах технической фонетики, таких, как автоматическое распознавание, членение и синтез звуков речи — существенное значение имеет установление новых физических признаков, которые могут служить для различения звуков. В настоящей работе проведены поиски особенностей звуков речи в инфразвуковом частотном диапазоне, который в этом отношении до сих пор не был исследован.

По техническим причинам (из-за значительного уровня фона инфразвуковых помех и малой интенсивности сигнала) наблюдения в частотном диапазоне от 5 до 12 гц не дали определенных результатов. На более высоких частотах вплоть до 80 гц были обнаружены инфразвуковые компоненты, принадлежащие взрывным (*с*) смычным согласным звукам $n, n', t, t', k, k', b, b', d, d', g, g'$, (штрихи обозначают мягкие палатализованные звуки). В этом диапазоне (12—80 гц) гласные звуки не обладают заметными компонентами. Из двух согласных только шипящие дают слабые низкочастотные составляющие.

Для звуков n, n', t, t' и других взрывных обнаружены составляющие в окрестности 60 гц, а также инфразвуки с частотами около 20 гц. Такие частоты находятся в соответствии со средними продолжительностями шумовых импульсов (взрывов) порядка 15—20 мсек и, вероятно, с переходной областью, продолжительность которой порядка 50 мсек.



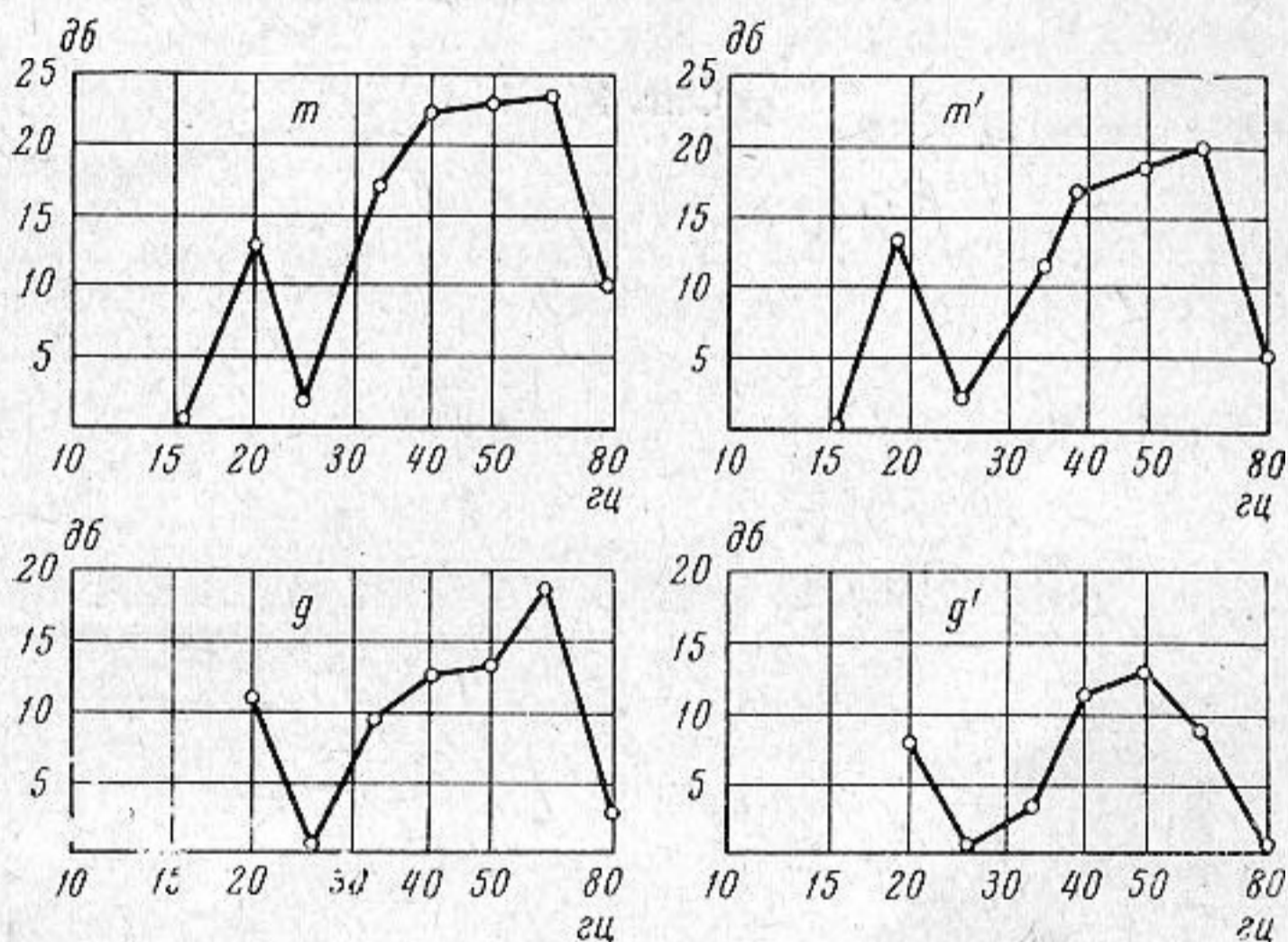
Фиг. 1

В комплект измерительной аппаратуры входили: 1) инфразвуковой приемник, собранный на кристаллах сегнетовой соли, пригодный для работы в области частот от 5 до 1000 гц, с защитным колпаком, предохраняющим его от действия воздушных потоков, возникающих при произнесении фонем; 2) калиброванный усилитель; 3) инфразвуковой спектрометр, построенный по принципу спектрометра Фрейштедта, с 22 треть-октавными фильтрами, перекрывающими диапазон частот от 1 до 130 гц с воспроизведением изображения спектра на экране электронной трубки; 4) самописец типа Неймана. Нуль децибел шкалы уровней спектрометра соответствовал усилению по напряжению на 55 дб, при входном уровне звукового давления $2 \cdot 10^{-5}$ н/м². Результаты спектрального анализа (т. е. отсчеты по электронной трубке) представлены на фиг. 1—3, причем 0 дб соответствует звуковому давлению около 0,01 н/м².

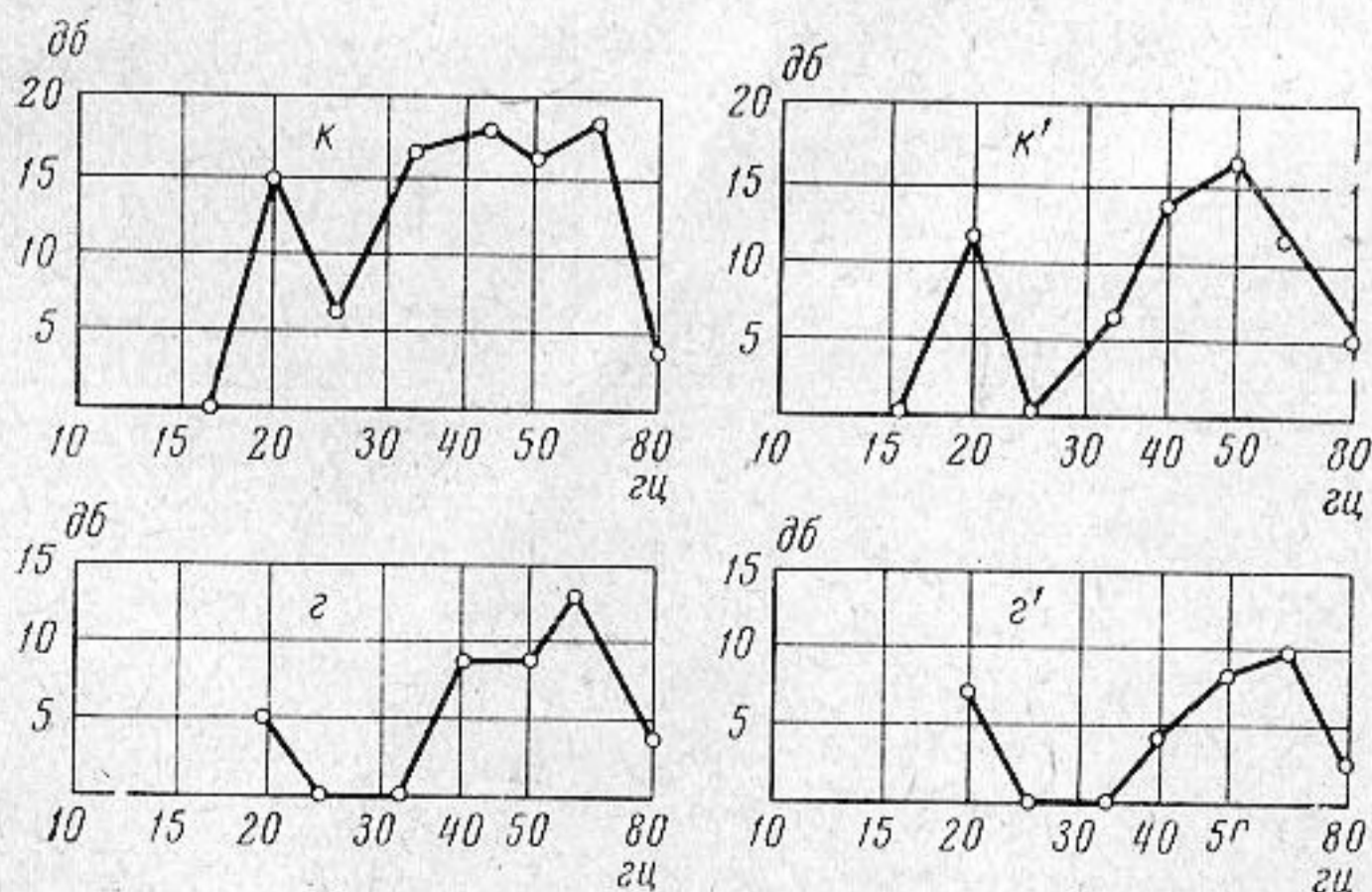
На фиг. 1 приведены спектры губных, где низкочастотные компоненты наиболее интенсивны, а на фиг. 2 — спектры переднеязычных зубных. Как правило, мягкие звуки имеют несколько пониженные уровни и у некоторых из них спектральные компоненты сдвинуты в сторону низких частот (*б'*, *д'*). Существует также различие в распределении звукового давления по спектру: у переднеязычных *т*, *д*, *т'* участок 30—40 гц несколько поднят по сравнению с губными *п*, *п'*, *б*, *б'*.

На фиг. 3 представлены данные для заднеязычных. Они отличаются более низким уровнем звуковых давлений. Твердые *к*, *г* здесь также отличаются от мягких *к'*, *г'* более сильно выраженной областью, лежащей от 30 до 40 гц. У *к'* заметен сдвиг компоненты в сторону более низких частот (50 гц).

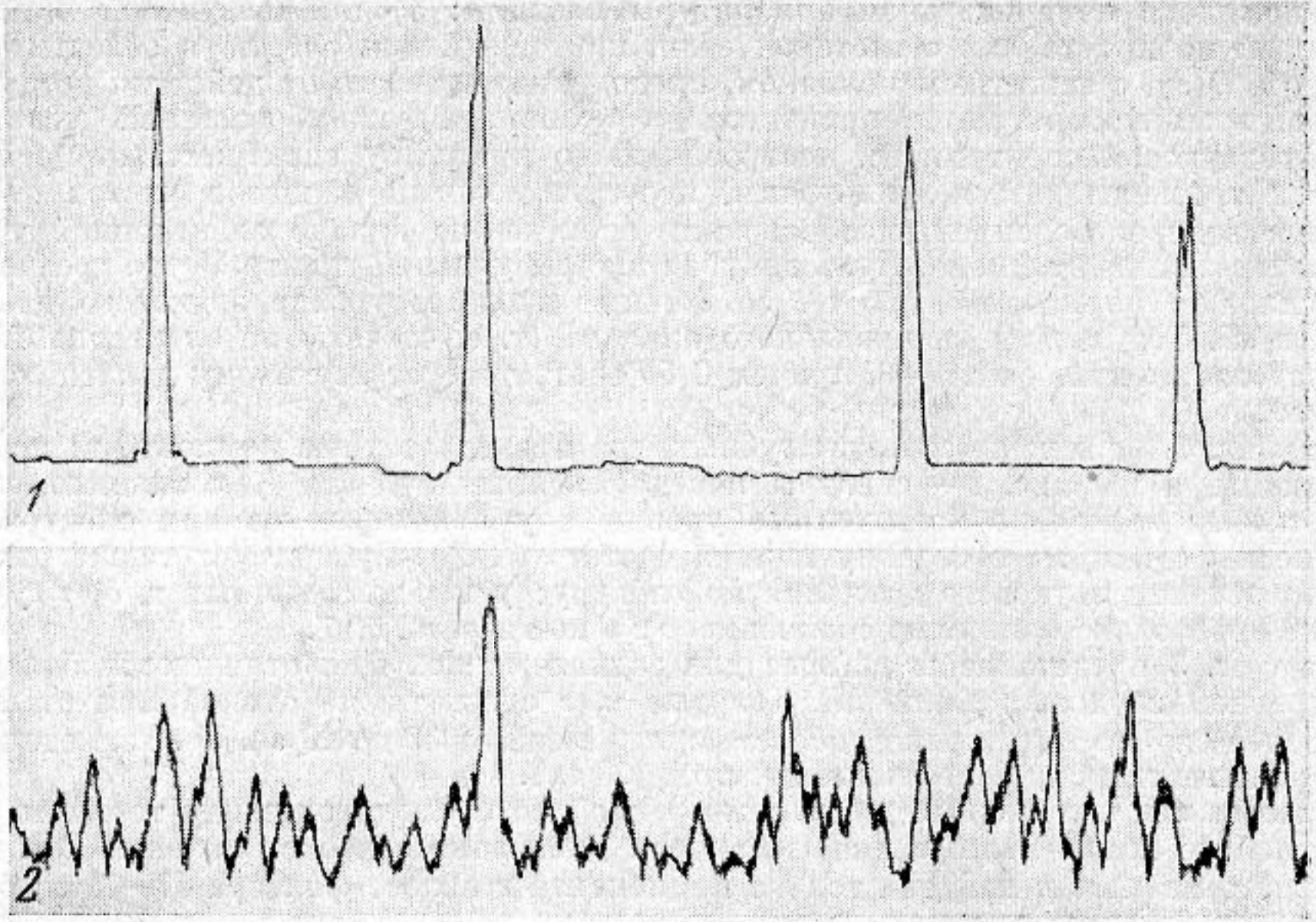
Диаграммы, полученные для звуков *р* и *р'*, не приводятся ввиду того, что у них спектральные компоненты в инфразвуковом диапазоне слабо выражены. Для переднеязычного небного *р* найдена только компонента в полосе фильтра 57—71 гц с уров-



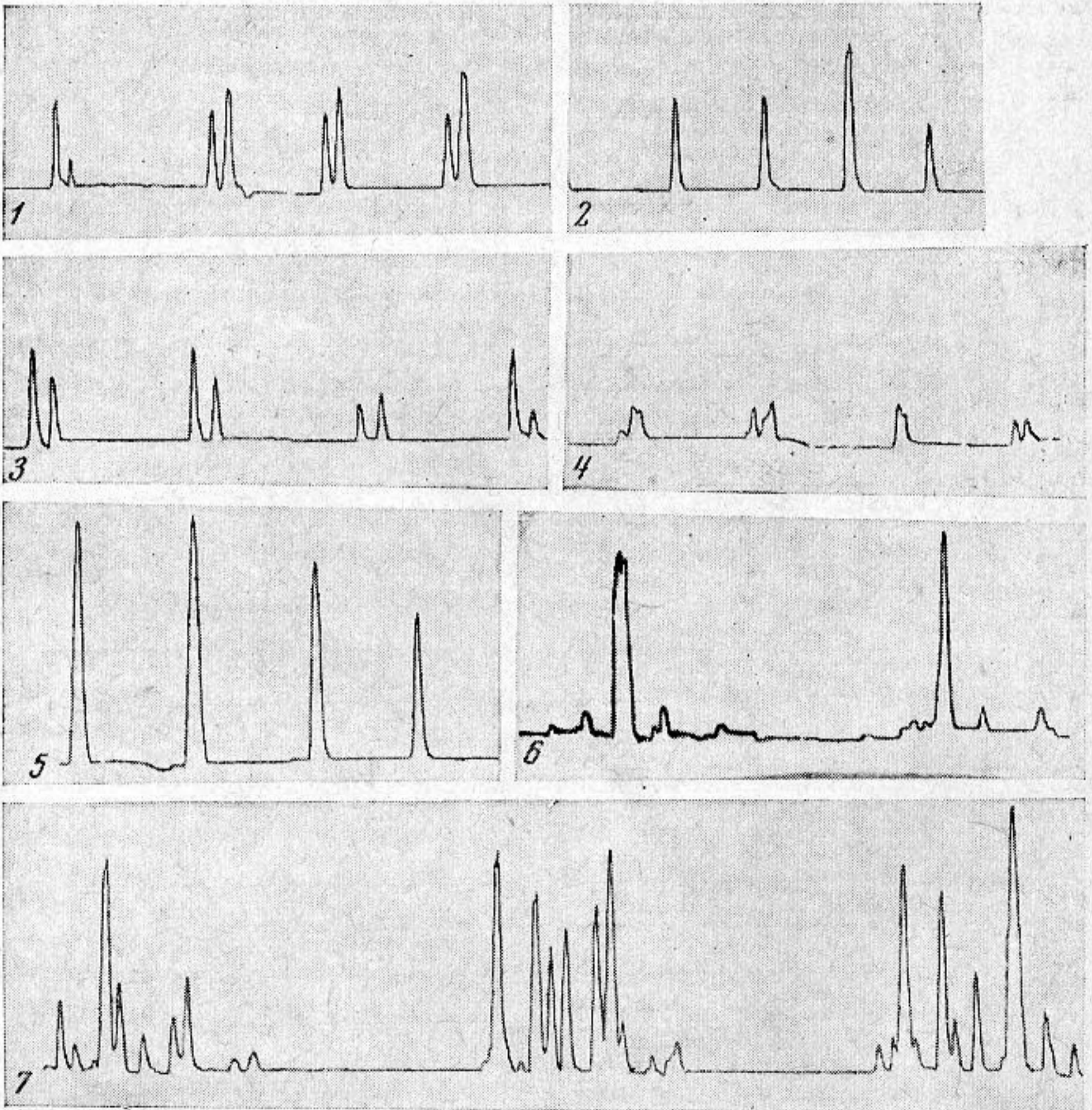
Фиг. 2



Фиг. 3



Фиг. 4



Фиг. 5

нем 3 дб, а для мягкого p' две компоненты: компонента в полосе 35—45 гц 11 дб, а в области 57—71 гц 3 дб.

Гласные не обладают заметными инфразвуками. Шипящие фриктивные s , s' , $ш$, $з$, $ж$ дали в полосе пропускания фильтра 57—71 гц отклонения порядка 1—3 дб. Несколько сильнее выделяются ϕ и ϕ' : уровень звукового давления доходит для них до 8 дб. Заметим, что и шипящие дают низкочастотные компоненты только при форсированном произнесении, когда, по-видимому, образуется шумовой импульс продолжительностью в 15—20 мсек в начале фонемы. В окрестностях частот 20 и 32 гц для фонемы ϕ наблюдались слабые компоненты (3 дб). Инфразвуковые составляющие у других шипящих обнаружены не были.

Приведенные на фиг. 1—3 данные получены на основании многочисленных тестов, произведенных при участии 5 дикторов, причем каждый из обследованных звуков речи был произнесен не менее 100 раз; взрывные произносились с дополнением гласной $ы$ ($ны$, $ты$ и т. д.). Интервальная оценка погрешности уровней спектральной плотности звукового давления с 95% доверительным интервалом дает не более ± 3 дб. Максимумы низкочастотных компонент приходятся на частоты 20 ± 3 и 60 ± 8 гц.

Путем применения соответствующих фильтров в схеме динамического анализатора [1] или в иных схемах для объективного распознавания и членения речи [2] можно производить отделение по инфразвуковым составляющим взрывных от других фонем. Между взрывными получается разделение губных и передне- и заднеязычных; кроме того, выделяются заднеязычные и некоторые мягкие взрывные согласные. Представляет интерес исследование инфразвуковых сегментов в словах в целях распознавания слов по признаку вхождения в них инфразвуковых, а также ультразвуковых [3] сегментов.

На фиг. 4 представлены нейманограммы взрывной согласной, произнесенной одним и тем же диктором по 4 раза (фильтр 57—71 гц) и нейманограмма того же звука речи (n), произнесенного 3 раза с фильтром 18—22 гц. Скорость ленты 3 мм/сек. На последней нейманограмме заметны значительные инфразвуковые помехи.

Ширина каждой полосы на нейманограммах составляет 5 дб. Уровень начала записи лежит выше уровня $2 \cdot 10^{-5}$ н/м² на 53 дб.

На фиг. 5 изображены нейманограммы некоторых слов: 1 — поп, 2 — тот, 3 — папа, 4 — сос, 5 — суп и фраз: 6 — от топота копыт пыль по полю несется, 7 — редеет облаков летучья гряда.

ЛИТЕРАТУРА

1. Е. Н. Мясникова. Некоторые вопросы автоматического распознавания звуков речи. Тр. Ленинградского механического института, 1965, 51, 2, 134—144.
2. М. А. Сапожков. Речевой сигнал в кибернетике и связи. М., Связьиздат, 1963.
3. Е. Н. Мясникова, Л. Л. Мясников, Н. П. Громов. Ультразвуковое членение речи в методах объективного распознавания. Тр. Ленинградского механического института, 1967, 62, 3, 121—126.

Ленинградский кораблестроительный институт

Поступило в редакцию
14 июля 1967 г.

УДК 534.29

ВЫСОКОТЕМПЕРАТУРНЫЙ МАТЕРИАЛ ДЛЯ АКУСТИЧЕСКОЙ СКЛЕЙКИ

Я. А. Обуховский, Л. А. Сысоев, М. Ш. Файнер

Ультразвуковые линии задержки находят широкое применение в современной технике. В качестве материалов звукопровода в этих линиях служат плавный и кристаллический кварц, монокристаллы NaCl, KCl, LiF и др. Созданные в последнее время высокочастотные пьезопреобразователи [1] позволяют более эффективно использовать линии задержки на частотах 50—200 Мгц. Однако их практическое применение часто сдерживается отсутствием подходящего материала для промежуточного слоя (склейки).