

## ЛИТЕРАТУРА

1. Л. Н. Галкин, Л. М. Луничкина. Температурные флуктуации в водной толще при резко выраженной стратификации. Изв. АН СССР. Сер. Физика атмосферы и океана, 1971, 7, 2, 232—236.
2. O. S. Lee. Effect of an internal wave on sound in the ocean. J. Acoust. Soc. Amer., 1961, 33, 5, 677—681.
3. О. П. Галкин, Л. В. Трезубова. Распространение звука при наличии внутренних волн. Тр. Акуст. ин-та, 1970, в. XIII, 107—113.
4. В. А. Полянская. О влиянии высокочастотных внутренних волн на звуковое поле точечного источника в океане. Акуст. ж., 1974, 20, 1, 95—102.

Акустический институт  
Академии наук СССР  
Сухумский филиал

Поступила  
28 июня 1973 г.

УДК 534.78

### СТАТИСТИЧЕСКОЕ ИССЛЕДОВАНИЕ ПАУЗ РЕЧИ ДИКТОРА ДЛЯ НЕКОТОРЫХ РОМАНСКИХ И ГЕРМАНСКИХ ЯЗЫКОВ

А. А. Глухов

Развитие техники звукового вещания требует разработки унифицированных систем контроля качества работы вещательных трактов и их резервирования. Одним из параметров речевого сигнала, необходимых для такой разработки, является статистическое распределение речевых пауз. Сведения такого рода могут быть полезны также при решении проблемы компрессии и синтеза речи, во многих аспектах лингвистики, при идентификации звуковых образов и т. п. Более того, средняя длительность и распределение длительности пауз речи могут оказаться признаком, несущим информацию о языке.

Представим временную функцию фонетических уровней звуков речи известным выражением

$$(1) \quad |p(t)| = |E(t)| |K(t)|,$$

где  $E(t)$  — генераторная функция гортани и  $K(t)$  — передаточная функция голосового тракта.

Так как минимумы функции (1) могут иметь место и при произношении некоторых звуков речи, паузами называют временные отрезки с нулевыми значениями амплитуд звукового давления, если при этом на некотором отрезке времени  $t_i$  соблюдается неравенство  $t_i > T_n$ , где  $T_n$  — «период» самой низкой частоты речевого сигнала. Отметим, что такие минимумы функции  $p(t)$  не поддаются обнаружению из-за влияния помех. Поэтому для правильного определения моментов пауз выбирают фонетические уровни звуков речи, так чтобы они превышали уровень аддитивных помех ( $N_n$ ).

Существует несколько типов пауз (между фонемами, словами и т. п.), по-разному связанных с особенностями языка, индивидуальностью диктора и типом сообщения. Характеристики длительности этих пауз, как и темп речи, являются признаком индивидуальных особенностей произношения данного диктора. По этой причине была подвергнута анализу речь достаточно представительной группы дикторов. При проведении исследования, в соответствии с суточным расписанием передач на иностранных языках, сигнал с выхода микрофона студии поступал на вход тракта формирования программ. Аппаратурный анализ проводился на выходе этого тракта, параметры которого нормированы. Звуковым материалом служили тексты, типичные для передач вещания на иностранных языках. Было подвергнуто анализу всего 66 передач длительностью от 10 до 30 мин в исполнении 24 дикторов (4 диктора для каждого языка).

Речевое сообщение содержит некоторое множество пауз  $M$ . В интервале времени анализа  $T$  это множество конечно и образовано из элементов  $\{x_1, x_2, x_3, \dots, x_i\}$ , каждый из них  $x_i \in M$ . Анализатор, использованный для экспериментальных исследований, измерял не длительность существования  $i$ -го элемента множества  $M$ , а число или вероятность наличия  $P(x)$  пауз, длительность которых попадает в той или иной временной интервал  $t_1, t_2, t_3, \dots, t_i$ . Когда пороговый уровень сигнала, при котором



Язык (длительность анализа, час, мин)	Вероятностная характеристика	Значения характеристик при интервалах длительности пауз, мсек							Сумма
		50-150	150-300	300-500	500-1000	1000-2000	2000-3000	>3000	
Французский (4.00)	$F(x)$ $m/T$	0,377 286	0,544 127	0,687 109	0,831 147	0,974 70	0,987 10	1,0 10	759
Итальянский (5.40)	$F(x)$ $m/T$	0,382 217	0,526 83	0,664 79	0,870 117	0,969 56	0,985 9	1,0 8	569
Испанский (8.00)	$F(x)$ $m/T$	0,330 300	0,540 186	0,720 153	0,930 194	0,990 53	0,995 4	1,0 7	897
Португальский (3.50)	$F(x)$ $m/T$	0,373 430	0,538 185	0,752 238	0,902 167	0,974 80	0,975 13	1,0 9	1122
Английский (6.30)	$F(x)$ $m/T$	0,360 322	0,537 155	0,680 128	0,873 173	0,976 92	0,995 16,5	1,0 8,3	895
Немецкий (7.50)	$F(x)$ $m/T$	0,365 280	0,554 145	0,687 102	0,883 152	0,971 68,5	0,984 10	1,0 12,2	770

определяются моменты пауз, больше уровня помех  $N_{пр}$ , вероятность появления паузы равна

$$(2) \quad P(x) = \begin{cases} 0 & N_j > N_{пр} \text{ при любом } t_i \\ N_j < N_{пр} & \text{при } t_i \leq T_0 \\ 1 & N_j < N_{пр} \text{ при } t_i > T_0, \end{cases}$$

где  $N_j$  — текущий уровень сигнала в момент анализа,  $N_{пр}$  — пороговый уровень сигнала, равный  $-40 \text{ дБ}$  (относительно нулевого уровня, при котором определяются моменты пауз),  $T_0$  — разрешающая способность анализатора пауз,  $t_i$  — возможные значения длительностей пауз для  $x_i \in M$ .

Постоянная времени усреднения речевого сигнала в анализаторе была равна  $40 \pm 5 \text{ мсек}$ . Она выбрана меньше времени восстановления коэффициента передачи в компандерной системе и больше времени интеграции в измерителях уровня, используемых в каналах иновещания.

Полученные в соответствии с формулой (2) экспериментальные данные для романской группы языков (французского, итальянского, испанского, португальского) и германских языков (английского, немецкого) приведены в сводной таблице. Интервалы длительностей пауз в таблице указаны без допусков, которые не превышают  $\pm 5\%$ . Использование пауз длительностью менее  $50 \text{ мсек}$  для целей контроля в вещании сопряжено с рядом трудностей. Поэтому такие паузы не учитывались.

Представительность выборок в разрядах была достаточно велика. Это позволяет для характеристики пауз рассчитать одномерную функцию распределения  $F(x)$ , среднее число пауз в час  $m/T$ , где  $m$  — общее число пауз, и среднее время семантических пауз за все время анализа  $T$ .

Анализ экспериментальных данных позволяет прийти к заключению о существовании единой закономерности распределения скважности речи диктора для большинства исследованных групп языков. Из таблицы видно, что влияние особенностей различных языков на одномерную функцию распределения пауз сравнительно невелико.

Максимальная вероятность появления пауз для всех исследованных языков при выбранном усреднении речевого сигнала приходится на интервал от  $50$  до  $150 \text{ мсек}$ .

В среднем в речи диктора португальского и испанского языков за каждые  $3$  и  $4 \text{ сек}$ , а французского и итальянского языков за  $5$  и  $6 \text{ сек}$  появляется пауза со средней длительностью больше  $50 \text{ мсек}$ . Сведения о паузах английской и немецкой речи мало отличаются от данных для испанской и французской речи соответственно.

Расчеты показывают, что минимальная средняя статистическая длительность семантических пауз имеет место для итальянского языка ( $2,6\%$ ), а максимальная ( $4,2\%$ ) для португальского языка. Для английского и испанского языков средняя статистическая длительность пауз не превышает  $4,0$  и  $3,6\%$ , а для французского и немецкого языков почти одинакова и составляет  $3,4\%$  от общего времени передачи.

Поступила  
14 марта 1974 г.