



СПЕКТРАЛЬНЫЙ АНАЛИЗ ЗВУКОВ РЕЧИ И ИНТОНАЦИЯ

ИЗДАТЕЛЬСТВО МГУ
1969 г.

Редакционная коллегия:

Главный редактор В. А. Артемов; редактор статей, выполненных в ЛЭФ ИВЯ при МГУ М. К. Румянцев. Члены редколлегии — К. К. Барышникова, В. А. Васильев, О. А. Норк, Д. С. Светлышев.

Настоящий сборник состоит из двух частей. В первой его части публикуются материалы симпозиума по спектральному анализу звуков речи, состоявшемуся в июне 1966 г. в 1 МГПИИЯ им. Мориса Тореза. Во второй части представлены работы, главным образом, по словесному ударению и интонации, выполненные в лаборатории экспериментальной фонетики Института восточных языков при МГУ.

НЕКОТОРЫЕ ВОПРОСЫ СПЕКТРАЛЬНОГО АНАЛИЗА ЗВУКОВРЕЧИ

1. Спектральный анализ звуков речи возник из настоятельной потребности ряда наук и практики иметь адекватную акустическую характеристику звукового состава языка.

2. Языкознание нуждается в адекватном акустическом описании звукового состава различных языков, в частности, звуковой структуры фонемы, слога, слова, фразы, живого потока речи. Особенно существенно провести достаточно тонкий адекватный анализ звуков речи, так как только таким путем можно решить вопрос о соотношении фонологического и фонетического анализа речи, о реализации фонемы в потоке речи, о различных признаках фонем, о соотношении фонемы, фонемного инварианта и звука речи. Для удовлетворения этих нужд возникла и успешно развивается экспериментальная фонетика, в которой спектральный анализ является одним из основных методов научного исследования.

3. Психологии крайне необходимы адекватные спектральные характеристики речи для изучения процессов мышления, чувств и воли, а также для наиболее надежного и экономного определения типа высшей нервной деятельности человека. Как показали исследования ЛЭФИПР, именно путем электроакустического анализа речи возможно раскрыть все еще таинственную природу эмоционально-волевого поведения человека, дать классификацию чувств и воли. В этих целях необходимо электроакустическое исследование не только интонации речи, но ее тембраль-

ной окраски, что, видимо, потребует спектрального анализа речевого сигнала. Необходимы эти спектральные характеристики и для изучения механизмов речи и процессов ее филогенетического и онтогенетического становления (развития).

Как показали исследования ЛЭФИПР совместно с Львовским научно-исследовательским институтом охраны материнства и детства для физиологии весьма существенным является электроакустический анализ речи, в частности, спектральный анализ ее фонетического состава.

5. Патологии, медицине, дефектологии спектральный анализ речи дает надежные индикаторы распознавания различных видов расстройства речевой и психической деятельности, в частности, при диагностировании и лечении афазии, алексии, дизлексии и других видов заболеваний и дефектов речи, а также душевных заболеваний. Спектральные исследования звуков речи оказались весьма существенными для раннего распознавания рака гортани и для изготовления протезов артикуляционного аппарата.

6. Методика обучения родному и иностранным языкам не может обойтись без сопоставительного спектрального анализа системы фонем родного и вновь изучаемого языка. В частности, без этих исследований невозможно программированное обучение звуковому составу языка и изготовление машин обучения при научении произношению на иностранных языках.

7. Всем ясна роль спектральных исследований речи для техники и теории связи, в частности, для решения вопросов автоматического распознавания, идентификации и компрессии речевых сигналов, а также для их имитации.

8. Проблема устного вывода в кибернетические устройства решается и может быть решена на основе спектральных исследований звуков речи. Тем самым спектральные исследования оказываются безусловно необходимыми для изучения и организации процесса коммуникации человек-машина.

9. Разнообразные и многочисленные задачи телеуправления также невозможно разрешить без надежных акустических характеристик речевых сигналов.

10. Культура речи и искусство выразительного слова, а также театральное творчество нуждаются в точных акустических характеристиках речи в их соотношении с языковыми, психологическими и коммуникативными особенностями речи, в их образной выразительности.

11. Таким образом, спектральные исследования звуков речи имеют колоссальное значение, не только для ряда наук, но и в теории и практике построения современной техники, коммуникации, овладения космосом, в обучении, медицине, искусстве и т.д.

12. Все это обязывает специалистов в области фонетических знаний значительно усовершенствовать спектральный анализ звуков речи, для чего прежде всего необходимо обсудить ряд существенных вопросов этого анализа.

Они следующие:

А. Система понятий и проблем.

Б. Аппаратура.

В. Методика исследования.

Г. Использование результатов исследований.

Рассмотрим эти вопросы.

13. В систему понятий и проблем спектрального анализа звуков речи входят:

А/1/ Явление, понятие, структура речевого спектра. Это колоссальная проблема современных фонетических наук. Она включает вопросы о взаимозависимости фонологии и фонетики, о фонеме и фонемном инварианте, о спектральной полиинформации (языковой, речевой, типологической, персональной), о формантах и резонансных частотах, об изменении во времени акустической структуры спектра, о статистическом анализе и математическом выражении спектральной структуры, о возможности судить о спектральном составе звуков речи по восприятию синтезированных акустических сигналов, об информативности огибающей спектра в ее целом и некоторые другие.

А/2/ Различительные и различающие признаки. Это очень сложная проблема современного структурного анализа единиц языка, как попытка свести всю сложность языкового кода к бинарным противопоставленным признакам. Причем состав бинарно-противопоставленных пар-признаков объявляется общим для всех языков мира. Следует экспериментально исследовать вопрос о дифференциальных различительных признаках реализующих информацию о языке, речи, авторе и т.д. Необходимо определить информативную цену каждого отдельного акустического признака спектра и их целостной структуры.

А/3 В то же время это вопрос о принципе коркового управления программной артикуляции фонем и фонемных инвариантов. Видимо, есть два плана этой программы: сигнификативный и коммуникативный. Первый, нужно думать, реализуется в системе фонем, второй - в системе фонемных инвариантов.

А/4/ Проблема преобразования нестационарного процесса речи в квазистационарный. Проблема объема и энтропии спектральной информации.

А/5/ Понятие фонетической функции (А.А.Пирогов), как попытка преодолеть ~~механистическое~~ понимание членения потока речи на фонемы и в то же время признание того факта, что мельчайшей единицей языка (не артикуляции!) служит фонема, а не слог.

Вопрос о создании более совершенной аппаратуры требует исследовать:

1. Степень ~~аппроксимации~~ анализаторов и синтезаторов речи (частотной, силовой, временной).

2. Членение общей полосы частот на отдельные полосы и вопрос о конструкции фильтров.

3. Форманты и резонансные частоты.

4. Экспериментальная критика.

5. Визуальные представления спектральной картины или же анализ импульсов.

6. Использование счетно-вычислительных машин.

7. Аналоговые цифровые устройства.

8. Фонемный синтезатор.

9. Спектрокоррелометр.

14. Методные исследования содержат следующие вопросы:

В/1/ Экспериментальный материал (звуки, слоги, слова, фразы, тексты, его изоморфность системе языка, экспериментальная экономичность, репрезентативность узкой задаче исследования.

В/2/ Дикторы. Наиболее вероятный диктор для данной задачи исследования. Математическое определение необходимого и достаточного числа дикторов.

Случайные дикторы и дикторы-профессионалы. Проверка голоса диктора (работы его артикуляционного аппарата).

В/3/ Контроль за экспериментальной установкой и ее подготовка к опытам.

В/4/ Составление инструкции дикторам (общие положения и их специальный вариант для данной задачи исследования).

В/5/ Проведение опытов со строжайшим соблюдением всех правил. Номенклатура этих правил.

В/6/ Студия, ее тип и технические требования.

В/7/ Тип микрофона, микшерского пульта, магнитофона, ленты. Скорость протяжения ленты.

В/8/ Первичное прослушивание и требования к записи на магнитофон при ее слуховом восприятии. Техническая проверка записи на магнитофон.

В/9/ Подбор auditors. Случайные аудиторы или профессионалы. Подготовка и тренаж auditors.

В/10/ Требования к проведению аудирования. Громкоговорители или наушники. Требования к специальной студии для прослушивания.

В/11/ Запись экспериментального материала на спектроанализатор.

В/12/ Обработка данных спектрального анализа. Три варианта этой обработки: с использованием ЦВМ, без их использования и смешанный тип анализа (первичное считывание вручную, составление перфокарт, машинная обработка).

В/13/ Требования к графическому изображению полученных данных.

В/14/ Правила структурного анализа и описание его результатов.

В/15/ Использование синтеза и требование к нему.

15. Использование результатов исследования:

С/1/ Использование экспериментальных данных в семантическом и прагматическом исследовании.

С/2/ Сочетание фонологического и экспериментально-фонетического плана анализа.

С/3/ Статистическое и репрезентативное исследование.

С/4/ Теоретическое использование данных исследования.

С/5/ Практическое их использование.

Prof. Artemov V.A.

Laboratory of Experimental Phonetics
& Psychology of Speech at Moscow
Institute of Foreign Languages

SOME QUESTIONS OF SOUND SPEECH SPECTRUM ANALYSIS

Sound speech spectrum analysis if adequately done, is of essential significance not only for communications and communication of speech via language and cybernetics but also for solution of problems of automatic speech identification, and oral input of commands into machines of "brain work" and also for linguistics, physiology, psychology and pathology of speech as well as for developing the science of the methods of teaching foreign languages, the culture of speech and the art of artistic performance.

To be referred to the most essential aspects of speech spectrum analysis are, to our mind, the following:

the system of concepts and problems, the use of apparatus, the method of investigation employed, and the application of results obtained, as well as the understanding of speech spectrum.

This problem involves such questions as the interrelation between phonetics and phonology; questions on the sound, phonemic invariant and phoneme; questions on the polyformatism of speech spectrum; on formants and resonance frequencies; on time variation of acoustic spectrum structure; on statistical analysis and mathematical interpretation of the spectrum structure; on the possibility of judging on the sound speech spectrum composition by perception of synthesized acoustic signals; on the

informative nature of envelope curves of spectrum in the whole and on some other factors.

Distinctive and identifying characteristics

In modern acoustic and statistical analysis of speech units this is a very complex problem. In the study of the acoustic speech signal important for determination are the difference and interlinking of distinctive properties of language, speech and individual - typological as well as personal characteristics. The problem is found complex because these characteristics make up not a sum but a specific structure of grey matter control of articulation of phonemic invariants, phonemes and their position variants. Two programs of this control seem to be apparent: significative program used for identification of things through the use of word symbols and communicative one necessary for the purpose of speech communicative act, and in particular, for identification of words and their meanings.

An important issue here is the problem of transformation of the speech spectrum from the non-stationary process into a quasi-stationary process. And in addition, we cannot fail to mention such problem as the problem of the amount and entropy of spectral information necessary and sufficient for the purposes of communication.

And at last, also important from the point of view of modern science is the understanding of the role of phonetic function (A.A.Pirogov) - an attempt to overcome a purely mechanical understanding of speech stream and the process of articulation. This approach does not only recognize the fact that speech is divided into phonemes but also endeavors to take account of the that the smallest unit of language is a phoneme and not a syllable.

To construct up-to-date analitik and synthesizing equipment it is necessary to find answers for certain questions including:

- 1) degree of approximation of phonetic analyzers and synthesizers;
- 2) dividing general frequency bands into separate bands and the problem related to filter construction;
- 3) formants and resonance frequencies;
- 4) interpretation of the visual spectral pictures and analysis of impulses;
- 5) analogue digital computers;
- 6) use of computers;
- 7) phoneme synthesizer;
- 8) correlation spectrometer;

Improvement of the methods of spectrum analysis and synthesis of phonemic composition of languages requires the knowledge on how to prepare the experimental material, how to select the announcers, how to evaluate the acoustic characteristics of the studio, i.e. to know the characteristics of the microphone, tape-recorder, control panel, and to provide proper selection of auditors and ensure auditor analysis as well as to answer questions of recording the sound from the tape recorder on to the panthograph and its structural analysis including the technique of statistical analysis and graphic representation of the material obtained.

The use of obtained data includes the questions on the correlation and peculiarities of semantic and pragmatic investigations, on correlation between experimental phonetic and phonological approaches to investigation, on peculiarities of representative and statistical investigations, on the theoretical and practical application of the obtained results and factors.

СПЕКТРАЛЬНЫЙ АНАЛИЗ И ПРОБЛЕМА СОКРАЩЕННОГО ОПИСАНИЯ РЕЧЕВЫХ СИГНАЛОВ

В данной статье обсуждаются некоторые общие методы решения задачи минимизации речевых сигналов. Отмечаются преимущества нелинейных разложений.

Трудность решения задачи в общем виде вынуждает обратиться к эвристическим методам. Из физических соображений следует, что описание речевого сигнала посредством задания параметров эквивалентного источника более экономно, чем, например, осциллографическое описание. Проблема сокращенного описания может быть сведена, таким образом, к нахождению наилучших оценок неизвестных параметров эквивалентного источника речевых сигналов. Спектральный анализ, выявляющий основные свойства источника (артикуляционного аппарата), рассматривается как один из методов получения сокращенного описания речевых сигналов. Отмечаются недостатки спектральных представлений нестационарных звуков. Анализ эквивалентной передаточной функции артикуляционного тракта свободен от некоторых недостатков спектрального анализа. Статистическими методами решается задача синтеза субоптимального анализатора параметров модели источника речевых сигналов.

Одна из основных задач, решаемых специалистами по рече-преобразованиям, состоит в следующем: дан речевой сигнал

$f(t)$, требуется описать этот сигнал минимальным числом двоичных единиц, но так, чтобы по этому описанию можно было восстановить сигнал $f^*(t)$, отличающийся от $f(t)$ не

больше, чем на некоторую величину ε , причем степень отличия функции $f(t)$ от $f^*(t)$ определяется функцией потерь $\mathcal{L}[f(t), f^*(t)]$. Эта задача имеет прямое отношение к проблеме выделения информативных признаков речи, к проблеме эффективной передачи речевых сигналов по каналам связи и к проблеме автоматического опознавания речи. До последнего времени задача эта решалась в основном эвристическими методами. Здесь мы обсудим некоторые аналитические методы решения этой задачи, а также укажем на трудности, связанные с применением аналитических методов.

Решение рассматриваемой задачи в самом общем виде сводится к построению минимальной ε -сети [1]. Если задано число n элементов сети, то задача состоит в нахождении n функций $f_1(t), f_2(t), f_3(t), \dots, f_n(t)$ из условия (I):

$$\int_{\Omega} P[f(t)] \cdot \left\{ \min_i \mathcal{L}[f(t), f_i(t)] \right\} df(t) \rightarrow \min_{f_1(t), \dots, f_n(t)}. \quad (I)$$

Здесь Ω - пространство реализаций функций $f(t)$

$P[f(t)]$ - функционал вероятностей реализаций речевого сигнала $f(t)$.

Таким образом, при нахождении элементов $f_i(t)$ из условия (I) учитываются как расстояния реализаций речевых сигналов до ближайших элементов сети, так и вероятности различных реализаций.

Для определения элементов $f_i(t)$ по формуле (I) необходимо знать функции $\mathcal{L}[f(t), f_i(t)]$ и $P[f(t)]$. Обе эти функции в настоящее время известны лишь весьма приблизительно. В большинстве случаев приемником речевого сигнала $f^*(t)$

является слуховой аппарат человека, поэтому функция потерь

$\mathcal{L}[f(t), f^*(t)]$ определяется свойствами слуха. Отсутствие единой теории слухового восприятия не позволяет в настоящее время установить окончательное выражение для

$\mathcal{L}[f(t), f^*(t)]$. Фонетическая функция А.А.Пирогова [2] может служить основой для построения $\mathcal{L}[f(t), f^*(t)]$.

Выражение для функционала вероятностей $P[f(t)]$ в настоящее время также неизвестно, известны лишь одномерные и двумерные распределения вероятностей речевых сигналов./

Правда, элементы минимальной ϵ - сети можно найти и не зная указанных функций, если воспользоваться методами случайного поиска [3], корректируя направление поиска по субъективным реакциям слушателей. Предварительная оценка показывает, что ввиду большой размерности m пространства функций $f(t)$ окажется необходимым производить огромное число итераций, с другой стороны, использование субъективных реакций слушателей исключает возможность эффективного применения быстродействующих вычислительных машин для поиска оптимума.

Порядок величины m можно установить из следующих соображений. Ширина спектра речевого сигнала $f(t)$ составляет несколько кГц. Длительность T анализируемого участка функции $f(t)$ должна быть не меньше длительности звука (в противном случае сегментация с последующим анализом каждого сегмента по отдельности может приводить к потере информации) и составлять, по крайней мере, десятки миллисекунд. Если теперь представить речевой сигнал $f(t)$ на рассматриваемом сегменте рядом Котельникова, то нетрудно под-

считать, что число членов m такого разложения исчисляется сотнями. Это значит, что для представления функции $f(t)$ в виде вектора, координатами которого являются отсчеты $f(t)$, выбранные в соответствии с теоремой Котельникова, нам потребуется задать сотни координат. Столь большая размерность пространства векторов, описание которых надлежит минимизировать, является основным препятствием применения аналитических методов к решению задачи сокращенного описания речевых сигналов.

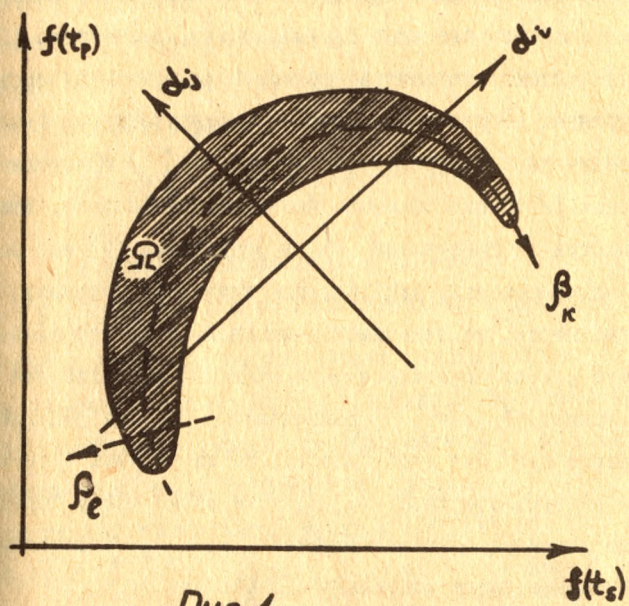
Разложение речевого сигнала $f(t)$ в ряд Котельникова не является оптимальным, поскольку координаты такого разложения оказываются коррелированными. Необходимое число членов разложения можно уменьшить до m_1 , если перейти к некоторой оптимальной системе базисных функций $\varphi_i(t)$:

$$f(t) \approx \sum_{i=1}^{m_1} a_i \varphi_i(t), \quad 0 \leq t \leq T. \quad (2)$$

Коэффициенты a_i описывают функцию $f(t)$. Методы нахождения оптимального (в смысле минимума среднеквадратичной ошибки аппроксимации (2)) базиса $\varphi_i(t)$ известны [4], функции $\varphi_i(t)$ вычисляются как собственные функции с наибольшими собственными значениями λ_i однородного интегрального уравнения

$$\int_0^T R(t, \tau) \varphi_i(\tau) d\tau = \lambda_i \varphi_i(t), \quad (3)$$

ядром которого служит корреляционная функция $R(t, \tau)$ речевого сигнала $f(t)$.



Представление (2) далеко не исчерпывает возможности сокращенного описания речевых сигналов. Дело в том, что малость среднеквадратичной ошибки аппроксимации (2) является достаточным условием хорошего воспроизведения речевых сигналов, но не является условием необходимым. Хорошо известно, что значительные искажения формы речевого сигнала в ряде случаев не приводят к существенным изменениям слухового восприятия этого сигнала. Таким образом, разложение (2) позволяет частично устранить статистическую избыточность, но не устраняет избыточность психо-физиологическую, обусловленную ограниченной (не больше 50 $\frac{\text{дБ.ед.}}{\text{сек.}}$ [5]) пропускной способностью слухового аппарата человека. Кроме того, некоррелированность коэффициентов α_i в разложении (2) еще не означает их независимости. Декорреляция позволяет устранить линейную зависимость, но зависимость нелинейная не устраняется. Например, если совместное распределение вероятностей для координат α_i и α_j равномерно в области Ω (см. рис. I) и равно нулю вне этой области, то из рисунка I очевидно, что система координат α_i, α_j не оптимальна, переход к криволинейной системе координат β_k, β_c позволил бы получить еще более сокращенное описание.

В качестве наглядного примера нелинейной зависимости между параметрами разложения (2) можно привести случай разложения спектральной функции $S(\omega)$ в ряд

$$S(\omega) \approx \sum_{i=1}^m \alpha_i \varphi_i(\omega)$$

причем базисные функции $\varphi_i(\omega)$ выбраны таким образом, что коэффициенты α_i некоррелированы. Поскольку $S(\omega) \geq 0$, коэффициенты α_i не могут принимать любых значений, их область ограничена гиперповерхностью

$$\sum_{i=1}^m \alpha_i \varphi_i(\omega) \geq -C, (0 < \omega < \omega_c) \quad (5)$$

где C — малая положительная величина;

ω_c — верхняя частота анализируемого спектра.

На этой особенности распределений коэффициентов основывается предложенный недавно А.А.Пироговым способ нелинейного восстановления первого коэффициента α_1 по высшим коэффициентам разложения спектральной функции.

Рассмотрим другой пример. Пусть $f(t) = A \sin(\omega t + \varphi)$, где $A = \text{const}$, ω и φ — независимые случайные величины с равномерными распределениями $P(\omega) = \frac{1}{\omega_c}, (0 < \omega < \omega_c)$, $P(\varphi) = \frac{1}{2\pi}$

$(0 < \varphi < 2\pi)$. Пусть длина интервала наблюдения $T \gg \frac{2\pi}{\omega_c}$. В этом случае функция автокорреляции процесса $f(t)$, полученная усреднением $f(t)f(t-\tau)$ по множеству реализаций, $R(\tau) = A^2 \frac{\sin \omega_c \tau}{2\omega_c \tau}$. Собственные функции интегрального уравнения (3) в этом случае есть функции вида $\frac{\sin(\omega_c t - i\pi)}{\omega_c t - i\pi}$, таким образом, оптимальным разложением типа (2) в рассматриваемом случае является разложение

Котельникова. Если взять две какие-нибудь координаты такого разложения и построить распределения этих координат при различных реализациях сигнала $f(t)$, то получим фигуру Лиссажу. Из всех возможных сочетаний координат в этом случае оказывается допустимой лишь малая часть; между координатами имеет место жесткая нелинейная зависимость, поэтому даже оптимальное линейное разложение (2) может существенно уступать нелинейным разложениям. В случае сигнала $f(t) = A \sin(\omega t + \varphi)$ такой результат очевиден: функция $f(t)$ полностью определяется всего двумя случайными величинами ω и φ , которые, однако, не выражаются линейно через отсчеты $f(t)$. Число координат при оптимальном линейном разложении $\frac{T \cdot \omega_0}{\pi} \gg 2$.

Как же найти оптимальное нелинейное преобразование m координат $\alpha_1, \alpha_2, \dots, \alpha_m$ в функцию $f(t)$? Широкий класс нелинейных преобразований может быть аппроксимирован суммами следующего вида:

$$f(t) \approx \sum_{i=1}^m \alpha_i \varphi_i(t) + \sum_{i=1}^m \sum_{j=1}^m \alpha_i \alpha_j \varphi_{ij}(t) + \underbrace{\sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^m \alpha_i \alpha_j \dots \alpha_k \varphi_{ijk}(t)}_P \quad (6)$$

Оптимальные функции $\varphi_1(t), \varphi_2(t), \dots, \varphi_m(t), \varphi_{11}(t), \dots, \varphi_{mm}(t), \varphi_{12}(t), \dots, \varphi_{pm}(t)$

могут быть найдены из условий минимизации интеграла

$$\int_0^T P[f(t)] \left\{ \min_{\alpha_{ij}} \int_0^T [f(t) - \sum_{i=1}^m \alpha_i \varphi_i(t) - \sum_{i=1}^m \sum_{j=1}^m \alpha_i \alpha_j \varphi_{ij}(t) - \dots \sum_{i=1}^m \sum_{j=1}^m \dots \sum_{k=1}^m \alpha_i \alpha_j \dots \alpha_k \varphi_{ijk}(t)]^2 df(t) \right\} = \min_{\varphi_{ij}} \quad (7)$$

$i = 1, 2, \dots, m \quad k = 1, 2, \dots, m$
 $j = 1, 2, \dots, m$
 $\dots \dots \dots$

Из условий (7) следует, что функции $\varphi_{ij} \dots$ однозначно определяются статистическими свойствами речевых сигналов, задаваемыми функционалом вероятностей $P[f(t)]$. Линейное разложение (2) можно рассматривать как частный случай нелинейного разложения (6); для этого частного случая из (7) можно получить уравнение (3). В частном случае линейных разложений, в соответствии с уравнением (3), достаточно задать вероятностные свойства процесса $f(t)$ моментами до второго порядка включительно. В более общем случае нелинейных разложений (6) потребуется знание моментов высших порядков. Экспериментальное измерение моментов выше второго для координат векторов, размерность которых исчисляется сотнями, представляет задачу чрезвычайной сложности. Интересно исследовать возможность вычисления функций $\varphi_{ij} \dots (t)$ методами стохастической аппроксимации [7], не требующими задания функционала $P[f(t)]$. Кроме того, вычисление координат α_i по заданной реализации $f(t)$ в соответствии с формулой (6), а также синтез функции $f(t)$ по заданным координатам представляют серьезные технические трудности. Поэтому от такого общего подхода к решению задачи минимизации описания речевых сигналов приходится пока отказаться. Аналитический подход может оказаться полезным при сравнительно малом числе m координат α_i (m порядка десяти), например, в случае аппроксимации спектральных функций. Другой областью применения аналитических методов является анализ речи. К рассмотрению некоторых вопросов нахождения структуры субоптимальных анализаторов речевых сигналов мы и переходим.

Трудность решения задачи минимизации описания в общем виде заставляет обратиться к эвристическим методам. Из физических соображений следует, что сокращение избыточности можно получить при переходе от непосредственного описания речевого сигнала $f(t)$ к описанию характеристик источника речевых сигналов. Представляя, как обычно, источник в виде линейного фильтра с переменными параметрами (импульсная реакция $h(t, \tau)$), возбуждаемого по входу последовательностью коротких импульсов¹⁾, следующих с частотой основного тона, получим следующую модель звонкого звука:

$$f(t) \approx \sum_{j=1}^n h(t-t_j, t_j) \quad (8)$$

Здесь через n обозначено число импульсов основного тона на рассматриваемом интервале наблюдения $0 < t \leq \tau$; t_j - момент прихода j -го импульса возбуждения.

При аппроксимации эквивалентной импульсной реакции артикуляционного тракта суммой типа (2), модель звонкого звука представится в виде

$$f(t) = \sum_{i=1}^m \sum_{j=1}^n \alpha_i(t_j) \varphi_i(t-t_j) + f(t), \quad (9)$$

где случайные функции $\alpha_i(t)$ описывают эквивалентную импульсную реакцию артикуляционного тракта. Поскольку всякая модель речевого сигнала является приближительным и упрощен-

1) Импульсы возбуждения артикуляционного тракта могут иметь более сложную форму. Такие импульсы можно представить как результат прохождения коротких импульсов через линейный формирующий фильтр. В дальнейшем мы будем считать, что источник речевого сигнала представлен последовательным соединением формирующего фильтра и артикуляционного тракта. Эту систему мы будем называть эквивалентным артикуляционным трактом. Эквивалентный артикуляционный тракт на звонких звуках возбуждается по входу короткими импульсами.

ным описанием действительного сигнала, необходимо ввести случайную функцию $\xi(t)$, учитывающую отличия модели от реального процесса. $\xi(t)$ мы будем условно называть шумовой составляющей. Распределение вероятностей для $\xi(t)$ нам неизвестно; имеет смысл воспользоваться минимаксной стратегией и принять для $\xi(t)$ наихудшее распределение, в данном случае мы будем считать $\xi(t)$ нормальным процессом с равномерным спектром. К нормальному распределению шумовой составляющей можно прийти и иначе, исходя из центральной предельной теоремы, ибо присутствие $\xi(t)$ обусловлено большим числом разнообразных и не учтенных в модели факторов примерно одного порядка малости.

Представление (9) напоминает формулу (2), функция $f(t)$ линейно выражается через коэффициенты $\alpha_i(t_j)$. Но следует указать и на отличие разложений (2) и (9). В разложении (2) базисные функции $\varphi_i(t_j)$ фиксированы и рассчитаны под средние статистические характеристики речевого сигнала $f(t)$, в разложении (9) базисные функции фиксированы не полностью, но частично подстраиваются (сдвиг по оси времени на случайные величины t_j) на основании анализа локальных свойств анализируемого сигнала.

Формула (10) дает другую аппроксимацию звонких звуков $f(t)$:

$$\sum_{i=1}^m \beta_i(t) f^{(i)}(t) = \sum_{j=1}^n \delta(t - t_j) + \xi(t) \quad (10)$$

Здесь речевой сигнал $f(t)$ представляется как нелинейная функция коэффициентов дифференциального уравнения, описывающего поведение эквивалента артикуляционного аппарата.

Представление (10) удобно для аппроксимации узкополосных процессов. Так, для аппроксимации рассмотренного выше сигнала $f(t) = A \sin(\omega t + \varphi)$ достаточно всего двух членов в левой части разложения (10).

Модели (9) и (10) задают речевой сигнал $f(t)$ через параметры эквивалентного источника $\alpha_i(t), \beta_i(t)$ и t_j . Задача анализа состоит в том, чтобы по заданной реализации $f(t)$ определить параметры эквивалентного источника. До последнего времени эта задача решалась в основном эвристическими методами. Так, для определения передаточной функции артикуляционного тракта обычно рассчитывают огибающую модуля мгновенного спектра. При анализе нестационарных звуков такой подход имеет ряд недостатков. Дело в том, что огибающая модуля мгновенного спектра нестационарных звуков есть нечто весьма неопределенное и условное, зависящее как от объективных свойств речевого сигнала, так и от выбранного определения мгновенного спектра и от способа построения его огибающей. При изменениях периодов следования возбуждающих импульсов мгновенный спектр источника возбуждения неравномерен [11], поэтому огибающая модуля мгновенного спектра речи определяется не только передаточной функцией артикуляционного тракта, но и характеристиками источника возбуждения, в результате спектральная картина усложняется. Для нахождения спектральных характеристик обычно используются анализаторы на резонаторах, но для таких анализаторов справедливо соотношение неопределенностей: чем выше разрешающая способность по частоте, тем больше время анализа, тем хуже передаются временные характеристики речи. Если учесть, что ширина пер-

вой форманты составляет около 40 гц, а длительность взрывных звуков - несколько миллисекунд, то получим, что для правильной передачи спектральных и временных характеристик речи произведение разрешающей способности анализатора по частоте Δf на разрешающую способность по времени Δt должно быть много меньше единицы:

$$\Delta f \Delta t \ll 1$$

(II)

Удовлетворить условию (II) в обычном анализаторе спектров на резонаторах принципиально невозможно, однако было бы неверно думать, что вообще принципиально невозможно построить анализатор, сочетающий высокую разрешающую способность с высоким быстродействием. Такие анализаторы существуют, слуховой аппарат человека может служить примером анализатора с $\Delta f \Delta t \ll 1$. Применением нелинейных методов обработки в сочетании с принципом адаптации можно преодолеть ограничение соотношения неопределенностей.

Как же построить анализатор, сочетающий высокую разрешающую способность с большим быстродействием? До последнего времени работы по созданию новых систем анализа шли главным образом по пути изобретательства. Недавно наметилась другая тенденция. Все чаще появляются описания анализаторов речи, моделирующих процессы переработки информации в слуховом аппарате человека. Оба указанных подхода имеют свои достоинства и недостатки. Здесь мы кратко рассмотрим третий возможный подход к решению задачи о нахождении оптимального анализатора речевых сигналов.

Структуру оптимального анализатора можно найти математически, если задаться моделью речевого сигнала и критерием

качества. В общем виде задача сводится к нахождению наилучших оценок векторов A и T , описывающих передаточную функцию эквивалентного артикуляционного тракта и характеристики возбуждения (для звонких звуков). Если воспользоваться Байесовой стратегией [7], то можно показать, что наилучшие оценки A^* и T^* должны находиться из условия минимизации среднего ожидаемого значения функции потерь $\mathcal{L}(A, T, A^*, T^*)$, задающей штрафа за ошибочные измерения параметров A^* и T^* вместо истинных A и T :

$$\overline{\mathcal{L}(A^*, T^*)} = \int_{\Omega} \mathcal{L}(A, T, A^*, T^*) P_{pr}(A, T) P[f^{(t)}/A, T] d\Omega = \min_{A^*, T^*}$$

Здесь Ω - пространство векторов A и T ;

$P_{pr}(A, T)$ - априорная плотность вероятностей векторов A, T ;

$P[f^{(t)}/A, T]$ - функционал вероятностей реализаций речевого сигнала $f(t)$ при заданных значениях параметров A и T .

Для заданной реализации $f(t)$ из (II) можно получить оптимальные оценки A_0^* и T_0^* , минимизирующие $\mathcal{L}(A^*, T^*)$. Алгоритмы преобразования $f(t)$ в A_0^* и T_0^* удалось получить для случая квадратичной аппроксимации функции потерь. Кроме того, пришлось ввести целый ряд дополнительных упрощающих предположений. Например, предполагалось, что анализатор не производит обработку принимаемого сигнала одновременно по всему интервалу наблюдения, а производит последовательный анализ, причем информация о форме предыдущих импульсов основного тона используется как априорная для анализа после-

дующих импульсов; использовалась гауссовская аппроксимация P_{pe} (A,T). При использовании этих (и некоторых других) упрощений для модели (9) методами теории статистических решений получена схема субоптимального анализатора, близкая к известной схеме адаптирующегося фильтра [8], предназначенного для приема сигналов неизвестной формы. Для модели (10) алгоритм обработки сводится к вычислению достаточных статистик вида $\int_0^T f^{(i)}(t) f^{(k)}(t) dt$ и к решению системы линейных уравнений, коэффициентами которой служат указанные достаточные статистики, а неизвестными — коэффициенты β_i уравнения (10). В систему уравнений входят также параметры априорных распределений и члены вида $\beta_i f^{(i)}(t_j)$. Таким образом, для определения свойств эквивалентного артикуляционного тракта оказывается необходимым знать характеристики возбуждения (t_j). Аналогичный вывод получен и для модели (9). Схема анализатора, оценивающего моменты t_j для модели (10) оказалась близкой к известной схеме выделения основного тона по методу обратной фильтрации [9]. Основное отличие полученного алгоритма от алгоритма [9] состояло в появлении дополнительного канала квадратичной обработки сигнала на выходе обратного фильтра, причем роль канала с квадратичной обработкой велика в начале работы, когда имеется существенная неопределенность относительно характеристики обратного фильтра. В частном случае одноформантных звуков полученный алгоритм субоптимальной обработки совпадает с известным методом выделения частот формант по моментам спектра [10].

Л и т е р а т у р а

1. А.Г.Витушкин. Оценка сложности задачи табулирования. Государственное издательство физико-математической литературы. Москва. 1959 г.
2. А.А.Пирогов. К вопросу о фонетической функции речи. Электросвязь, № 5, 1967г.
3. Л.А.Растринг. Случайный поиск в задачах оптимизации многопараметрических систем. Издательство "Зинатне". Рига. 1965 г.
4. В.И.Куля. Экспериментальное исследование корреляционных соотношений в спектре речи и сравнение некоторых вариантов ортогональных вокодеров. Электросвязь. № 4. 1964 г.
5. K. Kufmuller. Informationsverarbeitung durch den Menschen. Nachrichtentechnische Zeitschrift 1959, №2.
6. Я.З.Цыпкин. О восстановлении характеристики функционального преобразователя по случайно наблюдаемым точкам. Автоматика и телемеханика. 1965 г., т. XXVI, № II.
7. Д.Миддлтон. Введение в статистическую теорию связи. Издательство "Советское радио". Москва, 1962 г., т. П.
8. С.М.Флейшер. Основы теории и расчета одного класса самоорганизующихся приемных устройств. Диссертация кандидатская. Ленинград. ЛЭИС. 1965..

9. G. Faut. Sound Spectrography. Proceedings of the Fourth International Congress of Phonatic Sciences. 1962.
10. J. Suzuki a.o. Formant - Frequency Extraction by the Method of Moment Calculation. The Journal of the Acoustical Society of America. 1963 v. 35 N 9.

S.P. Baronin

Scientific Research Radio Institute.
MoscowSPECTRUM ANALYSIS AND PROBLEM OF ABBREVIATED DESCRIPTION
OF SPEECH SIGNALS

Summary

Described in the paper are some general methods of solving a minimization problem of speech signals. The paper shows the advantage of non-linear resolution. Since it is difficult to solve a general problem we resort to heuristic methods.

Guided by physical considerations one can say that speech signal description via setting parameters of equivalent source is more economical than, for instance, by means of an oscillograph finding the optimal evaluations of unknown parameters of speech signal equivalent source. The spectrum analysis which shows the signal equivalent source. The spectrum analysis which shows the basic properties of the source (articulation basis) is considered to be one of the methods enabling to obtain an abbreviated speech signal description. Some faults of spectrum representation of non-stationary sounds are observed. Analysis of the equivalent transit function of articulation channel is free of certain faults pertaining to spectrum analysis technique. The problem of synthesis of suboptimal analyzer of the model parameters of speech signal source is solved by means of statistical approach.

К ВОПРОСУ О ФОНЕТИЧЕСКОМ КОДИРОВАНИИ ТЕЛЕФОННОГО СИГНАЛА

Синтетическая телефония достигла значительных успехов в создании вокодеров, т.е. систем, кодирующих речевой сигнал на передающей стороне и декодирующих на приемной таким образом, что по одному телефонному каналу передается одновременно несколько десятков разговоров вместо одного. Достаточно сказать, что гармонические вокодеры позволяют передавать речь с помощью 2-3 тысяч двоичных единиц информации в секунду в натуральном масштабе времени, тогда как пропускная способность телефонного канала достигает 40-50 тысяч дв.ед./сек. При этом соблюдаются необходимые требования к качеству воспроизведения речи на приемной стороне [2].

Возникает естественный вопрос: каковы пределы сжатия информации, заключенной в речевом сигнале? Окончательный ответ на этот вопрос будет найден только при достаточно высоком уровне развития психолингвистики. Тем не менее уже сейчас можно сказать, что в системах, использующих смысловую запись, будет необходимо присутствие, по крайней мере, большого словаря, включающего все слова, встречающиеся в телефонных разговорах. В системах, ориентированных на прием и переработку звучащей речи хранение всех слов такого словаря в виде их спектрограмм (или осциллограмм) представляет значительные трудности (на 2-3 порядка больше, чем запись в памяти этих слов в виде фонетических знаков).

Таким образом, мы приходим к необходимости хранения, а, следовательно, и автоматического поиска в речевом сигнале более коротких единиц речи: слогов, фонем (звукотипов)*), или еще более мелких, чем фонемы, отрезков речевого сигнала.

Предположим, что мы сумели построить фонемный вокодер, т.е. систему, передающую вместо телефонного сигнала номера произнесенных фонем (и снова синтезирующую речь по этой транскрипции на приемной стороне). Подсчитаем объем информации, передаваемой таким образом за 1 сек. по каналу связи. Возьмем систему звуков русского языка; в ней содержится около 70 звукотипов [3]. Для обозначения номера каждого из них потребуется 7 двоичных знаков:

$$\log_2 70 \approx 7 \frac{96 \text{ б.ег.}}{\text{фон.}}$$

Средняя скорость произнесения фонем — примерно 10 фонем/сек. Отсюда средняя скорость передачи информации по каналу связи в фонемном вокодере

$$V = 7 \frac{96 \text{ б.ег.}}{\text{фон.}} \cdot 10 \frac{\text{фон.}}{\text{сек.}} = 70 \frac{96 \text{ б.ег.}}{\text{сек.}},$$

что на порядок меньше объема информации, передаваемой за 1 сек. по каналу связи в гармоническом вокодере.

Снова может возникнуть вопрос: каким образом и с каким успехом слова и фразы, произнесенные разными людьми в различных акустических условиях перейдут в фонемном вокодере в одинаковую транскрипцию при кодировании на передающем устройстве? Иными словами, что представляет собой та ин-

*) Вернее, переходов между ними (об этом см. дальше).

формация, которую мы выбрасываем при переходе от записи звучащей речи в виде осциллограмм или спектрограмм к записи звучащей речи в виде последовательности фонетических знаков.

Ответ на этот вопрос зависит от того, на какие единицы речи и с какой точностью мы делим при кодировании телефонный сигнал. Во всяком случае, выброшенная информация представляет собой шум, ненужный при данном способе кодирования. Например, при чисто фонемном кодировании нам не важна индивидуальная характеристика диктора так же, как и особенности акустических условий (что применимо в коммерческой телефонии). Для частной же телефонной связи придется передавать дополнительно к номерам фонем индивидуальную характеристику диктора, по-прежнему абстрагируясь от акустических условий.

Различные методы обработки речевой информации могут опираться на два представления речевого сигнала:

- а) в виде непосредственной осциллографической кривой,
- б) в виде мгновенного спектра речи.

Первый способ предусматривает обработку гораздо большего количества исходной информации, чем потребуется по второму способу. Кроме того, осциллограммы речи весьма сложны по своей форме. При классификации образов соответствующие им области могут обладать значительно более сложными границами, чем области, получаемые для мгновенных спектров речи. В осциллограмме речи имеются повторяющиеся участки, от которых можно освободиться при взятии мгновенного спектра речи, медленно меняющегося во времени на каждой частоте. Мож-

но думать, что это свойство мгновенного спектра отражает медленные изменения артикуляционного аппарата при речевых поступках.

Представляя речевой сигнал в виде непрерывно меняющегося во времени мгновенного спектра, А.А.Пирогов [1] предлагает считать шум, налагаемый на речевой сигнал различными акустическими условиями и индивидуальными особенностями диктора, состоящим из 2-х частей - мультипликативной и аддитивной:

$$S(\omega, t) = K(\omega, t) Q(\omega, t) + N(\omega, t) \quad (1)$$

где $S(\omega, t)$ - мгновенный спектр зашумленной речи;
 $K(\omega, t)$ - мультипликативная помеха;
 $Q(\omega, t)$ - "истинный" спектр речи;
 $N(\omega, t)$ - аддитивная помеха.

Если считать далее, что мультипликативная помеха $K(\omega, t) \approx k(\omega)$, т.е. достаточно медленно меняется во времени, например, диктор не очень сильно меняет свое положение относительно микрофона, а аддитивная помеха $N(\omega, t)$ достаточно мала, то мы можем избавиться от мультипликативной помехи, например, с помощью преобразования (см. [1]):

$$I_1(\omega, t) = \frac{1}{T} \int_0^T e^{-\frac{j}{T}\tau} \ln \frac{|S(\omega, t)|}{|S(\omega, t-\tau)|} d\tau \quad (2)$$

где T - постоянная времени интегрирования;
 $S(\omega, t)$ - мгновенный спектр зашумленной речи в момент времени t ;
 $S(\omega, t-\tau)$ - мгновенный спектр в момент времени $t-\tau$;
 τ - переменная интегрирования.

Действительно, учитывая формулу (1) и считая

$$N(\omega, t) = N(\omega, t - \tau) = 0$$

$$K(\omega, t) = K(\omega, t - \tau) = k(\omega)$$

для всех τ , мы сможем сократить $k(\omega)$ в числителе и знаменателе дроби интеграла (2). Полученная в результате интегрирования функция $I_1(\omega, t)$ (названная в [1] "фонетической функцией") не зависит от медленно меняющейся характеристики $k(\omega)$ при достаточно малом шуме $N(\omega, t)$. Указанным свойством обладает слуховое восприятие человека, понимающего речь независимо от акустических условий, если только они не меняются слишком часто по сравнению со скоростью артикуляции говорящего.

Аналогичную функцию, инвариантную относительно акустических условий, мы получим, действуя по следующей формуле:

$$I_2(\omega, t) = \ln \frac{|S(\omega, t)|}{\frac{1}{T} \int_0^T e^{-\frac{\omega \tau}{2}} |S(\omega, t - \tau)| d\tau}, \quad (3)$$

где берется логарифм отношения мгновенного спектра в момент времени t к его среднему значению за весь предшествующий период.

Обратим внимание читателя на то обстоятельство, что обе функции (2) и (3) обращаются в нуль при неизменном спектре, т.е. при

$$S(\omega, t) = S(\omega, t - \tau) = S(\omega),$$

* Это значит, что перенос эксперимента в другие акустические условия (с другой характеристикой $k(\omega)$) может существенно изменить мгновенный спектр речи, но несколько не затронет фонетическую функцию.

независимо от значений спектра $S(\omega)$ на разных частотах. Иными словами, согласно гипотезе о "фонетической функции", постоянный во времени мгновенный спектр не несет никакой полезной информации, так как фонетическая функция в этом случае не зависит от распределения мощности речевого сигнала по частотам. Соответствует ли этот вывод действительности? Обратимся к экспериментальным данным.

Опыты по идентификации звуков, проводившиеся Л.В.Бондарко и Л.Р.Зиндером [4], показали важность окружения звука для его правильного опознания. "Ядро" звука, оторванное от своего окружения, часто идентифицировалось аудиторами неадекватно, а воспроизведенное вместе с окружением распознавалось верно. Отсюда следует, что в "ядре" фонемы, которому соответствует квазипостоянный участок мгновенного спектра речи, содержится далеко не вся информация о произносимой фонеме.

К аналогичному выводу об информационной бедности квазипостоянных участков спектра приводят опыты по пересадке "ядра" фонемы, произнесенной в одних акустических условиях, в другие слова, где "ядро" начинало звучать как фонема, отличная от первой [5 - 6]. К выводу о важности переходных участков между фонемами, в отличие от их "ядра", приходят многие лингвисты [7-9], [12-14]. Но отсюда непосредственно следует, что бессмысленно требовать минимума средне-квадратичного отклонения спектров одного слова (или фразы), произнесенного в разных акустических условиях, т.к. сам по себе спектр может быть каким угодно, в зависимости от частотной характеристики $k(\omega)$, и наоборот, равенство некото-

рых участков двух спектров еще не означает совпадения фонетических значений этих участков.

Исходя из вышеизложенного, мы можем заключить следующее.

1. Заслуживают внимания только переходы из одной фонемы в другую.

2. Нельзя сравнивать спектры речи по среднеквадратичному (или ему подобному) критерию.

3. Фонетические функции, инвариантные относительно частотной характеристики тракта, можно сравнивать по среднеквадратичному критерию.

4. При желании обойтись без предварительного преобразования спектра в фонетическую функцию необходимо учитывать п.2, иначе говоря, критерий непосредственного сравнения двух спектров должен исключать их частотные характеристики (в общем случае отличные друг от друга).

Чтобы пояснить п.4, приведем пример критерия непосредственного сравнения двух спектров, учитывающего п.2. Пусть $S'(\omega, t)$ и $S''(\omega, t)$ суть два спектра, сравниваемых на участке времени длительностью τ . Пусть $k'(\omega)$ и $k''(\omega)$ — неизвестные нам частотные характеристики, входящие в спектры в качестве мультипликативных помех. Пусть имеются, кроме того, аддитивные помехи $N'(\omega)$ и $N''(\omega)$, также неизвестные нам и не зависящие от времени:

$$\left. \begin{aligned} S'(\omega, t) &= k'(\omega) Q'(\omega, t) + N'(\omega), \\ S''(\omega, t) &= k''(\omega) Q''(\omega, t) + N''(\omega), \end{aligned} \right\} \quad (4)$$

где $Q'(\omega, t)$ и $Q''(\omega, t)$ - "истинные" спектры, степень близости которых нам нужно узнать.

Зададимся частотой $\omega = \text{const}$ и будем сравнивать спектры на этой частоте как функции времени:

$$\left. \begin{aligned} S'(t) &= k'Q'(t) + N' \\ S''(t) &= k''Q''(t) + N'' \end{aligned} \right\} \quad (5)$$

Тогда при равных истинных спектрах

$$Q'(t) \equiv Q''(t), \quad (6)$$

где тождество по предположению выполняется на всем участке τ , наши мгновенные спектры будут линейно зависимы между собой. Действительно, умножим I-е уравнение (5) на k'' и вычтем из него 2-е, умноженное на k' . Тогда:

$$k''S'(t) - k'S''(t) = k''N' - k'N'', \quad (7)$$

где "истинные" спектры взаимно уничтожились благодаря тождеству (6). Уравнение (7) и показывает линейную зависимость мгновенных спектров, т.к. коэффициенты при них и правая часть (7) суть числа, не зависящие от времени.

Предположим теперь, что данные мгновенные спектры линейно зависимы на отрезке времени τ :

$$AS'(t) + BS''(t) = C \quad (8)$$

где A и B одновременно не равны нулю:

$$A^2 + B^2 \neq 0 \quad (9)$$

Тогда можно подобрать (на данной частоте ω) такие помехи k, k', N, N'' , что "истинные" спектры будут тождественно равны (см. (6)) на отрезке τ . В самом деле, подставим в (8) равенства (5):

$$Ak'Q(t) + AN' + Bk''Q(t) + BN'' = C. \quad (10)$$

Так как нам нужно добиться тождества в (10), то, учитывая (6) и приравнявая коэффициенты в левой и правой частях (10), получим:

$$Ak' + Bk'' = 0, \quad (11)$$

$$AN' + BN'' = C. \quad (12)$$

В (11) можно положить с точностью до множителя при "истинном" спектре:

$$k' = B, \quad k'' = -A. \quad (13)$$

Для нахождения аддитивных помех будем считать необходимым достижение минимума суммы их энергий:

$$\min [(N')^2 + (N'')^2] \quad (14)$$

при соблюдении зависимости (12).

Выражение (14) представляет собой квадрат расстояния на плоскости в Евклидовой метрике от некоторой точки прямой (12) до начала координат. Минимальное расстояние (14) дает точка прямой (12) с координатами:

$$N' = \frac{AC}{A^2 + B^2}, \quad N'' = \frac{BC}{A^2 + B^2}. \quad (15)$$

Таким образом, мы нашли все обещанные помехи (13) и (15) и, стало быть, можем записать, согласно (5):

$$\left. \begin{aligned} S'(t) &= BQ'(t) + \frac{AC}{A^2 + B^2} \\ S''(t) &= -AQ''(t) + \frac{BC}{A^2 + B^2} \end{aligned} \right\} \quad (16)$$

Проверим тождество (6) "истинных" спектров. Для этого подставим мгновенные спектры (16) в соотношение (8) их линейной зависимости:

$$ABQ'(t) + \frac{A^2 C}{A^2 + B^2} - ABQ''(t) + \frac{B^2 C}{A^2 + B^2} = C$$

откуда получим:

$$ABQ'(t) - ABQ''(t) = 0$$

и

$$Q'(t) \equiv Q''(t) \quad (6)$$

при условии

$$AB \neq 0 \quad (17)$$

Последнее условие, необходимое для выполнения тождества (6), означает неравенство нулю обоих коэффициентов A и B линейной комбинации (8). Посмотрим, что означает равенство нулю одного из них, например: $A = 0$.

Тогда (8) примет вид

$$BS''(t) = C,$$

что при $B \neq 0$ может быть записано как

$$S''(t) = \frac{C}{B},$$

т.е. постоянство второго мгновенного спектра на отрезке времени T . Отсюда следует, что постоянный спектр линейно зависит с любым спектром $S'(t)$. При этом (16) превращается в

$$S'(t) = BQ'(t)$$

$$S''(t) = \frac{C}{B}$$

т.е. постоянный спектр $S''(t)$ принимается просто за аддитивную помеху, а переменный спектр $S'(t)$ совпадает с "истинным"

с точностью до постоянного множителя. Очевидно, что ни о каком общем "истинном" спектре в этом случае не может быть и речи.

Итак, нами доказана следующая

Теорема. Для проверки гипотезы о том, что два мгновенных спектра $S'(\omega, t)$ и $S''(\omega, t)$ имеют на отрезке времени τ равные "истинные" спектры:

$$Q'(\omega, t) = Q''(\omega, t),$$

необходимо и достаточно обнаружить существование невырожденной линейной зависимости между данными спектрами на отрезке времени τ при каждом значении частоты ω :

$$A(\omega) S'(\omega, t) + B(\omega) S''(\omega, t) = C(\omega) \quad (8')$$

$$A(\omega) B(\omega) \neq 0 \quad (17')$$

Из доказанной теоремы следует, что критерием для непосредственного измерения близости двух мгновенных спектров на отрезке времени τ при данном значении частоты ω может служить величина, характеризующая степень линейной зависимости между этими спектрами на частоте ω . Так как в спектре могут присутствовать некоторые нестационарные шумы, то указанная величина должна быть статистической мерой линейной зависимости. Такую роль может играть, например, коэффициент корреляции, если вычесть его из 1, с тем чтобы получить 0 при полной линейной зависимости спектров, т.е. при нулевом "расстоянии" между ними:

$$r_{\tau}(\omega) = 1 - \frac{\int [S'(\omega, t) - M'(\omega)] [S''(\omega, t) - M''(\omega)] dt}{\sqrt{\int [S'(\omega, t) - M'(\omega)]^2 dt} \sqrt{\int [S''(\omega, t) - M''(\omega)]^2 dt}}, \quad (18)$$

где

$$M'(\omega) = \frac{1}{\tau} \int_{\tau} S'(\omega, t) dt, M''(\omega) = \frac{1}{\tau} \int_{\tau} S''(\omega, t) dt \quad (19)$$

— средние значения спектров на частоте ω за время τ .
Формула (18) близости двух мгновенных спектров на частоте ω исключает влияние частотных характеристик, наложенных на эти спектры в виде множителей или слагаемых, не зависящих от времени на отрезке τ , тем самым удовлетворяя требованиям (1) — (4) гипотезы о существовании фонетической функции.

Это свойство формул типа (18) позволяет применять их в качестве критерия близости двух мгновенных спектров без предварительного преобразования исходных спектров в фонетические функции.

Критерием же близости двух фонетических функций, как было сказано ранее, служит просто среднеквадратичное "расстояние" между ними:

$$r_{\tau}(\omega) = \frac{1}{\tau} \int_{\tau} [I'(\omega, t) - I''(\omega, t)]^2 dt, \quad (18')$$

которое обратится в 0 на частоте ω , если эти функции полностью совпадают в течение времени τ на данной частоте.

Дальнейшее изложение будет относиться равным образом к обеим формулам (18) и (18'). Под словом "речь" будет при этом подразумеваться одно из выражений: "мгновенный спектр" или "фонетическая функция". Теперь, исходя из формул типа (18) или (18'), мы можем определить расстояние между двумя отрезками речи, усреднив, например, величину $r_{\tau}(\omega)$ по всей полосе частот Ω :

$$R_{\tau} = \frac{1}{\Omega} \int_{\Omega} r_{\tau}(\omega) d\omega \quad (20)$$

Таким образом, в любых двух данных словах или фразах мы можем найти наиболее совпадающие места, сдвигая их относительно друг друга и варьируя длительность τ времени сравнения. Опишем этот процесс математически.

Пусть M_1 и M_2 - сравниваемые отрезки речи длительностью соответственно T_1 и T_2 . Будем считать за начало отсчета времени в каждом отрезке речи его начало. Пусть τ - некоторая длительность, заданная для сравнения двух участков данных отрезков речи.

Тогда для любых двух моментов времени t_1 и t_2 , не больших соответственно $T_1 - \tau$ и $T_2 - \tau$, мы можем найти по формуле (20) расстояние $R_\tau(t_1, t_2)$ между участками длины τ , начинающимися с моментов t_1 и t_2 .

Зафиксируем теперь длительность τ и момент t_1 . Тогда, пройдя по всем моментам t_2 от 0 до $T_2 - \tau$, мы выберем значение t_2^0 , которое даст участок речи из M_2

$$[t_2, t_2 + \tau]$$

наиболее близкий данному участку

$$[t_1, t_1 + \tau]$$

из отрезка речи M_1 :

$$R_\tau(t_1, t_2^0) = \min_{0 \leq t_2 \leq T_2 - \tau} R_\tau(t_1, t_2). \quad (21)$$

Сдвигая затем момент t_1 начала участка из M_1 от 0 до $T_1 - \tau$ и находя для каждого значения t_1 по формуле (21) момент t_2^0 , дающий наименьшее расстояние между сравниваемыми участками, получим момент t_1^0 , минимизирующий вместе с найденным для него моментом t_2^0 расстояние между любыми

двумя участками длительности τ , взятыми из отрезков речи M_1 и M_2 :

$$R_{\tau}(t_1^0, t_2^0) = \min_{0 \leq t_i \leq T_i - \tau} R_{\tau}(t_i, t_i^0). \quad (22)$$

Пусть задана нижняя грань Θ значений длительности τ сравниваемых участков речи. Верхнюю грань Θ можно также задавать заранее, но можно и определять каждый раз по длительностям T_1 и T_2 , если последние не очень велики, а именно:

$$\Theta = \min_{i=1,2} T_i. \quad (23)$$

Изменяя τ в указанных пределах и отыскивая при каждом значении τ ближайшие друг к другу участки M_1 и M_2 по формулам (21) и (22), можно найти оптимальную для двух данных отрезков речи M_1 и M_2 длительность τ^0 наиболее совпадающих участков:

$$R_{\tau^0}(t_1^0, t_2^0) = \min_{\Theta \leq \tau \leq \Theta} P(\tau) R_{\tau}(t_1^0, t_2^0) \quad (24)$$

где $P(\tau)$ - некоторая неубывающая весовая функция от длительности τ .

Итак, мы умеем найти в двух данных отрезках речи M_1 и M_2 наиболее совпадающие (в выбранной нами метрике) участки, т.е. моменты t_1^0, t_2^0 их начала в каждом из этих отрезков речи, их оптимальную длительность τ^0 , а также даваемое ими минимальное расстояние

$$R^{1,2} = R_{\tau^0}(t_1^0, t_2^0). \quad (25)$$

Здесь верхние индексы 1,2 при R означают номера сравниваемых отрезков речи M_1 и M_2 . Как можно заметить из

предыдущего, расстояние (25) не меняется от перестановки верхних индексов:

$$R^{1,2} = R^{2,1} \quad (26)$$

Теперь встанут на очередь четыре вопроса:

1. Что считать за эталон для двух найденных одинаковых участков (выбор эталона необходим, так как расстояние между этими участками не всегда будет равно 0)?

2. Какое расстояние от эталона должно считаться допустимым*) для отождествления с ним таких же участков других отрезков речи при анализе по готовым эталонам?

3. В каких отрезках речи (словах и фразах) имеет смысл искать общие участки?

4. Какому сочетанию звуков или части звука нужно считать соответствующим найденный эталон?

Последний вопрос может интересовать скорее с точки зрения распознавания речевых образов, так как для узких целей синтетической телефонии неважно, что означают найденные эталоны, лишь бы их было конечное и не очень большое число. Тогда с их помощью можно будет провести анализ передаваемой речи, затем пустить по каналу связи соответствующие номера эталонов и, наконец, составить на приемной стороне синтезированную речь из эталонов по принятым номерам. Однако положительное решение 3-го вопроса поневоле приведет нас к решению 4-го.

*) Такое расстояние необходимо (при наличии неполного набора эталонов) для ответа "не знаю" в тех местах анализируемого отрезка речи, которые далеки от всех имеющихся в момент анализа эталонов.

На 3-й вопрос можно было бы ответить отрицательно, т.е. не подбирать для нахождения эталонов специальные пары слов или фраз, а сравнивать между собой все имеющиеся у нас отрезки речи без разбора.

Против такого решения вопроса можно возразить следующее:

а) при сравнении двух совершенно различных отрезков речи будет упорно искажаться и найдется "эталон" с большим минимальным расстоянием, не имеющим никакого значения для анализа речи;

б) из-за такого беспорядочного поиска будет проделано во много раз больше бесполезной (см. выше), чем необходимой для отыскания действительных эталонов работы;

в) нет смысла полностью игнорировать достижения фонетики в такой близкой к ней области, как фонетическое кодирование.

Считаясь с возражениями а) - в), мы приходим к необходимости задавать наряду со спектрами отрезков речи также фонетическую транскрипцию, благодаря которой мы сможем выбирать для поиска эталонов те пары отрезков речи, в транскрипции которых имеются одинаковые последовательности символов и не сравнивать спектры бессмысленных паросочетаний.

Отсюда, как побочный (!) продукт, получается положительный ответ на 4-й вопрос: каждому выбранному эталону будет ставиться в соответствие та последовательность фонетических знаков, которая послужила основанием для поиска эталона.

Приведем систему звукотипов русского языка, взятую нами из [3]. Согласные:

б в г γ д ж з л м н р д̂з - -

б' в' г' γ' д' ж' з' л' м' н' р' - д̂'ж'

п ф к х т ш с (л) (м) (н) (р) ц - -

п' ф' к' х' т' ш' с' (л') (м') (н') (р') - ч' -

всего 49 единиц.

Гласные:

ы, у о а э я ы³

ы' у' о' а' э' ь -

и у о а е ь и^е

и̂ у̂ о̂ а̂ е̂ - -

всего 25 единиц.

Учитывая требование I) гипотезы о существовании фонетической функции, мы должны искать эталоны, соответствующие хотя бы двум соседним звукотипам. Например, в словах (мисквА) и мАлѣш) это будет пара (мА).

Однако, сколько бы звукотипов ни было в каждом эталоне, по каналу связи всегда можно передать только необходимые звукотипы (без повторений). Например: МА-АСК-СКВ-ВА будет передано как М-А-С-К-В-А. А на приемном конце можно заново разбить на части принятую транскрипцию согласно имеющимся эталонам для синтеза речи.

Использование фонетической транскрипции не только значительно ускорит нахождение эталонов, но и позволит эффективно контролировать работу по их отысканию с помощью автоматического анализа речи, состоящей из уже найденных звуко сочетаний, но заданной без транскрипции.

Заметим, что, подключаясь к каналу связи в его середине, мы получим транскрипцию передаваемой речи, т.е. попутно (!) подойдем к решению еще одной задачи - построению автоматической пишущей машины (диктофона). Правда, текст будет печататься в таком виде [3]:

Нъ другóј дѣн' вѣст'
лплжари° рѣзні°спас'
пѣсѣи°му лкллѡтку

Причем не все звуки можно без словаря безошибочно перевести в буквы (например, безударный гласный ь или глухой согласный ф). Но это уже другая проблема - перевод транскрипции в орфографически правильный текст. Во всяком случае, вполне возможно прочесть транскрипцию без всякого перевода и использовать ее для контроля при выработке эталонов.

Для нетривиального ответа на I-й вопрос необходимо рассматривать, по крайней мере, три отрезка речи:

$$M_1, M_2 \text{ и } M_3$$

с одним и тем же звуко сочетанием. Запишем формулу минимального расстояния (25) в общем виде:

$$R^{ij} = R_{\tau^0}(t_i^0, t_j^0). \quad (25')$$

Будем считать (см. дальше), что длительность τ^0 одинакова для всех найденных участков.

Определим меру ρ^i дальности i -го участка речи как среднее расстояние этого участка до двух остальных:

$$\rho^i = \frac{1}{2} \sum_{\substack{j=1 \\ j \neq i}}^3 R^{ij}, \quad i=1,2,3 \quad (27)$$

При $i = 1$ формула (27) будет выглядеть так:

$$\rho^1 = \frac{1}{2} (R^{1,2} + R^{1,3}).$$

Тогда за эталон будет принят участок с наименьшей дальностью:

$$\rho^{i_0} = \min_{i=1,2,3} \rho^i \quad (28)$$

Формулы (27) и (28) можно обобщить на любое количество сравниваемых отрезков речи

$$M_1, M_2, \dots, M_K,$$

в том числе на $K = 1$, если взять среднее расстояние по всем j , включая $j=i$, с учетом свойства расстояния:

$$R^{ii} = 0$$

Тогда мера дальности i -го участка примет вид:

$$\rho_i = \frac{1}{K} \sum_{j=1}^K R^{ij}, i=1, \dots, K. \quad (27')$$

Эталон будет искажаться по формуле:

$$\rho^{i_0} = \min_{1 \leq i \leq K} \rho^i \quad (28')$$

где i_0 - номер участка с минимальной*) дальностью ρ^{i_0} , которую мы назовем средним радиусом искомого эталона. Последний находится в отрезке речи M_{i_0} , начиная с момента $t_{i_0}^0$ до $t_{i_0}^0 + \tau^0$, где τ^0 - найденная одновременно для всех K отрезков речи, оптимальная длительность эталона.

Назовем радиусом действия R_d эталона расстояние до самого удаленного от эталона участка:

$$R_d = \max_{1 \leq i \leq K} R^{i i_0} \quad (29)$$

х) Может при этом оказаться более выгодным выбор двух или более эталонов, а для расстояния по корреляции - также усредненных спектров (см. [15]).

формула (29) радиуса действия R_d эталона дает нам возможность ответить на 2-й вопрос: если при сравнении анализируемой речи с эталоном получится расстояние, не большее радиуса действия данного эталона, то можно отождествить сравниваемый участок речи с этим эталоном.

Итак, мы ответили на все 4 вопроса. Обобщим теперь формулы (21), (22) и (24) на любое количество сравниваемых отрезков речи.

Пусть $M_1, M_2, \dots, M_k$ -

некоторые отрезки речи, в транскрипции которых встречается одно и то же звуко сочетание. Пусть $T_1, T_2, \dots, T_k$ - соответствующие длительности данных отрезков речи. Зададимся нижней гранью θ изменения длительности τ сравниваемых участков. За верхнюю грань Θ изменения τ примем

$$\Theta = \min_{1 \leq i \leq k} T_i, \quad (23)$$

или заранее известную величину Θ , если только она не больше (23').

Для всякого значения длительности τ ($\theta \leq \tau \leq \Theta$) и набора значений моментов

$$\bar{t} = (t_1, t_2, \dots, t_k); 0 \leq t_i \leq T_i - \tau; i = 1, \dots, k,$$

мы можем построить матрицу расстояний

$$D(\tau, \bar{t}) = \|R_\tau(t_i, t_j)\|, i, j = 1, 2, \dots, k, \quad (30)$$

для соответствующих участков речи с границами:

$$[t_i, t_i + \tau] \quad i = 1, 2, \dots, k.$$

Усредняя матрицу (30) по строкам, найдем дальность каждого участка:

$$\rho_{\tau}^i(\bar{t}) = \frac{1}{K} \sum_{j=1}^K R_{\tau}(t_i, t_j), i=1, \dots, K, \quad (27'')$$

и средний радиус $\rho_{\tau}^{i_0}(\bar{t})$ данного множества участков:

$$\rho_{\tau}^{i_0}(\bar{t}) = \min_{1 \leq i \leq K} \rho_{\tau}^i(\bar{t}), \quad (28'')$$

где i_0 - номер "центрального" участка.

Средний радиус (28'') зависит от набора моментов $\bar{t}$ и длительности τ . Если зафиксировать последнюю, то, пользуясь формулами (30), (27'') и (28''), можно найти такой набор моментов

$$\bar{t}_0 = (t_1^0, t_2^0, \dots, t_K^0),$$

при котором средний радиус (28'') будет минимален:

$$\rho_{\tau}^{i_0}(\bar{t}_0) = \min_{\substack{0 \leq \tau \leq T_j - t \\ 1 \leq i \leq K}} \rho_{\tau}^{i_0}(\bar{t}) \quad (31)$$

Нам осталось лишь выбрать оптимальное значение длительности τ сравниваемых участков, которое автоматически (согласно только что изложенному методу) даст нам эталон (т.е. "центральный" участок при оптимальном τ) и его средний радиус:

$$\rho_{\tau_0}^{i_0}(\bar{t}_0) = \min_{0 \leq \tau \leq \Theta} \rho_{\tau}^{i_0}(\bar{t}_0) \quad (32)$$

При формуле (29) вычислим для найденного эталона радиус его действия R_a , необходимый для поиска и отождествления с этим эталоном подходящих участков речи при ее анализе.

Литература

1. А.А.Пирогов. К вопросу о фонетическом кодировании речи. "Электросвязь". № 5, 1967.
2. А.А.Пирогов. Синтетическая телефония. Связьиздат. Москва, 1963 г.
3. Н.С.Валгина, Д.Э.Розенталь, М.И.Фомина, В.В.Цапужевич. Современный русский язык. Издание 2-ое, Изд. Высшая школа. М., 1964.
4. Л.В.Бондарко, Л.Р.Зиндер, Л.А.Вербицкая, А.П.Павлова. Воспринимаемые звуковые единицы русской речи. "Механизмы речеобразования и восприятия сложных звуков". М.-Л. "Наука". 1966 г.
5. Н.И.Дукельский. Принципы сегментации речевого потока. Изд-во АН СССР. М.-Л. 1962.
6. Э.Петрович. Может ли пересадка звуков (на киноплёнке или на магнитной ленте) решать фонологические вопросы?
7. Л.В.Златоустова. Фонетическая структура слова в потоке речи. Казань, 1962.
8. Л.А.Варшавский. Об автоматическом распознавании звуков речи. "Вопросы радиоэлектроники". Техника проводной связи. Вып.2, 1964 г.
9. Л.Р.Зиндер. Влияние темпа речи на образование отдельных звуков. "Вопросы фонетики". Уч.зап.ЛГУ им.Жданова. Серия филологических наук. № 325 вып.69. Ленинград, 1964.

10. А.А.Харкевич. Спектры и анализ. Гостехиздат, 1957.
11. М.А.Сапожков. Речевой сигнал в кибернетике и связи.
М.1963.
12. P. Delattre, A. Liberman and F. Cooper, JASA, 1965, 27.
13. K. Stevens and A. House, JASA, 1956, 28.
14. A.I. Fourcin, JASA, 1960, 32.
15. Г.С.Себестиан. Процессы принятия решений при распозна-
вании образов. Изд. "ТЕХНИКА". Киев,
1965.

Slutzker G.S.

ON THE PROBLEM OF PHONETIC CIPHERING OF SPEECH

Résumé

The article treats of the properties of the hypothetic 'phonetic speech function', invariant with respect to slow distortions of the transmitting track frequency characteristic. Different versions of transformation of spectral-band signals are suggested for excluding the constant frequency characteristic enabling the transition to a search for a sound-combination standard in the form of sections of a two-dimensional frequency versus time function.

The 'phonetic function' hypothesis offers a basis on which the requirements to the metrics in speech signal space for the most reliable automatic discernment of phonetic speech structure can be formulated.

Инженер В.Н.СОБОЛЕВ

ВСЕСОЮЗНЫЙ ЗАОЧНЫЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ
ИНСТИТУТ СВЯЗИ

АНАЛИЗ СПЕКТРОВ РЕЧИ
ПРИ ПОМОЩИ ЦВМ

Поясняется методика спектрального анализа речи при помощи ЦВМ. Предлагаются ускоренные алгоритмы анализа. Обсуждаются вопросы его применения для выделения основных параметров речевого сигнала.

1. ВВЕДЕНИЕ

Электронная вычислительная техника, постоянно расширяющая сферу своего применения, проникает и в область исследования речевого процесса. Цифровые вычислительные машины (ЦВМ) с успехом используются для математического моделирования систем и принципов сжатия спектра речи [1-4], для изучения свойств речевого сигнала и исследования методов автоматического распознавания звуковых образов [5]. Большие возможности современных цифровых вычислительных машин - высокая точность обработки информации, большое быстродействие, гибкость программного управления, возможность быстрой перестройки их работы - делают ЦВМ мощным средством исследования речевого процесса.

Известно, что речевая информация заключена в изменении спектральных распределений. Поэтому при проведении указанных выше исследований особое значение имеет спектральный анализ речевого сигнала. Здесь можно идти двумя путями. Первый из них - это ввод в память машины спектральных распределений речи, полученных при помощи существующих аналоговых анализаторов. Второй путь заключается в выполнении спектрального анализа при помощи самой ЦВМ.

Современные спектральные анализаторы обладают определенными погрешностями, что исключает возможность проведения целого ряда точных исследований. Кроме того, для ввода в ЦВМ спектральных распределений при соблюдении большой разрешающей способности одновременно по времени и по частоте требуется сложная и громоздкая аппаратура. При таком

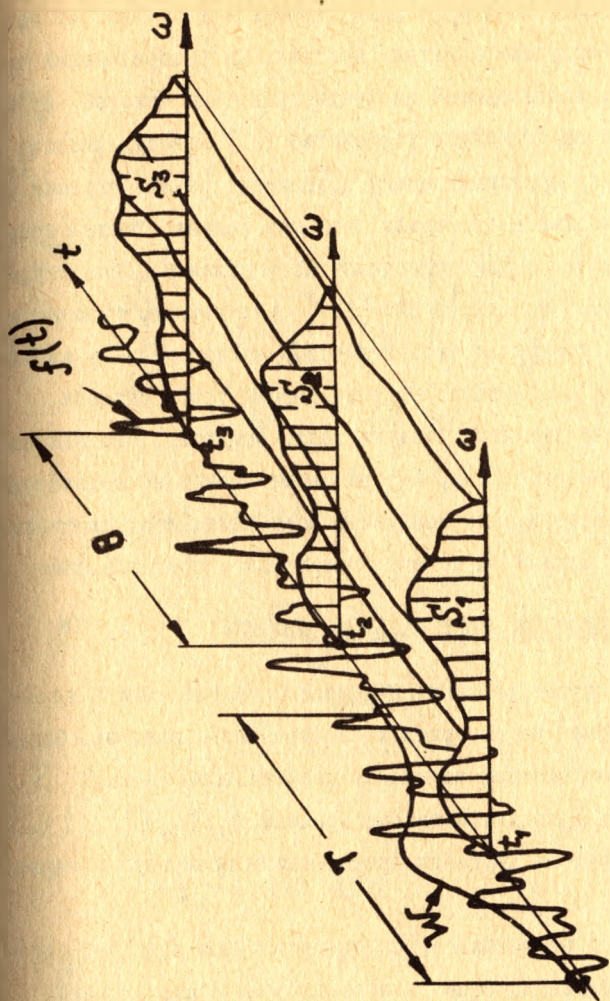


Рис. 1

способе ввода теряется также информация о фазовых соотношениях в сигнале. Поэтому в ряде случаев представляется целесообразным получение картин спектральных распределений при помощи ЦВМ, использующей дискретизированный речевой сигнал, введенный в запоминающее устройство (ЗУ) машины. Емкость внешних ЗУ существующих машин позволяет хранить отрезки речи длительностью в несколько десятков секунд. Такие отрезки вполне достаточны для проведения исследований. Аппаратура ввода речевого сигнала в ЦВМ [6,7] относительно проста, а тенденция к повышению быстродействия вычислительных машин указывает на возрастание рентабельности этого способа спектрального анализа. Следует также отметить, что некоторые характеристики речи, интересующие обычно исследователя, в ряде случаев удобнее получать, используя непосредственно сам речевой сигнал, а не его спектральное представление.

2. АЛГОРИТМЫ СПЕКТРАЛЬНОГО АНАЛИЗА

Ввиду перспективности цифрового анализа речи остановимся подробнее на алгоритмах спектрального анализа. Задачей алгоритма считаем получение спектральных сечений $S_1, S_2, S_3, \dots, S_j, \dots$ для моментов времени $t_1, t_2, t_3, \dots, t_j, \dots$ отстоящих друг от друга на произвольный временной интервал θ (рис.1).

Возможно несколько способов построения подобных алгоритмов. Наиболее простыми и эффективными являются алгоритмы, базирующиеся на теории гармонического синтеза.

Дискретные (по частоте) значения модуля мгновенного спектра

$$S(\omega, t') = \int_{-\infty}^{\infty} f(t) e^{-j\omega t} w(t, t') dt$$

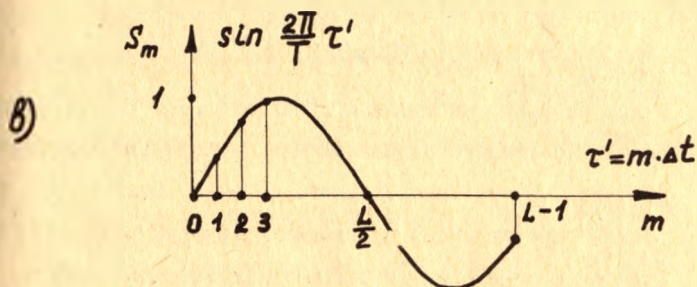
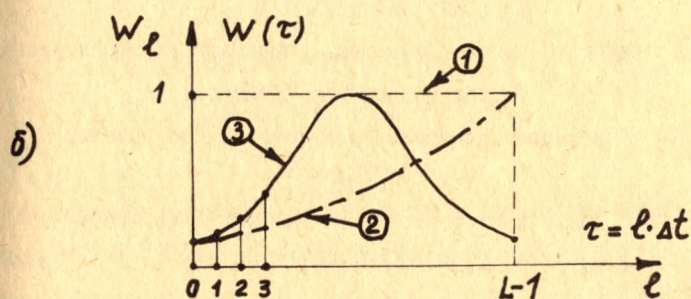
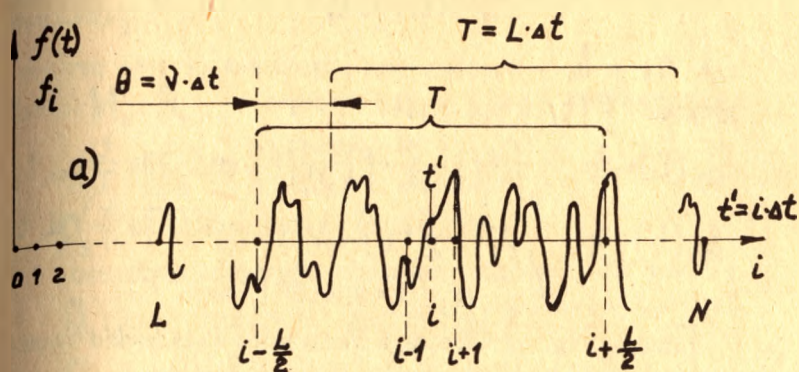


Рис. 2

вычисляются по формуле:

$$C_k = \sqrt{a_k^2 + b_k^2}$$

где a_k и b_k - коэффициенты разложения отрезка речевой функции $f(t)$ в ряд Фурье:

$$a_k(t') = \frac{2}{T} \int_0^T f(t' - \frac{T}{2} + t) \cdot w(t) \cdot \cos(2\pi k \frac{t}{T}) dt;$$

$$b_k(t') = \frac{2}{T} \int_0^T f(t' - \frac{T}{2} + t) \cdot w(t) \cdot \sin(2\pi k \frac{t}{T}) dt;$$

Здесь: $k = 1, 2, 3 \dots K$; $K = \text{entier}(f_{\max} \cdot T)$;

$f_{\max}$ - верхняя частота исследуемого частотного диапазона;

T - период разложения;

$w(\tau)$ - весовая функция, определяющая эффективный интервал наблюдения процесса.

Возможно определение и фазовых соотношений

$$\varphi_k = \arctg \frac{b_k}{a_k}.$$

Вводя дискретизацию по времени (рис.2,а):

$$t = l \cdot \Delta t, t' = i \cdot \Delta t, T = L \Delta t, dt = \Delta t, i = 1, 2, 3 \dots N,$$

получаем окончательные расчетные формулы:

$$a_k(i) = \sum_{l=0}^{L-1} f(i - \frac{L}{2} + l) \cdot w(\frac{l}{L}) \cos(2\pi k \frac{l}{L})$$

$$b_k(i) = \sum_{l=0}^{L-1} f(i - \frac{L}{2} + l) \cdot w(\frac{l}{L}) \sin(2\pi k \frac{l}{L}) \quad (I)$$

Здесь отброшен масштабный множитель $\frac{2}{L}$. Из (I) видим, что расчет коэффициентов разложения, определяющих дискретные значения высоты спектрального сечения, заключается в суммировании произведений мгновенных значений речевой функции (взятых с определенным весом) и вычисленных машиной значений синусоидальной величины.

Логическая схема алгоритма, представленная на абстрактном уровне, имеет вид^{*}):

$$I_0 Q_1 \downarrow Q_2 \downarrow Q_3 \downarrow F_4^{kn} P_5 \downarrow F_6^k P_7 \downarrow H_8^j F_9 P_{10} \downarrow A_n \quad (2)$$

здесь: I_0 и A_n - операторы начала и конца вычислительного процесса.

Операторы P проверяют условия:

$$P_5: l = L; P_7: k = K + 1; P_{10}: i > N - \frac{L}{2}.$$

Остальные операторы выполняют следующие действия^{**}):

$$Q_1: j := 1; i := L + \frac{L}{2}; \beta := \frac{T}{L} \cdot Q_2: k := 1.$$

$$Q_3: a_k := 0; b_k := 0; \alpha := \frac{2\pi}{L} k; n := i; l := 0.$$

$$F_4: a_k := a_k + f_{n-\frac{L}{2}} \cdot W(\beta l) \cdot \sin(\alpha l + \frac{\pi}{2});$$

$$b_k := b_k + f_{n-\frac{L}{2}} \cdot W(\beta l) \cdot \sin(\alpha l); n := n + 1;$$

$$l := l + 1. F_6: c_k := \sqrt{a_k^2 + b_k^2}; k := k + 1.$$

$$F_9: j := j + 1; i := i + \nu.$$

Оператор H_8 осуществляет обработку результатов анализа. Под ним, в частности, можно понимать остальную часть программы моделирования или программу статистической обработки, вывод информации о спектре на печать или подвод информации к специальным устройствам индикации.

Подобный алгоритм для случая $\Theta = T$ применялся нами и ранее [2]. Он имеет три цикла: по числу рассчитываемых спектральных сечений γ (цикл I), по числу дискретных

*) Для описания логических схем здесь и далее применяется операторный метод [8].

**) Значком := обозначен элементарный оператор присваивания.

значений высоты спектрального сечения K (цикл 2) и по числу дискретных значений речевой функции на интервале наблюдения процесса L (цикл 3).

Значения синусоидальной величины и весовой функции рассчитываются здесь каждый раз внутри всех трех циклов, вследствие чего алгоритм обладает малым быстродействием. Подавляющая часть машинного времени при анализе расходуется на вычисление значений $\sin x$ и e^x . Эти значения рассчитываются по стандартным подпрограммам, использующим разложение в ряд Тейлора. Для достижения необходимой точности обычно учитывается 6 - 10 членов ряда.

Временной масштаб, под которым понимают отношение времени обработки сигнала (в данном случае - времени анализа) к реальному времени течения процесса, при быстродействии машины порядка десятков тысяч операций в секунду составляет обычно несколько тысяч. При таком масштабе (учитывая высокую стоимость машинного времени) цифровой анализ следует признать нерентабельным. Необходимо искать пути повышения быстродействия алгоритма.

Некоторое ускорение дает применение известной теоремы Винера-Хинчина. Спектральная плотность рассчитывается по кратковременной автокорреляционной функции процесса, причем расчет пары коэффициентов a_k и b_k заменяется расчетом непосредственно C_k . Однако из-за относительно большого времени вычисления автокорреляционной функции выигрыш в быстродействии составляет менее двух раз.

Несколько больший выигрыш можно получить, пользуясь суммарной и разностной функциями [9] :

$$\left. \begin{aligned} A(t, \tau) &= \frac{1}{2} [f(t+\tau) + f(t-\tau)] \\ B(t, \tau) &= \frac{1}{2} [f(t+\tau) - f(t-\tau)] \end{aligned} \right\} \quad (3)$$

здесь возможно раздельное определение коэффициентов A_k и B_k по каждой из функций, причем интервал T сводится к интервалу $\frac{T}{2}$, за счет чего и получается выигрыш во времени примерно в два раза. Такой алгоритм несколько эффективнее алгоритма, основанного на использовании теоремы Винера-Хинчина, т.к. пара функций (3) рассчитывается значительно быстрее, чем автокорреляционная функция.

Более эффективным приемом является размен занимаемой емкости оперативного ЗУ на быстроедействие алгоритма. Нет необходимости каждый раз вычислять значения синусоидальной величины и весовой функции. Достаточно рассчитать их однажды, вне трех указанных циклов, и хранить в памяти машины.

Очевидно, что для таблицы значений весовой функции (например, экспоненциальной (2, рис.2,б) или колоколообразной (3, рис.2,б)) необходим участок памяти в L ячеек. Значениями функции вне интервала T можно пренебречь за их малостью. При использовании прямоугольной весовой функции (1, рис.2,б) необходимость в таблице отпадает. Введение такой весовой функции осуществляется конечными пределами суммирования.

С первого взгляда кажется, что для хранения матрицы значений $\sin(2\pi \frac{m}{L})$, имеющей размерность $K \times L$, необходим большой объем памяти. Однако более внимательное рассмотрение наводит на мысль об использовании лишь одного

периода этой функции (рис.2,в). Таблица $\sin(2\pi \frac{m}{L})$ длины L содержит все значения функции, необходимые для расчета любого коэффициента C_K . Достаточно лишь замкнуть таблицу в кольцо и, циркулируя по нему, выбирать для расчета K -го коэффициента значения таблицы, отстоящие друг от друга на K точек. Математически эту процедуру можно описать выражением:

$$S_\ell := S_{m \pmod{L}} ,$$

где

$$m = K\ell, \ell = 0, 1, 2, \dots$$

Программно эта операция может быть выполнена, например, посредством проверки условия $m \geq L$ и соответствующего формирования команд выборки значений S_m .

Логическая схема алгоритма, использующего описанный прием, имеет вид*):

$$\begin{array}{c}
 U_0 Q_{12} \downarrow F_{13}^{m\ell} P_{14} L Q_{15} \downarrow Q_{16} \downarrow Q_{17} \downarrow P_{18} \downarrow F_{19}^{m\ell} \downarrow F_{20}^{Knm\ell} P_5 L \\
 F_6^K P_7 L H_8^j F_9 P_{10} L A_{11}. \quad (4)
 \end{array}$$

Здесь (кроме ранее раскрытых операторов) условные операторы проверяют:

$$P_{14}: m = L; \quad P_{18}: m \geq L.$$

*). Во избежание пояснения ранее введенных операторов применена единая нумерация операторов для всех логических схем, приводимых в работе.

Остальные операторы выполняют следующие действия:

$$Q_{12}: m := 0; \ell := 0; \alpha := \frac{2\pi}{L}; \beta := \frac{T}{L}.$$

$$F_{13}: S_m := \sin(\alpha m); W_\ell := W(\beta \ell); m := m + 1;$$

$$\ell := \ell + 1. Q_{16}: j := 1; i := L + \frac{1}{2}; m := 0.$$

$$Q_{17}: a_k := 0; b_k := 0; n := i; \ell := 0.$$

$$F_{19}: m := m - L. F_{20}: a_k := a_k + f_{n-\frac{1}{2}} \cdot W_\ell \cdot S_{m+\frac{1}{4}};$$

$$b_k := b_k + f_{n-\frac{1}{2}} \cdot W_\ell \cdot S_m; n := n + 1;$$

$$\ell := \ell + 1; m := m + k.$$

Операторы P_{18} и F_{19} осуществляют описанное выше замыкание таблицы функции $\sin x$ в кольцо на периоде 2π .

Алгоритм (4) обладает быстроедействием примерно на порядок более высоким, чем алгоритм (2). Однако при частом получении спектральных сечений (при значении θ , близком к Δt) и такое быстроедействие оказывается недостаточным. Высокая частота получения спектральных портретов речи может потребоваться, например, при изучении динамики спектра взрывных звуков. Здесь эффективнее алгоритм, построенный на основе рекуррентных соотношений.

Положим $\theta = \Delta t$ и введем прямоугольную весовую функцию. Заметим, что значения одноименных коэффициентов разложения для соседних спектральных сечений могут отличаться лишь двумя членами суммы. Например, для первых коэффициентов справедливы соотношения:

$$\left. \begin{aligned} a_1(i+1) &= a_1(i) + \left[f\left(i + \frac{1}{2}\right) - f\left(i - \frac{1}{2}\right) \right] \cdot \cos\left(\frac{2\pi}{L} i\right) \\ b_1(i+1) &= b_1(i) + \left[f\left(i + \frac{1}{2}\right) - f\left(i - \frac{1}{2}\right) \right] \cdot \sin\left(\frac{2\pi}{L} i\right) \end{aligned} \right\} (5)$$

Равенства, аналогичные (5), определяют простой алгоритм нахождения последующего спектрального портрета по предыдущему. Необходимо лишь в начале последовательности дискретных значений речевой функции (рис.2,а) предусмотреть нулевой участок в L ячеек и положить равными нулю все коэффициенты разложения, определяющие первое спектральное сечение.

Логическая схема алгоритма, построенного на основе поясненных рекуррентных соотношений, имеет вид:

$$\begin{aligned} & I_0 Q_{12} F_{13}^{ml} P_{14} L Q_{21} L Q_{22} P_{72} L Q_{23} L Q_{24} L Q_{25} L Q_{26}^k P_{18} L \\ & F_{19}^m L F_{25}^{km} P_{75} L F_{26} P_{27} L H_8 F_{28} P_{29} L Y_{11}. \end{aligned} \quad (6)$$

Здесь (кроме ранее раскрытых операторов) условные операторы проверяют:

$$P_{27}: \gamma = 1; \quad P_{29}: i > N - \frac{L}{2}.$$

Остальные операторы выполняют следующие действия:

$$\begin{aligned} & Q_{21}: j := 1; \quad i := \frac{L}{2}; \quad k := 1. \quad Q_{22}: a_k := 0; \\ & b_k := 0; \quad \mu_k := 0; \quad k := k + 1. \quad Q_{23}: \gamma := 0. \\ & Q_{24}: m := \mu_k \cdot F_{25}: a_k := a_k + s_{m+\frac{1}{4}} (f_{i+\frac{1}{2}} - f_{i-\frac{1}{2}}); \\ & b_k := b_k + s_m (f_{i+\frac{1}{2}} - f_{i-\frac{1}{2}}); \\ & c_k := \sqrt{a_k^2 + b_k^2}; \quad \mu_k := m + k; \quad k := k + 1. \\ & F_{26}: i := i + 1; \quad \gamma := \gamma + 1. \quad F_{28}: j := j + 1. \end{aligned}$$

Алгоритм	А-I (2)		А-II (4)		А-III (6)	
0	III мсек	22 мсек	III мсек	22 мсек	III мсек	22 мсек
Временной масштаб для М-20	500.000	2.500	70.000	350	800	450
Временной масштаб для БЭСМ-6	10.000	50	1.400	7	16	9
Выигрыш во времени	I	-	7,14	-	625	-
	-	I	-	7,14	-	5,55

Алгоритм допускает применение лишь прямоугольной весовой функции. Для выполнения алгоритма (кроме группы ячеек памяти под таблицу значений S_m) требуется группа ячеек для хранения информации об адресах M_k выбираемых значений синусоидальной величины. Эта емкость обычно не превышает емкости отсутствующей здесь таблицы весовой функции. Кажущегося на первый взгляд накопления погрешностей не происходит, однако алгоритм чувствителен к случайным сбоям в работе машины. Такой сбой может привести к тому, что, начиная с определенного момента (соответствующего сбою), результаты дальнейшего вычисления "мгновенного" спектра окажутся неверными. Поэтому в случае невысокой надежности применяемой ЦВМ и большой длительности анализируемого отрезка речи необходимо осуществлять периодическую проверку результатов счета. Ее можно производить, например, путем сравнения результатов анализа со спектром, рассчитанным по алгоритму (4).

Выигрыш во времени при применении алгоритма (6) по сравнению с (4) зависит от соотношения $\alpha = \frac{T}{\theta} = \frac{L}{V}$. При $\alpha = 1$ быстродействия алгоритмов примерно одинаковы. При $\alpha < 1$ предпочтительнее алгоритм (4), при $\alpha > 1$ - алгоритм (6). Наибольший выигрыш достигается при $\theta = \Delta t$, когда исключается цикл по V (цикл 4).

В таблице I представлены результаты сравнительных испытаний на быстродействие программ, составленных в соответствии с логическими схемами (2), (4), (6).

Машина рассчитывала спектральные портреты речи в диапазоне частот от 0 до 3,5 кГц. Интервал наблюдения процес-

са T был выбран равным 22,2 мсек, что при шаге квантования $\Delta t = III$ мсек определило число точек на интервале наблюдения $L \approx 200$. Дискретные значения рассчитываемого спектра отстояли друг от друга на $\Delta f \approx 45$ гц. Использовалась прямоугольная весовая функция.

Испытания производились на цифровой вычислительной машине М-20, обладающей быстродействием порядка 20.000 операций в секунду. Таблица содержит также значения временных масштабов, пересчитанных для машины БЭСМ-6, обладающей быстродействием порядка 1.000.000 операций в секунду. Из таблицы видно, что выигрыш во времени алгоритма (4) по сравнению с алгоритмом (2) не зависит от частоты получения спектральных сечений и при указанных выше параметрах составляет примерно 7 раз. Алгоритм (6) при $\theta = \Delta t = III$ мсек эффективнее алгоритма (2) в 625 раз. Описанная методика позволяет на существующих ЦВМ производить спектральный анализ речи в масштабе времени, близком к натуральному.

3. СПЕКТРАЛЬНЫЕ ПОРТРЕТЫ РЕЧИ

На рис.3,а представлен спектральный портрет части второго звука "а" в слове "равнина", произнесенном высоким женским голосом (частота основного тона $f_0 \approx 350$ гц).*) Отчетливо видны гармоники основного тона (ОТ).

На рис.3,б изображен фазовый портрет того же звука. Следует отметить, что во всех фазовых портретах (как и в изображенном) наблюдались резкие скачки фазы на гармониках ОТ.

*) Масштаб по оси ординат - логарифмический. Параметры цифровой модели анализатора: $T \approx 22,2$ мсек, $\Delta f \approx 45$ гц.

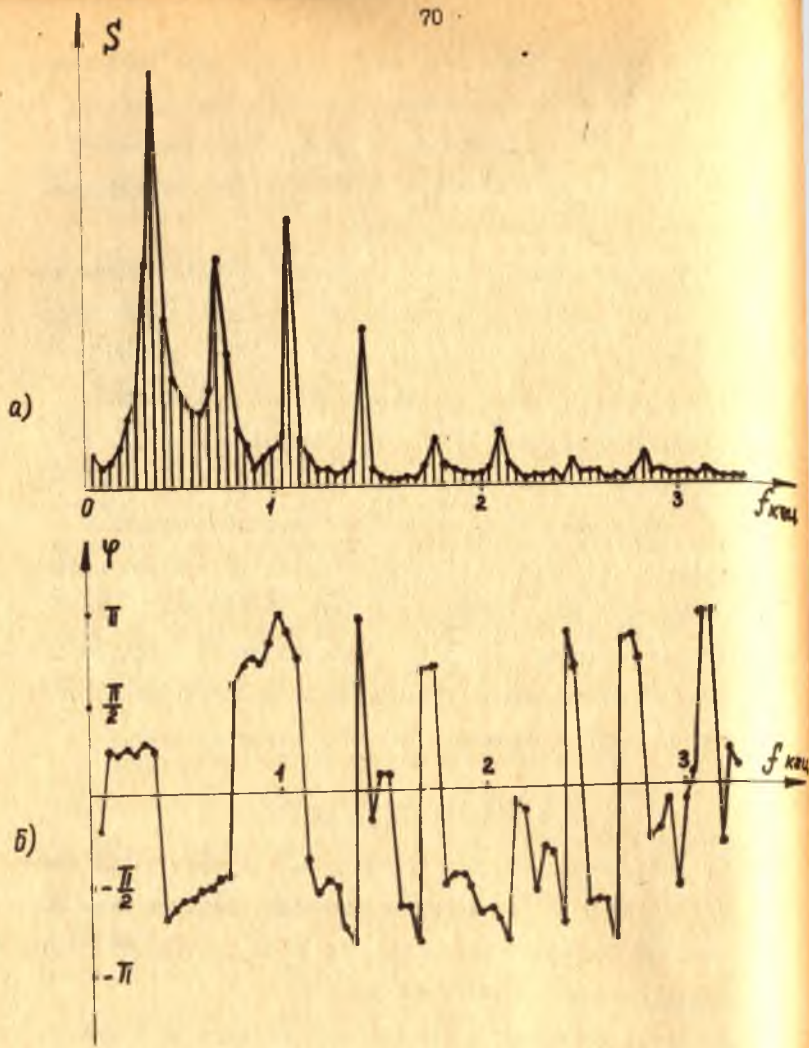


Рис.3

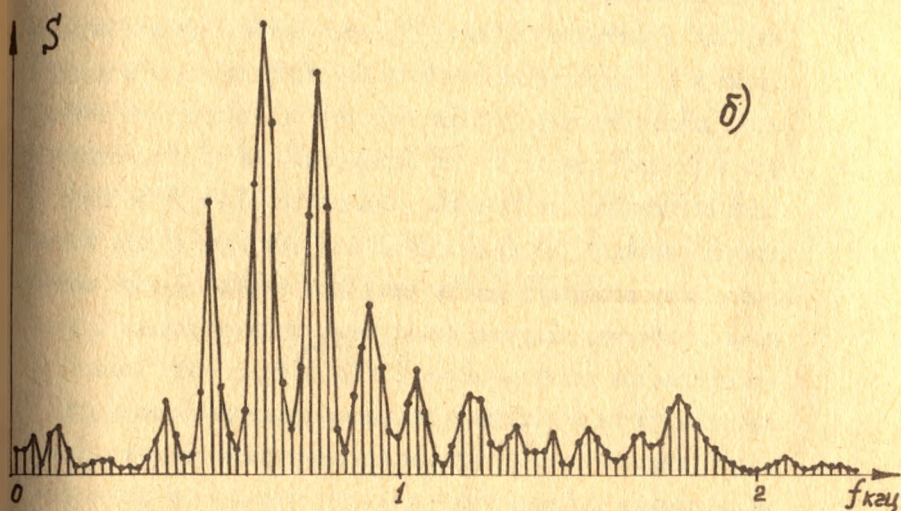
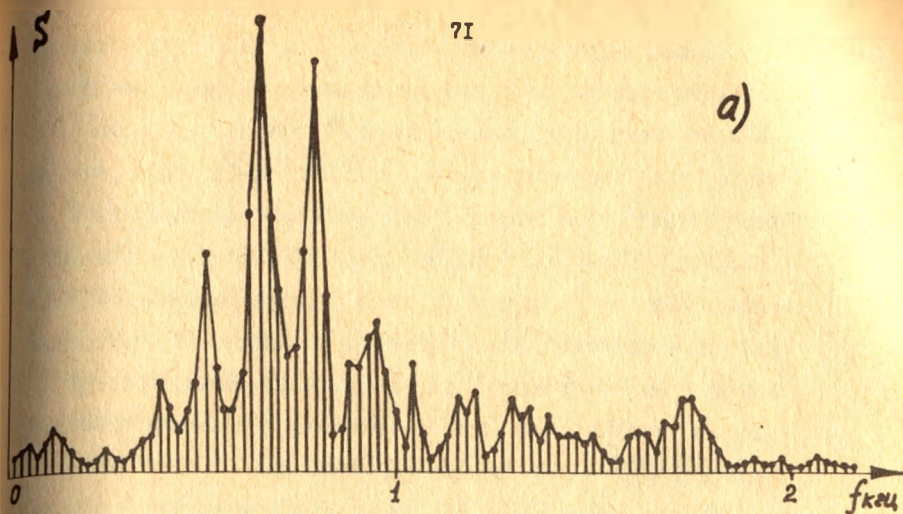


Рис. 4

Спектр, представленный на рис.3,а, рассчитан с использованием прямоугольной весовой функции. На рис.4 представлены фрагменты спектральных картин*) для части гласного звука, произнесенного мужским голосом ($f_0 \approx 130$ гц). Сопоставление спектральных картин, рассчитанных с прямоугольной (рис.4,а) и колоколообразной (рис.4,б) весовыми функциями, показывает, что в первом случае в спектре наблюдается большая "изрезанность". Дополнительные экстремумы в спектральной картине, полученной при использовании прямоугольной весовой функции, очевидно, объясняются наложением "хвостов" функции $\left| \frac{\sin f}{f} \right|$ от мощных гармоник ОТ.

Рассмотрение серии последовательных спектральных портретов позволяет изучить как изменение высоты основного тона, так и движение формант. Возможно автоматическое определение и регистрация положения и высоты формантных максимумов, ширины формант, а также частоты основного тона речи. Здесь следует отметить, что определение частоты основного тона по спектру для сигнала, прошедшего через телефонный тракт, очевидно, не должно встречать тех трудностей, которые обычно возникают при решении этой задачи другими методами (например, пиковым, фильтровым, корреляционным). Возможен простой дихотомический алгоритм подстройки модели анализатора на максимумы, соответствующие гармоникам ОТ. Модель такого адаптирующегося анализатора может быть полезна не только для выделения основного тона, но и для точного определения (по пикам гармоник ОТ) огибающей спектра

*) Масштаб по оси ординат - логарифмический. Параметры цифровой модели анализатора: $T \approx 44,4$ мсек, $\Delta f \approx 23$ гц.

речи, сведения о которой в той или иной форме передаются во всех вокодерных системах уплотнения.

4. ЗАКЛЮЧЕНИЕ

Рациональное построение алгоритмов анализа речи при помощи ЦВМ позволяет производить его в практически приемлемом масштабе времени.

Большие возможности современных цифровых вычислительных машин дают возможность использовать принцип адаптации, существенно повышающий точность анализа процесса.

Сочетание спектрального анализа речи при помощи ЦВМ с последующей машинной обработкой результатов этого анализа позволяет выделить все основные параметры речевого сигнала.

Изложенная методика может быть использована при исследовании звуков речи, цифровом моделировании систем передачи речевого сигнала (в частности, вокодеров) и исследовании методов автоматического распознавания звуковых образов.

Предлагаемые ускоренные алгоритмы могут быть использованы для спектрального анализа на ЦВМ не только речевого сигнала, но и случайного процесса любой природы.

ЛИТЕРАТУРА

- 1 Р.Д.Лейтес. Методика математического моделирования систем передачи речевого сигнала. "Электросвязь", 1963, № 8.
- 2 Р.Д.Лейтес, В.Н.Соболев. Принципы цифрового моделирования вокодеров. "Электросвязь", 1966, № 7.

- 3 Р.Д.Лейтес, В.Н.Соболев. Эксперимент по цифровому моделированию гармонического вокодера. Сборник трудов Гос. НИИ Мян. связи СССР, 1966, вып. 3 (43).
- 4 Г.А.Коротаев, В.Г.Михайлов. Современное состояние техники параметрического компандирования речи. "Зарубежная радиоэлектроника", 1966, № 3.
- 5 Н.Г.Загоруйко. Об обмене устной информацией между человеком и вычислительными системами. "Вычислительные системы". Сборник трудов Института математики СО АН СССР, 1964, вып. 10.
- 6 Ю.А.Лейбман, В.Н.Соболев. Преобразователь аналог-цифра для ввода речевого сигнала в вычислительную машину. "Электросвязь", 1963, № 8.
- 7 В.Н.Соболев. Вопросы ввода речевой информации в цифровую вычислительную машину. "Труды учебных институтов связи", 1965, вып. 25.
- 8 Н.А.Кривяцкий, Г.А.Миронов, Г.Д.Фролов. Программирование. Справочная математическая библиотека. "Физматгиз", 1963.
- 9 К.Ланцош. Практические методы прикладного анализа. "Физматгиз", 1961, стр. 226.

V.N. Sobolev

Moscow Communications Institute

SPEECH SPECTRUM ANALYSIS PERFORMED BY MEANS OF
DIGITAL COMPUTER

Summary

The present paper is devoted to investigating the possibility of employment of digital computers for carrying out speech signal spectrum analysis.

The author invents and draws comparison between miscellaneous logical circuits of spectrum analysis algorithms. The paper demonstrates that by means of rational use of algorithms the machine time can be drastically reduced. This device makes it possible to carry out the speech spectrum analysis with up-to-date highly efficient machines in time scale approximating the natural one.

Analysis of spectrum speech pictures (obtained by means of computer) shows the possibility to study and to automatically single out (with the computer available) the basic speech signal parameters.

The spectrum analysis algorithm described in the paper can be very good for the study of speech sounds, for digital analog formation of speech signal transmission systems (vocoders in particular) and for the study of the methods of automatic identification of sound images.

Лобанов Б.М.
Минский радиотехнический
институт

МНОГОФУНКЦИОНАЛЬНЫЙ АНАЛИЗАТОР СПЕКТРОВ
РЕЧЕВЫХ СИГНАЛОВ "СПЕКТР-1"

Анализатор спектров речевых сигналов типа "Спектр-1" разработан в научно-исследовательской лаборатории Минского радиотехнического института. Прибор работает по схеме параллельного спектрального анализа [1] с использованием 50 полосовых фильтров.

В процессе разработки ставилась задача создания многофункционального прибора, пригодного как для лингвистических исследований, так и для исследований в области компрессии и распознавания речи. Имея в виду лингвистическое применение, основное внимание уделялось разработке методов улучшения основных параметров анализатора (погрешность, время анализа и т.д.), а также вопросам удобства считывания результатов анализа (амплитудные и частотные метки, калибровка и др.). Имея в виду применение прибора для исследований в области компрессии и распознавания речи, основное внимание было уделено вопросам устранения избыточной (несмысловой) информации непосредственно в процессе анализа. При этом осуществляется промежуточная обработка на различных уровнях, а именно: всего сигнала (клиппирование, динамическая компрессия), каждой спектральной составляющей (логарифмирование и вычитание среднего значения) и огибающей спектра (фильтрация верхних частот).

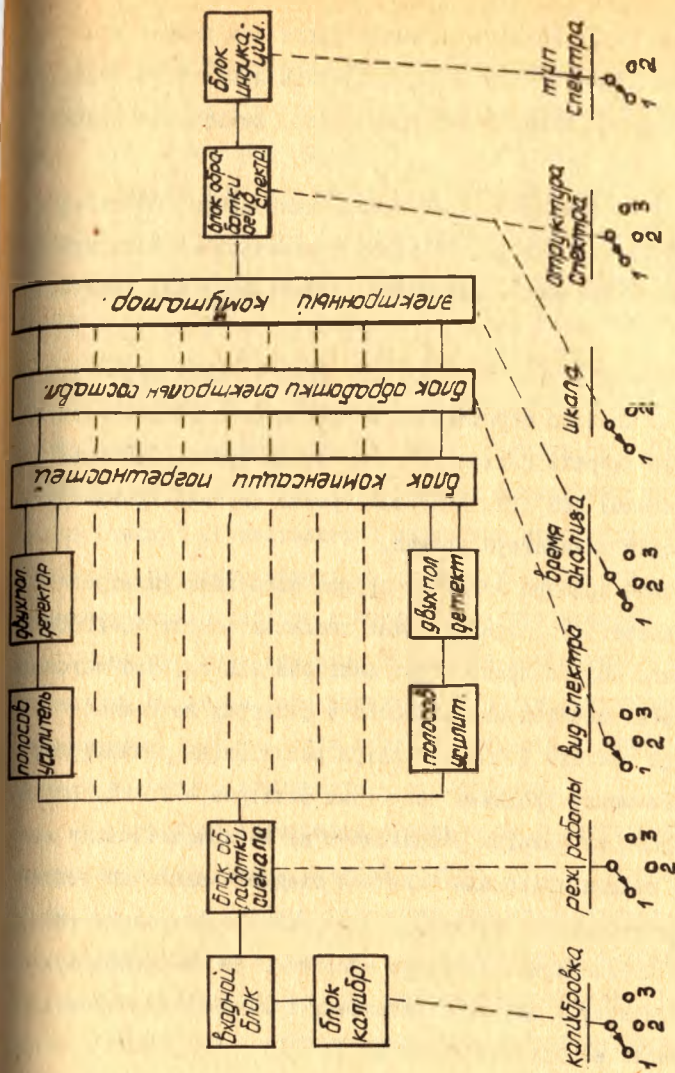


Рис. 1

Кроме того, большое внимание уделялось разработке улучшенных вариантов некоторых ответственных узлов анализатора таких, как электронный коммутатор, фильтровые ячейки и др.

На рис. I приведена блок-схема анализатора "Спектр-1".

Исследуемое напряжение может подаваться на вход прибора от трех источников: микрофона, магнитофона или сегментатора.

Усиленный входной сигнал поступает на блок обработки сигнала. С помощью переключателя "Режим работы" исследователь может выбрать режимы: клиппированная речь, усиление с АРУ и усиление без АРУ. Контроль уровня сигнала производится с помощью индикатора уровня.

Основой прибора является группа полосовых фильтров (в количестве 50), перекрывающая диапазон частот 60-10000 гц. Минимальная полоса пропускания фильтров 300 гц, что обуславливает время переходных процессов в фильтрах не более 5 мсек. До частоты 1000 гц фильтры расположены равномерно, расстояние между средними частотами фильтров - 75 гц. После частоты 1000 гц фильтры расположены по логарифмическому закону. За каждым полосовым фильтром следует канальный усилитель и двухполярный детектор. Конструктивно канальный усилитель и двухполярный детектор выполнены на полупроводниковых элементах и вместе с канальным полосовым фильтром смонтированы на отдельной плате.

Для уменьшения амплитудной погрешности, возникающей за счет отличия реальных характеристик фильтров от идеальных П-образных, а также за счет взаимного перекрытия сосед-

них фильтров, разработан блок компенсации погрешностей (БКП) [2]. При этом предусматривается возможность работы как с БКП, так и без него.

С выхода БКП (или минуя его) сигнал поступает на блок обработки составляющих спектра сигнала. Блок состоит из логарифматоров и фильтров нижних частот, коммутируемых специальным образом. С помощью переключаемых фильтров нижних частот (переключатель "Вид спектра") возможно получение "мгновенного" и "среднего" спектров, т.е. спектров, усредненных во времени на различных уровнях, а также "разностного спектра". В трех положениях переключателя "разностный спектр" осуществляется выделение информации о быстрых процессах речеобразования из медленных процессов, характеризующих индивидуальные отличия источника речи [3].

Предусмотрена возможность вывода напряжений с выхода канальных фильтров, что значительно расширяет возможности анализатора, так как позволяет исследовать и анализировать каждую частотную компоненту в отдельности.

Электронный коммутатор выполнен на коммутаторных декадах А-102. Время анализа регулируется в пределах 0,01, 0,02 и 0,04 сек. Скоммутированный сигнал поступает на блок обработки огибающей спектра, состоящий из логарифмического усилителя и регулируемых фильтров нижних частот. Предусмотрена возможность (переключатель "Шкала") пропускать сигнал через логарифмический усилитель (децибелная шкала выходного индикатора) или минуя его (линейная шкала). Переключателем "Структура спектра" изменяются постоянные времени регулируемого фильтра, т.е. имеется возможность по-разному

усреднять спектр по частоте, устраняя его мелкие детали. В приборе предусмотрена возможность получения двух типов спектров (переключатель "Тип спектра"): двухмерного и трехмерного. Для получения трехмерного спектра сигнал с регулируемого фильтра нижних частот подается на модулирующие пластины осциллографа. На вертикальные пластины осциллографа подается напряжение развертки, синхронное с частотой электронного коммутатора, а на горизонтальные пластины — напряжение вспомогательной и низкочастотной развертки (1-3 гц), необходимое для наблюдения изменений спектра исследуемого сигнала во времени.

Для получения двухмерного спектра возможно использование как электронного, так и шлейфового осциллографа. В случае электронного осциллографа исследуемый сигнал подается на вертикальные отклоняющие пластины. На горизонтальные отклоняющие пластины подается напряжение развертки, синхронное с частотой электронного коммутатора. Для использования шлейфового осциллографа анализатор снабжен усилителем мощности постоянного тока. Изображение с экранов электронных осциллографов может быть снято с помощью кинокамеры. Для удобства анализа спектральных картин выходное изображение снабжается амплитудными и частотными метками.

Предусмотрена калибровка прибора (переключатель "Калибровка") сплошным или дискретным спектрами.

Питание устройства осуществляется от стабилизированного блока.

Литература:

1. А.А.Харкевич "Спектры и анализ". Физматгиз, 1963 г.
2. Б.М.Лобанов "Исследование одного метода улучшения основных параметров параллельного спектрального анализа". Сб.Радиотехнические измерения. Минск, 1965.
3. А.А.Пирогов "К вопросу о фонетическом кодировании речи". Электросвязь № 5, 1967.

B.M. Lobanov

Minsk Radiotechnical Institute

SPEECH SIGNAL SPECTRUM MULTIFUNCTIONAL ANALYZER

"SPECTRUM" I

Summary

There are a few types of specialized spectrum analyzers designed for the study of speech signal. When constructing the described analyzer a problem was set to build a multifunctional device fitted for both linguistic research and research in the field of speech compression and identification.

Firstly, primary attention was paid to development of techniques and circuits improving the basic parameters of spectrum analyzer and, secondly, removing excess information non-comprehensive in the process of analysis.

Amplitude error compensation circuits have been developed and characteristics of certain important analyzer units improved. Use is made of the circuits effecting an intermediate interpretation of signal, each spectral constituent and envelope curve of spectrum.

The described analyzer can find employment in many aspects of speech signal research. The obtained information confirms the possibility of circuit realization of certain error reduction methods of analysis and method of removal of the effect of non-comprehensive information elements upon the pictures of "visible speech".

Как известно, любая система линейных алгебраических уравнений может быть представлена в форме графа. На рис.2-а изображен граф системы (1). Ветви исходного графа направлены от S_j к U_i . Для того, чтобы найти выражение передачи от U_i к S_j , инвертируем пути, соединяющие узлы с одинаковыми индексами. В результате получим граф, изображенный на рис.2-б. Аналитическое выражение передачи от каждого узла U_i к заданному узлу S_j будет искомым результатом.

На рис.2-б обозначены передачи ветвей инвертированного графа. Введем обозначения:

L_i - передача i -го контура, касающегося сверху i -го узла, внизу $i-1$ -го.

P_{ij} - путь от узла U_i к узлу S_j

Например:

$$L_1 = \frac{K_{12} K_{21}}{K_{11} K_{22}}, P_{12} = -\frac{K_{12}}{K_{11} K_{22}}.$$

Используя эти обозначения, запишем выражение для передачи от U_i к S_j -му узлу.

$$S_j = \frac{\sum_{i=1}^n (-1)^{i+j} \cdot U_i P_{ij} \prod_{k=1}^n (1-L_k)^*}{\prod_{k=1}^n (1-L_k)^*} \quad (2)$$

Знак * означает, что при умножении коэффициентов внутри скобок мы приравняем к нулю любой член, который содержит произведение передач двух контуров или произведение передач пути и контура, касающихся друг друга в графе.

В целях упрощения выражения (2) введем некоторые ограничения, удовлетворяющие многим практическим случаям. Будем считать, что $K_{ii} = K_{jj} = 1$. Это соответствует случаю анализатора с равномерной частотной шкалой, в котором

применены избирательные системы с одинаковыми частотными характеристиками. Равенство единице означает факт нормировки коэффициентов. Для этого же случая справедливо и второе условие $K_{ij} = m < 1$, где $i \neq j$. Учитывая эти условия, а также тот факт, что знаменатель выражения (2) имеет одинаковую величину для различных S_j , с точностью до постоянного множителя можно записать

$$S_j = \sum_{i=1}^n (-1)^{i+j} m^{|i-j|} U_i (1-m^2)^{n*}$$

Распишем выражение $(1-m^2)^{n*}$, используя формулу бинома Ньютона.

$$(1-m^2)^{n*} = 1 - C_n^{1*} m^2 + C_n^{2*} m^4 - C_n^{3*} m^6 + \dots$$

Учитывая, что $m < 1$, а практически для реальных анализаторов $m < 0,5$, ограничимся рассмотрением четырех членов разложения. Вся сложность состоит в нахождении значений C_n^{i*} . Здесь мы сталкиваемся с вопросом определения числа сочетаний, на которые накладываются фиксированные ограничения. Подобный вопрос в известной нам литературе не освещен. Математически задача формулируется следующим образом. Найти количество сочетаний из n элементов по m , если элементы сочетания пронумерованы индексами от 1 до n и искомые сочетания удовлетворяют следующим ограничениям:

- а) элементы, входящие в сочетание, не должны иметь индексов, отличающихся менее чем на две единицы;
- б) элементы, входящие в сочетание, не должны иметь первых p индексов.

Обозначим количество искомых сочетаний через Γ_{np}^m .

Тогда очевидно, что $\Gamma_{np}^1 = n - p$. Для нахождения Γ_{np}^2 применим следующий прием. Выпишем все сочетания для элементов C_i из 8 по 2.

C_1C_2	C_1C_3	C_1C_4	C_1C_5	C_1C_6	C_1C_7	C_1C_8
C_2C_3	C_2C_4	C_2C_5	C_2C_6	C_2C_7	C_2C_8	
C_3C_4	C_3C_5	C_3C_6	C_3C_7	C_3C_8		
C_4C_5	C_4C_6	C_4C_7	C_4C_8			
C_5C_6	C_5C_7	C_5C_8				
C_6C_7	C_6C_8					
C_7C_8						

Легко убедиться, что количество сочетаний, входящих в полный треугольник, соответствуют $\Gamma_{8,0}^2$, в треугольник без верхней строки - $\Gamma_{8,1}^2$, в треугольник без двух верхних строк $\Gamma_{8,2}^2$ и т.д. Общее выражение можно записать в виде:

$$\Gamma_{n,p}^2 = C_n^2 - \sum_{i=0}^p (n-i-1)$$

Воспользуемся тем же приемом для нахождения Γ_{np}^3 .

Выпишем опять все сочетания из 8 по 3 в виде треугольников.

$C_1C_2C_3$	$C_1C_2C_4$	$C_1C_2C_5$	$C_1C_2C_6$	$C_1C_2C_7$	$C_1C_2C_8$
$C_1C_3C_4$	$C_1C_3C_5$	$C_1C_3C_6$	$C_1C_3C_7$	$C_1C_3C_8$	
$C_1C_4C_5$	$C_1C_4C_6$	$C_1C_4C_7$	$C_1C_4C_8$		
$C_1C_5C_6$	$C_1C_5C_7$	$C_1C_5C_8$			
$C_1C_6C_7$	$C_1C_6C_8$				
$C_1C_7C_8$					

$C_2C_3C_4 \quad C_2C_3C_5 \quad C_2C_3C_6 \quad C_2C_3C_7 \quad C_2C_3C_8$
 $C_2C_4C_5 \quad C_2C_4C_6 \quad C_2C_4C_7 \quad C_2C_4C_8$
 $C_2C_5C_6 \quad C_2C_5C_7 \quad C_2C_5C_8$
 $C_2C_6C_7 \quad C_2C_6C_8$
 $C_2C_7C_8$
 $C_3C_4C_5 \quad C_3C_4C_6 \quad C_3C_4C_7 \quad C_3C_4C_8$
 $C_3C_5C_6 \quad C_3C_5C_7 \quad C_3C_5C_8$
 $C_3C_6C_7 \quad C_3C_6C_8$
 $C_3C_7C_8$
 $C_4C_5C_6 \quad C_4C_5C_7 \quad C_4C_5C_8$
 $C_4C_6C_7 \quad C_4C_6C_8$
 $C_4C_7C_8$
 $C_5C_6C_7 \quad C_5C_6C_8$
 $C_5C_7C_8$
 $C_6C_7C_8$

Нетрудно видеть, что количество сочетаний, входящих во все треугольники, соответствует $\Gamma_{8,0}^3$, в треугольники без одного верхнего - $\Gamma_{8,1}^3$, в треугольники без двух верхних - $\Gamma_{8,2}^3$ и т.д. Общее выражение для $\Gamma_{n,p}^3$ запишется в виде:

$$\Gamma_{np}^3 = 2 \sum_{i=1}^{n-4} i - \sum_{j=1}^p \sum_{i=1}^{n-j-3} i$$

Выражение для Γ_{np}^2 также может быть переписано в подобной форме, тогда:

$$\Gamma_{np}^2 = \sum_{j=1}^{n-2} i - \sum_{j=1}^p (n-j-1)$$

Используя выражение для суммы арифметической прогрессии, запишем окончательные выражения для Γ_{np}^m .

$$\left. \begin{aligned} \Gamma_{np}^1 &= n-p \\ \Gamma_{np}^2 &= \frac{(n-1)(n-2)}{2} - \frac{p(2n-p-3)}{2} \\ \Gamma_{np}^3 &= \frac{(n-4)(n-3)}{2} - \frac{p^2[(n-3)(n-4) + (n-p-2)(n-p-3)]}{4} \end{aligned} \right\} \quad (3)$$

Учитывая введенные ограничения, а также тот факт, что число p определяется количеством контуров, касающихся пути P_{ij} , перепишем общее выражение (2) к виду, удобному для практических расчетов.

$$S_j = \sum_{i=1}^n (-1)^{i+j} m^{|i-j|} U_i \left[1 + \sum_{k=1}^3 (-1)^k \Gamma_{ij-i+k}^k m^{2k} \right] \quad (4)$$

Рассмотрим пример, иллюстрирующий результаты применения расчетных формул (3), (4). Рассчитаем систему компенсации для десятиканального анализатора ($n = 10$) с коэффициентом перекрытия характеристик, равным 0,32 ($m = \frac{aU}{a_{i \min}} = 0,32$). Возможный вид частотных характеристик фильтров приведен на рис. I. Распишем выражение передачи для одного из узлов (например, S_5)

$$\begin{aligned} S_5 &= m^4 U_1 (1 - m^2 \Gamma_{10,5}^1 + m^4 \Gamma_{10,5}^2 - m^6 \Gamma_{10,5}^3 - \\ &- m^3 U_2 (1 - m^2 \Gamma_{10,5}^1 + m^4 \Gamma_{10,5}^2 - m^6 \Gamma_{10,5}^3 + \\ &+ m^5 U_{10} (1 - m^2 \Gamma_{10,7}^1 + m^4 \Gamma_{10,7}^2 - m^6 \Gamma_{10,7}^3) \end{aligned}$$

Наибольшие значения для Γ_{np}^m будут у слагаемого с коэффициентом U_5 .

Рассчитаем их по формулам (3).

$$\Gamma_{10,2}^1 = 9, \Gamma_{10,2}^2 = 21, \Gamma_{10,2}^3 = 6$$

$$m^2 \Gamma_{10,2}^1 = 0,9, m^4 \Gamma_{10,4}^2 = 0,21, m^6 \Gamma_{10,2}^3 = 0,006$$

Так как слагаемое, содержащее $\Gamma_{10,2}^3$, даже для наибольшего значения Γ_{np}^3 получилось значительно меньшей величины, чем остальные слагаемые, то в дальнейшем им можно пренебречь. Вычислив сумму и разность в скобках и подставив $m = 0,32$, получаем:

$$S_s = 0,0063U_1 - 0,017U_2 + 0,05U_3 - 0,14U_4 + 0,41U_5 -$$

$$- 0,14U_6 + 0,05U_7 - 0,017U_8 + 0,0063U_9 - 0,00022U_{10}$$

Произведем нормировку относительно U_5 , разделив все члены на 0,41.

$$S_{SH} = 0,015U_1 - 0,041U_2 + 0,12U_3 - 0,34U_4 + U_5 -$$

$$- 0,34U_6 + 0,12U_7 - 0,041U_8 + 0,015U_9 - 0,005U_{10} \quad (5)$$

Рассматривая последнее выражение, можно заметить, что коэффициенты при U быстро падают по мере удаления от слагаемого U_5 . Практически достаточно, в зависимости от требуемой точности, ограничиться двумя или четырьмя соседними с U_5 . Учитывая это обстоятельство, полная схема компенсации для рассматриваемого случая описывается следующей системой:

$$\begin{cases} S_1 = U_1 - 0,34U_2 + 0 + 0 \dots \dots \dots + 0 \\ S_2 = -0,34U_1 + U_2 - 0,34U_3 + 0 \dots \dots + 0 \\ S_3 = 0 \quad -0,34U_2 + U_3 - 0,34U_4 \dots + 0 \end{cases}$$

В заключение заметим следующее. Рассматривая выражение (5), нетрудно заметить, что оно с достаточной точностью может быть заменено следующей формулой:

$$S_{SH} = m^4 U_1 - m^3 U_2 + m^2 U_3 - m U_4 + U_5 - m U_6 + \dots - m^5 U_{10}$$

Или, ограничиваясь соседними с U_i членами, можно записать

$$S_i \approx -m U_{i-1} + U_i - m U_{i+1} \quad (4a)$$

Последнюю формулу можно рекомендовать для инженерного расчета с достаточной степенью точности, особенно при малых m .

Порядок расчета.

1. Рассчитывается коэффициент перекрытия

$$m = \frac{K_{ii}}{K_{i,i+1}}$$

где K_{ii} - площадь под кривой частотной характеристики i -го фильтра в i -й полосе усреднения

$K_{i,i+1}$ - площадь под кривой i -го фильтра в $i+1$ полосе усреднения.

2. Рассчитывается величина сопротивлений

$R_{i,i+1}$ при заданном R_{ii} (I)

$$R_{i,i+1} = \frac{R_{ii}}{m}$$

3. Собирается схема рис.3.

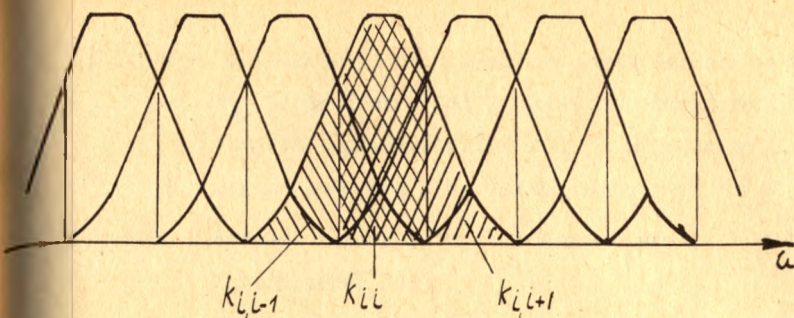


рис 1

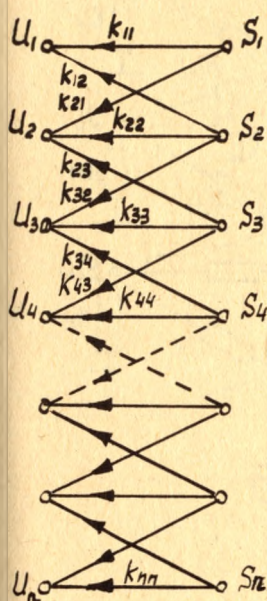


рис 2a

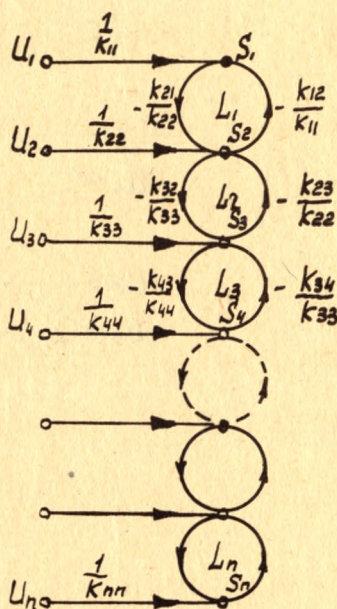


рис 2б

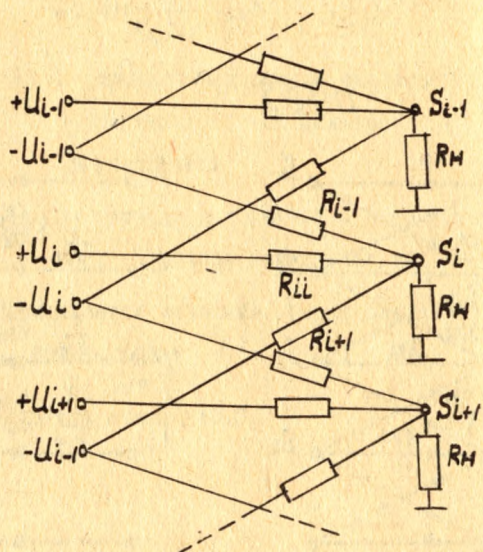


рис 3

Литература

1. Б.М.Лобанов "Исследование одного метода улучшения основных параметров параллельного спектрального анализа". Сб.Радиотехнические измерения под ред.Богдановича Б.М. Минск, 1965.
2. Л.Робишо, М.Буавер, Ж.Робер. "Направленные графы". Изд.Энергия, 1964.
3. С.Мезон, Г.Циммерман "Электронные цепи, сигналы и системы". ИИЛ, 1963.

AMPLITUDE ERROR COMPENSATION CIRCUIT FOR SPECTRUM
WITH TYPICAL FREQUENCY CHARACTERISTICS

(Summary)

The frequency filter characteristics in spectrum analyzers are known to differ usually from the required II-type ones. This gives rise to an error in determining the spectral density in the given band due to the effect of the neighbouring values of spectral density. This paper elaborates the given error compensation method earlier proposed by the author. The paper deals with the compensation circuit designed for analyzer with frequency filter characteristics approximating the resonance characteristics. The basic system of analyzer equations is solved by means of linear-directed curves. This approach permitted the author to obtain easily visible results from the point of view of the evaluation of the effect of roughly approximated values of the spectral density on the accuracy of the analysis involving the use of compensation technique.

The applied method made it possible to obtain a sufficiently accurate expression convenient for carrying out practical calculations of the compensation circuit. Besides, in the case when the frequency filter characteristics differ from the II-type ones to a minor degree, we obtain a simple expression good for engineering calculations.

At the end of the paper the author deals with the calculation procedure of the compensation circuit on the basis of the derived expression. The data obtained confirm the usefulness of application of the error compensation system in actual analyzers.

The comparatively easy and readily calculable circuit enables to considerably diminish the requirements for rectangularity of the filter characteristics. Moreover, this technique makes it possible to easily standardize the information of the analyzers provided with various types of filter systems.

РЕЧЕВОЙ СИГНАЛ И ФОНЕМНЫЙ ОБРАЗ.

(Канд.фил.наук-Л.П.БЛОХИНА, канд.фил.наук-Р.К.ПОТАПОВА)

Речевой сигнал, преобразованный в дискретные спектральные картины, представляет собой единичную во времени и индивидуальную в артикуляторно-акустическом плане реализацию; множество подобных реализаций составляют "образ", то есть локальное множество точек в n -мерном выборочном пространстве, где вдоль осей этого пространства отложены значения признаков (параметров), характеризующих данный образ. В результате анализа спектральных картин единичных реализаций определяем с помощью статистико-математического аппарата параметры, на основании которых можно найти фонемный образ.

Для ответа на вопрос о возможности нахождения фонемного образа на акустическом уровне было проведено исследование предударных гласных, произнесенных двумя группами дикторов: носителями литературной нормы и представителями вжновеликорусской группы диалектов русского языка.*)

В качестве экспериментального материала были использованы гласные первого слога в словах: "сестра", "пятак" (ж) произнесенных в коротких фразах.

*) В эксперименте участвовало 60 дикторов.

ж) Согласно данным диалектологии гласные "э" и "а" после мягких согласных в литературной норме совпадают в одном варианте - "и". Для южных диалектов для указанной позиции характерно произнесение гласных "э" и "а" или как звука "а" (в орфографии "я"), или как звука "э" (в орфографии "е").

В результате проведенного анализа был выделен ряд параметров, на основании которых были определены фонемные образы.

К числу этих параметров были отнесены:

1. f_1 - частота первой форманты.
2. f_2 - частота второй форманты.
3. f_3 - частота третьей форманты.
4. f_2/f_1 - отношение частот второй и первой формант.
5. f_3/f_1 - отношение частот третьей и первой формант.

Для количественной оценки данных, полученных в результате первичного считывания спектрограмм (переведения данных распределения энергии по полосам частот в цифровые значения) была составлена программа для ЭВМ "Стрела". Были запрограммированы задачи для следующих вычислений.

1. Амплитуды, усредненные по сечениям устойчивой части отдельных реализаций экспериментальных звуков, корректировались в соответствии с частотной характеристикой усилительного тракта, задаваемой в дБ.*)

А i умножалась на соответствующую величину $\frac{1}{K_i}$, где K_i - коэффициент усиления i - фильтра, предварительно переведенный из децибельной в линейную величину.

Следовательно, дальнейшей машинной обработке подвергалась величина $A_i \cdot \frac{1}{K_i}$

2. Для каждой экспериментальной единицы определялись номера фильтров, соответствующие максимальному значению.

$$A_m(N_m) = (A_i \cdot \frac{1}{K_i}) \max$$

3. Полученные номера фильтров N_m с помощью частотных таблиц переводились в значение частот f_m для каждой

*) В спектрографе системы ЛЭФИПР для высоких частот введена коррекция - усиление 3-4 дБ на октаву.

реализации экспериментальных звуков.

4. Далее по величине f_m для всех реализаций экспериментального звука находились:

1) $\bar{x}$ - среднее значение

$$\bar{x} = \frac{1}{N} \sum x_i; i=1.$$

2) G - среднее квадратичное отклонение.

$$G = \sqrt{\frac{1}{N} \sum n_i (x_i - \bar{x})^2}$$

3) v - коэффициент вариации.

$$v = \frac{G}{\bar{x}}$$

4) $G_{\bar{x}}$ - стандартное отклонение от среднего значения.

$$G_{\bar{x}} = \frac{G}{\sqrt{n}} \text{ где } G - \text{среднее квадратичное отклонение,}$$

n - число вариант .

5) U_p - доверительный интервал.

$$U_p = \bar{x} \pm 2 G_{\bar{x}}, \text{ где } p = 95,5\%$$

6) И, наконец, степень надежности разграничения анализируемых фонемных образов определялась по критерию Стюдента, а именно, вычислялась величина

$$t = \frac{\bar{x}_k - \bar{y}_e}{\sqrt{G_k^2 + G_e^2}},$$

где k - речевая группа говорящего;

e - литературная норма.

По значениям n (число вариантов) и t определялась p - вероятность того, что разность средних значений, рассматриваемых параметров не случайна, а отражает объективно существующую картину явления - наличие параметров для разграничения инвариантных фонемных образов речевого сиг-

нала в двух рассматриваемых группах: литературной норме и
вжных диалектов русского языка.

В результате статистической обработки значений параметров $f_1, f_2, f_2/f_1$ *) были получены следующие данные:

Параметр f_1

Литерат.	Ср. значение	Коефф. вариаци.	Доверит. интер. при $p=95\%$
норма	320	12%	$\pm 11,4$
вжные говоры	400	9%	± 15

Параметр f_2

Литерат. норма	1900	8%	± 85
вжные говоры	1650	12%	± 102

Параметр f_2/f_1

Литер. норма	5,9	10%	$\pm 0,24$
вжные говоры	4,3	15%	$\pm 0,3$

Анализ данных по параметру f_1 показал, что преобладающее число речевых сигналов, полученных в результате реализации единичных произнесений экспериментальных слов носителями литературной нормы и вжных говоров, характеризуются значениями f_1 , лежащими в зоне доверительного интервала среднего значения соответствующей группы.

*) Значения по параметрам f_3 и f_3/f_1 не приводятся, т.к. имеющиеся данные по этим параметрам недостаточны для статист. обраб.

Процент попадания значений f_1 в зону доверительного интервала для слова "пятак" равен 80%, для слова "сестра" - 70%. Таким образом, параметр f_1 можно считать достаточно надежным в целях разграничения по этому параметру фонемных образов для речевого сигнала, полученного от двух групп дикторов: носителей литературной нормы и южных говоров.

Параметр f_2 является менее индикативным для нахождения фонемных образов исследованных речевых сигналов.

Средние значения f_2 имеют довольно большие зоны перекрытия. Разграничение фонемных образов по параметру f_2 может быть произведено только с очень малой степенью надежности (равной приблизительно 30-40%).

Наиболее индикативным параметром является, по данным нашего исследования, параметр f_2/f_1 . Сравнение значений f_2/f_1 реализаций произнесений двух исследованных групп дикторов обнаружило наименьшую зону перекрытия. Разграничение фонемных образов речевого сигнала по параметру f_2/f_1 может быть произведено с большой степенью надежности (равной приблизительно 85%).

Высокую степень надежности параметра f_2/f_1 подтверждают также данные, полученные в результате применения критерия Стьюдента ($p=95\%$).

L.P.Blochina & P.K. Potapova
Laboratory of Experimental
Phonetics and Psychology of
Speech, Moscow Institute of
Foreign Languages

SPEECH SIGNAL AND PHONEMIC PICTURE

(Summary)

The present paper represents an initial stage of research of the problem of interconnection of the number of specific realizations (speech signals) and parameters (features) obtained as a result of the analysis of spectral pictures of single realizations by applying mathematical and statistical approach. The main problem is to find such parameters which, when taken together (or selectively), would make it possible to determine the sought-for-phonemic image and, moreover, would enable to distinguish between certain phonemic images with a sufficient degree of probability.

The experimental data show that the areas of selected space corresponding to different phonemic images overlap to one degree or another. The degree of overlapping would depend on the kind of parameter which is taken as a basic one when making a distinction between phonemic images.

О РЕЧЕВЫХ ВАРИАНТАХ КАК О ФОРМЕ СУЩЕСТВОВАНИЯ ЯЗЫКОВОГО ИНВАРИАНТА¹⁾

М.Г.Каспарова, В.Д.Лаптев

I. Постановка вопроса и методика эксперимента

В своем обзоре Линдгрэн (I) указывает на существование двух основных направлений в решении проблемы объективного распознавания речи. Представители первого направления формулируют свою задачу следующим образом: "... в окончательном виде распознающая машина будет в некотором виде адаптивной, т.е. подстраивающейся под новых дикторов" (Хендал, Мичиганский ун-т). Другой путь намечает Мартин (R SA): "Вместо контролирования дикторов мы хотим лучше найти истинные инварианты (звуков речи)".

В лаборатории экспериментальной фонетики и психологии речи I МГПИИЯ находит свое развитие третий путь, который, на наш взгляд, является органическим сочетанием описанных выше двух направлений, а именно - нахождение истинных вариантов для описания языкового инварианта. Мы не имеем в виду варианты звука, которые определяются различной фонетической позицией (начальной, конечной, срединной, ударной, безударной и т.д.) или влиянием предшествующих и последующих звуков последовательности. Речь идет о разнoвариантном нормированном произнесении звука различными людьми.

1) Исследование выполнено в ЛЭФИПР под руководством проф. В.А.Артемова.

Известно, что акустический эффект речевой деятельности человека определяется анатомо-физиологическими особенностями его речедвигательного аппарата и артикуляции. Если на основе антропометрических измерений создать классификацию анатомических групп людей и на основе изучения типа высшей нервной деятельности разработать классификационную систему артикуляционных движений, то открылась бы возможность описать произносительные группы людей. Добавим, что для этого нам пришлось бы еще учесть (и не в последнюю очередь) различия в произношении звуков и манеры говорить в зависимости от культуры речи и характера говорящего.

Однако такой подход к решению проблемы явился бы весьма не эффективным, т.к. факторы, определяющие форму речевой волны, необычайно сложны, многочисленны, непостоянны и не все еще известны. Кроме того, пришлось бы проверять многочисленные взаимозависимости между ними. Например, несомненно, анатомические особенности тракта в известной мере формируют звучание. Но говорящий при корректирующем контроле слухового аппарата с помощью обратной связи отрабатывает и отбирает такие движения, которые позволяют создать звук, максимально приближающийся к языковой норме. Этой особенностью формирования звуков объясняется тот факт, что 67% больных с разного рода аномалиями зубочелюстной системы имеют нормальное произношение (2). Такая взаимокомпенсаторная связь маскирует в какой-то степени характеристики тракта и артикуляции и нарушает однозначную корреляцию между каждым из них, с одной стороны, и звуковой волной, с другой. Их особенности суммарно отражаются в звуковой волне.

В связи с этим мы полагаем, что изучение акустической формы речи – наиболее короткий путь к выявлению произносительных групп людей. Такие работы на интонационном уровне ведутся в ЛЭФИПР в течение нескольких лет и были освещены в печати (В.И.Ильина, 3). В спектральном плане небольшая попытка в этом направлении была сделана нами. Изучалось произнесение звука (ш) в одной фонетической позиции (на слове "большой"). Выбор звука определяется следующим. В литературе отмечено, что фрикативные звуки являются более стабильными, нежели гласные, формантные характеристики которых весьма индивидуальны (1). Тем интереснее было проследить варианты фрикативного в нормативном произнесении.

Было отобрано 53 диктора (мужчины, 22–25 лет) с правильным московским произношением. В первой серии 35 дикторов произнесли слово "большой" в предложении "Большой театр приехал в Париж", по 3 раза каждый. Во второй серии другие 18 дикторов произнесли изолировано слово "большой" по 20 раз каждый. Анализ был произведен на речевом спектрографе типа С-48, имеющем 48 параллельных каналов с частотным диапазоном 60–12000 гц. Детальному изучению подвергался отрезок в 15 мсек. из стационарной части звука. На пяти звуках каждого диктора исследуемые отношения были прослежены на протяжении всей длительности звука.

Обработка результатов измерений проводилась в направлении нахождения не средних, а предельных значений характеристик.

2. Результаты эксперимента и выводы

Спектральная огибающая, описывающая пространственный контур звука (φ), совершенно определенно вырисовывает три сегмента:

- I сегмент расположен в области 60-500 гц;
- II сегмент ("нулевая область") -" 500-1000 гц;
- III сегмент ("формантная область") -" 1000-12000 гц.

Типичная спектральная огибающая звука (φ) в указанной выше позиции приведена на рис. I.

Такова общая конфигурация русского (φ). Она может в теоретическом плане рассматриваться как инвариант (φ).

Однако в конкретной реализации количественные характеристики этих трех областей значительно варьируют, сохраняя определенную стабильность внутри группы дикторов. Ниже приводятся групповые вариации трех сегментов.

а) Групповые вариации I сегмента (полоса до 500 гц).

Можно наметить три группы по вариациям в реализации данного сегмента.

I группа. У дикторов этой группы наблюдается увеличение энергии в области от 270-320 гц, образующее максимум высотой в 4-5 усл. единиц.¹⁾

II группа. Дикторов данной группы характеризует отсутствие максимумов в этом сегменте. На спектрограмме наблюдается ровное плато с амплитудой в 1-2 усл.ед.

¹⁾ В спектрографе С-48 выходы частотных фильтров имеют линейные детекторы. Масштаб по оси амплитуд также линейный. Условные единицы численно соответствуют высоте максимумов, измеренных по ленте шлейфного осциллографа.

III группа. В данную группу объединяются дикторы, у которых описываемый сегмент сливается с II сегментом ("область нулей"). Это означает, что на спектрограмме амплитуда всей полосы до 1000 гц равна 0-0,5 услов.ед.

б) Групповые вариации II сегмента (500-1000 гц).

Уровень интенсивности частот данного диапазона незначителен (от 0 до 0,5 усл.ед.), и с перспективной точки зрения этот сегмент правомерно было бы назвать "область нулей".

Дикторы различаются между собой областью локализации II сегмента. По данному параметру наметились четыре группы.

I группа. Нижняя граница этой области - 500 гц.

Верхняя граница доходит до 1000 гц.

II группа. Нижняя граница - 500 гц.

Верхняя граница доходит - 1500-1700 гц.

III группа (малочисленная) располагает областью нулей в пределах от 500 до 2500 гц.

IV группа. В данной группе нижняя граница области проходит не на 500 гц, а с 90-120 гц, т.е. по существу эта группа совпадает с 3-й группой по I-му сегменту. Иначе говоря, у дикторов данной группы имеется не три, а только два сегмента в спектральной картине (ш).

Кроме того, по данному признаку можно наметить группу, в которой указанный признак имеет фиксированное значение: колебание границ II сегмента незначительно. Вторая группа по стабильности признака характеризуется большим разбросом

верхней границы этого сегмента, которая может составлять от 1130 до 2500 гц.

в) Характеристика III сегмента.

Данный сегмент является по существу "формантной областью" изучаемого звука. Г.Фант пишет, что "тонкая структура картины, дающей интенсивность в функции времени и частоты для фрикативных согласных и других глухих звуков представляется на широкополосной спектрограмме в виде вертикальной штриховки. Положение этих штрихов случайны и, как правило, не связаны с движением речевых органов. Они просто отражают статистические свойства шумового источника" (4). В данном положении заключено, по крайней мере, две неточности. Во-первых, усиление энергии данной частотной области является не случайным явлением, а языковой нормой и прослеживается абсолютно у всех дикторов московской группы. Естественно, что подобный акустический эффект формируется в результате нормированного произношения. Во-вторых, энергетические характеристики отдельных частей III сегмента изменяются в процессе развития звука, что, несомненно, связано с движением органов речи. За рамки этих общих закономерностей не выходят эффекты, на которые указывает Г.Фант.

Для удобства рассмотрения разделим данный сегмент на три части (см.рис.I):

1-я часть - 1100-2800 гц

2-я часть (центральная) - 2800-6500 гц.

3-я часть - 6500-12000 гц.

Анализ показал, что III сегмент характеризуется четкой энергетической структурой, являющейся инвариантной для дикторов данной группы: отношение суммы амплитуд всего III сегмента к той же характеристике его 2-й части (центральной) лежит в пределах 1,5-2,5 (только у одного диктора это отношение оказалось в пределах 2-3). Это означает, что центральная часть III сегмента спектрального сечения является самой "тяжелой". Такая структура сохраняется почти стабильной на протяжении всей квазистационарной части звучания звука. Начальные и конечные 20-40 мсек. произнесения характеризуются уменьшением веса центральной части сегмента. Таким образом, в отличие от английских, польских и шведских фрикативных, где общие для всех дикторов признаки лежат в диапазоне от 5000-13000 гц (Ясsem, 5), общий признак для группы русских дикторов лежит в диапазоне от 2800-6500 гц.

Амплитудный вес I-й и 3-й части сегмента занимает гораздо меньшее место (по сравнению с центральной частью) в общем весе формантной области (ш). Описываемая характеристика (отношение амплитудной суммы всего III-го сегмента к амплитудной сумме одной из его трех частей) приобретает различные значения в процессе развития звука. Выделяется группа дикторов, у которых вес нижней области (I части) сегмента стремительно увеличивается по мере произнесения звука. Другая группа дикторов отличается стабильным весом нижней области. Отмечается даже некоторая потеря веса в центральных кадрах произнесения.

Необходимо отметить, что абсолютные количественные характеристики внутри формантной области весьма вариативны

даже в пределах произнесения одного диктора. Релевантным остается грубая структура признака.

Интересные данные получены в результате рассмотрения частотной структуры III сегмента. Всего очевиднее, что локализация формантных максимумов внутри III сегмента не несет языковой информации: от диктора к диктору максимумы незакономерно перемещаются в диапазоне от 5 кгц до 10 кгц и такая мобильность не приводит к искажению звука (ш).

Однако локализация максимума несет информацию о произносительной группе диктора. По этому признаку различились три группы:

- 1) низкочастотное (ш) – максимум в области
1100-3800 гц (11%)
- 2) среднечастотное (ш) – "– 3500-6000 гц (61%)
- 3) высокочастотное (ш) – "– 5500-9000 гц (28%).

Можно было бы провести групповой анализ по количеству максимумов, их ширине и т.д. Но нам представляется, что приведенные выше данные в достаточной мере иллюстрируют мысль о том, что реальностью языкового инварианта являются его нормированные варианты. Сам же языковой инвариант является некоторой грубой структурой, лингвистическое изучение которой полезно в теоретико-информационном плане.

Систематическое изучение речи в направлении нахождения правильных вариантов произнесения на спектральном и интонационном уровнях даст возможность создать классификацию произносительных групп людей, что явится большим подспорьем в вопросе подготовки программы для машинного распознавания звучащей речи.

Схема русского "ш"

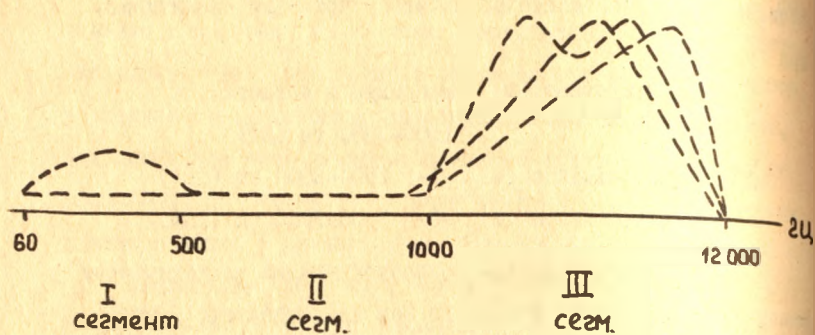


Рис.1

Литература

1. Линдгрен Ж. "Зарубежная радиоэлектроника", машинное распознавание речи. 1966, 3.
2. Омельченко Н.А. Косноязычие и его связь с аномалиями зубочелюстной системы и слуха. Медгиз, 1961.
3. Ильина В.И. Ж. "Вопросы психологии", 1965, № 6.
К вопросу об изучении вариантов речи.
4. Фант Г. Акустическая теория речеобразования.
Изд. Наука, М. 1964.
5. Jasson W. Noise Spectra of Swedish English and Polish Fricatives Speech Communication Seminar, Stockholm, 1962.

M.G. Kasparova, V.D. Laptev
Laboratory of Experimental Phonetics
and Psychology of Speech
Moscow Institute of Foreign Languages

SPEECH VARIANTS AS FORM OF EXISTENCE OF LANGUAGE INVARIANT

Summary

The paper is devoted to the problem of objective speech identification. The language invariant is proposed to be considered as statistically interpreted aggregate of normalized variants. By analyzing the spectrum curve of Russian (III) the authors show that both common and group features can be identified in pronunciation of various people.

Преподаватель кафедры французской
филологии Тбилисского государствен-
ного университета

Р.Г. МЕГРЕЛИШВИЛИ

СПЕКТРАЛЬНЫЙ АНАЛИЗ ЩЕЛЕВЫХ СОГЛАСНЫХ ФРАНЦУЗСКОГО ЯЗЫКА В СОЧЕТАНИИ С ГЛАСНЫМИ (∂, i, α) В ПОЗИЦИИ СТ^I)

В данной работе мы ставили себе задачей выделить устой-
чивые акустические характеристики щелевых согласных ($f-v, j-z,$
 $\int-z$) посредством сопоставления спектров согласных в
позиции СТ со спектром изолированных согласных.

Рассмотрение интересующей нас группы согласных в соче-
тании с гласными приводит к необходимости изучения спект-
ральных характеристик, наблюдающихся в момент перехода от
согласного к гласному. Таким образом, мы исследуем спектр
согласного как в момент наибольшей устойчивости звука, так
и в процессе разворачивания спектра при переходе от соглас-
ного к гласному. При выборе гласных мы руководствовались
следующим: гласный (∂) является характерным оттенком
французских согласных, f - картина которого имеет прибли-
зительно такой же рисунок, что и F - картина нейтраль-
ного гласного с резонансными пиками на частотах
приблизительно в 500 гц, 1500 гц, 2500 гц и т.д. Гласные
(i) и (α) берутся как самый диффузный и самый
компактный гласные.

1) Руководителем работы является профессор К.К. Барышникова.

При изучении данной группы согласных в позиции СГ исследовались: частотные характеристики спектра, интенсивность частотных составляющих спектра, сдвиги формант гласных в переходные моменты и длительность переходных процессов.

Экспериментальный материал записан на спектрографе конструкции лаборатории I МГПИИЯ и наговорен двумя дикторами — мужчинами, произношение которых соответствует произносительным нормам литературного французского языка.

Спектрограммы изученных сочетаний позволяют получить следующие данные:

- 1/ о расположении резонансных частот в спектре согласных,
- 2/ о ширине спектра (диапазон разброса резонансных частот),
- 3/ об областях максимальной концентрации энергии и границы спада интенсивности частотных составляющих спектра согласного,
- 4/ о соотношении интенсивности резонансных частот глухого согласного и его звонкой пары,
- 5/ о сдвигах формант гласных в момент перехода от согласного к гласному.

Экспериментальные данные о частотной характеристике составляющих спектра щелевых согласных, а также данные о частоте основного тона и длительности согласных представлены в суммарном виде в таблицах I-8. В приложении на рисунках (I-6) представлены огибающие спектров, т.е. функциональные зависимости амплитуды составляющих спектра от частоты. (Огибающая спектра строится по наиболее характерному

кадру спектрограммы, т.е. берется один из кадров выдержки (согласного).

В приложении приведены также графики функциональной зависимости частотных составляющих от времени в момент перехода от согласного к гласному (рис.7-24).

I. Шелевые согласные (f-v, j-z, ʃ-ʒ) с гласным (a, i, e) в позиции СГ

I.I. Частотная характеристика согласных:

а) Сочетания (f) + (a, i, e).

Таблица I

Сочетание	Согласный	Диктор	Основной тон	Нижний и верхний пределы частот (гц)	Резонансные частоты (гц)	Длительность мсек.
[f]	I			800-12100	1100, 1300, 1700, 1950, 2200, 3000, 4150, 5750, 6800, 7600, 8380, 9200, 10100	640
[f+a]	[f]	I		800-10100	950, 1100, 1300, 2200, 3000, 5750, 6800, 7600, 8380	220
[f+i]	[f]	I		800-10100	1100, 5750, 6800	140
[f+e]	[f]	I		800-10100	1100, 5750, 6800	140
	[f]	II		800-9200	950, 1100, 1300, 2200, 3200, 5750, 6300, 6800, 7600, 8380	580
[f+a]	[f]	II		800-12100	950, 1100, 2200, 3200, 5400, 6800, 9200, 10100	220

Сочетание	Согласный	Диктор	Основной тон	Нижний и верхний пределы частот (гц)	Резонансные частоты (гц)	Длительность мсек.
[f+i]	[f]	II		800-9200	950, 1100, 1300, 3200, 5750, 6800	120
[f+a]	[f]	II		800-8200	1100, 5750, 6800 ...	120

Приведенные данные показывают, что в спектре согласного (f) нижний предел частот 800 гц, верхний предел частот более низок в произношении I диктора. Вместо 12100 гц при изолированном произнесении согласного (f), имеем в позиции СГ - 10100 гц, в произношении П диктора верхний предел частот - 9200 гц; в соседстве с гласным верхняя граница смещается в сторону высоких частот и равна 12100 гц.

Плотность спектра согласного меняется в зависимости от соседнего гласного и возрастает в последовательности

$$[f+a] - [f+i] - [f+e]$$

б) Сочетания (v) + (a, i, α)

Таблица 2

Сочетание	Согласный	Диктор	Основной тон	Нижний и верхний пределы частот (гц)	Резонансные частоты (гц)	Длительность мсек
[v]	I	I00 200	800-10100	1100, 2200, 3000, 4150, 5750, 6300, 6800, 9200	920	
[v+a]	[v]	I	I00 200	800-6800	950, 1100, 1300, 5750	200
[v+i]	[v]	I	I00 200	800-6800	950, 1100, 5750.....	120

Соче- тание	Сог- лас- ный	Дик- тор	Ос- нов- ной тон	Нижний и верхний пределы частот (гц)	Резонансные частоты (гц)	Дли- тель- ность мсек
[v+ɑ]	[v]	I	100 200			120
	[v]	II	100 200	800-9200	950, 5750, 6800, 8380	400
[v+ə]	[v]	II	100 200	800-12100	950, 2200, 5400, 6800, 9200, 10100	200
[v+i]	[v]	II	100 200	800-6800	950, 2200, 5400	100
[v+ɑ]	[v]	II	100 200			100

Сопоставление данных, изложенных в таблице 2 позволяет заключить, что в спектре согласного (v) в сочетаниях ($v + ə$) и ($v + i$) резонансная частота в 800 гц остается устойчивой, меняется верхний предел частот. В произношении I диктора, верхний предел частот изолированного согласного (v) равен 10100 гц, в группе ($v + ə$) и ($v + i$) - 6800 гц; в произношении II диктора верхний предел частот изолированного (v) равен 6800 гц. В соседстве с гласным ($ɑ$) в спектре согласного (v) отсутствуют высокие резонансные частоты. Поэтому, в данной позиции согласный (v) можно охарактеризовать, в основном, по переходному моменту.

Частота основного тона звонкого щелевого согласного колеблется в диапазоне 100 - 200гц.

в) Сочетания (j) + (j, i, d)

Таблица 3

Сочетание	Согласный	Ос-нов-ной тон	Диктор	Нижний и верхний пределы частот (гц)	Резонажные частоты (гц)	Длительность мсек
[j]			I	4600-12100	6300, 6800, 7600, 8380, 9200, 10100	900
[$j+d$]	[j]		I	4600-12100	6300, 6800, 8380, 9200, 10100	320
[$j+i$]	[j]		I	4150-10100	6300, 6800, 8380	200
[$j+d$]	[j]		I	6300-9200	6800	140
		[j]	II	5400-1000	6300, 6800, 7600, 9200	540
[$j+d$]	[j]		II	4150-11100	5750, 6800, 7000, 8380, 9200, 10100	220
[$j+i$]	[j]		II	4150-11100	5400, 5750, 6300, 7600, 9200	200
[$j+d$]	[j]		II	4150-11100	5400, 6300, 6800, 7600, 9200	140

На основании данных, приведенных в таблице 3, можно заключить, что в спектре согласного (j) нижний предел частот в названных сочетаниях, в произношении II диктора, сдвигается в сторону низких частот и лежит на частоте в 4150 гц, в произношении I диктора, в сочетании ($j+d$) остается устойчивым (4600 гц), в сочетании ($j+i$) равен 4150 гц, а в сочетании ($j+d$) равен 6300 гц. Верхний предел частот в названных сочетаниях остается устойчивым, в произношении II диктора (11100), а в произношении I диктора меняется: в сочетании ($j+i$) равен 10100, в сочетании

($\beta + \alpha$) равен 9200 гц и остается устойчивым в сочетании ($\beta + \partial$). Плотность спектра согласного уменьшается в последовательности ($\beta + \partial$) $\rightarrow$ ($\beta + i$) $\rightarrow$ ($\beta + \alpha$).

г) Сочетания (β) + (∂, i, α)

Таблица 4

Сочетание	Согласный	Диктор	Ос-нов-ной тон	Нижний и верхний пределы частот (гц)	Резонансные гласные (гц)	Длительность мсек
	[β]	I	100 200	4600-12100	6300, 6800, 7600, 8380, 101000	840
[$\beta + \partial$]	[β]	I	100 200	4600-10100	6300, 6800, 7600	220
[$\beta + i$]	[β]	I	100 200	6300-6800		140
[$\beta + \alpha$]	[β]	I	100 200	6300-9200	6800	120
	[β]	II	100 200	5750-11100	6800, 7600, 8380, 9200, 10100	740
[$\beta + \partial$]	[β]	II	100 200	4150-11100	5400, 5750, 6800, 8380, 9200, 10100	160
[$\beta + i$]	[β]	II	100 200	4150-11100	5750, 6300, 7600, 9200	220
[$\beta + \alpha$]	[β]	II	100 200	4150-11100	5750, 6300, 8380, 9200	140

Анализ данных показывает, что в спектре согласного (β) в сочетании ($\beta + \partial$) резонансная частота 4600 гц остается устойчивой, в соседстве с гласными (i, α) смещается в сторону высоких частот и равна 6300 гц (I диктор).

В произношении II диктора нижний предел частот в названных сочетаниях смещается в сторону низких частот и располо-

жен на частоте в 4150 гц. Верхний предел частот остается устойчивым и равен 11100 гц (П диктор), в произношении I диктора понижается в последовательности: $(\bar{z} + \bar{a}) \rightarrow (\bar{z} + \bar{u}) \rightarrow (\bar{z} + \bar{i})$.

Плотность спектра повышается в последовательности $(\bar{z} + \bar{i}) \rightarrow (\bar{z} + \bar{u}) \rightarrow (\bar{z} + \bar{a})$, в произношении I диктора. Максимальная плотность спектра согласного $(\bar{z})$ наблюдается в группе $(\bar{z} + \bar{a})$, а минимальная в группах $(\bar{z} + \bar{i})$ и $(\bar{z} + \bar{u})$ (П диктор).

В области низких частот наиболее устойчивая резонансная частота расположена на высоте 200 гц.

д) Сочетания $(\int \bar{z} + (\bar{a}, \bar{i}, \bar{u}))$

Таблица 5

Сочетание	Согласный	Диктор	Ос-нов-ной тон	Нижний и верхний пределы частот (гц)	Резонансные частоты (гц)	Длительность мсек
$[\bar{z}]$	I			1950-6800	2200, 3000, 4600, 5750	720
$[\bar{z} + \bar{a}]$	$[\bar{z}]$	I		1950-6800	2200, 2400, 3000, 4150, 5400	240
$[\bar{z} + \bar{i}]$	$[\bar{z}]$	I		2200-6800	3000, 4150, 5750	180
$[\bar{z} + \bar{u}]$	$[\bar{z}]$	I		2200-6800	3000, 3200, 5750	180
	$[\bar{z}]$	II		2200-11100	2650, 3000, 3582, 4150, 5750, 6800, 8380	440
$[\bar{z} + \bar{a}]$	$[\bar{z}]$	II		1950-6800	2200, 2650, 3000, 3582, 4150, 5400, 5750, 6300	180
$[\bar{z} + \bar{i}]$	$[\bar{z}]$	II		2200-6800	2400, 3000, 3200, 4600, 5700, 5750, 6300	200
$[\bar{z} + \bar{u}]$	$[\bar{z}]$	II		2200-11100	2400, 3000, 3582, 4600, 5400, 5750, 6300, 8380, 9200	160

Данные анализа спектра согласного ($\int$) показывают, что нижний предел частот изменяется от 1950 гц до 2200 гц, верхний предел частот - от 6800 гц до 11100 гц. В спектре согласного ($\int$) в названных сочетаниях имеются две области усиления энергии¹⁾. Доминирующая резонансная частота каждой области остается устойчивой (для I области доминирующая резонансная частота расположена на высоте 3000 гц, для II области - 5750 гц). Диапазон резонансных частот изменяется (таблица 6).

Таблица 6

I область		II область	
1950 гц - 3582 гц	($\int$ + ∂)	4150 гц - 6800 гц	
2200 гц - 3200 гц	($\int$ + i)	4150 гц - 6800 гц	
2200 гц - 3582 гц	($\int$ + α)	4600 гц - 11100 гц	

Наибольшей плотностью частотных составляющих характеризуется спектр согласного ($\int$) в сочетании ($\int$ + ∂).

е) Сочетания ($\int$) + (∂, i, α)

Таблица 7

Сочетание	Согласный	Диктор	Основной тон	Нижний и верхний пределы частот (Гц)	Резонансные частоты (Гц)	Длительность мсек
	[з]	I	100 200	1950-6800	2200, 2600, 3000, 4600, 4700, 5750	840
[з+а]	[з]	I	100 200	2200-6800	3000, 4150, 4600, 5750	200

1) В некоторых случаях две области образуют одну широкую полосу спектра (формантную зону).

Соче- таение	Сог- лас- ный	Дик- тор	Ос- нов- ной тон	Нижний и верхний пределы частот (гц)	Резонансные частоты (гц)	Дли- тель- ность мsec
[j+i]	[j]	I	100 200	2200-6800	3000, 4150, 5750	120
[j+λ]	[j]	I	100 200	2200-6800	3000, 5750	100
	[j]	П	100 200	2200-10100	2650, 3000, 3200, 3582, 4150, 5400, 5750, 6800, 8380	420
[j+ə]	[j]	П	100 200	2200-6800	3000, 4150, 5750	140
[j+i]	[j]	П	100 200	2650-8380	3000, 3582, 4600, 5750	120
[j+λ]	[j]	П	100 200	2200-6300	2400, 3000, 3582, 4600, 5750	140

Мы видим, что в спектре согласного (*з*) в зависимости от последующего гласного, нижний предел изменяется от 2200 гц до 2650 гц, **верхний** предел частот - от 6300 гц до 8380 гц.

В спектре согласного (*з*) в названных сочетаниях имеются две области усиления энергии. Доминирующая резонансная частота каждой области остается устойчивой (I область - 3000 гц, П область - 5750 гц), меняется диапазон резонансных частот.

Таблица 8

I область		П область	
2200 гц - 8000 гц	(<i>з</i> + <i>ə</i>)	4150 гц - 6800 гц	
2200 гц - 3582 гц	(<i>з</i> + <i>i</i>)	4150 гц - 8380 гц	
2200 гц - 3581 гц	(<i>з</i> + <i>λ</i>)	4600 гц - 6800 гц	

Плотность спектра повышается в последовательности $(\text{з} + \alpha) \rightarrow (\text{з} + i) \rightarrow (\text{з} + \alpha)$ (в произношении I диктора) и в последовательности $(\text{з} + \alpha) \rightarrow (\text{з} + i) \rightarrow (\text{з} + \alpha)$ в произношении II диктора.

В спектре согласного (з) появляются низкие частоты 100–200 гц, которые указывают на звонкость данного согласного.

I.2. Характеристика интенсивности частотных составляющих согласных

Характеристика интенсивности частотных составляющих в спектре щелевых согласных зависит от последующего гласного.

В сочетании $(f) + (\text{э}, i, \alpha)$ в спектре согласного максимальный уровень интенсивности наблюдается в соседстве с гласным (э) , минимальный – в соседстве с гласным (α) .

В сочетании $(\text{в}) + (\text{э}, i, \alpha)$ для согласного (в) характерно увеличение амплитуды интенсивности низких резонансных частот и уменьшение амплитуды интенсивности высоких резонансных частот. Интенсивность низких частотных составляющих в спектре согласного возрастает в последовательности: $(\text{в} + \alpha) \rightarrow (\text{в} + \text{э}) \rightarrow (\text{в} + i)$ (Рис. I, 2).

В сочетании $(\text{г}) + (\text{э}, i, \alpha)$ интенсивность частотных составляющих согласного возрастает в последовательности $(\text{г} + \alpha) \rightarrow (\text{г} + \text{э}) \rightarrow (\text{г} + i)$ в произношении I диктора и в последовательности $(\text{г} + \alpha) \rightarrow (\text{г} + i) \rightarrow (\text{г} + \text{э})$ в произношении II диктора.

В сочетании $(\text{з}) + (\text{э}, i, \alpha)$ для согласного (з) характерно увеличение амплитуды интенсивности низких резонансных частот.

нальных частот (100 гц, 200 гц) и уменьшение амплитуды интенсивности высоких резонансных частот (Рис.3,4).

Амплитуда интенсивности низких резонансных частот возрастает в последовательности $(z+i) \rightarrow (z+a) \rightarrow (z+\alpha)$, а амплитуда интенсивности высоких резонансных частот - в последовательности $(z+i) \rightarrow (z+\alpha) \rightarrow (z+a)$.

В сочетании $(\int) + (a, i, \alpha)$ амплитуда интенсивности частотных составляющих возрастает в последовательности $(f+i) \rightarrow (f+\alpha) \rightarrow (f+a)$. Соотношение амплитуды интенсивности доминирующих резонансных частот I и II областей меняется в зависимости от соседних гласных. В сочетании $(\int+a)$ амплитуды интенсивностей I и II областей располагаются на одном уровне в 30 дб, в сочетании $(\int+\alpha)$ и $(\int+i)$ амплитуды интенсивностей частотных составляющих I области резко понижаются.

В сочетании $(\cdot \int) + (a, i, \alpha)$ соотношение амплитуды интенсивностей частотных составляющих I и II областей показывает, что интенсивность частотных составляющих II области выше интенсивности частотных составляющих I области.

В сочетании $(\int+a)$ амплитуды интенсивности резонансной частоты в 200 гц ниже амплитуды интенсивности доминирующей резонансной частоты II области и выше амплитуды интенсивности доминирующей резонансной частоты II области и выше амплитуды интенсивности доминирующей резонансной частоты I области.

В сочетаниях $(\int+i)$ и $(\int+\alpha)$ амплитуды интенсивности резонансной частоты в 200 гц совпадают с амплитудой интенсивности доминирующей резонансной частоты II области

превышают доминирующую резонансную частоту I области
рис.5,6).

1.3. Длительность щелевых согласных

Наибольшая длительность щелевых согласных наблюдается
соседстве с гласным (э).

Длительность согласных меняется в пределах:

для согласного (ф) от 220 мсек до 120 мсек

для согласного (в) от 220 мсек до 100 мсек

для согласного (ш) от 320 мсек до 140 мсек

для согласного (з) от 220 мсек до 120 мсек

для согласного (ж) от 240 мсек до 160 мсек

для согласного (з) от 200 мсек до 100 мсек

Таблица 8

Образец таблицы анализа формант гласных (a, i, α)
в процессе перехода, в позиции СТ

Сочета- ние	Звук	Длительность в мсек			F_1 (гц)			F_2 (гц)			F_3 (гц)			диапа- зон смеще- ния
		пере- хода	выдерж- ка	глас- ного	при пере- ходе	выдерж- ка	диапа- зон смеще- ния	при пере- ходе	выдерж- ка	диапа- зон смеще- ния	при пере- ходе	выдерж- ка		
[f+a]	[ə]	60	60		300	536	236	800	1500	700	2200	2400	200	
[f+i]	[i]	40	100		200	300	100	2200	2400	200	3000	3200	200	
[f+α]	[α]	80	100		536	800	264	1100	1500	400	2400	3000	600	
[v+a]	[ə]	60	60		200	536	336	800	1500	700	2200	2400	200	
[v+i]	[i]	60	100		200	300	100	2200	2400	200	3000	3200	200	
[v+α]	[α]	60	100		200	800	600	950	1500	550	2400	3000	600	
[s+a]	[ə]	40	180		536	536	-	1500	1500	-	2400	2400	-	
[s+i]	[i]	40	60		200	300	100	2200	2400	200	3000	3200	200	
[s+α]	[α]	60	100		536	800	264	1500	1500	-	3000	3000	-	
[z+a]	[ə]	20	120		260	536	336	1500	1500	-	2400	2400	-	
[z+i]	[i]	40	100		200	300	100	2200	2400	200	3000	3200	200	
[z+α]	[α]	60	100		200	800	600	1500	1500	-	3000	3000	-	
[ʃ+a]	[ə]	40	80		300	536	236	1960	1500	450	2200	2400	200	
[ʃ+i]	[i]	20	180		-	300	-	2200	2400	200	3000	3200	-	
[ʃ+α]	[α]	60	100		536	800	336	1950	1500	450	3000	3000	-	
[ʒ+a]	[ə]	40	140		200	536	336	1950	1500	450	2200	2400	200	
[ʒ+i]	[i]	20	120		200	300	100	1950	2400	450	3000	3200	200	
[ʒ+α]	[α]	20	100		200	800	600	1952	1500	450	3000	3000	-	

2. Анализ переходных процессов

При анализе переходных процессов в сочетаниях СГ в подавляющем большинстве случаев наблюдаются две возможности:

1) реализация согласного полностью прекращается к моменту перехода. Проникновение частотных составляющих спектра согласного в переходную часть наблюдается в редких случаях;

2) на границе между гласным и переходными кадрами наблюдается перерыв, равный примерно 20 мсек, который четко отделяет спектр глухого целевого согласного от переходных кадров.

Исходя из анализа развернутого спектра в сочетаниях СГ, нам кажется целесообразным отнести переходные кадры к спектру гласного, так как в этот отрезок времени в спектре наблюдается четкий рисунок сдвигов формант гласных к частотам локуса соответствующего согласного.

При исследовании переходных процессов мы применяем трехформантный анализ гласных. Данные о расположении формант гласных в момент перехода, выдержка гласных, а также диапазон смещения формант в длительность переходных процессов приведены в таблице 8.

Следует отметить, что общим для всех исследованных случаев являются два момента:

во-первых, смещение f_2, f_3 гласных имеет одинаковый характер при переходе от глухого согласного и его звонкой пары (т.е. характер перехода один и тот же в сочетаниях $(f + \Gamma)$ и $(\nu + \Gamma)$; $(j + \Gamma)$ и $(z + \Gamma)$ и т.д.

При переходе от согласного к гласному F гласного сдвигается в сторону высоких частот - переход всегда положительный. При переходе от глухого согласного к гласному F , гласного незначительно сдвигается в сторону высоких частот или остается без изменения ;

во-вторых, в процессе перехода от звонкого щелевого согласного к последующему гласному амплитуда интенсивности частоты в 200 гц (основной тон) согласного повышается, если частота F , гласного близко расположена к частоте основного тона согласного (что наблюдается в сочетаниях согласных + [i] ; в том числе, когда частота F , гласного значительно превышает частоту основного тона согласного, при переходе наблюдается снижение амплитуды интенсивности частоты в 200 гц (основной тон), например в сочетаниях С + / э / и / э /.

Таким образом, характер смещения формант гласных в момент перехода в суммарном виде можно представить следующей схемой, в которой знаком "+" обозначается положительный переход форманты, знаком "-" - отрицательный переход и знаком "o" - нейтральный переход.

Сочетание	Характер перехода формант гласных			Рис.
	F ₁	F ₂	F ₃	
[f, v] + [ə]	+	+	+	7, 10
[ɹ] + [ə]	0	0	0	13
[z] + [ə]	+	0	0	16
[ʃ, ʒ] + [ə]	+	-	+	19, 22
[f, v] + [i]	+	+	+	8, 11
[ɹ, ʒ] + [i]	+	+	+	14, 17
[ʃ] + [i]	0	+	+	20
[ʒ] + [i]	+	+	+	23
[f, v] + [a]	+	+	+	9, 12
[ɹ, ʒ] + [a]	+	0	0	15, 10
[ʃ, ʒ] + [a]	+	-	0	21, 24

ВЫВОДЫ

Основные результаты, полученные при анализе спектрограмм щелевых согласных ($f-v, 1-z, f-3$) в изолированной позиции и в позиции СГ в соседстве с гласными (a, i, α) сводятся к следующему:

1. Щелевые согласные характеризуются высокочастотным спектром. Верхний предел частот спектра согласных ($f, 1, z$) в большинстве случаев достигает 12000 гц.

2. Ширина спектра щелевых согласных (диапазон разброса резонансных частот в спектре согласных) следующая:

для ($f-v$) - 800 гц - 12100 гц

для ($1-z$) - 4150 гц - 12100 гц

для ($f-3$) - 1950 гц - 6800 гц

3. Интенсивность частотных составляющих в спектре щелевых согласных располагается в убывающей последовательности от ($f-3$) $\rightarrow$ ($1-z$) $\rightarrow$ ($f-v$). В спектре согласных ($f-v, 1-z$) интенсивность прогрессивно повышается от низких к высоким частотам. В спектре согласных ($f-v$) наибольшая интенсивность частотных составляющих расположена в полосе от 5750 - 6800 гц. Спад интенсивности частотных составляющих в спектре согласного (f) наблюдается на частоте 10100 гц, а в спектре согласного (v) - на частоте 7000 гц.

Для согласных ($1-z$) наибольшая интенсивность частотных составляющих расположена в полосе от 7600 до 9200 гц, спад интенсивности наблюдается на частоте в 10100 гц.

Интенсивность частотных составляющих спектра согласных (f-з) расположена в полосе 2000 гц - 3000 гц и в полосе 5750 гц - 6800 гц. Для звонких щелевых согласных характерным является повышение интенсивности низких резонансных частот и понижение интенсивности высоких резонансных частот.

4. Сопоставление спектров выдержки щелевых согласных в изолированной позиции и в позиции СГ показывает, что структура спектра остается устойчивой, а также не наблюдаются изменения в формантной структуре выдержки гласных (таблица 8).

5. Спектр изолированных согласных характеризуется наибольшей плотностью, интенсивностью частотных составляющих и длительностью. (Таблицы I - 7 и рис. I - 6).

6. В позиции СГ в сочетании с гласными (э, и, а), в спектре щелевых согласных области усиления остаются устойчивыми; однако, изменяется плотность и интенсивность частотных составляющих. Четкий рисунок спектра щелевых согласных сохраняется в соседстве с гласным (э).

7. Сравнение спектров глухих и звонких щелевых согласных показывает, что спектр глухих согласных характеризуется большей плотностью и интенсивностью частотных составляющих.

8. Основной тон постоянно присутствует в спектре звонкого согласного. Гармонический источник звука или одновременно накладывается на источник шума, или предшествует источнику шума.

9. Анализ спектрограммы показывает, что для щелевых согласных в позиции СГ характерным является увеличение интенсивности в середине или к концу звука.

10. Наблюдается разница в длительности звуков: звонкие щелевые более кратки, чем соответствующие глухие. Длительность звонких щелевых меняется в пределах 100–200 мсек, глухих щелевых – 120 – 320 мсек.

11. Переход от звонкого щелевого согласного к гласному плавный, или непрерывный, а от глухого щелевого согласного к гласному в большинстве случаев наблюдается перерыв, равный примерно 20 мсек. В большинстве случаев в процессе перехода от согласного к гласному реализация согласного полностью прекращается.

12. В момент перехода от согласного к гласному форманты гласных подстраиваются к локусу согласных.

13. Степень зависимости спектра щелевых согласных от переходных процессов в позиции СГ следующая: согласные (f, v) больше зависят от переходного момента, чем согласные ($1, z$) и ($f, 3$). Частотные составляющие спектра согласных (f, v) слабо выражены на спектрограммах, поэтому, в основном, данные о характеристике спектра согласных (f, v) можно получить при анализе переходных процессов, пользуясь данными о движениях формант гласных. Относительно большая интенсивность частотных составляющих спектра согласных ($1-z, f-3$) дает возможность получить достаточную информацию об акустической картине данных звуков.

14. Длительность переходных процессов меняется в пределах от 20 мсек до 80 мсек. Наименьшей длительностью пере-

да характеризуются согласные, локусы f_1 , f_2 которых близко расположены к f_1 и f_2 соседних гласных, как например, в сочетаниях ($j + i$) и ($3 + i$).

15. При переходе от глухого щелевого согласного к гласному наблюдается незначительный сдвиг f_1 гласного, который объясняется глухостью предшествующего согласного.

Полученные результаты дают нам возможность перейти к следующему этапу исследования согласных, а именно, к изучению сочетаний ГС - ГСГ.

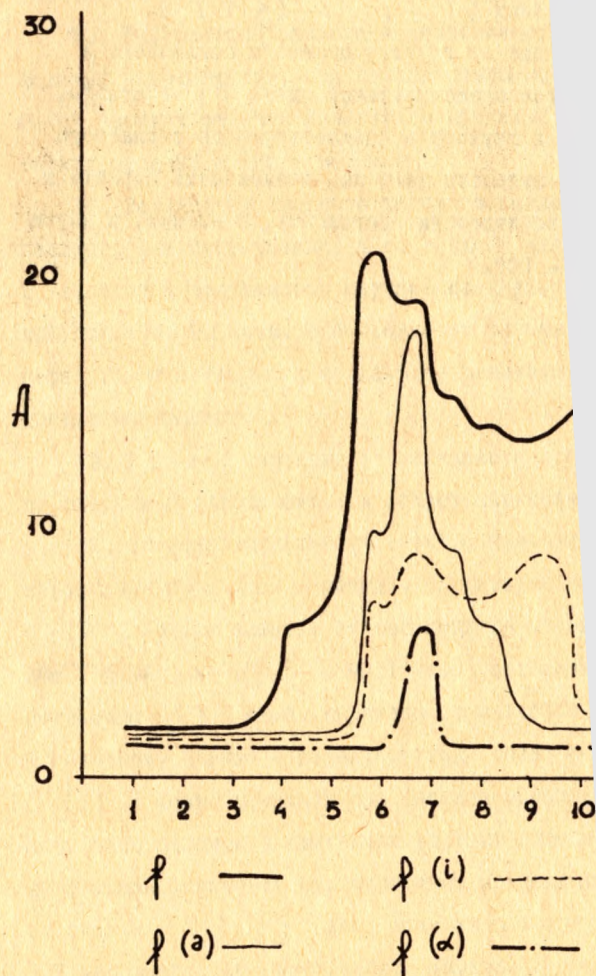


Рис. 1

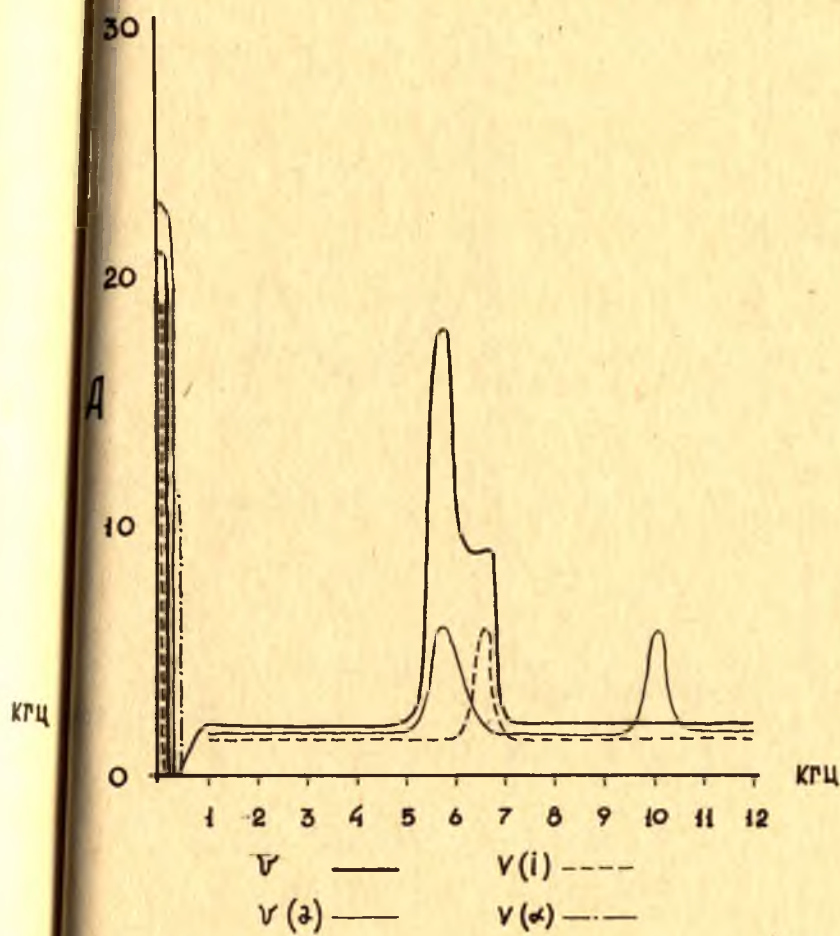


Рис. 2

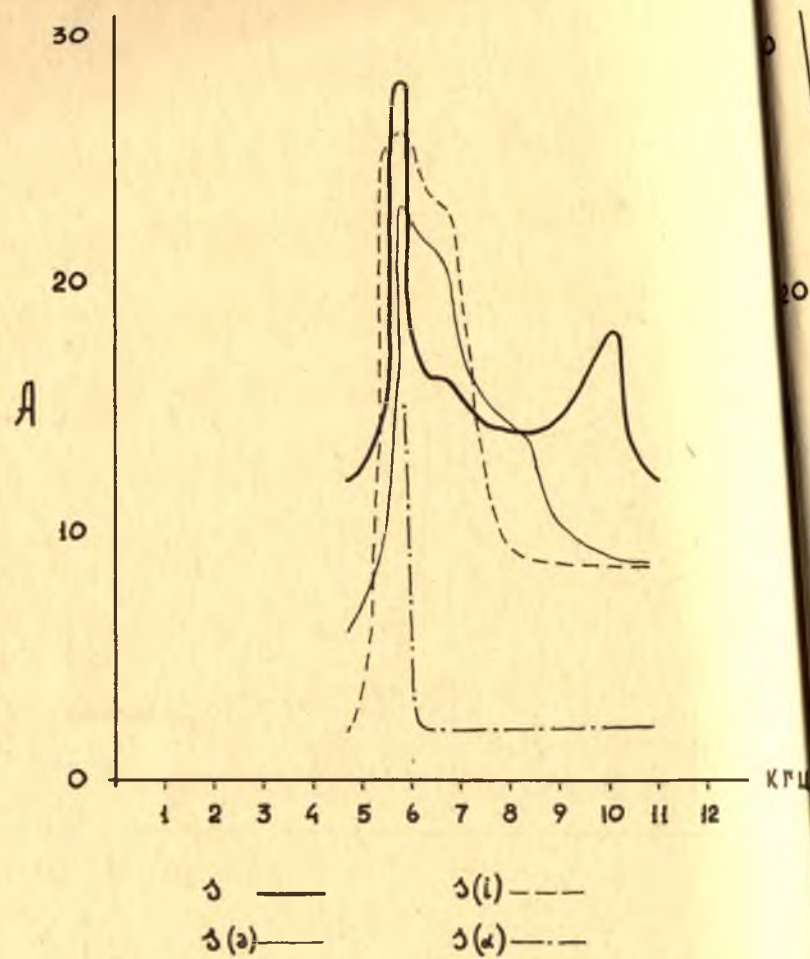


Рис. 3

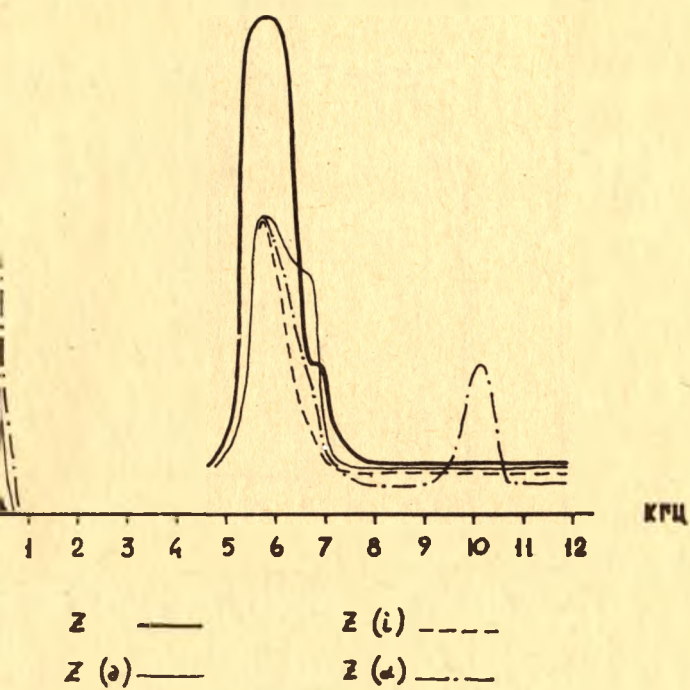
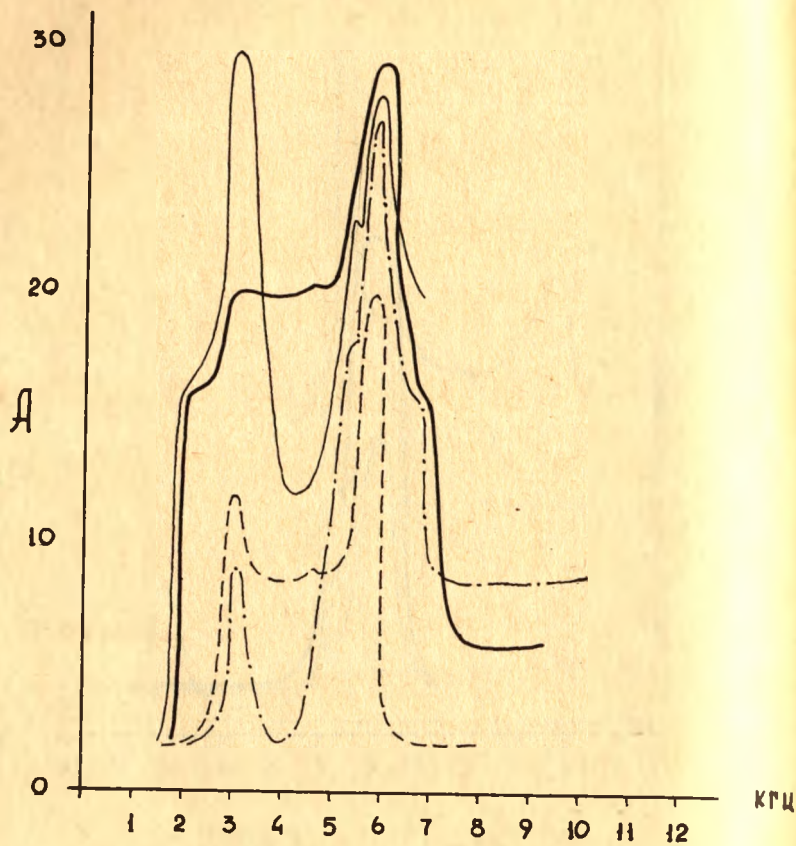
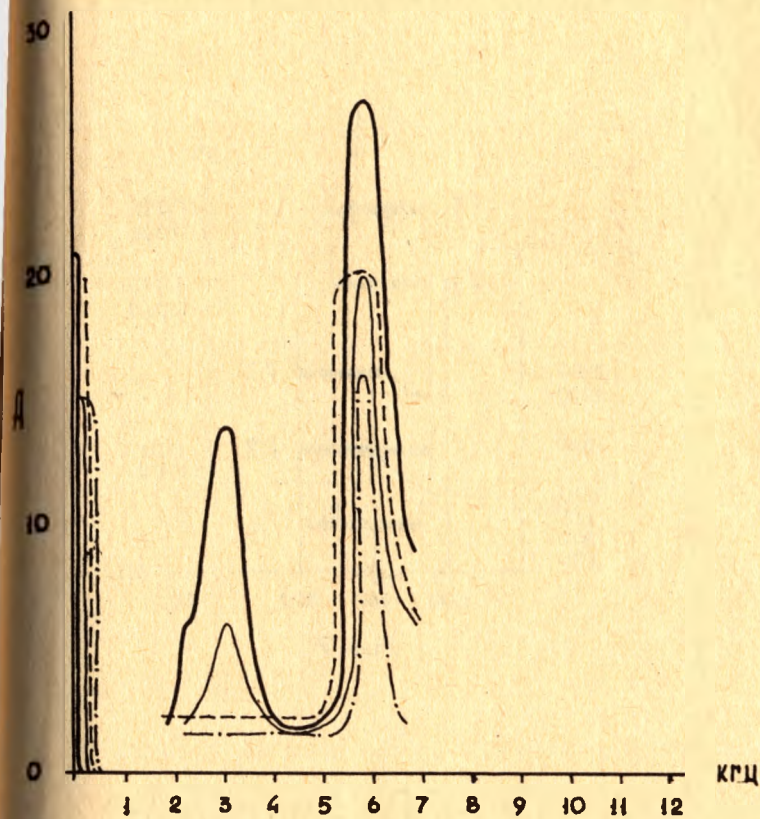


Рис. 4



$\int$ — $\int(i)$ ---
 $\int(\delta)$ — $\int(\alpha)$ —

Рис. 5



ζ — $\zeta(i)$ ----
 $\zeta(a)$ — $\zeta(i)$ — —

Рис. 6

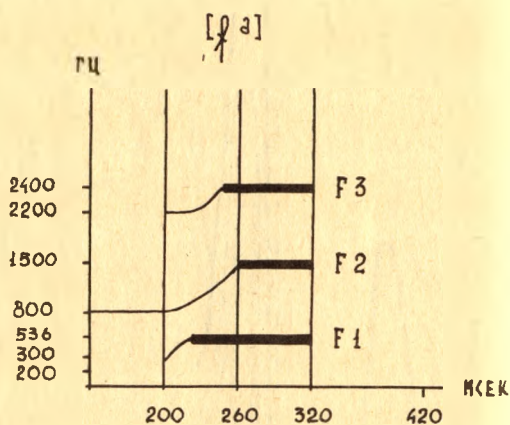


Рис. 7

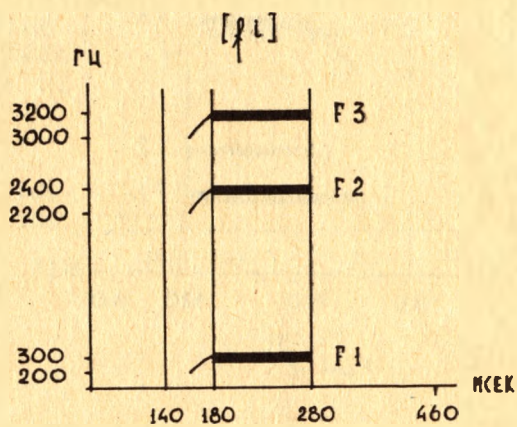
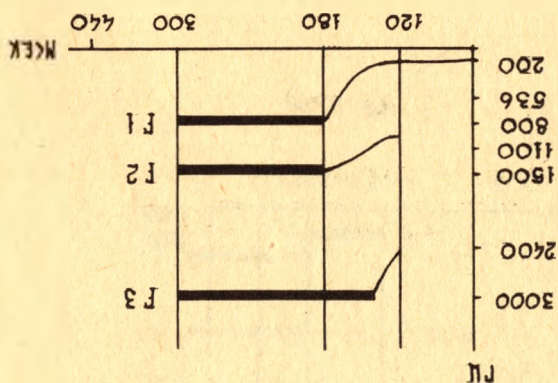


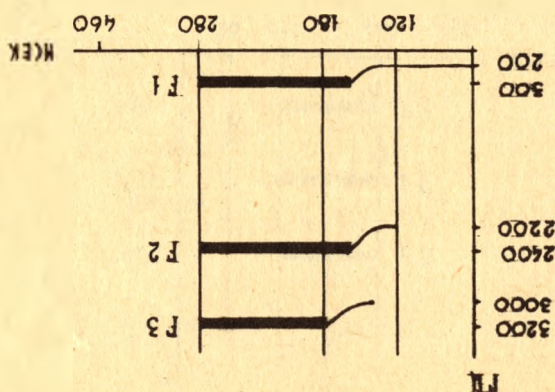
Рис. 8

Puc. 12



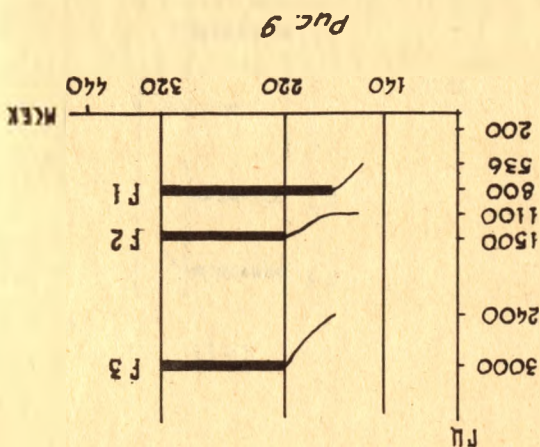
[V x]

Puc. 11

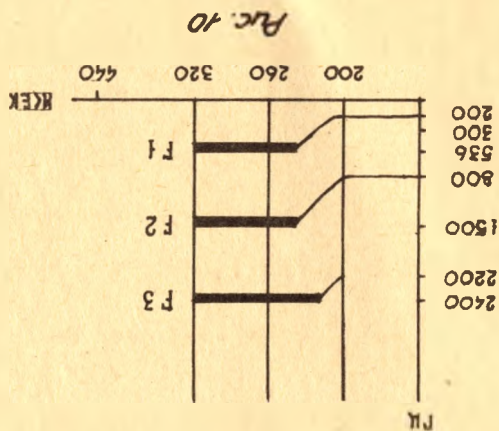


[V 1]

[82]



[V3]



[33]

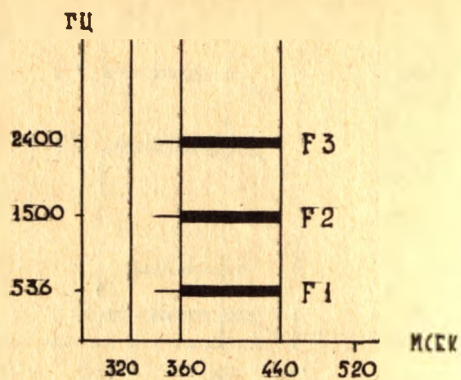


Рис. 13

[34]

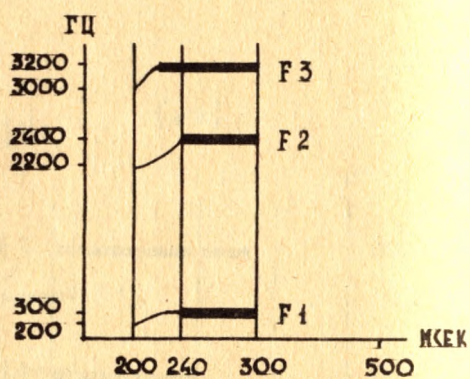
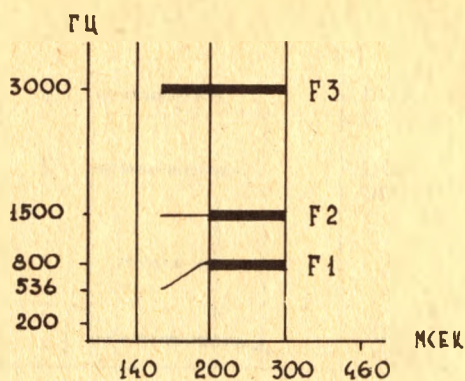
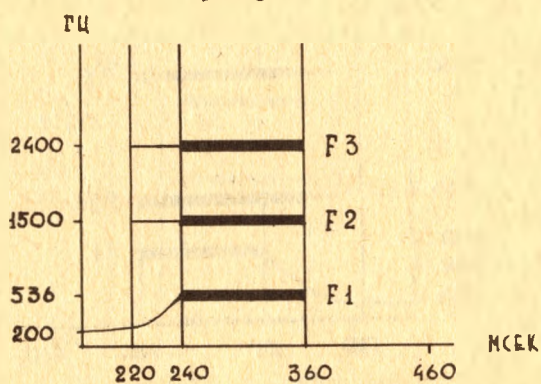


Рис. 14

[D α]

Puc. 15

[Z β]

Puc. 16

[ZL]

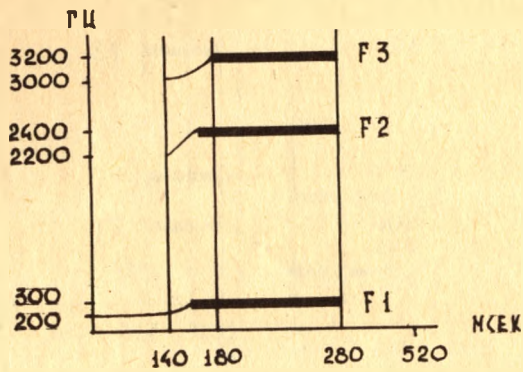


Рис. 17

[Zα]

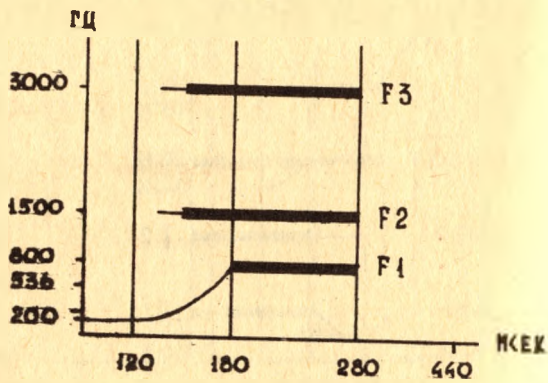


Рис. 18

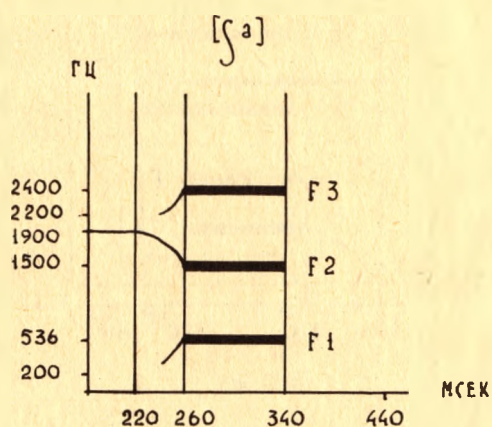


Рис. 19

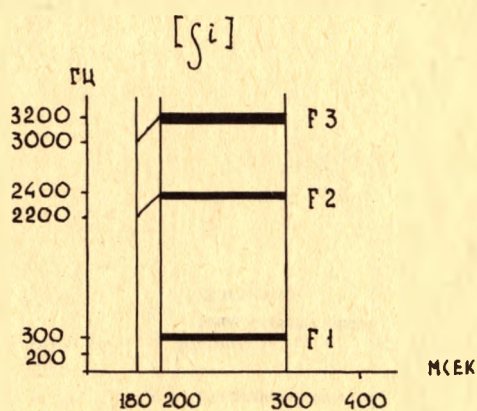


Рис. 20

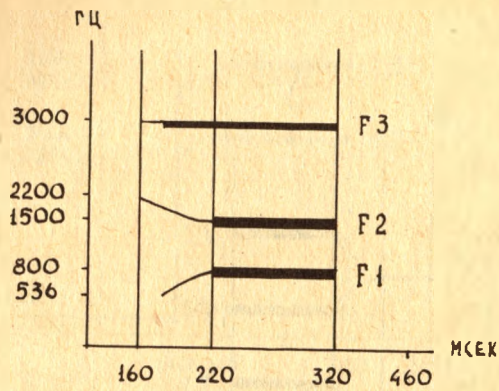
[γ^*]

Рис. 21

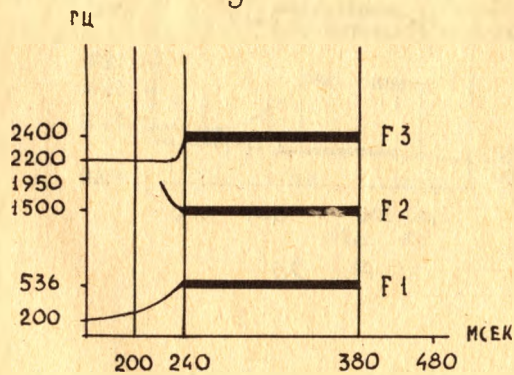
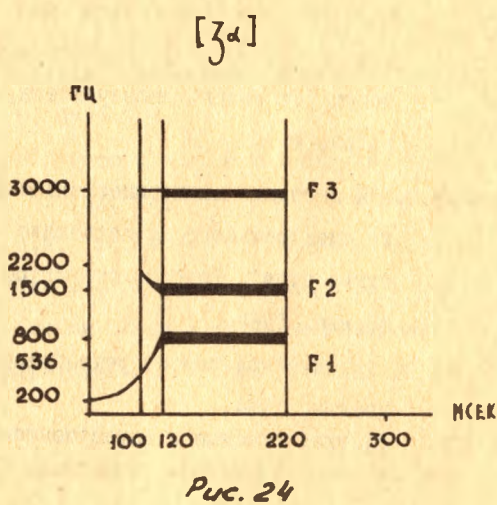
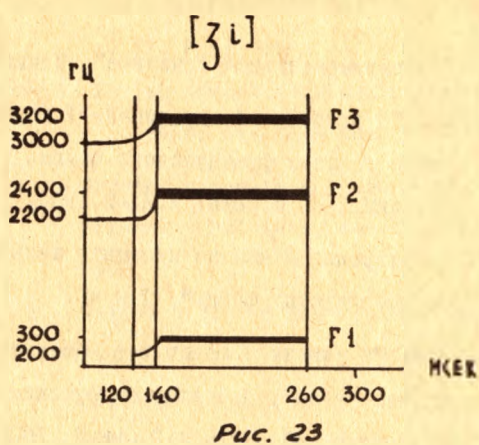
[γ^*]

Рис. 22



ЛИТЕРАТУРА

1. Артемов В.А. "Экспериментальная фонетика". Москва, 1956.
2. Барышникова К.К. "К вопросу об акустической характеристике гласных французского языка. Ученые записки I-го МГПИИЯ, т. XX-M. 1960.
3. Блохина Л.П. "Спектральный анализ немецких нелабиализованных гласных фонем". Там же.
4. Богородицкий А.К. "К учению о гласных в общерусском произношении". Известия по русскому языку и словесн. АН СССР, т. 2, кн. 2, 1929.
5. Бондарко Л.В., Зиндер Л.Р. "О некоторых дифференциальных признаках русских согласных фонем". В.Я., № 1, 1966.
6. Мегрелишвили Р.Г. "Спектральный анализ изолированных целевых согласных французского языка ($f-v$; $s-z$; ʃ-ʒ). Труды Тбилисского государственного университета, серия филологии, №
7. Соколова В.С. "К методике экспериментально-фонетического спектрального исследования вокалического стыка слов". Вопросы структуры языка. Москва, 1964.
8. Фант Г. Акустическая теория речеобразования. Москва, 1964.
9. Щерба Л.В. Русские гласные в качественном и количественном отношении. Ленинград, 1912.

11. Delattre P.C., Liberman A.M., Cooper F.S. "Acoustic Loci and Transitional Cue for Consonants, JASA, v. 27, N 4, 1955.
12. Delatre P. Le jeu transitions deiformants et la perception des consonnes. Proceedings of the Fourth International Congress of Phonetic Sciences. Helsinki, 4-9-IX-1961. The Hague, 1962.
13. Halle M., Hughes G.W. and Radely Jean-Pierre A. On Acoustic cues for stop Consonants JASA, v. 28, N 1, 1956.
14. Halle M., Hughes G.W. and Radely J.P.A. "Acoustic Properties of Stop Consonants JASA, v. 29, N 1, 1957.
15. Hughes G.W. and Halle M. "Spectral Propertie of Fricative Consonants", JASA, vol. 28, N 2. 1956.
16. Hughes G.W. and Halle M. "Spectral Properties of Stop and Fricative Consonants, JASA, vol. 27, N 5, 1955.
17. Heinz I.M. and K.N. Stevens "On the Properties of Voiceless Fricative Consonants", JASA, v.33, N 5, 1961.
18. Ingemann F. "Formants as a Voicing Cue in the Perception of /z/. JASA, vol. N 32, N 11, 1960.
19. Liberman A.M., Delattre P.C., Gersman L.I. and Cooper F.S. "Tempo of Frequency Change as a Gue for Distinguishing Classes of Speech Sounds". Journal of Experimental Psychology. vol. 52, N 2, 1956.

20. Liberman A.M., Delattre P.G. Cooper F.S. "Some Cues for the Distinction between Voiced and Voiceless Stops in Initial Position JASA, vol. 29, N11, 1957.
21. Liberman A.M., Ingeman F., Lisker L., Delattre P. and Cooper F.S. "Minimal Rules for Synthesizing Speech". JASA, vol. 31, N11, 1959.
22. Pleasant R.K. Etudes su L'e muet. Timbre, duree. intensite, Lanteur musicale., Paris, 1956.
23. Potter R.K., Kopp G.A., Green H.C. Visible Speech, New-York, 1947.
24. Rousselot P.I. (abbe). Principes de phonetique experimentale. Tome I, Paris-Leipzig, 1897-1901.
25. Sholes G.W. "Synythesis of Final (Z) without "voicing", JASA, vol. 31, N11, 1959.
26. Stevens K.N. and House A.S. "Studies of Formant Transitions Using a Vocal Tract Analog, JASA, v.28, N 4, 1956.
27. Strevens P. "Spectra of Fricative Noise in Human Speech" Language and Speech, N 3, 1960.

ANALYSE SPECTRALE DES CONSONNES FRICATIVES FRANÇAISES
EN POSITION CV AVEC LES VOYELLES (ə, i, d)

Résumé

Cette expérience a pour but d'étudier les caractéristiques acoustiques (fréquence, intensité, durée) des fricatives françaises (f-v, s-z, ʃ-ʒ) par comparaison avec les spectrogrammes des consonnes isolées correspondantes et de donner l'analyse des transitions, c'est-à-dire des changements de fréquences entre la tenue d'une consonne et la tenue de la voyelle suivante.

L'analyse des spectrogrammes est faite d'après la méthode (analyse structurale du spectre) du professeur Artémov V.A.

Le matériel à étudier a été enregistré sur le spectrographe de type du Laboratoire de phonétique expérimentale et de psychologie de la parole. Le matériel enregistré est dit par deux speakers français.

Les résultats obtenus au moyen des spectrogrammes sont les suivants:

Les spectres des fricatives sont toutes caractérisées par une haute fréquence. Chaque spectre diffère des autres par la largeur des bandes des frictions:

pour la consonne (f) = 800 cycles-12100 cycles

pour la consonne (s) = 4150 cycles - 12100cycles

pour la consonne () = 1950cycles - 6800 cycles

Sur les spectrogrammes des consonnes sonores correspondantes on remarque une zone grave à 100 - 200 cycles, et une zone aigue pareille à celle des consonnes sourdes. Mais la fréquence

de cette dernière zone est d'une amplitude plus faible que celle de la zone aigue des consonnes sourdes correspondantes. Sur le spectrogramme de la consonne () on distingue deux zones isolables: la première zone s'étend de 1950 cycles à 3582 cycles et la deuxième zone de 4150 cycles à 6800 cycles.

Dans la première zone on relève un renforcement dominant à 3000 cycles, dans la seconde, on en relève un autre à 5400 cycles ou à 57500 cycles selon le speaker.

Pour les consonnes (f) et (s) la concentration d'énergie acoustique apparaît après 5000 cycles. Le spectre de la consonne (s) est caractérisé par une énergie acoustique supérieure à celle de la consonne (f). Des exemples analysés il ressort également que les durées des tenues des consonnes ne sont pas constantes: la durée des tenues des consonnes sourdes est plus longue que des tenues des consonnes sonores correspondantes. Dans la plupart des cas pendant la transition d'une consonne sourde à une voyelle, on note sur le spectrogramme une pause dont la durée ne dépasse pas 2cs. Cette pause est suivie des changements des trois premiers formants. La direction de chaque formant est conditionnée par le locus des consonnes voisines.

НЕКОТОРЫЕ ДАННЫЕ СПЕКТРАЛЬНОГО АНАЛИЗА СОНОРНЫХ СОГЛАСНЫХ ЛИТОВСКОГО ЯЗЫКА (РЕЗУЛЬТАТЫ СТАТИСТИ- ЧЕСКОЙ ОБРАБОТКИ)

Т.Е. ПЛАКУНОВА

Исследовались сонорные согласные литовского языка в произнесении двух дикторов. Для того, чтобы проверить, насколько надежны данные анализа двух реализаций каждого варианта фонемы – по одному от диктора, – были проведены сопоставление и статистическая обработка десяти реализаций фонем [m] и [m'] в произнесении и того, и другого дикторов. Фонемы взяты из слов *kamà* и *kimi*.

Запись экспериментального материала производилась в радиостудии на магнитофоне типа МЗЗ-28А при стандартных условиях. Весь материал был начитан в три приема с сохранением по возможности однотипной интонации – близкой к назывной. Спектрограммы были получены на быстродействующем спектрографе конструкции лаборатории экспериментальной фонетики и психологии речи I-го МГПИИЯ им.М.Тореза.

Из спектрограмм слов *kama* и *kimi* путем визуального сопоставления отдельных реализаций были выделены относительно устойчивые части $[m]$ и $[m']$. В среднем их длительность составляла 5-6 кадров, т.е. 75-90 мсек. В дальнейшем подвергалась анализу только устойчивая часть этих фонем.

Уже простое сопоставление отдельных реализаций этих фонем позволяет заметить, что, во-первых, от реализации к реализации в спектре звука оказываются усиленными в основном одни и те же составляющие; колебания в местоположении максимального усиления той или иной области частот находятся в пределах двух соседних фильтров. Так, в произнесении дикт.Т. (женский голос) среднечастотный максимум (он двухвершинный) колеблется в пределах полосы частот 890-1330 гц; в произнесении дикт.Б. (мужской голос) сходные колебания наблюдаем в области средневисоких частот - 2170-2980 гц. При этом следует отметить, что от произнесения к произнесению границы усиленной полосы частот в основном не изменяются - меняется лишь местоположение максимума внутри нее.

Во-вторых, наблюдаются колебания в местоположении низкочастотного максимума - то Р4200 гц, то Р4250 гц¹⁾, что

1) Под Р4 (резонансной частотой) будем понимать среднюю частоту полосы пропускания каждого фильтра; так Р4200 гц означает, что мы имеем дело с полосой пропускания 175-225 гц, т.е. четвертым фильтром (см.таблицу I).

можно объяснить тем, что энергетический максимум располагается, по-видимому, где-то между этими частотами.

В-третьих, смещение местоположения низкочастотного максимума, как правило, вызывает подобное же смещение и в области более высоких частот: если низкочастотным максимумом является $\text{P}\chi 200$ гц, то в среднечастотной области усиления максимально увеличенной оказывается амплитуда $\text{P}\chi 1088$ гц; при смещении же низкочастотного максимума на $\text{P}\chi 250$ гц в области средних частот усиливается амплитуда $\text{P}\chi 1180$ или $\text{P}\chi 1280$ гц. Иными словами, выявляется, что при изменении абсолютной величины частоты максимумов, структура сонорного остоаается постоянной: если в первом случае отношение $\text{P}\chi 200 : \text{P}\chi 1088$ можно представить в виде отношения $1 : 5$, то подобное же соотношение сохраняется и при смещении низкочастотного максимума на $\text{P}\chi 250$, т.е. $\text{P}\chi 250 : \text{P}\chi 1180$ (или $\text{P}\chi 1280$) равно примерно $1 : 5$. Появление в последнем случае усиления то на $\text{P}\chi 1180$ гц, то на $\text{P}\chi 1280$ гц можно объяснить тем, что в действительности составляющая оказывается, по-видимому, почти на грани этих резонансных частот.

В-четвертых, нерегулярно, спорадически появляются составляющие в области высоких частот; это подтверждает мысль о том, что высокочастотный шум не является характерным для сонорных высокочастотных. Незначительная же величина амплитуд составляющих (0,5–1 мм, что в 15–10 раз меньше низкочастотных составляющих, а после коррекции²⁾ – едва 0,01 мм) говорит и

2) Поскольку спектрограф ЛЭФИП усиливает высокочастотные составляющие – на последних полосах пропускания это усиление достигает 17–18 дб, – возникает необходимость во введении коэффициента коррекции, который снимает это усиление.

о слабости этого шума. Ср. замечание Р.И.Аванесова о шумовых составляющих русских сонорных: при образовании сонорных "слабый ток воздуха образует лишь незначительный, слабо локализованный шум, свойственный самой форме резонирующей полости рта и отдельно не слышимый" (1; 138).

Визуальный анализ был дополнен статистической обработкой полученного материала. С самого начала пришлось отказаться от нормирования ввиду значительной трудоемкости этой операции при ручной обработке.³⁾

Для каждого сонорного в произнесении и того и другого дикторов были определены:

1) среднее значение амплитуды интенсивности для каждой РЧ:

$$\bar{A}_i = \frac{1}{n} \sum_{i=1}^n A_i,$$

где A_i - величина амплитуды на данной РЧ в усредненном спектре сонорного от реализации к реализации (усреднение проводится лишь по реализациям, а не по дикторам);

n - количество реализаций сонорного данным диктором;

2) среднеквадратическое отклонение средней величины амплитуды интенсивности:

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (A_i - \bar{A}_i)^2}$$

3) доверительный интервал для оценки $\bar{A}_i$, отвечающий доверительной вероятности $n = 95\%$:

$$\Delta \bar{A}_i = \bar{A}_i \pm t_{0,95;n-1} \cdot \frac{\sigma}{\sqrt{n-1}},$$

3) См. подобное замечание, сделанное Л.А.Варшавским и Л.А.Чистович (4; 186), в котором авторы предполагают возможность обработки спектрограммы без нормирования.

где вместо σ (среднеквадратического отклонения) введена более точная величина – стандартная ошибка, которая приближенно равна $\frac{\sigma}{\sqrt{n-1}}$. При расчетах исходим из того, что распределение средних величин $\bar{A}_i$ каждой РЧ подлeжит t – распределению Стьюдента.

Уже простое сопоставление отдельных реализаций того или другого сонорного показываеt, что величина амплитуд интенсивности очень вариативна. Это подтверждается и статистическим анализом. Например, амплитуда РЧ250 гц в спектре $[m]$ (*kama*, дикт.Б) при $\bar{A}_i = 7$ мм изменяется в диапазоне от 4,8 мм до 10 мм. Средняя величина амплитуды колеблется в пределах 6,01 – 8,15 мм при среднеквадратическом отклонении $\sigma = 1,43$. Еще больше эти значения у амплитуды РЧ200, которая в четырех случаях из десяти в произнесении этого диктора является максимумом в спектре $[m]$. Средняя величина амплитуды равна 6,35 мм, диапазон изменения – от 4 мм до 10 мм при $\bar{A}_i = 4,75 - 7,95$ мм и $\sigma = 2,12$. Для наглядности приводим графики рассеивания средней величины амплитуд интенсивности в спектрах $[m]$ и $[m']$ (см.рис.1-4). Нижняя и верхняя границы возможных значений средней величины каждой амплитуды обозначены кружочками. Надежность определения этих границ для высокочастотных составляющих резко снижается из-за того, что последние появляются в спектре сонорного крайне нерегулярно. Приводимый на рис.1-4 разброс средних для высокочастотных составляющих может рассматриваться поэтому лишь как иллюстрация, но вряд ли отражает действительный диапазон колебаний средних величин этих амплитуд.

Приводимые графики показывают, что амплитуды отдельных резонансных частот изменяются в разных пределах. Однако сопоставление изменчивости средних величин амплитуд на основании лишь их диапазонов колебаний (или доверительных интервалов) представляется затруднительным. Также и величина среднеквадратического отклонения вряд ли может служить надежным параметром при оценке колебаний амплитуд интенсивности. Ср. 2,12 ($\bar{A}_i$ Р4250 гц) и 0,13 ($\bar{A}_i$ Р45275 гц) ([m] дикт.Т.) без соотнесения со средней величиной амплитуды трудно сказать, велики ли эти значения. Но и подобное сопоставление не позволяет дать оценку этим величинам, поскольку они разновелики. В таком случае удобно ввести еще один критерий оценки — так называемый коэффициент вариативности. Подобный анализ мы находим в статье М.Ф.Деркача "Роль стабильности спектра в объективном различении речевых звуков" (5).

Коэффициент вариативности представляет собой отношение среднеквадратического отклонения к средней интенсивности, в нашем случае — средней величине амплитуды интенсивности, т.е.

$$C_v = \frac{\sigma}{\bar{A}_i}$$

Тогда для $C_v = 2,12$ коэффициент вариативности равен 0,33, для $C_v = 0,13$ величина этого коэффициента равна 0,81, т.е., несмотря на то, что абсолютная величина C_v превышает таковую C_{38} более, чем в 16 раз, в действительности средняя величина амплитуды этой РЧ оказывается почти в 2,5 раза стабильнее средней величины амплитуды другой РЧ. Иными словами, коэффициент вариативности позволяет очень наглядно представить изменчивость амплитуды интенсивности, а тем самым

отыскать и наиболее стабильную и наиболее изменяющуюся части спектра.

Значения коэффициента вариативности для резонансных частот от 80 гц до 4400 гц в спектрах $[m]$ и $[m']$ представлены в таблице 2. Более высокие частоты не берутся из-за крайней нерегулярности появления составляющих на этих частотах.

Таблица 2

Коэффициент вариативности амплитудных значений
для резонансных частот в спектрах $[m]$ и $[m']$

№ фильтра	Средняя частота, гц	Значение коэффициента вариативности			
		$[m]$		$[m']$	
		д.Б.(м)	д.Т.(ж)	д.Б.(м)	д.Т.(ж)
I	80	0,13	0,21	0,16	0,30
2	110	0,13	0,32	0,18	0,42
3	150	0,15	0,32	0,16	0,35
4	200	0,33	0,37	0,33	0,40
5	250	0,20	0,72	0,29	0,22
6	300	0,14	0,29	0,15	0,25
7	350	0,12	0,30	0,10	0,33
8	400	0,10	0,64	0,22	0,52
9	450	0,06	0,60	0,22	0,67
10	500	0,26	0,37	0,36	0,51
11	550	0,33	0,25	0,39	0,45
12	600	0,26	0,70	0,38	0,55
13	655	0,34	0,55	1,00	0,67

№ фильтра	Средняя частота, Гц	Значение коэффициента вариативности			
		[m]		[m']	
		д.Б.(м)	д.Т.(ж)	д.Б.(м)	д.Т.(ж)
14	715	0,35	0,44	0,72	0,34
15	780	0,27	0,49	0,55	0,40
16	852	0,28	0,55	0,57	0,69
17	927	0,24	0,50	0,73	0,47
18	1005	0,16	0,65	1,00	0,50
19	1088	0,19	0,51	0,87	0,77
20	1180	0,20	0,54	0,62	0,47
21	1280	0,22	0,61	0,57	0,43
22	1390	0,29	0,50	0,87	0,46
23	1510	0,25	0,45	0,56	0,54
24	1635	0,24	0,38	0,53	0,40
25	1775	0,28	0,38	0,53	0,42
26	1925	0,38	0,52	0,72	0,53
27	2085	0,34	0,41	0,51	0,49
28	2255	0,36	0,53	0,51	0,53
29	2440	0,40	0,46	0,59	0,59
30	2645	0,48	0,50	0,58	0,53
31	2865	0,42	0,51	0,58	0,57
32	3100	0,69	0,46	0,49	0,49
33	3365	0,51	0,54	0,55	0,47
34	3675	0,45	0,61	0,55	0,48
35	4020	0,60	0,64	0,66	0,53
36	4400	0,67	0,57	0,70	0,61

Приведенные данные позволяют заметить, что небольшими величинами характеризуются низкочастотные составляющие, связываемые обычно с частотой основного тона диктора⁴⁾, - это РЧ80, II0, I50 гц. Коэффициент вариативности для амплитуд этих составляющих - 0,13 - 0,15([m]) и 0,16 - 0,18([m'])

у диктора - мужчины. Еще меньше этот коэффициент у составляющих в полосе частот 275-475 гц, что особенно отчетливо видно в спектре [m] дикт.Б. В спектрах [m] и [m'] дикт.Т. относительно небольшим коэффициентом вариативности характеризуются отдельные составляющие в полосе частот 275-375 гц и 685-815 гц. По величине амплитуд эти составляющие лишь незначительно отличаются от соседних. Однако, если вспомнить, что область частот 300-600 гц обычно связывают с носовым резонансом, можно предположить, что небольшая вариативность амплитуд интенсивности составляющих в этой области является чертой, характеризующей фонемы [m] и [m'] как носовые и, по-видимому, отличающей их от других, не носовых, фонем. В литературе уже отмечалось, что, возможно, фонемы отличаются друг от друга не всеми параметрами, а лишь каким-то одним (3, 5), о чем говорил в своем докладе и проф. В.А.Артемов⁵⁾. Для носовых сонорных - это относительная стабильность указанных выше составляющих, в чем и проявляется назальность. Эти составляющие непосредственно смыкаются

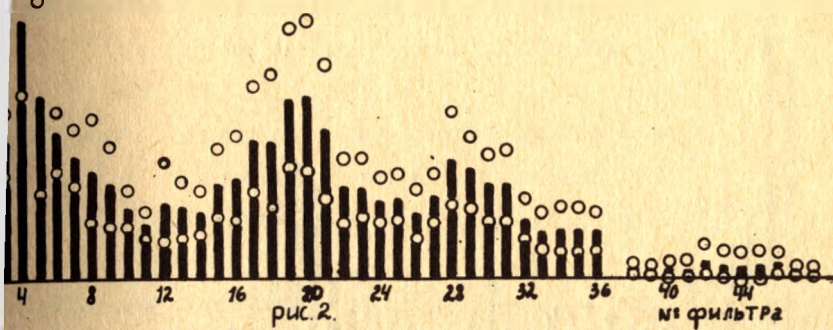
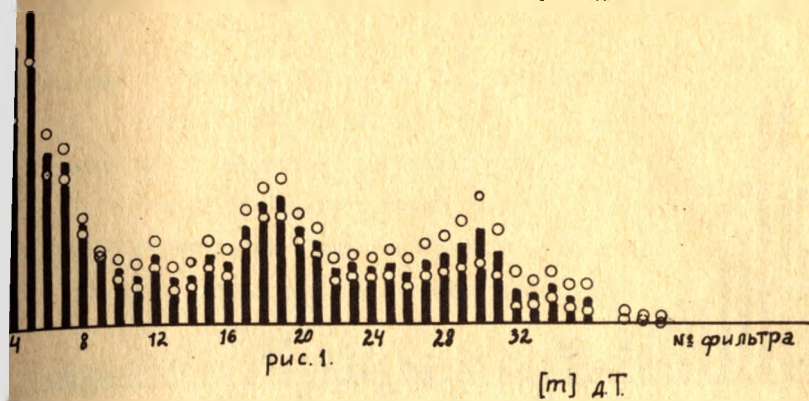
4) При анализе спектров фонем в произнесении разных дикторов коэффициент вариативности составляющих этой области, естественно, должен быть больше, т.к. диапазон колебаний частоты основного тона от диктора к диктору будет шире, чем подобный диапазон в речи у одного диктора.

5) См. настоящий сборник.

с составляющими РЧ200 и 250 гц, коэффициент вариативности которых колеблется в пределах 0,20 – 0,40 и которые являются максимально усиленными составляющими в спектре носовых сонорных. Эта область частот – от 200 до 450 гц в нашем случае и будет областью, которой отличается спектр носовых сонорных от спектров других фонем. Проверка различаемости внутри сонорных на основании показателей о наименее вариативных составляющих (с применением t – распределения Стьюдента) подтверждает, что наиболее стандартная область спектра характеризует носовые как группу – (Процент распознаваемости в среднем 48–52%, т.е. в звуке в равной мере может быть распознана как фонема [m], так и фонема [n]).

Таким образом, сопоставление десяти реализаций одного и того же сонорного в произнесении как того, так и другого диктора дает нам хорошее совпадение спектров по частотным максимумам при значительной вариативности их амплитуд. Следует также отметить, что незначительное смещение максимума внутри определенной полосы частот происходит при сохранении неизменными нижней и верхней границ этой полосы частот. Иными словами, важным оказывается усиление не одной резонансной частоты, а полосы частот. При этом характеризующей оказывается, по-видимому, лишь одна полоса частот в спектре фонемы.

[m] 45



[m'] д.Б.

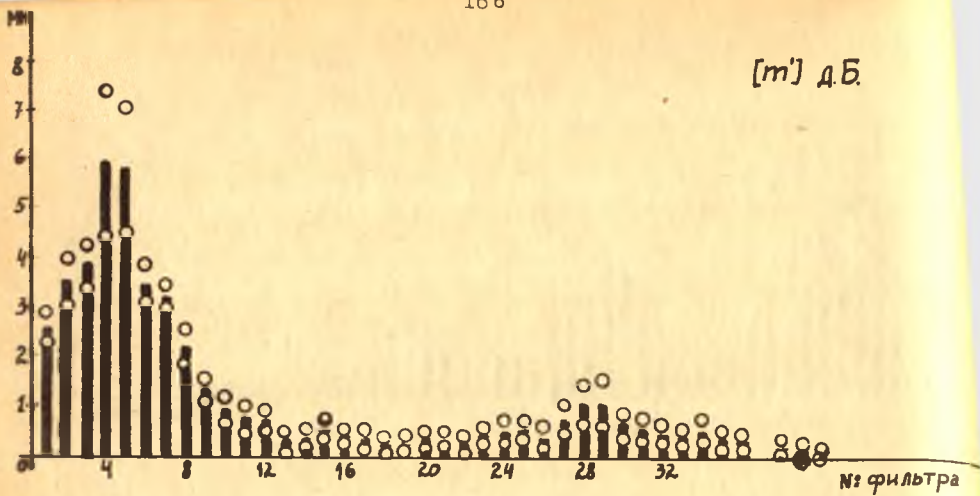


рис. 3

[m'] д.Т.

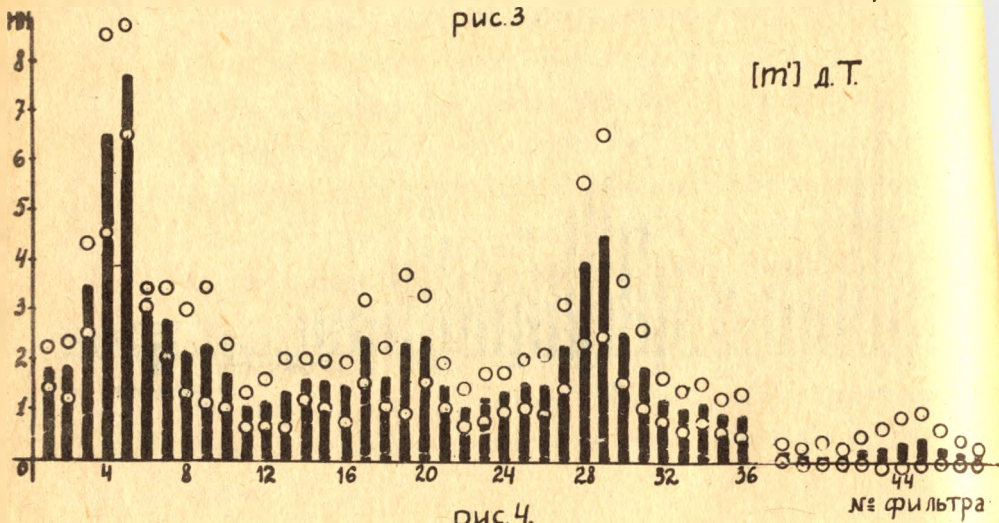


рис. 4.

№ фильтра

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Р.И.Аванесов. Фонетика современного русского литературного языка, изд.МГУ, 1956г.
2. В.А.Артемов. Метод структурного анализа речевых спектров, 1959-1962гг. (машинопись).
3. В.А.Артемов. Об интонеме и интонационном инварианте. Сб. "Интонация и звуковой состав", изд. МГУ, 1965г.
4. Л.А.Варшавский, Л.А.Чистович. Средние спектры русских гласных фонем "Проблемы физиологической акустики", 1959г., т.IV.
5. М.Ф.Деркач. Роль стабильности спектра в объективном различении речевых звуков. Сб. "Бионика", изд. "Наука", М., 1965г.
6. Н.А.Любимова. Спектральные характеристики русских сонантов. Вестник Ленинградского ун - та, № 2, серия истории, языка и литературы, вып.I, 1965г.
7. В.И.Романовский. Применение статистики в опытном деле. Гостехиздат, 1947г.
8. Н.В.Смирнов и др. Курс теории вероятностей и математической статистики, изд. "Наука", 1965г.
9. Г.М.Фант. Акустическая теория речеобразования, изд. "Наука", 1964г.
10. Л.А.Чистович. О методах обработки спектрограмм гласных звуков. "Биофизика", 1957г., т.2, вып.4.

T.E. Plakunova

Vilnius State University

CERTAIN DATA RELATING TO SPECTRUM ANALYSIS OF
SONOROUS CONSONANTS OF LITHUANIAN LANGUAGE
(Results of statistical Interpretation)

Summary

The subject matter of the paper is devoted to the results of visual and statistic analysis of nasal sonorous consonants of the Lithuanian language /m/ and /m'/ in a tenfold repetition. The statistic analysis was carried out with a view of checking to what extent the data of two realizations of each varian of the phoneme are applicable to linguistic research (by one realization from the announcer). Both visual and statistical analysis showed a good overlapping of the spectra by frequency maxiums with amplitudes varying to a considerable measure. A negligent shift of maximums, as a rule, does not go beyond one or another band of frequencies the lower and the upper limits of which remain unalterable. In addition, the statistical analysis makes it possible to detect the most stable portion of the spectrum which, in our case, characterizes the nasal sonorous consonants as a group. Thus, the author arrived at the conclusion that phonemes differ not by their whole spectrum but only by a certain frequency band in the spectrum, this band comprising 200-450 cps for nasal sonorous consonants.

НЕКОТОРЫЕ ДАННЫЕ СПЕКТРАЛЬНОГО АНАЛИЗА ДОЛГИХ ГЛАСНЫХ ЛИТОВСКОГО ЯЗЫКА В УДАРНОМ, ПРЕДУДАРНОМ И ЗАУДАРНОМ СЛОГАХ

Я. УРБЕЛЕНЕ

Безударному вокализму литовского языка до сих пор уделялось мало внимания. Экспериментально безударные гласные исследованы только в количественном отношении (1,2), и на основе этих исследований говорится о количественной редукции в литовском языке. Наличие же качественной редукции безударных гласных на основе слухового восприятия отрицается, т.е. указывается, что безударные гласные сохраняют качество ударных гласных (2-63).

Для получения объективных данных мы подвергли безударные гласные спектральному анализу, на основе которого мы смогли установить, с одной стороны, те общие акустические характеристики ударных и соответствующих безударных гласных, которые не нарушают их адекватного восприятия, и, с другой стороны, те индивидуальные свойства, которые характеризуют лишь данный оттенок в отличие от других.

Для этой цели мы записали на быстродействующем 48-канальном спектрографе конструкции лаборатории экспериментальной фонетики и психологии речи I-го МГПИИЛ с полосой частот 60 гц - 12000 гц слова с интересующими нас долгими гласными [a:] , [æ:] , [e:] , [i:] , [и:] и [o:] в ударном, предударном и заударном слогах в окружении согласного "р" в произношении двух дикторов (мужчины и женщины). Всего было получено 192 спектрограммы.

Данные спектрального анализа долгих ударных и безударных гласных, проведенного по методу, разработанному в ЛЭФИПР (4), дали нам возможность установить следующее:

Гласные [a:] и [æ:] сохраняют признак компактности во всех интересующих нас позициях - во всех спектрах данных гласных произносительная энергия сконцентрирована в центральной области спектра, в области 852-1180 гц. Наиболее сближенными на частотной шкале оказываются максимумы в 927 гц и в 1180 гц в ударном [a:] и в 852 гц и 1088 гц в предударном и заударном [a:] ; в гласном [æ:] близко друг к другу подходят максимумы в 852 гц и 1180 гц. Это повышает в известной мере интенсивность данного участка.

Однако на структуре спектров гласных [a:] и [æ:] сказывается положение гласного по отношению к ударению. Так, в предударном и заударном слоге по сравнению с ударным намечаются определенные изменения в расположении максимумов.

В предударном слоге максимум в 500 гц в гласных [a:] и [æ:] понижается до 450 гц, а в заударном слоге - до 350 гц ; максимум в 927 гц в спектре гласного [a:] в предударном и заударном слогах понижается до 852 гц ; максимум в 852 гц в гласном [æ:] понижается - только в заударном слоге - до 780 гц (см. таблицу I расположения максимумов).

Если, как говорит Г.Фант, сдвиг первой форманты в сторону возрастания выступает как показатель увеличения компактности (3-70), то сдвиги максимумов в 500 гц, 852 гц и 927 гц в сторону более низких частот говорят о том, что гласные [a:] и [æ:] безударные становятся менее ком-

пактными, чем ударные. Этот вывод подтверждается и уменьшением уровня амплитуды интенсивности составляющих (см. таблицу 2. Интенсивность максимумов).

На частотной шкале как в спектрах безударных [а:], так и в спектрах безударных [æ:] максимумы первой области отдалаются от максимумов второй области. Это значит, что в безударных слогах данные гласные можно характеризовать как несколько более высокотоновые.

Несмотря на вышеуказанные сдвиги в расположении максимумов, соотношения максимумов¹⁾ остаются относительно одинаковыми (см. табл. 3). Для гласного [а:], например, в произношении диктора Т. характерно следующее соотношение максимумов:

в ударном слоге: $I(200) : 2,5 : 4,6 : 5,9 : I_{2,2} : I_{6,8}$

в предударном слоге: $I(200) : 2,2 : 4,3 : 5,4 : I_{2,2} : I_{5,5}$

в заударном слоге: $I(200) : 1,8 : 4,3 : 5,9 : I_{0,4} : I_{5,5}$

Исследование спектров гласных [i:] и [e] показало, что во всех позициях - в ударном, предударном и заударном слогах - данные гласные сохраняют признак диффузности.

Местонахождение максимумов в каждой из позиций несколько иное. В гласном [i:] , например, максимум, расположенный на резонансной частоте в 500 гц, в предударном слоге понижается до 450 гц, а в заударном слоге - до 350 гц, в то время как максимум в 2255 гц не изменяется.

1) Для установления соотношения максимумов мы используем метод, предложенный В.А.Артемовым (4).

Аналогичная картина наблюдается и в спектрах гласного [e:]. Максимум на частоте в 927 гц понижается до 852 гц в предударном слоге, а в заударном слоге по отношению к предударному он не изменяется, в то время как максимум в 450 гц понижается до 400 гц в предударном и до 350 гц в заударном слогах.

Такие сдвиги в расположении максимумов свидетельствуют о том, что в безударных слогах данные гласные становятся более диффузными и более высокотоновыми.

Хотя, как видно, в расположении максимумов имеются определенные изменения, соотношение максимумов относительно одинаково (см. табл. 3). В гласном [i:], например, оно следующее:

в ударном слоге: I(250) : 2 : 9 : I6, I

в предударном слоге: I(200) : 2,2 : I0,4 : I8,3

в заударном слоге: I(200) : I,8 : II,2 : I8,3.

По сравнению с ударными [i:] и [e:] соответствующим предударным и в еще большей мере заударным гласным свойствен меньший уровень произносительной энергии, о чем говорят данные анализа общей интенсивности максимумов (см. табл. 2).

В отношении акустической характеристики гласных [u:] и [o:] можно заметить следующее.

Для спектральной структуры гласных [и:] и [о:] во всех позициях характерно преобладание низкого края спектра над высоким – произносительная энергия звуков сконцентрирована в низко- и среднечастотной областях спектра. Это говорит о сохранении признака диффузности в спектральной харак-

теристике данных гласных ударного, предударного и заударного слога.

Однако в спектрах гласных [i:] и [e:] предударных и заударных происходят определенные изменения. Это, во-первых, изменения в расположении максимумов. В гласном [u:] максимумы в 250 гц, 500 гц и 3675 гц в предударном, и особенно в заударном, слогах расположены в более низких частотах, чем в ударном слоге. В спектрах гласного [o:] предударного слога все максимумы понижаются; в заударном слоге по сравнению с предударным местоположение максимумов в 852 гц и в 3365 гц не изменяется, а максимум в 450 гц понижается до 350 гц. Такие изменения в частотном уровне максимумов характеризуют безударные [i:] и [e:] как более диффузные. Об этом же говорит и меньший уровень произносительной энергии, особенно в заударном слоге.

Наряду с этим можно наметить и некоторые сдвиги в признаке тональности данных гласных. В безударной позиции тональность данных гласных повышается.

Соотношение максимумов в гласном [и:] , например, относительно постоянно во всех позициях:

в ударном слоге: $I(250) : 2 : 3, I : 9,7 : 14,7$

в предударном слоге: $I(200) : 2,2 : 4,6$

в заударном слоге: $I(200) : 1,8 : 4,3$.

Для гласных [и:] , [о:] существенным является признак бемольности. Известно, что бемольность проявляется в сдвиге вниз нескольких или даже всех формант (5-195).

По нашим данным, в предударном слоге в спектре [и:] повышается максимум в 927 гц (вместо 780 гц в ударном слоге); в спектре предударного гласного [о:] в одной из реализаций не изменяется максимум в 927 гц, а в заударном слоге максимум в 2440 гц повышается (вместо 2255 гц в предударном слоге).

Итак, наше исследование показало, что в безударной позиции (в предударном и заударном слогах) противопоставления литовских долгих гласных [а:], [æ:], [е:], [i:], [и:] и [о:] по признакам компактный - диффузный, низкотоновый - высокотоновый, бемольный - простой не снимаются. Иначе говоря, ударные и безударные гласные имеют относительно одну и ту же структуру, т.е. безударные гласные не являются новым качеством по отношению к ударным. Но определенные изменения в акустической характеристике безударных гласных свидетельствуют о наличии качественной редукции в литовском языке. Причем заударные гласные несколько больше редуцированы, чем предударные.

РАСПОЛОЖЕНИЕ МАКСИМУМОВ В СПЕКТРАХ ГЛАСНЫХ
(В ПРОИЗНОШЕНИИ ДИКТОРА Т.)

175

гл. [а:]

ударный слог	200	500	927	1180	2440	3365
предуд. слог	200	450	852	1088	2440	3100
	200	450	852	1088	2440	2865
зауд. слог	200	350	852	1180	2085	3100
	200	350	852	-	2085	3365

гл. [æ:]

ударный слог	250	500	852	1180	2085	3675
предуд. слог	200	450	852	1088	2085	3100
	200	400	852	1180	1925	3100
зауд. слог	200	350	852	-	1775	2865
	200	300	780	-	1775	- 3675

гл. [е:]

ударный слог	200	450	927	2440	3365	
предуд. слог	200	450	927	2255	3675	- 4020
	200	400	852	2085		
зауд. слог	200	350	852	2255	3675	
	200	350	852	2255	3675	

гл. [i:]					
ударный слог	250	500	2255	4020	
предуд. слог	200 200	450 450	2085 2255	3675 4020	
заудар. слог	200 200	350 350	2255 2255	3675 3675	
гл. [и:]					
ударный слог	250	500	780	2440	3675
предуд. слог	200 200	450 450	927 927	3365	
заудар. слог	200 200	350 350	852 852		
гл. [о:]					
ударный слог	250	500	927	2440	3675
предудар. слог	200 200	450 450	852 927	2255	3365 3365
заудар. слог	200 200	350 350	852 852	2085 2440	3365 3365

АМПИТУДА ИНТЕНСИВНОСТИ ГЛАССИМОВ / В ПРЕДЕЛАХ ВЫПРЕКИ/
В ММ (ДИКТОР Т.)

Таблица 2

Глас- ный	Позиция	fmax1	fmax2	fmax3	fmax4	fmax5	fmax6	Суммарная интенсивн. глассимумов
[a:]	ударн.	7,66	3,94	9,64	8,39	3,34	5,07	38,04
	предуд.	4,95	4,-	10,5	7,96	1,37	1,39	30,17
	зауд.	4,39	3,33	10,11	7,-	0,87	0,75	26,45
[æ:]	ударн.	1,6	1,56	4,4	2,57	0,6	0,26	8,42
	предуд.	0,93	2,46	8,32	2,57	1,21	1,-	16,49
	зауд.	8,66	5,42	10,79	3,08	5,67	1,72	35,34
[e:]	ударн.	3,45	4,0	2,01	1,05	3,95	1,88	16,34
	предуд.	3,74	3,48	5,92	3,-	5,-	2,08	23,22
	зауд.	1,77	1,56	3,1	2,66	1,2	0,31	10,73
[i:]	ударн.	1,40	1,43	2,66	2,-	0,87	8,36	20,30
	предуд.	4,82	7,12	0,61	4,67	3,08		
	зауд.	4,21	6,98	0,03	2,81	1,25	15,28	14,05
[u:]	ударн.	2,6	5,-	0,3	0,83	0,24	8,97	10,05
	предуд.	2,1	5,-	0,25	1,9	0,8		
	зауд.	18,75	1,45	2,26	2,54		25,-	
[o:]	ударн.	5,61	1,51	1,39	0,79		9,30	5,57
	предуд.	3,74	0,73	0,73	0,37			
	зауд.	2,8	5,3	0,63	0,43		9,16	12,36
[y:]	ударн.	4,7	6,-	1,2	0,46		31,07	
	предуд.	16,52	3,62	10,51	0,24	0,18		
	зауд.	4,67	2,72	0,51	0,09		7,90	9,24
[e:]	ударн.	5,61	3,03	0,51	0,09		8,91	11,70
	предуд.	2,1	6,-	0,81				
	зауд.	3,4	7,1	1,2			27,24	
[i:]	ударн.	7,48	6,52	11,28	1,60	0,36		
	предуд.	5,61	6,82	7,69	0,13	0,09	20,21	22,70
	зауд.	6,54	8,33	7,61	0,13	0,09	14,20	14,46

Таблица 3

СООТНОШЕНИЯ МАКСИМУМОВ В СПЕКТРАХ ДОЛГИХ ГЛАСНЫХ
ЛИТОВСКОГО ЯЗЫКА (ДИКТОР Т.)

[a:]	уд.	I(200) : 2,5 : 4,6 : 5,9 : 12,2 : 16,8
	предуд.	I(200) : 2,2 : 4,3 : 5,4 : 12,2 : 15,5 I(200) : 2,2 : 4,3 : 5,4 : 12,2 : 14,3
	зауд.	I(200) : 1,8 : 4,3 : 5,9 : 10,4 : 15,5 I(200) : 1,8 : 4,3 : 5,9 : 10,4 : 16,8
[æ:]	уд.	I(250) : 2 : 3,4 : 4,7 : 8,2 : 14,7
	предуд.	I(200) : 2,2 : 4,3 : 5,4 : 9,6 : 15,5 I(200) : 2,3 : 4,3 : 5,9 : 10,4 : 15,5
	зауд.	I(200) : 1,8 : 4,3 : - : 8,8 : 18,3 I(200) : 1,5 : 3,9 : - : 8,8 : 14,3
[e]	уд.	I(200) : 2,2 : 4,6 : 12,2 : 16,8
	предуд.	I(200) : 2,- : 4,3 : 10,4 : 20,01 I(200) : 2,2 : 4,6 : 11,2 : 18,3
	зауд.	I(200) : 1,8 : 4,3 : 11,2 : 18,3 I(200) : 1,8 : 4,3 : 11,2 : 18,3
[i:]	уд.	I(250) : 2 : 9 : 16,1
	предуд.	I(200) : 2,2 : 10,4 : 18,3 I(200) : 2,2 : 11,2 : 20,01
	зауд.	I(200) : 1,8 : 11,2 : 18,3 I(200) : 1,8 : 11,2 : 18,3
[u:]	уд.	I(250) : 2 : 3,1 : 9,7 : 14,7
	предуд.	I(200) : 2,2 : 4,6 : - : 16,8 I(200) : 2,2 : 4,6 : - : -
	зауд.	I(200) : 1,8 : 4,3 : - : - I(200) : 1,8 : 4,3 : - : -
[o:]	уд.	I(250) : 2 : 3,7 : 9,7 : 14,7
	предуд.	I(200) : 2,2 : 4,3 : 11,2 : 16,8 I(200) : 2,2 : 4,6 : - : 16,8
	зауд.	I(200) : 1,8 : 4,3 : 10,4 : 16,8 I(200) : 1,8 : 4,3 : 12,2 : 16,8

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. R.Ekblom Quantitat und Jntonation im zentralen Hochlitauischen. Uppsala, 1925.
2. Lietuviu kalbos gramatika. I tomas. Fonetika ir morfologya. Lietuvos TSKS Mokslo, Akademya Lietuviu kablos ir literaturos Jntitutas 1965.
3. Г.Фант. Акустическая теория речеобразования. Перевод с английского Л.А.Варшавского и В.И.Медведева. М., 1964.
4. В.А.Артемов.Метод структурного анализа речевых спектров. М., 1958-1962 (машинопись).
5. Р.Якобсон, Г.М.Фант и М.Халле. Введение в анализ речи. Различительные признаки и их корреляты. Сб. Новое в лингвистике, вып. II, 1962.
6. J. Dambrauskaite Lietuviu kalbos fonemines balsiu sistemas nustatymas. VVPJ Mokslo darbai, 3t., 1957.
7. К.К.Барышникова. Спектральный анализ неударных гласных фонем французского языка. Интонация и звуковой состав. Материалы коллоквиума по экспериментальной фонетике и психологии речи. Изд. Московского университета, 1965.
8. Б.И.Свечевичус. Простые гласные (монофтонги) современно-го литовского литературного языка. Канд.дисс. Вильнюс, 1964.

Urbelene Y.C.

Vilnius State Pedagogical Institute

SOME DATA RELATED TO SPECTRUM ANALYSIS OF LONG
VOWELS OF LITHUANIAN LANGUAGE IN ACCENTED, PRE-TONIC
AND AFTER-TONIC SYLLABLES

Summary

Investigations relating to unstressed vocalism of the modern Lithuanian language describe only the quantitative reduction, while the presence of qualitative reduction is questioned.

To check this, we put under study such vowels as [a:], [æ:], [e:], [i:], [u:], [o:].

When evaluating the spectrogram we determined the frequency of maximum constituents. Structural analysis technique was also made use of (V.A.Artemov).

The experiment carried out has proved that the contrast between compact and diffusial lowpitch and highpitch and simple sound quality of the investigated vowels still remains in the unstressed position. In other words, Lithuanian stressed and corresponding unstressed vowels are of relatively the same structure.

However the acoustic factor of unstressed vowels as compared with the stressed ones undergoes certain modifications, namely the maximum frequency level becomes lower and the sound production energy decreases, especially in the after-tonic syl-

table. In unstressed syllables the dense vowels [a:], [ɛ:] lose some of their density (and become more highpitched), diffusive vowels [i:], [e:] become more diffuse (and, more highpitched, vowels [u:], [o:] become more highpitched). This is a clear evidence of qualitative reduction in the Lithuanian language.

Reduction which is noticeable at the acoustic level (and, for sure, at the articulation level) is not observed at the phonological level since the Lithuanian unstressed vowels as well as stressed vowels carry an identical phonological load.

ТОН И УДАРЕНИЕ В КИТАЙСКИХ ДВУСЛОЖНЫХ И ТРЕХСЛОЖНЫХ СЛОВАХ

Вопрос о наличии, характере и месте ударения в китайском двусложном и многосложном слове до сих пор является малоисследованным и спорным.

Китайские исследователи признают, что в словах, состоящих только из тонированных морфем, слоги могут быть различны по степени ударенности, но подчеркивают доминирующую роль тона слога. По их мнению, в китайском языке прежде всего противопоставляются слоги полного образования, тонированные (ударные) и редуцированные, утратившие этимологический тон (безударные) (1). Акустической природы словесного ударения в китайском языке они почти не касаются.

В трудах советских китаистов словесное ударение экспериментально почти не исследовалось. Специально этому вопросу посвящена статья Н.А.Спешнева, опубликованная в 1958 г. (2). Н.А.Спешнев считает, что в китайском языке имеется три степени ударенности слога: сильноударенный, слабоударенный и безударный, характеризующийся полной редукцией тона. На основании экспериментальных данных он приходит к заключению, что акустическая природа китайского ударения количественная, а не силовая; ударный слог характеризуется большим временем звучания, а не большей интенсивностью. Исследования Т.П.Задюченко, проведенные в 1965 г., показали, что словесное ударение китайского языка характеризуется совокупностью нескольких выделительных средств: длительностью слога, его

интенсивностью и качественной характеристикой структурных элементов слога (3).

О точке зрения большинства советских и западных синологов можно судить лишь по их практическим пособиям.^I В ряде словарей тоны обозначаются только в тех слогах, на которые падает ударение, однако, не указываются критерии, на основании которых выделяется тот или иной слог (4-8).

Что касается тонального контура трехсложных слов, то, если судить по данным словаря И.М.Ошанина, средний слог всегда теряет свой этимологический тон (6). По данным словаря Б.С.Исаенко средний слог трехсложного слова может быть тонированным, т.е. ударным, и может утрачивать свой этимологический тон, т.е. быть безударным (4). Китайские составители словарей, как правило, отмечают сохранение тона всех слогов трехсложного слова, за исключением слов, образованных от двусложных с конечным редуцированным слогом, например, *boli* "стекло" - *bolibei* "стакан". Сведений об ударении ни один китайский словарь не дает.

1. УДАРЕНИЕ В КИТАЙСКОМ ДВУСЛОЖНОМ СЛОВЕ

Целью исследования является выяснение следующих двух вопросов:

1) получает ли китайское слово с двумя полными тонами только одно ударение, или же в словах такого типа оба слога являются ударными (равноударными или с ударениями разной степени);

2) связано ли распределение ударения с тональным рисунком слова.

I. Исключение составляют работы П.Кратохвила. См., например, его работу *Disyllabic Stress Patterns in Peking Dialect*, Archiv Orientalny, 1964, vol. 32, N3, pp.302-402.

В качестве исследуемого материала были взяты слова, сгруппированные по тональным рисункам (контурам) – всего 83 слова. Слова были выбраны так, чтобы в каждом из них финали обоих слогов были одинаковы или хотя бы близки по своему составу. Слова были наговорены пекинцами-преподавателями кафедры китайской филологии ИВЯ Жень Си-нин'ом и Лю Цзань-ли и записаны на интонографе в ЛЭФИПР I-го МГПИИЯ. Слова были прочитаны вне контекста с назывной интонацией, без явных эмоциональных оттенков. Цели исследования диктограм не были сообщены.

При обработке материала были измерены как длительность, так и интенсивность финали, т.е. гласной части слога. Затем эти слова и еще 200 других, наговоренных другим диктором, были предложены для слухового анализа 14-ти аудиторам – студентам и преподавателям китайского отделения ИВЯ с целью определения места и степени ударения¹. Таким образом, мы смогли сопоставить одни и те же слова в произнесении разных дикторов.

1. А н а л и з и н т е н с и в н о с т и

Чтобы нейтрализовать фактор тона, влияющего на интенсивность (2), мы подобрали для эксперимента такие слова, в которых не только финали обоих слогов одинаковы или близки по звуковому составу, но и тон один и тот же. Результаты измерений амплитуд интенсивности показаны в таблице 1.²

1. Для троих аудиторов китайский язык является родным, остальные аудиторы – русские, владеющие китайским языком. Показания аудиторов как китайцев, так и русских почти полностью совпадали.

2. Интенсивность условно дается в миллиметрах.

Таблица 1. Соотношение амплитуд интенсивности финалей
в словах типов I - I,2 - 2,4 - 4

Слова	Начальный слог	Конечный слог	Соотношение интенсивности
тип I-I			
pīnyīn	11 мм	14 мм	I : I,27
yīndāng	12 мм	14 мм	I : I,16
tōngzhī	3 мм	8 мм	I : 2,66
тип 2-2			
rénmín	14,5 мм	9 мм	I,6 : I
rúhé	11,5 мм	15 мм	I : I,3
jíshí	7,5 мм	9 мм	I : I,2
qíshí	11 мм	9 мм	I,22 : I
tóngxué	5 мм	6 мм	I : I,2
тип 4-4			
yùndòng	15 мм	12 мм	I,25 : I
lǐjù	10,5 мм	7,5 мм	I,4 : I
sùshè	13 мм	13 мм	I,07 : I
bàogào	16 мм	14 мм	I,15 : I

Результаты измерений показывают, что, несмотря на одинаковость финалей и тонов, интенсивность их в начальном и конечном слогах неодинакова: в одних примерах более интенсивной является финаль начального слога, в других - финаль конечного слога. Иногда обе финали оказываются почти равными по интенсивности.

Можно отметить также и то, что корреляционной зависимости между распределением интенсивности и тональным рисунком слова не наблюдается.

2. А н а л и з д л и т е л ь н о с т и

Данные наших интонограмм показывают, что в большинстве случаев финаль конечного слога продолжительнее финали начального. Так обстоит дело во всех словах типов 2-1, 3-1, 4-1, 3-2, 4-2, 2-3, 4-3. В некоторых из этих типов различие в длительности финалей начального и конечного слогов весьма значительно (см.табл.2).

Таблица 2. Соотношение длительности финалей
в словах типа 4-2^I

Слово	Начальный слог	Конечный слог	Соотношение длительности
shìshí	250	398	1:1,5
gòngtóng	314	406	1:1,3
dìtú	338	434	1:1,28
zìtiáor	272	446	1:1,7

Иначе обстоит дело в некоторых других типах.

В большинстве слов типов 1-1, 2-2, 1-4, 2-4, 3-4 наблюдается почти одинаковая длительность финалей начального и конечного слогов.

1. Длительность измерялась в миллисекундах.

Наконец, в типе 4-4 начальный слог гораздо продолжительнее конечного (см.табл.3).

Таблица 3. Соотношение длительности финалей
в словах типа 4-4

Слово	Начальный слог	Конечный слог	Соотношение длительности
yūdōng	488	276	1,8:1
lǐjù	346	258	1,4:1
sūshē	300	262	1,15:1
bāogāo	320	220	1,45:1

Таким образом, вся совокупность разбираемых двусложных слов распадается на две категории по признаку равенства и неравенства длительности финалей начального и конечного слогов.

Первую категорию образуют слова с равной длительностью финалей; это свойство может рассматриваться как показатель равноударности начального и конечного слогов в таких словах.

Вторую категорию образуют слова с неравной длительностью финалей. В отношении последней категории возникает вопрос, может ли это свойство рассматриваться как показатель равноударности таких слов, или оно свидетельствует лишь о разных степенях ударения.

Для того, чтобы разобраться в этом вопросе необходимо сопоставить слова рассматриваемой категории с двусложными

словами, конечный слог которых редуцирован, т.е. является безударным. Сравним, например, длительность финалей в следующих контрастных парах слов: *dōngxī-dōngxī*, *shēnghuó-shēnghuó*, *lǎodòng-lǎodòng*, *xiōngdì-xiōngdì*.

Таблица 4. Соотношение длительности финалей
в контрастных парах слов

Слова	Начальный слог	Конечный слог	Соотношение длительности
<i>dōngxī</i>	420	400	1,01:1
<i>dōngxī</i>	380	180	2,1:1
<i>shēnghuó</i>	450	460	1:1
<i>shēnghuó</i>	300	220	1,36:1
<i>lǎodòng</i>	480	440	1,09:1
<i>lǎodòng</i>	360	200	1,8:1
<i>xiōngdì</i>	500	400	1,25:1
<i>xiōngdì</i>	280	160	1,75:1

Как видно из приведенной таблицы, различие в длительности между финалью слога полного образования и слога редуцированного настолько резкое, что с ним ни в какое сравнение не может идти то различие, которое наблюдается в словах с двумя слогами полного образования. Это свидетельствует о том, что в словах второй категории неравенство финалей по длительности говорит лишь о разной степени ударности, а отнюдь не о наличии ударного и безударного слогов.

3. Соотношение длительности и интенсивности

В ходе исследования возник вопрос: какой из этих двух компонентов ударения - интенсивность и длительность - является главным. Для выяснения этого были сопоставлены соотношения слогов по интенсивности и длительности в словах с одинаковыми тонами (тип I-I, 2-2, 4-4, см. табл. 5).

Таблица 5. Соотношение длительности и интенсивности
в словах типов I-I, 2-2, 4-4

Слова	Соотношение длительности слогов	Соотношение интенсивности слогов
Тип I-I		
yīngāi	I:I	I:4,22
chūfā	I:I,04	I:16,5
tōngzhī	I,06:I	I:3,25
cānjiā	I:I,15	I:1,56
Тип 2-2		
huífa	I:I,02	I:I,05
tóngxué	I:I,06	I:1,85
Тип 4-4		
sùshè	I,3:I	I:3,66
bàogào	I,45:I	I:I,1
réliè	I,45:I	I:I,52
zhōngyào	I,27:I	I:4,43
huàju	I,66:I	10:I
jiěshào	I,4:I	I:1,06

Как видно из таблицы 5, в словах типа I-I по интенсивности преобладает конечный слог. В словах типа 2-2 интенсивность обоих слогов в одном из слов одинакова, а в другом преобладает конечный слог. Что же касается слов типа 4-4, в трех из них преобладает по интенсивности конечный слог, в двух интенсивность обоих слогов почти одинакова, а в одном слове интенсивность начального слога намного больше конечного. Здесь возникает вопрос: не связана ли эта различная интенсивность с воспринимаемыми на слух различиями в степени ударенности. Этот вопрос решается отрицательно и притом не только на основании данных слухового анализа. Об этом говорит различие в распределении интенсивности в словах одного и того же тонального рисунка, которые диктор произносил повторно среди других слов (3-4 слова, затем повторная реализация слов одного тонального рисунка, см. табл. 6).

Таблица 6. Соотношение интенсивности в повторно произнесенных словах

Слова	Соотношение интенсивности	Слова	Соотношение интенсивности
Тип I-I		Тип 4-4	
pīnyīn	I:I,27	yùndòng	I,25:I
pīnyīn	I:I	yùndòng	I:I,08
Тип 2-2			
rénmín	I,6:I		
rénmín	I:I,2		

Показательно также и то, что по мнению аудиторов, в слове *zùshè* "обсуждение" начальный слог сильнее, тогда как его интенсивность оказывается меньше. В слове *yùndòng* "движение" при первом произнесении интенсивнее начальный слог, при втором - конечный (см.табл.6), а аудиторы одинаково воспринимали начальный слог как более сильный.

Анализ длительности слогов в словах типа I-I, 2-2, 4-4 показывает, что во всех словах типа I-I и 2-2 длительность обоих слогов одинакова. Во всех словах типа 4-4 начальный слог значительно продолжительнее конечного. В слове *yùndòng* (см.табл.3) начальный слог почти в два раза продолжительнее конечного. Во всех словах типа 3-2 продолжительнее конечный слог.

Таблица 7. Соотношение длительности и интенсивности слогов в словах с разными тонами

Слова	Соотношение длительности слогов	Соотношение интенсивности слогов
Тип I-2		
<i>shēngcí</i>	I:I,05	I:I,16
<i>gāngcái</i>	I:I,18	I:I,6
<i>fēicháng</i>	I:I,14	I:I,41
<i>huānyíng</i>	I:I	6,56:I
Тип 3-2		
<i>lǐtáng</i>	I:I,45	I:6,5
<i>zhùxí</i>	I:I,25	I:3,44
<i>yǎnyuán</i>	I:I,15	I:I,30
<i>yóumíng</i>	I:I,22	I,22:I

Как показывает таблица 7, во всех словах типа 1-2 длительность обоих слогов почти одинакова. Что же касается интенсивности, то в двух случаях преобладает конечный слог, в одном - начальный, а в одном оба слога имеют почти равную интенсивность. В словах типа 3-2 по длительности преобладает конечный слог. В трех словах интенсивность конечного слога также больше, а в одном из слов начальный слог преобладает по интенсивности.

Таблицы соотношения слогов в словах с одинаковыми и разными тонами показывают, что в словах одного и того же тонального рисунка наблюдается закономерность в соотношении слогов по длительности. В отношении интенсивности такой закономерности не наблюдается.

Аудиторы, как правило, воспринимали слоги с большей длительностью как сильно ударенные.

Совокупность этих данных дает, по-видимому, право считать длительность основным компонентом ударения в китайском языке. Что касается интенсивности, то в словах с двумя полными тонами она не является устойчивым показателем разной степени ударения.

Соотношение слогов по длительности и интенсивности показывает, что преобладание того или иного слога по интенсивности не всегда совпадает с преобладанием по длительности. Нередко бывает так, что более интенсивный слог уступает по длительности менее интенсивному и, наоборот. (см. табл. 5, 7, слова *huàjū* *lǐtáng* *zhōuxí*). Возможно, что в этих случаях происходит компенсация одного компонента ударения другим.

В ы в о д ы:

1. В китайском слове с двумя полными тонами оба слога получают ударение; при этом в одной категории слов оба слога оказываются равноударными, в другой категории наблюдается преобладание одного из слогов.

2. В отличие от двусложных слов с конечным безударным слогом, где преобладание ударного слога постоянно выражено не только длительностью финали, но и ее интенсивностью, в словах с двумя полными тонами, где оба слога получают ударение, противопоставление слогов по степени ударенности выражается в основном длительностью. Интенсивность же в данном случае является лишь факультативным выделительным фактором.

3. В двусложных словах с двумя полными тонами корреляцию между распределением ударения и тональным рисунком удалось обнаружить лишь приблизительно, как показано в таблице 8.

Таблица 8. Распределение ударения в двусложных словах с двумя полными тонами

А. Слова равноударные	Б. Слова разноударные
Равная длительность финалей в обоих слогах	Неравная длительность финалей в обоих слогах
Тип 2-2	І. С преобладанием начального слога - только тип 4-4
	ІІ. С преобладанием конечного слога
	Сочетания 2-го и 3-го тонов; 3-2, 2-3 (в том числе и 3-3)
Все, кончающиеся на 4-й тон (кроме 4-4)	Все, начинающиеся на 4-й тон (кроме 4-4)
Все, начинающиеся на І-й тон	Все, кончающиеся на І-й тон (кроме І-І)

Если обозначить слог буквой S , каждый из четырех тонов соответствующими цифрами (1,2,3,4), а совокупность всех тонов - буквой n , то получим следующую схему (см.табл.9).

Таблица 9.

А. Слова равноударные	Б. Слова разноударные	
$S^2 - S^2$	$S^4 - S^4$ (преобладает начальный слог)	
	$S^3 - S^2, S^2 - S^3$ (в том числе) $S^3 - S^3$	
----- $S^n - S^4$ (кроме $S^4 - S^4$) $S^1 - S^n$	$S^4 - S^n$ (кроме $S^4 - S^4$)	(преобладает конечный слог)
	$S^n - S^1$ (кроме $S^1 - S^1$)	

ТОН И УДАРЕНИЕ В ТРЕХСЛОЖНЫХ СЛОВАХ

Целью исследования является выяснение следующих двух вопросов:

1. Определить тональный рисунок трехсложных слов, а именно, выяснить сохраняются ли тональные контуры морфем в трехсложном слове.
2. Определить степень ударенности слогов в трехсложных словах.

Для слухового анализа были записаны 400 слов в произнесении двух дикторов и 500 слов в произнесении одного диктора. Слова группировались по сочетаемости тонов. 28 слов

разного тонального рисунка были отобраны для записи на интонограф. Эти слова были подобраны так, чтобы в них были одинаковые (или близкие) по составу финалей слоги, в различной позиции и с разными тонами.

По интонограммам была измерена длительность финали слогов и частота колебаний основного тона. Экспериментальный материал был прослушан 10-ю аудиторами — студентами и преподавателями китайского отделения ИВЯ. Задача аудиторов состояла в том, чтобы определить место и степень ударения в трехсложных словах.

1. Д л и т е л ь н о с т ь ф и н а л е й

Как показало исследование двусложных слов, длительность играет первостепенную роль в выделении ударного слога, поэтому, определяя место и степень ударения в трехсложных словах китайского языка, мы основывались на данных измерения длительности каждого из слогов.

Наши измерения показывают, что в большинстве случаев финаль конечного слога продолжительнее финалей начального и среднего слогов. При сравнении длительности финалей конечного и начального слогов, выяснилось, что, как правило, финаль конечного слога в среднем в 1,5-2 раза продолжительнее финали начального слога (см.табл.10). Из 28 слов исключение составляют 3 слова (*dàzìbào*, *dàshìguān*, *shēnghuófài*), в которых конечная и начальная финали почти равны по длительности, и 2 слова (*piànmǎnxìng*, *mǎnxìngbìng*), где начальная финаль несколько превосходит конечную.

Наиболее короткой является финаль среднего слога.

Особенно заметно различие в длительности средней и конечной финалей. В большинстве случаев финаль конечного слога оказывается в 2-3 раза, а иногда даже в 4-5 раз продолжительнее финали среднего слога.

Цифровые данные показывают также, что и начальная финаль продолжительнее средней, хотя различие здесь не так велико, как при сравнении с конечной финалью. В некоторых же случаях длительность средней финали оказывается равной начальной финали (см. табл. 10, слова *gōngchéngshǐ fādiàn-jī gònghéguó*), а в одном из 28-ми примеров (*guófángjūn*) длительность средней финали даже больше начальной.

Таблица 10. Соотношение по длительности слогов
трехсложных слов

Слова	Начальный слог	Средний слог	Конечный слог	Соотношение финалей по длительности		
				Н	С	К
1. <i>tōngzhīshū</i>	248	130	320	1,9	: 1	: 2,04
2. <i>xīngqīsān</i>	220	86	414	2,55	: 1	: 4,81
3. <i>gōngnóngbīng</i>	242	148	196	1,63	: 1	: 2
4. <i>gōngchéngshǐ</i>	234	220	332	1,06	: 1	: 1,50
5. <i>shēnghuófèi</i>	258	176	260	1,48	: 1	: 1,49
6. <i>xībēijiǎo</i>	192	180	430	1,06	: 1	: 2,38
7. <i>dōngbēijiǎo</i>	280	168	366	1,66	: 1	: 2,17
8. <i>cānkǎoshū</i>	238	202	348	1,17	: 1	: 1,72
9. <i>shēngqīngshū</i>	258	212	318	1,18	: 1	: 1,48

Слова	Началь- ный слог	Сред- ний слог	Конеч- ный слог	Соотношение финалей по длительности		
				Н	С	К
10. fādiānjī	226	228	352	I	: I	: I,47
11. gōngkēbiāo	232	168	440	I,34	: I	: 2,6
12. bōlìzhǐ	252	104	368	2,42	: I	: 3,53
13. bōlìbēi	186	98	452	I,89	: I	: 4,6I
14. bōlìbān	160	100	458	I,6	: I	: 4,58
15. guófāngjūn	182	214	314	I	: I,18	: I,9I
16. wémuèjiǎ	256	152	358	I,68	: I	: 2,35
17. fúzérén	152	300	398	I	: I,97	: 2,56
18. xíngróngcǐ	288	174	310	I,65	: I	: I,77
19. qīngyǔbiǎo	244	162	414	I,5	: I	: 2,55
20. huángǎngtǔ	228	166	400	I,37	: I	: 2,4
21. gōnghéguó	188	188	384	I,02	: I	: 2,3
22. mǎiguózéi	242	170	338	I,42	: I	: I,98
23. piānmǐanxíng	274	108	252	2,53	: I	: 2,33
24. mǎnxíngbǐng	256	172	204	I,48	: I	: I,18
25. dàzìbào	222	98	274	2,26	: I	: 2,78
26. fúxiězhǐ	188	150	400	I,25	: I	: 2,66
27. dàshíguān	224	110	248	2,03	: I	: 2,25
28. dòufunǎo	212	52	390	4,03	: I	: 7,5

В особую группу выделяются трехсложные слова, образованные от двусложных слов с конечным редуцированным слогом, например, bōlì "стекло" - bōlìbēi "стакан", dòufu "бобы" - dòufunǎo "бобовый сыр". Средний слог в таких трехсложных словах особенно короткий, что свидетельствует о редукции слога в составе трехсложного слова.

Но особенно интересным является тот факт, что такая же малая длительность средней финали наблюдается и в некоторых трехсложных словах, образованных от двусложных с конечным нередуцированным слогом (см. табл. IО, примеры: *xīngqī sān*, *fú zé rén*, *piān miàn jīng*, *dà zǐ bào*). Это говорит о том, что в трехсложных словах в серединной позиции редуцироваться может и такой слог, который в соответствующем двусложном слове не является редуцированным.

Таким образом, основываясь на измерениях длительности финалей, можно сделать следующие выводы относительно места и степени ударения в трехсложных словах.

Разная длительность финалей трехсложного слова свидетельствует о трех разных степенях ударения в слове. Основное, главное ударение, как правило, падает на конечный слог, характеризующийся наибольшей длительностью финали. Начальный слог является слабоударенным. Что же касается среднего слога, то он, в зависимости от длительности его финали, может быть таким же слабоударенным, как начальный, либо безударным.

Это подтверждается и данными слухового анализа.

2. Т о н а л ь н ы й р и с у н о к т р е х - с л о ж н ы х с л о в

Анализ наших интонограмм показывает, что начальный и конечный слог трехсложного слова сохраняют свой этимологический тон. Рассматривая графические изображения тонов в слогах трехсложных слов, мы можем заметить, что в начальной и конечной позиции этимологический тон сохраняет свои основные признаки: направление движения основного тона и частотный

диапазон, или высотный интервал между начальной и конечной точками тона (3).

Проследим, например, реализацию второго тона в начальном и конечном слогах трехсложного слова. Графическое изображение финали слога *xíng* в слове *xíngróngcí* - "имя прилагательное" (см. график I) показывает, что по направлению движения этот тон восходящий; частотный интервал между начальной и конечной точками тона довольно большой - чистая кварта (150-200 гц).

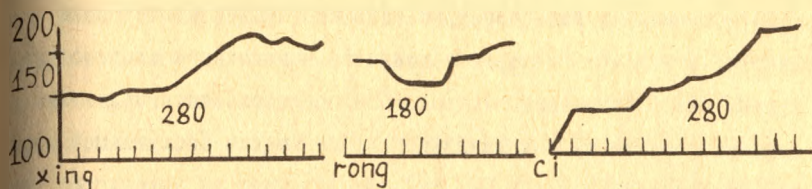


График I. Движение тона в слогах слова *xíngróngcí*

Все это говорит о том, что в начальном слоге *xíng* сохраняется этимологический второй тон. Это подтверждается и данными слухового анализа, тон начального слога довольно четко воспринимается на слух как восходящий.

Рассмотрим график конечного слога *cí* в том же слове *xíngróngcí* (см. график I). По направлению движения тон также восходящий. Частотный интервал между начальной и конечными точками равен октаве, что в два раза больше частотного интервала в начальном слоге *xíng*.

Таким образом, в конечном слоге 2-й тон выражен более четко, чем в начальном, что вполне согласуется с известным правилом акцентов в словах такого типа: конечный слог сильно-

ударенный - тон этого слога выражен более четко. Об этом же говорят и данные слухового анализа: тон конечного сильноударенного слога воспринимается на слух намного отчетливей, чем начального.

Сопоставим теперь 4-й тон в начальном и конечном слогах трехсложных слов. В начальном слоге 4-й тон сохраняет все свои характерные особенности, что видно при рассмотрении графика слога *mān* в слове *mānxīngbīng* "хроническая болезнь" (см. график 2). Тон нисходящий; интервал между начальной и конечной точками тона довольно большой - равен увеличенной кварте. Эти данные говорят о том, что в начальном слоге сохраняется этимологический 4-й тон. Это подтверждается и данными слухового анализа. Тон начального слога четко воспринимается как нисходящий.

Тональный контур конечного слога *bīng* в том же слове представлен на графике 2. Тон - нисходящий, интервал между начальной и конечной точками значительно больший, чем в начальном слоге *mān* - равняется октаве (200 гц - 100 гц).

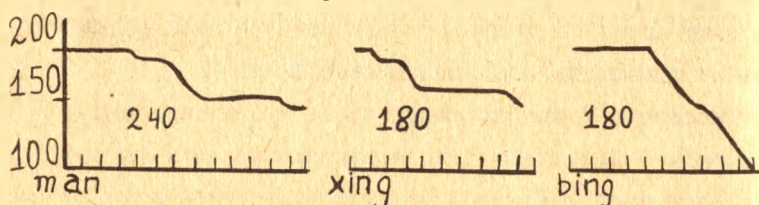


График 2. Движение тона слогов в слове *mānxīngbīng*

Таким образом, четвертый тон, как и рассмотренный выше второй тон, в конечном слоге выражен более четко, чем в начальном, что также связано с распределением ударения в трех-

сложном слове.^{х)} Это подтверждается и данными слухового анализа. Большинство аудиторов воспринимает тон конечного слога, как более четко выраженный 4-й тон, чем тон начального слога.

Тон среднего слога. На графике 3 мы видим, что второй тон слога *nóng* (в слове *gōngnóngbīng* солдат рабочей крестьянской армии) почти утратил восходящее направление. В течение первых 60 мсек звучания слога (из 160) – тон ровный, в середине слога частота колебаний повышается на малую терцию (от 150 гц до 175 гц), а затем снова падает (до 150 гц), причем падение тона происходит только в последние 40 мсек, а перед этим на протяжении 40 мсек фиксируется наивысшая точка в звучании тона. В слове *chéng* (слово *gōngchéngshī* "инженер", график 3) второй тон в первой (большей половине звучания) повышается на целую терцию (от 140 гц до 175 гц), а потом падает до 160 гц. Тональный контур слогов *nóng* и *chéng* в срединной позиции отличается от обычного контура этимологического второго тона тем, что он в значительной степени нивелирован: малый интервал повышения на небольшом участке звучания. Это связано с общим тональным рисунком слов этого типа, а именно, с тем, что второй тон находится в окружении двух первых тонов. Кроме того, это безусловно связано с тем, что средний слог наименее ударенный, что ведет к нивелированию высотного интервала тона.

х) Мы не даем сопоставления первого и третьего тонов в начальном и конечном слогах, поскольку 1-му тону вообще не свойственен такой признак как высотный интервал (начальная и конечная точка произносятся на одной и той же высоте); 3-й же тон имеет свою полную форму только в конечном слоге, а в начальном и среднем он по общему правилу превращается в ровный низкий.

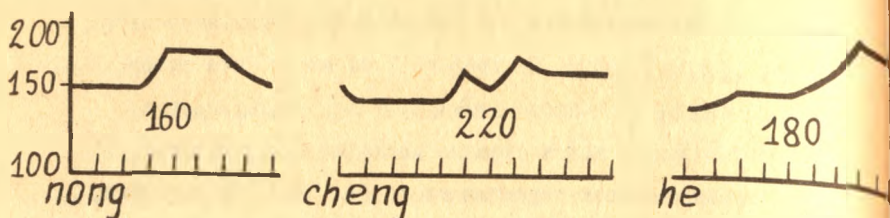


График 3. Движение тона в слогах: *nóng*, *chéng*, *hé*
(в словах) *gōngnóngbǐng*, *gōngchéngshǐ*,
gònghéguó).

В других случаях (в окружении других тонов) 2-й тон выражен более четко. Например, в слове *gōng* (график I) тон восходящий, частотный интервал равен большой терции. Сравнительно четок второй тон и в слове *hé* в слове *gònghéguó* "республика" (график 3). Частотный интервал равен кварте (140 гц - 200 гц).

Ч е т в е р т ы й т о н в с р е д и н н о й
п о з и ц и и . Сравним средний слог *xíng* в слове *máoxíngbǐng* (график 2) с начальным и конечным слогами этого слова. В среднем слове так же, как в начальном и конечном слогах, сохраняется нисходящее направление движения тона. Частотный интервал тона довольно большой (чистая кварта), но все же меньше, чем в начальном слове; интервал начального слога равен увеличенной кварте, а конечного - октаве. Это показывает, что в среднем слове четвертый тон, хотя и сохраняется, но оказывается менее четко выраженным, чем в конечном и даже в начальном слогах.

Редуцированный тон в середине трехсложного слова. Как известно, редуцированный тон произносится на различной высоте в зависимости от тона предыдущего слога. В частности, после первого тона редуцированный тон произносится очень низко (напр., в слове *bōlǐ* "стекло"). Мы предполагали, что также низко будет произноситься и редуцированный тон в словах *bōlibēi* "стакан", *bōlizhǐ* - "целлофан", *bōlibǎn* "стеклянная пластинка". Интонограммы же показывают, что в трехсложном слове редуцированный тон среднего слога *lǐ* производится на той же высоте, что и начальный первый тон, иначе говоря, происходит своего рода ассимиляция тонов (см. график 4).

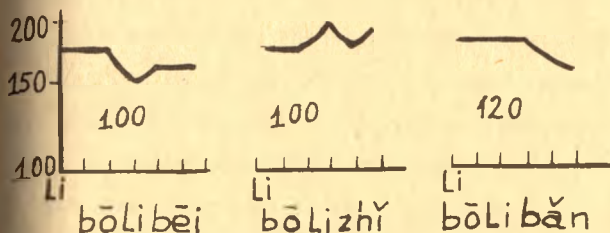


График 4. Движение тона слога *lǐ* в словах *bōlibēi*, *bōlizhǐ*, *bōlibǎn*

Анализ интонограмм показал также, что тон среднего слога может быть редуцированным не только в тех словах, которые образованы от двусложных слов с конечным редуцированным, но в некоторых случаях и в таких словах, которые образованы от двусложных слов с конечным полным тоном. Например: *xīngqī* "неделя" - *xīngqīsan* "среда", *tōngzhī* "сообщение" - *tōngzhīshì* "письменное извещение". Графики слога *qī* в

слове *xīngqísān* и слога *zhī* в слове *tōngzhīshū* (график 5) показывают, что тональный контур этих слов весьма похож на контур абсолютно редуцированного слога *lǐ* в словах *bōlibāi*, *bōlizhǐ*, *bōlibān*.

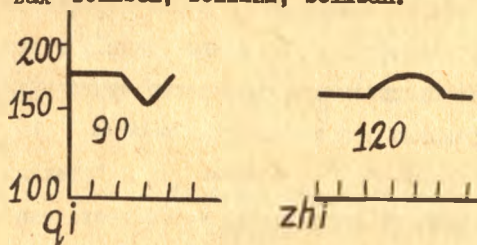


График 5. Движение тона слогов *qí* и *zhī* в словах *xīngqísān*, *tōngzhīshū*.

В ы в о д ы :

1. В китайском трехсложном слове возможны три степени ударения. Основное, главное ударение падает на конечный слог, характеризующийся наибольшей длительностью финали. Начальный слог является слабоударенным. Что же касается среднего слога, то он может быть таким же слабоударенным как начальный, либо безударным.

2. В трехсложных словах этимологические тоны морфем, как правило, сохраняются. В конечном, сильноударенном, слоге тон выражен более четко, чем в слабоударенных слогах (начальном и срединном). В слабоударенных слогах сокращается частотный интервал тона.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Luo Ghang PEI, Putonghua yinmu gangyao, Beijing, 1957.
Hanyu jiangyi, Beijing, 1958.
Xiandai Hanyu, Beijing, 1957.
2. Н.А.Спешнев, Акустическая природа словесного ударения в современном китайском языке. Вопросы корейского и китайского языкознания, Ленинград, ЛГУ, 1958, стр.140-145.
3. Т.П.Задоевко, "Фонетическая природа словесного ударения в современном китайском языке" сб.Материалы коллоквиума по экспериментальной фонетике и психологии речи, М, 1966, стр.127.
4. Б.С.Исаенко, Опыт китайско-русского фонетического словаря, Москва, 1957.
5. В.С.Колоколов, Китайско-русский словарь, М, 1956.
6. И.М.Ошанин, Предисловие к макету Большого китайского словаря, М., 1960.
7. E.H. Mathews, A Chinese-English Dictionary, Shanghai, 1931.
8. Chao Yan Ren, Hanyu cidian, Shanghai, 1953.
9. P. Kratochvil, Disyllabic Stress Patterns in Peking Dialect, Archiv Orientalny, 1964, vol. 32, N3, pp. 383-402.

I. Doumkowa

L.E.Ph., I.O.L. M.S.U.

THE STRESS AND TONES IN THE CHINESE TWO AND
THREESYLLABIC WORDS

The article concerns some specific traits of the Chinese word-stress and the specific nature of the stress and tones in the Chinese two - and threesyllabic words.

On the basis of experimental data the author comes to the conclusion that the main component of the stress in an isolated Chinese word is the duration of the finale of the stressed syllable. The intensity is a supplementary component. During the process of research a certain correlation was found between the tone-draw of the words and the position of the stress in Chinese two-syllabic words.

The article also concerns the stresses and the tonedraw of threesyllabic words. The experimental data show that there are three degrees of word-stress in a Chinese threesyllabic word. The most stressed is the last syllable. The initial syllable is weakstressed; as to the middle one it may be either weakstressed as the initial or without stress at all. As a rule, the syllables of three-syllabic words keep their etymological tones. There is a direct dependence between the degree of stress on the one hand and the tone of the syllable on the other.

А.М.КАРАПЕТЬЯНЦ

ОПЫТ ФОРМАЛЬНОГО ПОДХОДА К ИЗУЧЕНИЮ
ИНТОНАЦИИ ПУТУНХУА

О. В В Е Д Е Н И Е

Данная статья является продолжением и уточнением ранее написанной статьи [6], находящейся в настоящее время в печати¹⁾, поэтому здесь не будут повторяться уже данные определения и объясняться полученные прежде результаты. С другой стороны, часть определений и понятий здесь уточнена или введена заново.

I. М О Д Е Л Ъ В О С П Р И Я Т И Я И Н Т О Н А Ц И И

I.I. МОДЕЛЬ ВОСПРИЯТИЯ ИНТОНАЦИОННЫХ ХАРАКТЕРИСТИК
ЧЕЛОВЕКОМ

Нижеприведенную модель ни в коем случае не следует рассматривать как описание процессов, действительно имеющих место в человеческом организме; ее следует рассматривать лишь как весьма примитивную интерпретацию в физиологических терминах тех характеристик речевых явлений, которые изучаются в данной работе.²⁾

1) Краткое изложение ее см. в журн. *Phonetica*, vol. 12, № 3-4, Basel-New-York, 1965.

2) Вопрос о восприятии речевых характеристик до сих пор остается невыясненным до конца. Наиболее полные данные об этом см. в [9] , стр.82-84.

Предполагается, что барабанная перепонка (мембрана) совершает вынужденные колебания с частотой, равной частоте основного тона. Данные о частоте колебаний мембраны передаются нейронами в головной мозг, причем пока мозг не обработает первого показания частоты, он не воспринимает последующего (хотя нейрон может их передавать непрерывно). В свою очередь мембрана обладает определенной инерцией, что приводит к тому, что нейрон передает не мгновенные значения частоты, а усредненные за определенный промежуток времени.

В связи с этим представляется целесообразным ввести следующие термины. Время, проходящее между двумя последовательными восприятиями, обозначим как время X . Таким образом, время X является промежутком времени, проходящим между снятием показаний с мембраны в процессе восприятия речи. Время X может рассматриваться как единица времени в человеческом мозге, как единица отсчета времени своеобразных часов, работающих в нем.

С другой стороны, вполне естественно считать, что эти часы не работают непрерывно. Они включаются тогда, когда в мозг поступает первый речевой сигнал, причем отсчет времени X начинается через определенный промежуток времени после получения этого сигнала. Этот промежуток времени мы обозначим как время Y ($Y \geq 0$). Часы, будучи раз включенными, должны когда-нибудь выключиться. Естественно полагать, что они должны выключиться либо тогда, когда через время X не поступает никакого сигнала, либо тогда, когда поступает сигнал конца.

Наконец, инерция мембраны может быть выражена через тот промежуток времени, за который происходит усреднение колеба-

Этот промежуток времени мы обозначим как время K . Время K является характеристикой инерции мембраны — чем больше K , тем больше ее инерция. Все эти три характеристики — X , Y и K — можно показать на следующей схеме:



Вполне естественно, что совершенно не обязательно, чтобы, как на данной схеме, выполнялись следующие условия: $Y < X$, $K = Y$. В действительности X и Y являются совершенно независимыми друг от друга величинами. K также не зависит от X , но есть единственное условие, которое должно выполняться: $K \leq Y$. Впрочем, если предположить, что пуск часов не является одновременно восприятием характеристик сигнала, то и это условие отпадает. Однако при дальнейшем изложении мы будем исходить из предпосылки, что пуск часов является одновременно восприятием сигнала.

1.2. РЕГИСТРАЦИЯ ИНТОНАЦИОННЫХ ХАРАКТЕРИСТИК ИНТОНОГРАФОМ И ИХ СЧИТЫВАНИЕ

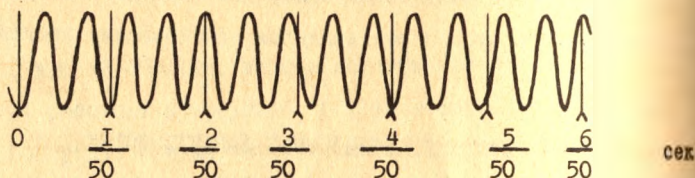
Запись интонационных характеристик на интонограф и последующая расшифровка его показаний представляет некоторую аналогию вышеизложенной модели восприятия интонационных характеристик человеком.

Интонограф дает осциллограмму колебаний, и ввиду того, что звуки речи весьма сложны по своему составу, форма этой

осциллограммы представляет собой результат наложения синусоиды основного тона на синусоиды гармоник. Однако синусоиду основного тона, как обладающую наибольшей интенсивностью, обычно достаточно легко выделить.

Как легко видеть, осциллограмма дает мгновенные значения частоты, которые невозможно измерить. Для определения частоты пользуются отметками времени, которые также даются интонографом таким образом, что единицей времени является $1/50$ секунды. За $1/50$ сек. подсчитывается число колебаний и определяется частота основного тона. Не трудно видеть, что в данном случае мы имеем дело с усреднением. Для интерпретации этого положения воспользуемся терминами § I. I.

На нижеприведенном чертеже показана интонограмма первого слога некоторой речевой последовательности (для простоты слог начинается с гласной и показана лишь синусоида основного тона).



Количество колебаний в единицу времени подсчитывается с точностью до $1/4$ колебания, таким образом, мы имеем следующие значения количества колебаний в $1/50$ сек: 2 ; $2 \frac{1}{2}$; $2 \frac{1}{4}$; $2 \frac{1}{4}$; $2 \frac{1}{4}$; $2 \frac{1}{4}$. Не трудно видеть, что здесь $X = Y = K = 1/50$ сек, так как слог начался в точке 0, а первый подсчет сделан в точке $1/50$ сек.

Запись $X = Y = K = 1/50$ сек является неточной. По определению X , Y и K величины объективно существующие и не определенные. В данном же случае X , Y и K нами заданы заранее, поэтому вышеприведенное равенство надо переписать в следующем виде:

$$x = y = k = 1/50 \text{ сек.}$$

Здесь вводятся новые обозначения: x , y и k , которые являются некоторыми значениями времени, задаваемыми нами самими, и в данном случае каждое из них равняется $1/50$ сек.

Из всего вышесказанного необходимо следует вывод: данные интонографа уже являются полученными на основании заранее заданных величин x , y и k , а так как X , Y и K не известны, x , y и k с равными основаниями могут быть заданы любыми числами.

Операцию задания x , y и k мы назовем квантованием потока речи. Так, квантованием потока речи интонографом является задание x , y и k равными $1/50$ сек.¹⁾

В заключение введем оставшиеся обозначения. Частотой (λ) называется отношение количества колебаний (m) за время k к времени k :

$$\lambda = \frac{m}{k}.$$

Количество значений частоты в определенном отрезке времени длительностью в t определяется целой частью следующего отношения:

$$n_{\lambda} = \frac{t + (x - y)^2}{x}.$$

1) $1/50$ сек в дальнейшем кратко называется "первоначальным квантом" (ПК).

2) Если пуск часов на является одновременно восприятием сигнала, то $n_{\lambda} = \frac{t - y}{x}.$

1.3. ПЕРЕКВАНТОВАНИЕ

Выше уже говорилось, что при квантовании потока речи интонографом $x = y = k = 1/50$ сек, однако мы имеем право задавать любые x , y и k . На первый взгляд кажется, что это можно сделать посредством изменения отметок времени на интонографе. Однако это приведет к многократному считыванию с интонографа одного и того же текста, что является весьма трудоемкой операцией; поэтому мне показалось целесообразным предложить процедуру получения данных нового квантования из данных исходного квантования потока речи интонографом ($x = y = k = 1/50$ сек), которая была названа **пере-квантованием**.

Предположим, при $x = y = k = 1/50$ сек мы получили следующие значения m : 2, 2 1/2, 2 1/4, 2 1/4. Так как величина k при данном квантовании постоянна, вместо λ мы можем непосредственно взять величину m , причем для последующей машинной обработки ее удобно выразить в четвертых долях; таким образом, $m = 08, 10, 09, 09$. Считая величину m частотой, предполагаем, что мгновенная величина m на протяжении всего отрезка времени, равного $1/50$ сек, постоянна, и определяем ее через каждую сотую секунды. Теперь $m = 08, 08, 10, 10, 09, 09, 09, 09$. Величину $1/100$ сек мы и примем как основу квантования, следовательно, как минимальную единицу времени.¹⁾

1) Основу квантования можно брать любой, однако взять ее меньше $1/100$ сек мы не имеем достаточных оснований, так как, во-первых, наши исходные данные не обладают нужной степенью точности, и, во-вторых, предположение о постоянном значении мгновенной частоты на протяжении всего отрезка в $1/50$ сек уже не допускает дробления этого отрезка на меньшие части (чем меньше части, тем больше различных значений частоты, соответствующих данному отрезку, мы получим при переквантовании).

Теперь будем задавать различные k в единицах, равных $1/100$ сек. При $k = 1$ мы получаем следующий ряд m :

08, 08, 10, 10, 09, 09, 09, 09 ;

при $k = 3$:

00, 00, 26, 28, 29, 28, 27, 27

и т.д.

Здесь следует оговорить, что при переквантовании мы складываем значения m влево от нужного числа, то есть для получения первого m при заданном k складываются числа, находящиеся в исходном ряду слева от первого m и значение первого m , причем если одно из слагаемых равно нулю, вся сумма считается равной нулю.

Нетрудно заметить, что при различных k λ получается из m в результате различных преобразований, однако, учитывая интересы машинной обработки (трудность обращения с дробными числами), а также взаимно-однозначные соответствия m и λ в каждом случае при сохранении простоты обращения с натуральными числами, представлялось целесообразным сохранить рассмотрение m как частоты. Теперь очевидно, что значение m следует рассматривать как обозначение вынужденных колебаний мембраны при инерции, которая задается значениями k .

Выразим y в тех же единицах, что и k , а под x будем понимать, через сколько значений воспринимаются значения частоты. Тогда при $k = 3$, $y = 1$, $x = 2$ берутся следующие значения m :

00, 29, 27.

Далее k , x и y будут задаваться в этих единицах; единица времени равняется $1/100$ сек.¹⁾

1.4. МОДЕЛЬ ВОСПРИЯТИЯ ИНТОНАЦИИ В ПОТОКЕ РЕЧИ

В настоящем параграфе нам предстоит обобщить все то, что было сказано в этой главе. Исходя из вышеизложенного, модель восприятия потока речи можно представить себе в следующем виде.

После того, как в мозг воспринимающего поступает сигнал о начале речевого сообщения, через определенное время (время y) начинают работать своеобразные часы, которые позволяют воспринимать и обрабатывать сигналы через определенные промежутки времени (время x). Эти часы не только как бы отбирают сигналы, которые должны поступать в мозг, но также выступают и как своеобразный счетчик времени, определяющий длительность отдельных частей потока речи. Естественно предположить, что этот счетчик времени, будучи однажды включенным, должен когда-нибудь выключиться. По поводу того, когда же именно он должен выключиться, можно сделать два предположения: либо он выключается лишь в конце сообщения, либо он включается и выключается неоднократно на протяжении данного сообщения. Более оправданно признать второе предположение.

Действительно, это предположение подтверждается рассмотрением вопроса о паузах. Из положения о времени x как

1) Было бы последовательно выражать x в тех же единицах, что k и y , однако в связи с тем, что уже были отлажены программы, в которых x понимался в вышеуказанном смысле, пришлось, к сожалению, сохранить это обозначение. Нетрудно видеть, что для выражения x в тех же единицах, что k и y , к его значению достаточно прибавить единицу.

счетчике времени следует, что время X является единицей измерения длительности пауз. То есть различные паузы различаются друг от друга по своей длительности, измеренной в единицах времени X . Однако внутри одного сообщения (текста) имеются паузы, длительность которых может иметь любую величину, превышающую определенную. Так, например, диктор, читающий текст, дойдя до конца фразы, может сделать очень долгую паузу, если он устал или ему необходимо перевернуть страницу написанного текста, или в случае устной речи, когда оратору необходимо выпить стакан воды и т.п. В этом случае нарушения интонационной структуры сообщения не наблюдается, в то время как если такая пауза была бы сделана в середине фразы, она была бы воспринята как явное нарушение интонационной структуры. Из этого следует вывод, что существуют паузы, длительность которых не воспринимается (подобно тому, как есть числа $1, 2, \dots$, а дальше тьма), их длительность не измеряется, на них и должен выключаться счетчик времени.

Каким же образом производится это выключение? Уместно предположить, что такое выключение производится после получения сигнала "конец макросегмента". Весь вопрос заключается в том, какую форму принимает этот сигнал. Сигнал этот может принимать как форму "длительность паузы превысила значение

n ", так и форму "поступил частотный сигнал, который может иметь место лишь в конце макросегмента". Представляется более экономичным второй вид организации выключения счетчика времени X . Сигналом этого рода может считаться появление КЧ I, которая, по наблюдениям над обработанными текстами, появляется именно в конце макросегмента.

Положение о выключении счетчика времени X подтверждается данными Л.А.Чистович с сотрудниками о временной организации синтагмы. На основании сравнения коэффициента вариации длительности синтагм и пауз между ними, ими было показано, что "исследование временной организации последовательности артикуляторных движений может производиться только в пределах синтагмы. Переход от одной синтагмы к другой не является определенным по времени" (см. [7] , стр.85).

1.5. ИНТОНЕМА

В соответствии с моделью восприятия интонации, предложенной в § I.I, при правильном квантовании воспринимаемые сигналы имеют величины лишь соответствующие КЧ. Эти величины по аналогии с фонетическим уровнем можно назвать интонемами. Как было уже показано в работе [6] , количество этих единиц (КЧ) не превышает шести. Функция интоном на интонационном уровне полностью соответствует функции фонем на фонетическом. Сами по себе интономы никакой смысловой нагрузки не несут. Смысловую нагрузку может нести лишь интонационное "слово", определенная последовательность интоном между двумя паузами. Следует подчеркнуть, что паузы также могут нести различную смысловую нагрузку, наподобие пробела просто, пробела с точкой, пробела между абзацами и т.п. Из этого следует важность КЧ как языка описания интонации, так как интоному можно еще определить и как "истинную КЧ".

2. МЕТОДИКА ИССЛЕДОВАНИЯ

2.1. ОСНОВНЫЕ ЭТАПЫ ИССЛЕДОВАНИЯ.

2.1.1. Как уже было указано в [6] поток речи должен делиться на части, называемые сегментами, которые отделяются друг от друга паузами определенной длины, то есть каждый вид сегмента отличается определенной КП.

Если сопоставить вышеуказанное с моделью §1.1, то вполне уместно предположить, что описанные там часы должны выключаться на время определенной КП, скорее всего КП, образующей макросегмент. Таким образом, КП нужны нам не только для выделения типов сегментов, но и для построения языка их описания - системы КЧ. Следовательно, первым этапом исследования должно являться выделение КП.

2.1.2. Вторым этапом исследования будет получение распределения при исходном квантовании; это распределение целесообразно принять в качестве некоторого эталона, с которым будут сравниваться распределения, получаемые в дальнейшем. После этого, задавая различные x и y , мы получим новые распределения, которые будем сравнивать с исходным распределением и между собой до тех пор, пока при определенных x и y не получим наибольшую контрастность.

2.1.3. После этого мы переходим к переквантованию. Получая исходные распределения для каждого k , мы сопоставляем их с исходным распределением и между собой, находя лучшее из них. Определяя лучшее, мы задаем те же x и y , что в § 2.1.2, и сравниваем полученное распределение с исходным. В том случае, если оно лучше исходного при данном k , и

лучше полученного в § 2.1.2, - можно считать, что заданные значения x , y и k являются истинными.

2.1.4. Как видно из вышеизложенного, алгоритм определения КЧ строится путем перебора почти всех логических возможностей. Это заколдованный круг: мы не знаем, каким образом части потока речи связаны с КЧ, мы не знаем точного членения потока речи на части и величин КЧ, и тем самым не можем определить однозначно ни того, ни другого. Поэтому нам и приходится перебрать почти все логические возможности, чтобы найти достаточно контрастное распределение частот. Но здесь есть одно обстоятельство, которое дает нам некоторый выход из этого круга. Действительно, если длительность определенных частей потока речи является релевантной, тогда она должна каким-то образом фиксироваться у слушающего и говорящего. Восприятие же длительности, исходя из модели § 1.1, сводится, очевидно, к подсчету значений, полученных мозгом, или, что одно и то же, к определению времени X каждого слога. Таким образом, изучение распределения длительностей слогов, например, может нам дать некоторые указания о времени X . В то же время, существование определенных КД, а также возможность иррелевантности длительности некоторых частей потока речи (слогов) нарушают связь распределения длительностей со временем X . Кроме того, возможна, хотя и очень мало вероятно, независимость определения длительности человеком от частотных характеристик.¹⁾

1) При построении алгоритма настоящего исследования я исходил из отрицания этой возможности.

2.1.5. Внутренний критерий правильности выбора данного квантования — получение наибольшей контрастности распределения для всех текстов при данном квантовании. Контрастность распределения длительности частей потока речи (слов), о котором говорилось в § 2.1.4, можно также рассматривать как внешний критерий правильности квантования, так как он с частотой не связан. Другими внешними критериями являются наибольшая контрастность распределения длительности пауз, наибольшая контрастность распределения интенсивностей (здесь в свою очередь возможны ее различные усреднения), наибольшая контрастность распределения промежутков времени между одноименными частями слогов и т.п.

2.1.6. После определения величин X , Y и K можно непосредственно перейти к исследованию интересующих нас текстов, имея в своем распоряжении язык для описания их интонационной структуры. Исследование проводится с учетом данных о делении потока речи на части, описанных в [6]. Благодаря этим данным мы можем определить КЧ внутри каждого слога, в том числе КЧ его начала и конца, определить связь этих КЧ с тоном слога, с его фонетическим составом и его положением в сегменте и т.п. Как легко видеть, заключительная часть исследования может проводиться на вычислительной машине. При достаточной степени формализации грамматики на машине также может быть произведено изучение зависимости интонационных характеристик какой-либо части потока речи от его грамматического значения и т.п.

2.2. АЛГОРИТМ ИССЛЕДОВАНИЯ

В настоящем параграфе я попытаюсь более точно сформулировать алгоритм определения истинного квантования в рамках §§ 2.1.1 - 2.1.3. Алгоритм будет изложен по пунктам, каждый из которых содержит ограниченное число операций. Если после определенного пункта специально не указано, к какому пункту необходимо переходить после его выполнения, - нужно переходить к следующему по порядку пункту. Цифровыми примечаниями (их номер дан в круглых скобках) к алгоритму являются части логической схемы, которые будучи заданы в алгоритмической форме, придадут изложению излишнюю громоздкость.

1. Нахождение распределения $\mu_{\alpha\beta}$ и определение КП.

2. Нахождение общего распределения при исходном квантовании.

3. Задание $a_1 \leq x \leq b_1$, и $a_2 \leq y \leq b_2$ ¹⁾. Нахождение распределения при всех x и y . Их упорядочение по возрастанию контрастности с отбрасыванием всех распределений, худших исходного (1). Взятие наиболее контрастного распределения.

1) Если не будут указаны пределы, то этот пункт никогда не будет выполнен. Разумеется, $\alpha \geq 0$ (то же относится к п.7).

4. Нахождение нового распределения при сбрасывании счета времени на большую КП. Сравнение его с лучшим распределением п.3. Если первое хуже, перейти к п.5. Если первое лучше, считать эту КП макросегментальной и переходить к п.7.

5. Сравнить величину предыдущей КП с минимально возможной. Если она минимальна, перейти к п.6. Если она не минимальна, взять следующую по значению КП и перейти к п.4.

6. Взять следующие по убывающей контрастности значения x и y из п.3 и перейти к п.4(2).

7. Нахождение общих распределений при $\alpha_3 \leq k \leq \beta_3$. Их упорядочение по возрастанию контрастности и отбрасывание всех распределений, худших, чем распределение п.2.

8. Задание x и y из п.4. Сравнение этого распределения с распределением п.7. Если первое лучше, процедура заканчивается. Если первое хуже, перейти к п.9.

9. Взять следующее по убыванию контрастности значение k и перейти к п.8 (3).

Примечания:

(1). Если все получаемые распределения хуже исходного, необходимо задать $a_3 \leq k \leq b_3$. При этом найти k , дающее лучшее распределение и проводить с ним далее операции как с исходным квантованием. Если таких k не найдется, считать $X = Y = K = k_{\text{исх.}}$.

(2). Если значений x и y больше нет, тогда надо переходить к п.7, беря в п.8 лучшие значения из п.3 и проверяя их на каждую КП. Если ни при каком k при сбрасывании не получится большей контрастности, взять лучшие значения x из п.3 и, считая $X = Y$, найти лучшее k . В этом случае следует полагать, что $X = Y = x$.

(3). Если значений k больше нет /ср.(2)/, следует перебрать остальные значения x и y из п.3 от лучших к худшим со всеми k не отброшенными в п.7. Если же это не даст результата, считать $K = k_{\text{исх.}}$, а X и Y взять из п.4.

Вышеизложенный алгоритм предусматривает, естественно, далеко не все логические возможности. Поэтому после его осуществления необходима проверка по явлениям, указанным в §§ 2.1.4 и 2.1.5. В случае отрицательного результата этой проверки следует расширить количество логических возможностей, например, не отбрасывать всех распределений хуже исходного и т.д.

2.3. РЕАЛИЗАЦИЯ АЛГОРИТМА ИССЛЕДОВАНИЯ НА ЦЭВМ.

Вычислительная часть настоящей работы производилась на ЦЭВМ "Гамма". Средняя скорость выполнения команд машиной "Гамма" приблизительно равна 1800-2000 команд в секунду. Ввод информации в машину производится с перфокарт со скоростью 150 перфокарт в минуту, а вывод - с помощью широкой печати со скоростью 150 строк в минуту¹⁾, либо с помощью перфорирующего устройства со скоростью 75 перфокарт в минуту, или того и другого. Существенным преимуществом машины "Гамма" является ее высокая надежность, что позволяет обходиться без двойного счета. (Счет на этой машине можно проводить непрерывно в течение многих часов).

Для осуществления алгоритма мною было написано и отлажено более десяти программ различной степени сложности. Следует заметить, что полное осуществление всего алгоритма исследования требует больших затрат машинного времени. Например, для текста средней длины²⁾ на осуществление первых пяти пунктов алгоритма требуется примерно 14 часов работы машины.

2.4. КОДИРОВАНИЕ ДАННЫХ

Полное изложение этого вопроса потребовало бы слишком много места; здесь я останавлиюсь лишь на кодировании фонетической структуры слога и его тона, так как эти кодовые обозначения будут употребляться и в дальнейшем.

1) Строка содержит 92 символа из алфавита, состоящего из латинских букв, цифр и точки.

2) См. табл. 7.

Китайский слог состоит из трех частей: инициали, медиали¹⁾ и финали, каждая из которых в системе чжуньцзыму может отсутствовать. Ниже (см. стр. 225) приведен список всех этих частей с использованными мною кодами и их алфавитными обозначениями, принятыми в настоящее время в КНР. Код данной части определяется сочетанием цифр в левом столбце и в верхней строке. Медиали кодируются одним числом.

Для лиц, не знакомых со структурой китайского слога, следует отметить, что финали о и е находятся в отношении дополнительной дистрибуции: после медиалей о и 2 всегда следует о, после медиалей I и 3 всегда следует е. Кроме того, инициали 4-ой и 5-ой строчек также находятся в отношении дополнительной дистрибуции — инициали четвертой строчки всегда встречаются перед медиалами 0 и 2, а инициали пятой строчки — перед медиалами I, 3.

Нарушением системы чжуньцзыму является в данном случае включение е редуцированного как отдельной фонемы и включение слога ег, выпадающего из общей системы, в систему финалей.

Все части обладают способностью составлять отдельный слог, но из инициалей этой способностью обладают лишь инициали 6X и 7X; слог, состоящий из этой инициали и следующего за ней специфического гласного *ĩ*, кодируется как слог, состоящий из одной инициали.

1) В китайстике обычно медиаль рассматривается как часть финали.

Инициали

	1	2	3	4
1	b	p	m	f
2	d	t	n	
3			l	
4	g	k		h
5	j	q		x
6	zh	ch	r	sh
7	z	c		s

Медали

1	i
2	u
3	ü

Финали

	0	1	2	3	4	5
1	a	ai	ao	an	ang	
2	o(e)	ei	ou	en	eng	er
3	е редуц.					

Следовательно, слог кодируется посредством пятизначного числа, два первых знака которого обозначают инициаль, третий знак - медиаль, а два последних - финаль, причем отсутствие той или иной части слога кодируется нулями в соответствующих позициях.

Тон слога кодируется следующим образом: нулевой тон обозначается нулем, четыре тона - своими номерами. Кроме слогов, обладающих постоянным тоном, имеется еще ряд слогов с переменным тоном, которые произносятся различными тонами (их тон определяется главным образом положением в сегменте и тоном последующего слога). У этих морфем кодируется тот тон, которым они должны произноситься в данном случае, причем к его номеру прибавляют 4. Так, тон слога с переменным тоном, который в данном месте должен произноситься четвертым, обозначается цифрой 8.

2.5. КОНТРОЛЬ

Явления, ставящие под сомнение объективность исследования, уже назывались в работе [6] ; основным из них следует считать совпадение КЧ с целым числом колебаний в отрезке. Там же были указаны причины, позволяющие считать это явление действительно присущим языку. Однако мне показалось целесообразным проверить причины более частого появления таких частот посредством точных критериев. Для этой цели была составлена программа, на результатах работы которой мне хотелось бы здесь остановиться.

Во-первых, программа печатает количество случаев, когда разность суммы колебаний в двух последующих ПК и количества

колебаний, подсчитанных в двух ПК, равна той или иной величине. Эти данные приведены в табл. I. Здесь содержатся данные по двум текстам (07 и 09), а также по частям текста 09 (в первом столбце указано, из каких перфокарт¹⁾ состоят эти части). Эта таблица показывает, что в среднем в 81% всех случаев эти числа совпадают, и лишь в 2% всех случаев разность превышает $1/4$ колебания. Эти данные свидетельствуют о том, что колебания не имеет смысла подсчитывать с точностью, превышающей $1/4$ колебания в ПК, и точность подсчета колебаний составляет $\pm 1/4$, хотя и данные, полученные с точностью до $1/4$ колебания, можно признать вполне удовлетворительными.

Во-вторых, программа позволяет получить данные о зависимости несовпадения суммы колебаний в двух соседних ПК от того, положено ли число колебаний в них равным целому числу, числу с $1/4$, числу с $1/2$ или числу с $3/4$, а также от величины одного из суммируемых чисел m , так как при больших m подсчитать число колебаний с заданной точностью труднее. Эти данные приведены в табл. 2-5. В них столбец под знаком равенства дает количественную оценку случаев, когда число m должно было быть таким, столбец под знаком " $>$ " дает количество случаев, когда оно должно было быть больше, а столбец под знаком " $<$ " - число случаев, когда оно должно было быть меньше.

Как получаются эти числа, лучше всего пояснить на примере. Допустим, что в некотором ПК число колебаний оказалось

1) На каждой карте содержатся данные о 20 ПК.

Таблица I

Отклонение	00	01	02	03	Σ
Т.07 общ.	824 80%	189 18%	10 1%	1 0%	1024 100%
Т.09 общ.	4501 81%	938 17%	99 2%	23 0%	5561 100%
I-64	569 76%	162 22%	19 2%	2 0%	752 100%
65-128	500 75%	146 22%	18 3%	5 0%	669 100%
129-192	578 78%	137 18%	19 3%	4 1%	738 100%
193-256	449 71%	153 24%	22 3%	3 1%	627 100%
257-320	596 86%	90 13%	8 1%	0 0%	694 100%
321-384	616 87%	87 12%	7 1%	1 0%	711 100%
385-448	535 85%	90 14%	3 0%	4 1%	632 100%
449-512	499 88%	64 11%	3 1%	2 0%	568 100%
513-528	156 92%	9 5%	0 0%	2 2%	167 100%
I-256	2096 75%	598 21%	78 3%	14 1%	2786 100%

равным 4, в следующем $4\frac{1}{4}$, а в двух вместе - 8. Это можно записать следующим образом: $m_1 = 16$, $m_2 = 17$, $m_3 = 36$. (Нетрудно видеть, что в табл. I этот случай прибавит единицу к столбцу с разностью 01). Это говорит о том, что для каждого числа с вероятностью $p = 2/3$ можно утверждать, что каждое из этих чисел определено правильно, а с вероятностью $p = 1/3$, что первое и второе число должны быть меньше, а третье - больше. Поэтому мы должны в таблице для целых чисел (табл.2) к столбцу " $=$ " прибавить 4 (у нас два целых числа, а к столбцам " $<$ " и " $>$ " прибавить по единице, а в таблице для чисел "с одной четвертью" (табл.3) к столбцу " $=$ " прибавить 2, а к столбцу " $<$ " прибавить 1. Приведем еще несколько примеров без объяснений.

$$1) \quad m_1 = 16, \quad m_2 = 16, \quad m_3 = 32;$$

$$2) \quad m_1 = 16, \quad m_2 = 18, \quad m_3 = 32;$$

$$3) \quad m_1 = 17, \quad m_2 = 18, \quad m_3 = 32.$$

	0			$1/4$			$1/2$			$3/4$		
	$>$	$=$	$<$	$>$	$=$	$<$	$>$	$=$	$<$	$>$	$=$	
1)		9										
2)	4	2					I	2				
3)	3					3			3			

Таким образом, при разности, равной нулю, для каждого числа мы прибавляем 3 в столбец " $=$ ", при разности, равной единице, мы прибавляем 2 в столбец " $=$ " и 1 в столбец " $>$ "

Таблица 2

	$m_1 \leq 20$				$m_1 > 20$				Итого			
	>	=	<	Σ	>	=	<	Σ	>	=	<	Σ
07 обш.	57 2%	2282 95%	97 4%	2436 100%	0	0	0	0	57 2%	2282 95%	97 4%	2436 100%
09 обш.	222 2%	10515 95%	379 3%	11116 100%	143 3%	4615 92%	234 5%	4992 100%	365 2%	15130 94%	613 4%	16108 100%
I-64	35 2%	1509 94%	58 4%	1602 100%	35 5%	634 88%	51 7%	720 100%	70 3%	2143 93%	109 5%	2322 100%
65-128	35 3%	1231 95%	60 4%	1326 100%	13 2%	577 92%	43 7%	633 100%	48 2%	1808 93%	103 5%	1959 100%
I29-192	41 2%	1576 93%	63 4%	1680 100%	13 3%	427 93%	22 6%	462 100%	54 2%	2003 93%	85 4%	2142 100%
I93-256	49 3%	1333 92%	76 5%	1458 100%	22 5%	429 88%	38 8%	489 100%	71 4%	1762 90%	114 6%	1947 100%
257-320	16 1%	1367 96%	42 4%	1425 100%	21 3%	773 93%	34 4%	828 100%	37 1%	2140 95%	76 3%	2253 100%
321-384	22 2%	1312 96%	37 3%	1371 100%	17 3%	557 96%	11 2%	585 100%	39 2%	1869 96%	48 2%	1956 100%
385-448	14 1%	1038 96%	25 2%	1077 100%	11 2%	549 96%	13 2%	573 100%	25 2%	1587 96%	38 3%	1650 100%
449-512	8 1%	850 97%	15 2%	873 100%	8 1%	553 96%	18 2%	579 100%	16 1%	1403 96%	33 2%	1452 100%
513-528	1 0%	281 99%	3 1%	285 100%	3 2%	116 94%	4 4%	123 100%	4 1%	397 97%	7 2%	408 100%
I-256	160 3%	5649 93%	257 4%	6066 100%	83 4%	2067 89%	154 7%	2304 100%	243 3%	7716 92%	411 5%	8370 100%

Таблица 3

	$m_1 \leq 20$				$m_1 > 20$				Итого			
	>	=	<	Σ	>	=	<	Σ	>	=	<	Σ
7 00м.	59 3%	1683 92%	88 5%	1830 100%	0	0	0	0	59 3%	1683 92%	88 5%	1830 100%
09 00м.	242 3%	7843 93%	348 4%	8433 100%	143 3%	4615 92%	234 5%	4992 100%	385 3%	12458 93%	582 4%	13425 100%
1-64	41 3%	1159 92%	63 5%	1263 100%	35 5%	634 88%	51 7%	720 100%	76 4%	1793 90%	114 6%	1983 100%
65-128	54 5%	1005 90%	57 5%	1116 100%	13 2%	577 92%	43 8%	633 100%	67 4%	1582 90%	100 6%	1749 100%
129-192	34 3%	1093 92%	58 5%	1185 100%	13 2%	427 94%	22 4%	462 100%	47 3%	1520 92%	80 5%	1647 100%
193-256	30 3%	874 91%	59 6%	963 100%	22 5%	429 88%	38 7%	489 100%	52 3%	1303 90%	97 7%	1452 100%
257-320	25 2%	919 95%	25 2%	969 100%	21 5%	773 87%	34 8%	828 100%	46 2%	1692 94%	59 3%	1797 100%
321-384	13 1%	933 96%	26 2%	972 100%	17 3%	557 95%	11 2%	585 100%	30 2%	1490 96%	37 2%	1557 100%
385-448	24 2%	860 94%	31 3%	915 100%	11 2%	549 96%	13 2%	573 100%	35 2%	1409 95%	44 3%	1488 100%
449-512	20 2%	698 95%	20 2%	738 100%	8 2%	553 95%	18 4%	579 100%	28 2%	1251 95%	38 3%	1317 100%
513-528	1 0%	296 97%	9 3%	206 100%	3 2%	116 94%	4 3%	123 100%	4 1%	412 96%	13 3%	429 100%
1-256	159 4%	4131 91%	237 7%	4527 100%	83 4%	2067 90%	154 6%	2304 100%	242 4%	6198 91%	391 7%	6881 100%

Таблица 4

	$m_1 \leq 20$				$m_1 > 20$				Итого			
	>	=	<	Σ	>	=	<	Σ	>	=	<	Σ
07 обш.	49 2%	2136 93%	101 4%	2286 100%	0	0	0	0	49 2%	2136 93%	101 4%	2286 100%
09 обш.	183 3%	6283 93%	290 4%	6756 100%	85 2%	3236 93%	101 6%	3522 100%	268 3%	9519 93%	491 5%	10278 100%
I-64	28 3%	826 93%	40 4%	894 100%	18 4%	380 88%	34 8%	432 100%	46 4%	1206 91%	74 5%	1326 100%
65-I28	28 4%	678 91%	38 5%	744 100%	10 4%	279 87%	23 9%	322 100%	38 3%	957 91%	61 6%	1056 100%
I29-I92	31 3%	1046 91%	75 7%	1152 100%	11 3%	419 88%	47 9%	477 100%	42 3%	1465 90%	122 7%	1629 100%
I93-256	34 4%	736 91%	43 6%	813 100%	14 5%	234 85%	25 10%	273 100%	48 5%	970 89%	68 6%	1083 100%
257-320	10 2%	731 95%	21 3%	762 100%	11 3%	372 94%	13 3%	396 100%	21 2%	1103 95%	34 3%	1158 100%
321-384	17 2%	833 95%	29 3%	879 100%	3 0%	520 95%	23 5%	546 100%	20 1%	1353 95%	52 3%	1425 100%
385-448	21 3%	639 93%	24 3%	684 100%	10 3%	395 91%	27 6%	432 100%	31 3%	1034 93%	51 4%	1116 100%
449-512	10 2%	551 96%	15 2%	576 100%	8 1%	551 97%	8 1%	567 100%	18 2%	1102 96%	23 2%	1143 100%
513-528	4 2%	243 96%	5 2%	252 100%	0 0%	86 99%	1 1%	87 100%	4 1%	329 97%	6 2%	339 100%
I-256	121 4%	3286 91%	196 5%	3603 100%	63 4%	1312 87%	129 9%	1504 100%	174 4%	4598 90%	325 6%	5094 100%

Таблица 5

	$m_1 \leq 20$				$m_1 > 20$				Итого			
	>	=	<	Σ	>	=	<	Σ	>	=	<	Σ
07 обм.	59 2%	2479 93%	126 4%	2664 100%	0	0	0	0	59 2%	2479 93%	126 4%	2664 100%
09 обм.	192 3%	6391 91%	404 6%	6987 100%	107 3%	3586 90%	225 6%	3918 100%	299 2%	9977 91%	629 6%	10905 100%
I-64	23 2%	877 91%	66 6%	966 100%	16 4%	309 86%	35 10%	360 100%	39 3%	1186 89%	101 8%	1326 100%
65-128	23 3%	818 88%	86 9%	927 100%	26 6%	359 86%	32 8%	417 100%	49 4%	1177 88%	118 9%	1344 100%
129-192	36 4%	722 90%	40 6%	798 100%	15 4%	345 91%	21 5%	381 100%	51 4%	1067 90%	61 6%	1179 100%
193-256	42 5%	792 88%	72 7%	906 100%	14 4%	323 86%	38 10%	375 100%	56 4%	1115 87%	108 9%	1279 100%
267-320	13 2%	745 94%	37 5%	795 100%	5 1%	480 96%	22 4%	507 100%	18 1%	1235 94%	59 4%	1302 100%
321-384	17 2%	834 95%	31 3%	882 100%	17 3%	531 92%	28 5%	576 100%	34 2%	1365 94%	59 4%	1458 100%
385-448	18 2%	832 94%	38 4%	888 100%	8 1%	551 94%	26 4%	585 100%	26 2%	1383 94%	65 4%	1473 100%
449-512	12 2%	532 92%	32 6%	576 100%	6 1%	586 96%	20 3%	612 100%	18 2%	1118 94%	52 4%	1188 100%
513-528	8 3%	236 96%	2 1%	246 100%	0 0%	102 97%	3 3%	105 100%	8 2%	338 96%	5 2%	351 100%
I-256	124 4%	3209 89%	264 7%	3597 100%	71 5%	1336 87%	126 8%	1533 100%	195 3%	4545 89%	388 8%	5128 100%

или " $<$ ", при разности, равной двум, мы прибавляем 1 в столбец " $=$ " или 2 в столбец " $>$ " или " $<$ ", и наконец, при разности, равной 3, мы прибавляем 3 в столбец " $>$ " или " $<$ ".

Данные таблиц 2-5 сведены в табл. 6, где указан процент отклонений и частное от деления случаев, когда число на самом деле должно было быть меньше, на число случаев, когда оно должно было быть больше. Данные этих таблиц убедительно показывают, что целые значения числа колебаний в кванте дают как раз наименьшие отклонения, кроме того, наблюдается некоторая тенденция к завышению значений m , особенно характерная для числа " $с\ 3/4$ ". Однако эта тенденция весьма слабая. Все вышеизложенное позволяет полностью снять предположение о возможной необъективности исследования в этом плане.

Таблица 6

	0		1/4		1/2		3/4	
	% отклон.	$\frac{<}{>}$	% от- клон.	$\frac{<}{>}$	% от- клон.	$\frac{<}{>}$	% от- клон.	$\frac{<}{>}$
07 общ.	5	1,7	8	1,5	7	2,1	7	2,1
09 общ.	6	1,7	7	1,5	7	1,8	9	2,1
I-64	7	1,6	10	1,5	9	1,6	11	2,6
65-128	7	2,1	10	1,5	9	1,6	12	2,4
129-192	7	1,6	8	1,7	10	2,9	10	1,2
193-256	10	1,6	10	1,9	11	1,4	13	1,9
257-320	5	2,1	6	1,3	5	1,6	6	3,3
321-384	4	1,2	4	1,2	5	1,6	6	1,7
385-448	4	1,5	5	1,3	7	1,6	6	2,5
449-512	4	2,1	5	1,4	4	1,3	6	2,9
513-528	3	1,7	4	3,2	3	1,5	4	0,6
I-256	8	1,7	9	1,6	10	1,9	11	2,0

3. ДАННЫЕ ИССЛЕДОВАНИЯ

3.1. МАТЕРИАЛ ИССЛЕДОВАНИЯ

Для исследования было избрано 24 связанных текста. Все эти тексты носят повествовательный характер и имеют минимальную эмоциональную окрашенность. Данные о них приведены в таблице 7. В первом столбце этой таблицы указан номер текста, во втором диктор. Здесь гласные обозначают дикторов-женщин и согласные - дикторов-мужчин. В третьем столбце указывается характер текста. Обозначения приняты следующие:

- "в" - выступление или речь,
- "д" - научно-исторический текст,
- "л" - научно-лингвистический текст,
- "п" - общественно-политический текст,
- "р" - объявление по радио,
- "с" - художественный текст (сказка),
- "т" - научно-технический текст,
- "х" - художественный текст (роман, рассказ).

Остальные обозначения указаны в шапке таблицы.

Источником получения данных о фонетическом составе и тоне слогов послужили [1] и [2]. Для слогов, не вошедших в [1], тон и фонетический состав проверялись по [3]. Для морфем с переменным тоном вопрос о их тонировании был решен в соответствии с положениями работы [4] с той лишь разницей, что морфема *yi*, соответствующая иероглифу — , перед первым тоном считалась произносящейся также первым тоном. Это было вызвано тем, что по предварительным данным тон морфемы *yi* перед слогом первого тона более напоминает первый. Следует отметить, что при машин-

Таблица 7

Текст	Диктор	Характер текста	Кол-во морфем	Длительность в ПК	Длительность в сек	Средняя скорость слог/сек
01	А	х	363	6133	123	2,9
02	Б	х	362	6978	140	2,6
03	Д	в	151	3992	80	1,9
04	В	п	153	3270	66	2,3
05	А	п	495	6739	135	3,6
06	А	и	279	4323	87	3,2
07	Б	с	119	2207	44	2,7
08	Б	с	127	2494	50	2,5
09	И	т	611	10538	211	2,9
10	И	л	820	13430	269	3,0
11	О	р	64	927	19	3,4
12	О	р	33	479	10	3,3
13	Ж	в	335	3999	80	4,2
14	У	в	410	6195	124	3,3
15	Г	р	43	630	15	2,8
16	З	в	397	5879	117	3,4
17	Д	р	111	1892	38	2,9
18	Я	в	771	7823	177	4,4
19	И	х	699	12565	251	2,8
20	Б	с	136	2327	47	2,9
21	В	х	322	5991	120	2,7
22	В	п	128	2240	45	2,8
23	Б	и	271	3655 ¹⁾	91	3,0
24	Б	и	120	1957	39	3,1
Итого			7320	116666	2400	-
Средняя			305	4861	99	3,1

1) В этом тексте ПК = 1/40 сек, таким образом это соответствует 4570 обычных ПК.

той обработке лучше иметь большее дробление, чем есть на самом деле, а не наоборот. Так, в данном случае можно составить программу так, чтобы машина читала пятерки как восьмерку, в то время как обратное невозможно или затруднительно (если бы тон морфемы y_i перед всеми тонами, кроме четвертого, кодировался восьмеркой, то для определения, перед слогом какого именно тона находится данная морфема, потребовалось бы одновременное обращение к следующей карте деления).

Машинной обработке по техническим причинам до настоящего времени были подвергнуты лишь тексты 07, 09 - 13, 19.

3.2. ДАННЫЕ ПО ПАУЗАМ

Как уже говорилось в § 2.2, первым пунктом построенного алгоритма является "нахождение распределения пауз и определение КП". Эта процедура была проделана со всеми текстами, подвергшимися машинной обработке (см. § 3.1). На ее результатах мы остановимся ниже, однако сначала следует остановиться на некоторых предварительных соображениях.

3.2.1. Как было определено ранее, КП являются контрастными совокупностями распределения длительностей пауз. Но, как известно, величина пауз колеблется в очень широких пределах от 0 до 150 ПК. С другой стороны, количество пауз длиннее 5ПК в тексте примерно на один порядок ниже числа морфем; таким образом, как правило, контрастное распределение практически не наблюдается из-за незначительного количества точек. В связи с этим встает вопрос об объединении данных о длительности пауз по каким-то разрядам. Мне представляется, что вопрос этот не только формальный. Как будет видно из

данных по длительности, приведенных в § 3.4, единица КД приблизительно соответствует одному ПК, но такая единица дает слишком большую дифференциацию пауз. В этой связи вполне естественно предположить, что счетчик времени X на протяжении паузы работает несколько иначе, чем во время поступления обычных сигналов, и само время X для паузы несколько иное. Обозначим это время через X_I ; тогда $X_I > X$. Определить время X_I можно, задавая данные x_I до получения наибольшей контрастности распределения пауз. Это время x_I и является объединением определенным способом величин длительностей пауз в ПК. Нетрудно видеть, что при $x_I \ll X_I$ мы получим слишком много контрастных совокупностей (или очень большой разброс точек); при $x_I \approx X_I$ количество контрастных совокупностей будет одного порядка с количеством сегментов (но не обязательно равно); при $x_I \gg X_I$ распределение потеряет свою дискретность и приблизится к нормальному.

С другой стороны, в связи с неточностью определения границ слога, в особенности начала некоторых инициалей, точность определения величины пауз, близких к нулю, кажется несколько сомнительной. Поэтому представляется целесообразным начинать отсчет времени x_I не с нуля, а с определенного количества ПК, которое по аналогии (подчеркивая, совершенно поверхностно!) обозначено через u_I .

Для нахождения длительностей пауз при заданных x_I и u_I была составлена специальная программа, с помощью которой были обработаны тексты 07, 09, 10, 13 и 19 (тексты 11 и 12 этим методом не обрабатывались ввиду крайней малочисленности пауз в них). Пример зависимости распределения длитель-

ностей пауз от величин x_I и y_I показан в табл.8. Здесь для наибольшего по объему текста 10 показаны данные при $x_I=y_I=0$ (первые два столбца), при этом таблица сокращена за счет отсечения всех пауз больше 45ПК (весь интервал между 46 и 142 ПК заполнен спорадически появляющимися единицами); при $x_I = 3$, $y_I = 1$ (вторые два столбца) в первом столбце обозначены лишь концы интервалов; так, в третьей строчке содержатся данные о паузах длительностью от 5 до 8 ПК; при $x_I = 9$, $y_I = 1$ (последние два столбца). Эта таблица довольно наглядно иллюстрирует вышеизложенные положения.

3.2.2. Наименее раздробленную из наиболее контрастных картину дало распределение длительностей пауз при $x_I = 2$ и $y_I = 1$. Эти данные приведены в табл.9, где в первом столбце указаны интервалы, а в остальных число пауз для данного текста. На этой таблице горизонтальными линиями приблизительно отделены контрастные совокупности для каждого текста (следует отметить, что для текста 13 эти совокупности выделяются довольно плохо). Как видно из этой таблицы, вместо трех контрастных совокупностей получилось четыре. Эти совокупности были обозначены соответственно как микросегментальная (M_i), мезосегментальная (M_e), макросегментальная 1 ($Ma\ 1$) и макросегментальная 2 ($Ma\ 2$) паузы. Забегая вперед, отметим (см. § 3.3.2), что и $Ma\ 1$ и $Ma\ 2$ отделяет именно макросегменты и разница между ними заключается, вероятно, в том, что $Ma\ 1$ являются минимальными макросегментальными паузами в системе противопоставления пауз (следует отметить, что величины интервалов M_i , M_e и $Ma\ 1$ приблизительно равны между собой), а $Ma\ 2$ являются макросегментальными паузами, полностью

Таблица 8

$X_1=0, Y_1=0$	$X_1=3, Y_1=1$	$X_1=9, Y_1=1$
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 522 33 59 46 21 19 7 6 4 3 4 1 2 2 1 4 5 4 5 4 2 1 4 3 3 2 4 0 4 1 2 1 2 3 1 2 2 2 0 2 0 1 1 1 1 0 0	0 4 8 12 16 20 24 28 32 36 40 44 48 52 56 60 64 68 72 76 80 84 88 92 96 100 104 108 112 116 120 124 128 132 136 140 144 522 159 36 10 12 15 11 10 6 4 4 3 2 3 2 1 2 2 0 0 2 1 1 3 0 1 0 0 0 1 0 0 0 0 0 0 1	0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 522 202 30 15 6 5 4 2 4 2 0 1 0 0 1

Таблица 9

Интервал	07	09	10	13	19
0 3 6 9	54 35 6	387 119 29	522 138 47 13	186 90 13	384 170 34 11
12 15 18 21 24 27 30 33 36 39 42 45 48 51 54 57 60 63 66 69 72 75 78 81 84 87 90 93 96 99 102 105 108 111 114 117 120 123 126 129 132 135 138 141 144	1 0 1 0 1 3 6 2 2 2 0 1 2 1 1	2 1 4 11 12 6 5 5 5 3 4 2 0 1 2 3 1 1 1 1 0 1 0 1 0 0 1 1 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1	7 7 14 7 10 6 7 6 5 4 2 1 2 5 1 1 1 1 1 3 0 0 0 1 1 1 1 2 1 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 1	7 6 3 4 4 2 2 4 1 1 0 5 3 2	2 4 16 5 12 6 6 5 2 2 2 4 1 3 1 2 5 1 2 0 0 1 2 0 1 2 0 1 0 1 0 0 0 0 0 1 0 0 1

разрушающими ритмическую структуру речи (диктор просто отдал)¹), и интервал Ма 2 на порядок больше интервалов всех остальных пауз.

Более точно границы между типами пауз были определены на основании распределения пауз при $x_I = y_I = 0$. Данные об этих границах приведены в табл.10, где в первом столбце обозначен тип паузы, а в остальных интервалы для текстов (удалось приблизительно определить величины пауз для текстов II и I2, причем в случае возможного двойного определения границы менее вероятное значение поставлено в скобки). Следует отметить, что интервалы одного вида паузы для разных текстов дают еще более близкие значения, чем в табл.9.

Таблица 10

Пауза	07	09	10	11	12	13	19
Ми	0-9	0-7 (0-12)	0-11	0-8 (0-11)	0-8	0-6(16)	0-13 (0-16)
Ме	14-26 (15-25)	12-27 (15-25)	13-30 (15-36)	20-21 (22-24)	17 (21)	8(18)-26 (10-26)	16-27 (17-31)
Ма I	(20-34) (32-37)	29-42(46) (39-46)	32-42 (39-44)	32 (35)	27-30 (28-31)	28-35	29-43 (32-44)
Ма 2	40-49 (43-53)	46(50)-143 (46-143)	46-142 (49-143)	43-68 (45-70)	- -	41-47	46-131 (48-131)

1) Возможно, также, что эти паузы отделяют большие куски потока речи (нечто вроде абзацев), собственной интонационной структурой не обладающие.

Следует заметить, что паузы считались до начала согласной части. Однако при осуществлении п.4 алгоритма возникают трудности в связи с тем, что в большинстве случаев согласные не дают значений частоты и пауза кажется больше, чем она на самом деле. В таком случае происходит лишнее сбрасывание y . Поэтому было изучено также распределение пауз и до гласной части, и в случае расхождения величин КП для каждого вида пауз они задавались обе. Данные о КП пауз до гласной части также приведены в табл.10 во второй строке каждого вида в скобках.

3.3. ДАННЫЕ ПО КП

3.3.1. Как уже упоминалось в работе [6], исходные данные, получаемые при $x = y = 0$ (непосредственные данные, снятые с интонографа), не всегда дают контрастное распределение.

Распределения частот в мелах для текстов, подвергнутых машинной обработке, показаны на рис.1 (т.07), 2 (т.09), 3 (т.10), 4 (т.11), 5 (т.12), 6 (т.13), 7 (т.19).

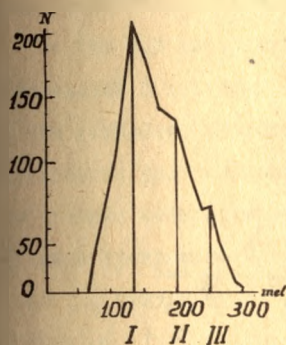


Рис.1

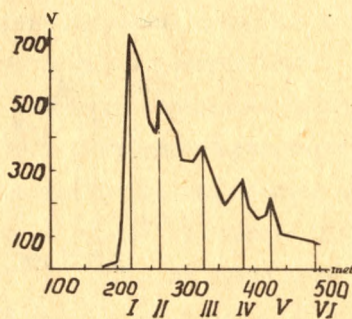


Рис.2

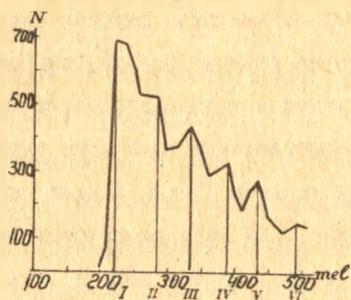


Рис.3

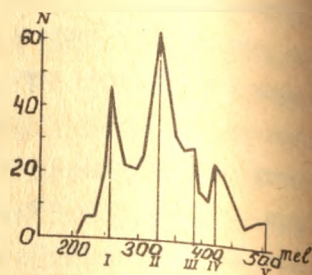


Рис.4

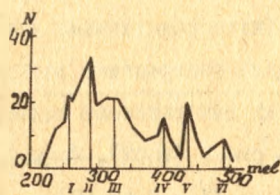


Рис.5

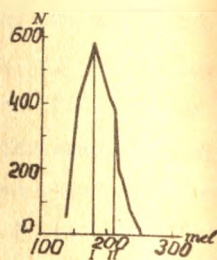


Рис.6

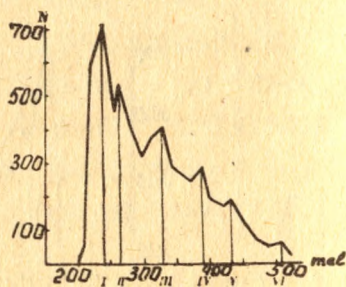


Рис.7

На них хорошо видно примерное равенство расстояний между КЧ и сами КЧ даже в тех случаях, когда последние не резко выражены. Однако, несмотря на различную степень контрастности, распределение частот в тексте в общем совпадает. Для определения характера распределения значений m при исходном квантовании мною был использован метод спрямления нормальной кривой, описанной в [10], стр.68-72/. Данные по распределению значений m для ряда текстов приведены на рис.8, где текст ОI обозначен треугольниками вершиной вниз, текст О4 - кружочками, текст О7 - кружочками с палочками, а текст О9 - квадратами. Масштаб и начало точки отсчета на оси абсцисс¹⁾ подбирался таким образом, чтобы кривые лучше всего совпадали. (Распределение значений m для текстов ОI-О4 было определено ранее вручную).

Рис.8 свидетельствует о том, что распределение значений m , близкое между собой для всех текстов, отличается от нормального, особенно среди значений m меньше среднего, для которых $\Psi_{2,1} \leq 0$, что говорит об отрицательном эксцессе.

Так как П.3 алгоритма предусматривает сравнение новых распределений с исходным и отбрасывание распределений, худших, чем исходное, возникает вопрос об определении степени контрастности распределений. В качестве предварительного критерия степени контрастности распределения мне кажется целесообразным предложить частное от деления количества значений m , приходящихся на КЧ ($\sum KЧ$), на общее количество значений m ($\sum m$), выражая его в процентах. Эта вели-

1) Соответствующий частотам, а не их логарифмам или медам.

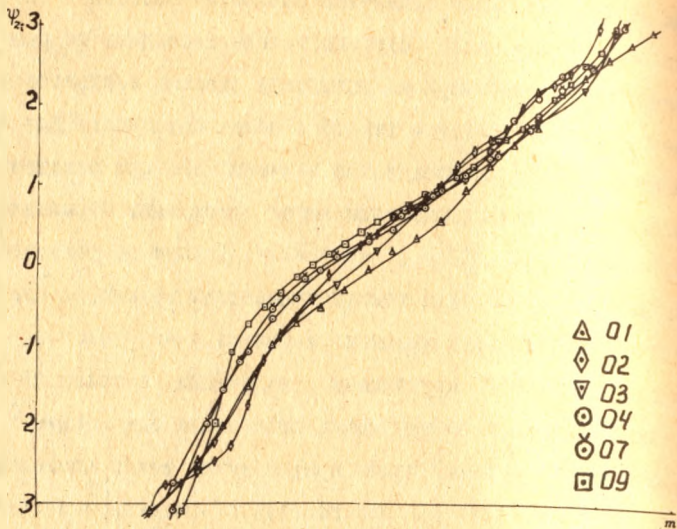


Рис.8

чина будет именоваться коэффициентом контрастности. Другой целесообразной величиной оценки является разность коэффициента контрастности и процента значений m , которые должны были бы приходиться на КЧ, если бы распределение было однородным. Все эти величины для обчисленных текстов, а также текстов 01, 02 и 04 (в исходном распределении текстов 03, 07 и 13 КЧ не выделяются), приведены в табл. II. Однако, разумеется, предложенных оценок степени контрастности далеко недостаточно: коэффициент контрастности не отражает степени стремления количества m , соседних с КЧ, к самой КЧ, а также особенностей распределения, которое отнюдь не является равномерным. Вопрос о построении достаточно надежной системы оценок контрастности еще предстоит решить, но и коэффициент контрастности позволяет сделать предварительные выводы о том, какое распределение из сравниваемых является лучшим.

Таблица II

Текст	Σm	КЧ	Коэффициент контрастности	Коэффициент контрастности при однородном распределении	Разность
01	2466	835	34%	17%	17%
02	2435	839	34%	18%	16%
04	2385	664	28%	15%	13%
09	6416	2072	32%	19%	13%
10	7873	2416	31%	19%	12%
11	252	86	34%	18%	16%
12	279	120	43%	30%	13%
19	6571	2164	33%	21%	12%

Данные о величине КЧ в m при исходном квантовании для всех изученных текстов, в которых они наблюдаются, приведены в табл. I2. С целью проверки и уточнения алгоритма были сняты алгоритмические ограничения пункта 4, и сбрасывание y на паузу производилось при всех значениях x и y . С самого начала был поставлен вопрос, влияют ли изменения x , y и M ¹⁾ сколь-нибудь существенно на распределение m , то есть получаются ли при этом различные распределения, иными словами, имеет ли эта операция какой-нибудь смысл. С этой

1) M - это величина паузы, при которой происходит выключение счетчика времени.

целью был использован критерий Колмогорова, при помощи которого сравнивались получившиеся распределения с исходным, рассматриваемым как теоретическое. Величину Λ критерия Колмогорова вместе с другими величинами см. в табл.13. Если учесть, что при принятии экспериментального распределения качестве теоретического больше возможность принятия разных распределений за одинаковые, чем наоборот, следует признавать, что операция задания x , y и M имеет смысл.

Таблица 12

КЧ \ Т	01	02	04	09	10	11	12	19
6	-	-	-	32	31	-	31	32
5	-	28	-	28	28	-	28	28
4	24	24	-	24	24	28	25	24
3	20	20	20	20	20	24	20	20
2	16	16	16	16	16	20	17	16
1	12	12	12	12	12	16	15	13

3.3.2. Для сравнения контрастности полученных распределений с контрастностью исходного можно воспользоваться как коэффициентом контрастности, так и некоторой модификацией критерия Колмогорова. Если в критерии Колмогорова заменить модуль разности между теоретической функцией распределения и соответствующей статистической функцией распределения на разность между теоретическими частотами p_i и статистическими частотами p_i^* , то отклонения каждого значения сравнивае-

распределения от исходного выразится величиной

$$\lambda' = (p_i - p_i^*) \sqrt{n},$$

n - величина сравниваемой выборки. Если сравниваемое распределение контрастнее исходного, то для $m = KЧ$ средняя величина $\lambda' > 0$, а для $m \neq KЧ$ средняя величина $\lambda' < 0$. Легко убедиться, что средняя величина λ' для $m = KЧ$ одного и того же распределения равна отрицательной величине λ' для $m \neq KЧ$ того же распределения.

Оба эти критерия (коэффициент контрастности и λ'), к сожалению, далеко не идеальны; о недостатках первого уже говорилось выше, а второй представлен лишь в тех случаях, когда $KЧ$ сравниваемого распределения совпадают с $KЧ$ исходного, что бывает далеко не всегда.

Значения этих величин для всех случаев квантования приведены в табл. I3, где в трех столбцах для каждого текста даны значения λ , коэффициента контрастности и λ' . Значения, помещенные в скобки, получены при делении на $KЧ$ по паузам до гласной части.

Данные табл. I3 свидетельствуют о том, что за исключением текстов I2 и I9 остальные дают максимальную контрастность при $M = MaI$ и в районе $x = 2$; текст I9 все же дает максимальную контрастность при $M = MaI$ (правда, $x = 3$), а у текста I2 крайне невелика выборка.

Все это позволяет сделать вывод, что время X приближается к 3/50 сек. и паузы $Ma2$ и MaI обе являются макросегментальными. Последнее обстоятельство показывает, как исходя из чисто структурных соображений, можно определить из них самих вещи, определяющиеся одна через другую.

11			12			13			19		
λ	к. к.	λ'	λ	к. к.	λ'	λ	к. к.	λ'	λ	к. к.	λ'
0,32	34	0,03	0,30	43	0,34	0,36	-	0,32	0,14	32	-0,14
0,32	35	-0,03	0,30	43	-0,33	0,36	-	-0,32	0,14	33	0,14
0,37	33	-0,33	0,62	45	0,83	0,25	-	-0,08	0,18	34	-0,27
0,48	39	-0,04	0,41	49	0,11	0,22	-	0,11	0,23	33	0,06
0,42	33	0,07	0,35	52	-0,93	0,20	-	-0,02	0,19	32	0,22
0,55	36	-0,21	0,41	36	0,72	0,11	-	-0,03	0,61	26	0,09
0,54	32	-0,16	0,53	34	0,06	0,34	-	0,11	0,45	32	0,45
0,48	29	0,36	0,40	38	0,02	0,60	-	0,48	0,58	35	-0,61
0,36	37	0,21	0,57	56	-0,21	0,31	48	-0,56	0,29	34	-0,27
0,35	34	0,22				0,15	-	0,29	0,16	33	-0,23
0,36	29	-0,22				0,15	-	-0,29	0,16	33	0,23
0,34	41	-0,25				0,32	-	-0,03	0,23	32	0,11
0,30	32	-0,12				0,32	-	0,77	0,31	33	0,33
0,42	33	0,38				0,31	48	-0,74	0,35	30	0,81
0,55	35	0,44				0,31	-	-0,34	0,33	34	-0,38
0,29	30	-0,43				0,30	-	-0,26	0,42	21	0,50
0,43	30	0,18				0,52	-	0,55	0,44	33	0,17
0,35	29	-0,18				0,29	-	-0,15	0,38	35	0,59
0,34	34	-0,10	0,21	46	-0,41	0,23	-	-0,28	0,39(0,42)	33(32)	0,09(0,10)
0,34	29	0,11	0,21	39	0,42	0,23	-	0,28	0,39(0,36)	33(33)	-0,09(-0,05)
0,36	40	-0,62	0,52	34	0,52	0,22	-	-0,22	0,30(0,31)	31(31)	0,00(0,47)
0,41	27	0,62	0,50	46	-0,27	0,25	-	-0,46	0,20(0,22)	33(33)	-0,18(0,01)
0,41	26	0,29	0,59	46	-0,25	0,40	-	0,72	0,33(0,30)	34(31)	-0,22(0,36)
0,52	34	0,59	0,33	47	-0,36	0,40	-	-0,37	0,43(0,38)	31(32)	0,39(0,05)
0,31	30	-0,64	0,40	45	0,29	0,48	-	-0,35	0,20(0,32)	33(33)	0,26(0,40)
0,29	37	0,09	0,31	42	-0,23	0,43	-	0,07	0,30(0,31)	33(33)	-0,41(0,16)
0,36	23	0,10	0,29	42	0,14	0,63	-	0,77	0,37(0,43)	33(32)	-0,48(-0,49)
0,34	26	0,00	0,21	46	-0,16	0,24	-	-0,30	0,44	33	0,23
0,34	35	-0,01	0,21	39	0,17	0,24	-	0,30	0,39	33	-0,19
0,50	29	-0,31	0,52	34	0,52	0,47	-	0,15	0,42	31	0,46
0,36	36	0,34	0,50	46	-0,27	0,29	-	-0,42	0,25	33	-0,20
0,37	-	-0,03	0,27	46	-0,25	0,28	-	0,23	0,33	31	-0,08
0,43	-	0,01	0,38	49	-0,53	0,19	-	-0,43	0,54	31	0,71
0,29	32	0,14	0,32	35	0,39	0,31	-	0,08	0,28	32	0,33
0,33	27	0,22	0,49	52	0,05	0,16	-	0,09	0,32	35	-0,44
0,33	31	-0,29	0,24	38	0,20	0,49	-	0,32	0,38	34	-0,39

x	y	07			09			10		
		λ	κ, κ	λ'	λ	κ, κ	λ'	λ	κ, κ	λ'
M=∞										
I	0	0,16	-	-0,32	0,65	33	0,22	0,33	3I	-0,04
I	I	0,16	-	-0,32	0,65	34	-0,22	0,33	3I	0,04
2	0	0,22	33	0,08	0,26	33	0,30	0,4I	3I	0,53
2	I	0,28	33	-0,04	0,4I	36	-0,68	0,23	3I	-0,4I
2	2	0,27	-	-0,05	0,29	33	0,40	0,43	32	-0,12
3	0	0,17	-	-0,25	0,49	34	-0,1I	0,37	3I	-0,19
3	I	0,45	-	0,10	0,87	34	-0,08	0,44	33	0,0I
3	2	0,20	27	-0,2I	0,72	32	0,42	0,29	3I	0,13
3	3	0,42	33	0,36	0,49	33	-0,22	0,49	3I	0,06
M=Na2										
I	0	0,23	-	-0,16	0,27	32	0,18	0,30	3I	0,10
I	I	0,23	-	0,16	0,27	32	-0,18	0,30	32	-0,1I
2	0	0,34	28	0,60	0,28	33	0,1I	0,48	3I	0,79
2	I	0,22	34	0,2I	0,30	33	-0,0I	0,37	3I	-0,14
2	2	0,24	37	-0,83	0,20	34	-0,08	0,35	32	-0,67
3	0	0,27	35	-0,38	0,44	32	0,03	0,49	3I	-0,08
3	I	0,27	34	-0,02	0,22	34	0,16	0,29	26	-0,49
3	2	0,25	-	0,15	0,38	32	0,22	0,33	30	0,23
3	3	0,4I	33	0,25	0,33	35	-0,42	0,49	32	0,35
M=NaI										
I	0	0,29	-	0,19	0,44(0,46)	33(33)	-0,55(-0,53)	0,46(0,3I)	25(3I)	-0,44(-0,29)
I	I	0,29	-	-0,18	0,44(0,46)	3I(3I)	0,55(0,53)	0,47(0,30)	3I(3I)	0,44(0,29)
2	0	0,89	27	0,38	0,25(0,27)	34(34)	0,04(0,28)	0,5I(0,47)	25(30)	0,99(0,77)
2	I	0,27	33	0,50	0,28(0,30)	35(30)	-0,10(-0,09)	0,52(0,43)	32(32)	-0,88(-0,34)
2	2	0,27	37	-0,88	0,16(0,17)	33(34)	0,06(-0,19)	0,30(0,39)	3I(3I)	-46(-0,32)
3	0	0,42	33	0,00	0,85(0,73)	33(34)	-0,62(-0,59)	0,5I(0,52)	3I(3I)	-0,13(-0,24)
3	I	0,18	34	-0,19	0,49(0,47)	33(33)	0,25(0,50)	0,55(0,59)	3I(3I)	0,00(0,2I)
3	2	0,38	34	0,26	0,55(0,44)	33(34)	-0,19(-0,16)	0,35(0,28)	25(25)	-0,52(-0,26)
3	3	0,30	-	-0,06	0,80(0,76)	33(35)	0,54(0,44)	0,34(0,40)	3I(3I)	0,49(0,40)
M=Ne										
I	0	0,27	33	-0,06	0,30(0,27)	34(35)	-0,42(-0,40)	0,27	24	-0,07
I	I	0,27	-	0,06	0,30(0,27)	33(33)	0,43(0,40)	0,29	3I	0,07
2	0	0,17	-	0,35	0,26(0,29)	32(32)	0,1I(0,25)	0,49	3I	0,5I
2	I	0,30	33	0,1I	0,22(0,26)	35(35)	-0,38(-0,35)	0,28	32	-0,3I
2	2	0,22	35	-0,48	0,23(0,20)	33(33)	0,27(0,09)	0,35	3I	-0,33
3	0	0,29	34	0,09	0,56(0,50)	33(33)	-0,14(-0,35)	0,69	3I	0,67
3	I	0,25	33	-0,06	0,34(0,43)	33(33)	0,15(0,17)	0,38	3I	-0,35
3	2	0,3I	36	-0,18	0,4I(0,39)	34(34)	-0,10(-0,22)	0,40	32	-0,64
3	3	0,26	-	0,13	0,80(0,83)	33(33)	0,44(0,39)	0,47	3I	0,47

В заключение рассмотрения КЧ следует подчеркнуть, что полученные данные еще не совсем точны, и таким образом еще не достигнута максимальная контрастность. Это связано с тем, что машина работала по одним частотным данным, не принимая во внимание деления потока речи на части, что приводит к возможности появления лишних частот и неточного учета КЧ при сбрасывании времени у .

3.4. ДАННЫЕ ПО ДЛИТЕЛЬНОСТИ

3.4.1. С помощью специальной программы были получены распределения длительностей всех слогов текстов 07, 09, 10, 11, 12, 13 и 19 с инициалами, без инициалей, с инициалами и без конечного согласного. Все эти данные здесь привести невозможно из-за недостатка места. Достаточно упомянуть, что они занесены в таблицы, где в столбцах записаны длительности в ПК, а в строках фонемы, причем имеются отдельные таблицы распределения по инициалам и по финалям, медиалам и тону. Каждая такая таблица (а их всего 56) имеет в среднем 26 столбцов и 35 строк. Кроме того, были получены итоговые значения, отражающие распределение фонем в тексте (в настоящей работе эти данные приведены не будут) и значения распределения для объединений, например, для всех инициалей, начинающихся с единицы (1х) и т.п. В связи с этим следует отметить, что инициали 33 тоже отнесены к 2х, а для финалей существует два различных х0: х0 без 30 (так как слог с финалью 30 всегда произносится нейтральным тоном) и х0₁ вместе с 30. Кроме того, были получены значения распределения для второго и шестого¹ тона

1) Тон, который кодируется шестеркой.

вместе (2 + 6), а также четвертого и восьмого¹⁾. На основании этих данных были высчитаны средние длительности для каждой строчки (то есть фонемы или их объединения).

3.4.2. Ввиду того, что длительности были определены с точностью до одного ПК, среднюю длительность не имело смысла определять с точностью, превосходящей 0,5 ПК. Но даже и такая точность чрезмерна. Следует считать более или менее достоверными лишь отклонения больше чем на единицу от целой величины. Кроме того, средняя величина считалась достоверной лишь в том случае, если она была найдена при суммировании десяти или более величин. Данные об отклонении средней величины длительности слогов с данной фонемой (или их объединением) от средней длительности для всего текста приведены в табл.14 - 17. Во всех этих таблицах в последней строчке даны средние длительности для всего текста (так, из этой величины можно получить среднюю длительность слогов с данной фонемой, прибавив к этой длительности величину отклонения), а в последних двух столбцах - мода (наиболее часто встречающееся значение) и средняя величина. Везде пустые места либо означают отсутствие в тексте данных фонем, либо присутствие их менее чем в 10 случаях. Табл.14 содержит отклонения длительностей слогов без инициалей по инициалам, табл.15 - соответственно слогов с инициалами по инициалам, табл.16 - соответственно слогов без инициалей по финалам и медиалам (медиали отделены горизонтальной линией и кодируются одной цифрой), а табл.17 - соответственно слогов с инициалами по финалам. Теперь нам предстоит, исходя из распределения этих отклонений, решить

1) Тон, который кодируется восьмеркой.

Таблица 14

	07	09	10	11	12	13	19	Мо	Ср.
11		-1,0	-0,5				-0,5	-0,5	-0,5
12							+0,5	-	-
13		-1,0	-0,5			0	0	-	-0,5
14			+0,5				-0,5	-	0
21	+1,0	-0,5	-0,5			-0,5	0	-0,5	0
22		-0,5	-0,5				-0,5	-0,5	-0,5
23			-1,0			+1,5	-1,5	-	-0,5
33	-1,5	-1,5	-0,5			0	0	-1,5	-1,0
41		+1,0	0			+0,5	0	-	+0,5
42			0			+0,5	+1,5	-	+1,0
44		+1,5	-1,0			-1,0	0	-1,0	0
51		0	-0,5			0	+1,0	-	0
52		-1,0	+0,5			+0,5	-2,0	+0,5	-0,5
54		+0,5	-0,5			-0,5	-0,5	-0,5	-0,5
61		0	0			0	-0,5	0	0
62		0	-0,5			+1,5	-0,5	-0,5	0
63		-1,0	-1,0					-1,0	-1,0
64	0	-3,0	-1,5			-0,5	-0,5	-0,5	-0,5
71		-1,0	+0,5			-1,0	-0,5	-1,0	-0,5
72		-1,5	-				+1,5	-	0
74		-2,0	-1,5			-2,0	-1,0	-2,0	-1,5
00	+2,5	+1,0	+1,5			+1,0	+1,5	+1,5	+1,5
1x	+1,5	-0,5	-0,5			0	0	-	0
2x	-0,5	-0,5	-0,5	0		0	-0,5	-0,5	-0,5
4x		+0,5	-0,5			0	+0,5	+0,5	0
5x	0	0	0	+1,0		0	0	0	0
6x	0	-0,5	-1,0			+0,5	-0,5	-0,5	-0,5
7x		-1,5	0			-1,0	-0,5	-	-1,0
x1	-0,5	0	0	+0,5	0	0	0	0	0
x2	+1,0	-0,5	0	-0,5		+0,5	0	-	0
x3	-0,5	-1,0	-0,5			+0,5	0	-0,5	-0,5
x4	0	-0,5	-1,0	-1,5		-1,0	-0,5	-1,0	-1,0
	9,5	10,5	9,5	7,5	8,5	6,0	9,0		

	07	09	10	11	12	13	19	Мо	Ср.
11		-1,5	-1,5				-2,0	-1,5	-1,5
12							+1,5	-	-
13		+1,0	+1,0			+1,0	0	+1,0	+1,0
14			+2,0				+0,5	-	+1,0
21	0	-1,5	-1,0			-1,5	-1,5	-1,5	-1,5
22		+1,5	+1,5				-0,5	+1,5	+1,0
23			-0,5			+0,5	-2,0	-	-1,0
33	-2,0	-0,5	+0,5			-0,5	0	-0,5	-0,5
41		0	-1,0			0	-1,5	-	-1,0
42			+1,5			+2,0	+2,5	-	+2,0
44		+2,5	+1,0			+0,5	+1,5	-	+1,0
51		+0,5	-0,5			0	+1,0	-	0
52		+2,0	+3,0			+2,0	-0,5	+2,0	+1,5
54		+2,0	+1,5			+1,0	+2,5	-	+1,5
61		0	-0,5			-0,5	-1,0	-0,5	-0,5
62		+1,5	+1,0			+3,5	+0,5	-	+1,5
63		0	-1,0					-	-0,5
64	+1,0	-0,5	+1,0			+1,0	+1,5	+1,0	+1,0
71		-0,5	+0,5			-1,0	-0,5	-0,5	-0,5
72		+2,0					+1,5	-	+1,5
74		+1,0	+1,0			-0,5	+0,5	+1,0	+0,5
00	0	-0,5	-0,5			-1,0	-1,0	-	-0,5
1x	+1,0	0	0	-0,5		+1,0	-0,5	+1,0	0
2x	-0,5	0	-0,5	0		-0,5	-0,5	-0,5	-0,5
4x		+1,0	+0,5			+0,5	+1,0	-	+0,5
5x	+1,5	+1,5	+1,0	+2,5		+0,5	+1,0	-	+1,5
6x	+0,5	0	0			+1,5	+0,5	+0,5	+1,0
7x		+0,5	+0,5			-0,5	0	+0,5	0
x1	-1,5	-0,5	-0,5	0	-0,5	-1,0	-1,0	-0,5	-1,0
x2	+1,5	+1,0	+1,5	+1,0		+2,0	+0,5	-	+1,0
x3	-1,0	0	0			+0,5	-0,5	-	0
x4	+1,5	+2,0	+1,0	+0,5		+0,5	+1,5	-	+1,0
	12,0	12,5	11,5	10,0	11,0	8,0	11,5		

Таблица 16

	07	08	10	11	12	13	19	Mo	Ср.
10	-0,5	+1,0	+0,5			+1,0	-1,5	+1,0	0
11		+0,5	+1,0			0	+1,5	-	+1,0
12	+2,0	+1,0	+1,0			+0,5	0	+1,0	+1,0
13	+3,5	+2,5	+2,0			+2,0	+3,0	+2,0	+2,5
14		+2,0	+2,0			+0,5	+3,0	+2,0	+2,0
20	-1,5	-2,0	-0,5	-1,0		-0,5	0	-0,5	-1,0
21		+1,5	+0,5				+0,5	+0,5	+1,0
22		0	+1,0			+0,5	0	-	+0,5
23		+1,0	+1,5			+0,5	+0,5	+0,5	+1,0
24		+1,0	+1,0			+1,0	+2,5	+1,0	+1,5
30		-2,0	-1,5			-3,5	-1,5	-1,5	-2,0
00	-2,0	-2,0	-2,0	-1,5		-1,0	-2,0	-2,0	-1,5
1x	+1,5	+1,5	+1,5	+1,0	0	+1,0	+1,0	-	+1,0
2x	+0,5	+0,5	+0,5	-0,5	0	+0,5	+0,5	+0,5	+0,5
x0	-1,0	-0,5	-1,5	-1,0	+0,5	-0,5	-0,5	-0,5	-0,5
x1		+1,0	+0,5			+0,5	+1,0	-	+0,5
x2	+2,0	+0,5	+1,0			+0,5	0	+0,5	+1,0
x3	+3,0	+1,5	+2,0	+2,5		+0,5	+2,0	+2,0	+2,0
x4	+3,0	+1,5	+1,5	+0,5		+1,0	+2,5	+1,5	+1,5
x0 ₁	-1,5	-1,0	-1,0	-1,5	0	-1,0	-0,5	-1,0	-1,0
0	-0,5	-0,5	-0,5	-0,5	-0,5	0	-0,5	-0,5	-0,5
1	+2,0	+0,5	0	+1,0		+0,5	0	+0,5	+0,5
2	-0,5	0	+0,5	-0,5	+0,5	+0,5	+1,0	+0,5	0
3		0	0			+0,5	+2,0	0	+0,5
	9,5	10,5	9,5	7,5	8,5	6,0	9,0		

Таблица 17

	07	09	10	11	12	13	19	Мо	Ср.
10	0	+1,5	+1,0			0	-1,5	0	0
11		+1,0	+0,5			-1,0	+1,5	-	+0,5
12	+1,0	+1,0	+0,5			+0,5	+0,5	+0,5	+0,5
13	+5,0	+2,5	+2,0	+1,5		+2,0	+2,5	-	+2,5
14		+2,5	+3,0			+1,5	+3,0	+3,0	+2,5
20	-2,0	-0,5	0	0		0	-0,5	0	-0,5
21		+1,5	0			0	0	0	+0,5
22		+0,5	+1,0			+0,5	0	+0,5	+0,5
23		+1,0	+1,0			+0,5	+0,5	-	+0,5
24		+1,5	+1,0			+1,5	+2,5	+1,5	+1,5
30		-2,5	-2,0			-3,0	-2,0	-2,0	-2,5
00	-1,5	-1,0	-1,5	-0,5		-0,5	-2,0	-	-1,0
1x	+2,0	+2,0	+1,5	+1,5	+1,0	+1,0	+1,5	+1,5	+1,5
2x	0	+0,5	+0,5	0	0	+0,5	+0,5	+0,5	+0,5
x0	-1,0	0	0	-0,5	+0,5	0	-0,5	-0,5	0
x1		+1,0	0			0	+1,0	-	+0,5
x2	+1,0	+0,5	+1,0			+0,5	+0,5	+0,5	+0,5
x3	+4,0	+2,0	+1,5	+1,0		+1,5	+1,5	+1,5	+2,0
x4	+2,5	+2,0	+1,5	+1,0		+1,5	+2,5	-	+2,0
x0 _I	-2,0	-0,5	-0,5	-1,0	0	-1,0	-1,0	-1,0	-1,0
0	0	0	0	0	-1,5	0	0	0	0
1	+1,5	+1,0	0	+1,0		+0,5	-0,5	+1,0	+0,5
2	-1,0	0	0	0	+1,0	0	+0,5	0	0
3		-0,5	-0,5				+1,0	-0,5	0
	12,0	12,5	11,5	10,0	11,0	8,0	11,5		

вопрос, какая длительность значима для языка: без инициалей или с инициалами?

В табл. I4 видно четкое и значительное положительное отклонение для нулевой инициали, а в табл. I6 такое же отрицательное отклонение для нулевой финали. Эти отклонения говорят о том, что длительность финалей слогов без инициалей больше, чем длительность финалей слогов с инициалами, то есть отсутствие инициали компенсируется за счет увеличения длительности финали, и наоборот - длительность гласной части слогов "без финалей" (то есть слогов, имеющих лишь медиаль или гласный $\dot{i}$) меньше длительности финалей слогов "с финалями". Таким образом, недостаток длительности финалей компенсируется за счет увеличения длительности инициалей. О компенсации этого рода свидетельствует и табл. I5 и I7, в которых вышеуказанные отклонения, однако, выражены значительно слабее.

Для окончательного решения вопроса о том, какая длительность воспринимается - всего слога или лишь его гласной части - , целесообразно прибегнуть к дополнительным критериям. Мы будем считать существенным для языка то распределение, которое дает наибольшую абсолютную величину отклонения и меньший разноразряд в знаках для разных текстов. Однако для того, чтобы эта оценка была полнее, следует изучить отклонения средних длительностей тонов. Эти данные приведены в табл. I8 (без инициалей) и в табл. I9 (с инициалами). Эти таблицы построены так же, как и предыдущие, с той лишь разницей, что они повернуты на 90° с целью экономии места. Для построения их были привлечены материалы ручной обработки текстов OI - O4. Из этих таблиц видно, что нулевой тон дает вполне определенное

Таблица 18

	0	1	2	3	4	6	8	2+6	4+8	
01	-2,0	+1,0	+0,5	+1,0	+0,5					7,0
02	-2,0	+0,5	0	+0,5	0					7,0
03	0	+0,5	+0,5	+1,5	+1,5					9,0
04	-3,0	+0,5	+0,5	+1,5	+0,5					9,0
07	-3,0	+1,0	+1,0	+1,5	-0,5					9,5
09	-0,5	+0,5	0	-0,5	+0,5	-2,5	-3,5	-0,5	0	10,5
10	-0,5	+0,5	+1,0	0	-0,5	-4,0	-2,5	+0,5	-0,5	9,5
11	-1,0	-1,0	+1,0	+0,5	0			+1,0	0	7,5
12	0	+0,5	0	-0,5	-0,5			0	-0,5	8,5
13	-1,0	+1,0	+1,0	+0,5	0			+1,0	0	6,0
19	-0,5	+0,5	+1,5	-0,5	0	-3,0		+1,0	0	9,0
Mo	-0,5	+0,5	+1,0	-	-	-	-	+1,0	-0,5	
Ср.	-1,0	+0,5	+0,5	+0,5	0	-3,0	-2,5	+0,5	0	

и достаточно большое отрицательное отклонение, а из четырех тонов самым длительным является второй тон, далее первый, третий и четвертый. Шестой и восьмой тоны дают резкое отрицательное отклонение, которое говорит в пользу предположения, что в связной речи они в действительности произносятся нейтральным тоном.

Данные, позволяющие окончательно решить, какие длительности являются значимыми, приведены в табл.20. Эта таблица состоит из трех частей: для инициалей, финалей и медиалей и

Таблица 19

	0	1	2	3	4	6	8	2+6	4+8	
01	-3,0	+1,0	+1,0	0	0					10,0
02	-2,5	+1,5	+1,0	+1,0	+0,5					9,0
03	-2,5	-0,5	+1,0	-0,5	+1,0					12,0
04	-3,0	+1,0	+2,0	+2,0	0					11,0
07	-4,0	+1,5	+1,5	+1,5	0			+1,5	0	12,0
09	-0,5	+1,0	+0,5	+0,5	+0,5	-4,0	-4,5	0	+0,5	12,5
10	-0,5	+1,5	+1,0	0	0	-4,5	-3,5	0	-0,5	11,5
11	-1,5	0	+2,5	+0,5	0			+2,5	0	10,0
12	-1,0	+0,5	+1,5	+0,5	-0,5			+1,5	-0,5	11,0
13	-2,0	+0,5	+1,5	+0,5	0			+1,5	0	8,0
19	-1,5	+0,5	+1,5	-0,5	0	-5,5		+1,0	0	11,5
Мо	-	-	-	+0,5	+0,5	-	-	+1,5	-0,5	
Ср.	-2,0	+1,0	+1,5	+0,5	0	-4,5	-4,0	+1,0	0	

тона. В первом столбце каждой части указаны фонемы, далее отклонения средней длительности без инициали и средней длительности с инициалью. Затем идет разность абсолютной величины отклонения средней длительности с инициалью и средней длительности без инициали (Δ). В столбце "знак", если в отличие от отклонений без инициалей все отклонения с инициалью имеют одинаковый знак, ставится +, если имеет место обратное, то ставится знак - . В последней строчке таблицы

Таблица 20

	б/ин	с/ин	Δ	знак		б/ин	с/ин	Δ	знак		б/ин	с/ин	Δ	знак
11	-0,5	-1,5	+1,0		10	0	0	0		0	-1,0	-2,0	+1,0	
12					11	+1,0	+0,5	-0,5	-					
13	-0,5	+1,0	+0,5		12	+1,0	+0,5	-0,5		1	+0,5	+1,0	+0,5	+
14	0	+1,0	+1,0	+	13	+2,5	+2,5	0						
					14	+2,0	+2,5	+0,5		2	+0,5	+1,5	+1,0	
21	0	-1,5	+1,5	+										
22	-0,5	+1,0	+0,5	-	20	-1,0	-0,5	-0,5		3	+0,5	+0,5	0	
23	-0,5	-1,0	+0,5		21	+1,0	+0,5	-0,5						
33	-1,0	-0,5	-0,5	-	22	+0,5	+0,5	0		4	0	0	0	+
					23	+1,0	+0,5	-0,5						
41	+0,5	-1,0	-0,5		24	+1,5	+1,5	0		6	-3,0	-4,5	+1,5	
42	+1,0	+2,0	+1,0											
44	0	+1,0	+1,0	+	30	-2,0	-2,5	+0,5		8	-2,5	-4,0	+1,5	
51	0	0	0		00	-1,5	-1,0	-0,5		2+6	+0,5	+1,0	+0,5	+
52	-0,5	+1,5	+1,0											
54	-0,5	+1,5	+1,0	+	1x	+1,0	+1,5	+0,5		4+8	0	0	0	-
					2x	+0,5	+0,5	0	+					
61	0	-0,5	+0,5											
62	0	+1,5	+1,5	+	x0	-0,5	0	-0,5						
63	-1,0	-0,5	-0,5		x1	+0,5	+0,5	0						
64	-0,5	+1,0	+0,5	-	x2	+1,0	+0,5	-0,5						
					x3	+2,0	+2,0	0						
71	-0,5	-0,5	0		x4	+1,5	+2,0	+0,5						
72	0	+1,5	+1,5	+										
74	-1,5	+0,5	-1,0	-	x0	-1,0	-1,0	0						
00	+1,5	-0,5	-1,0											
					0	-0,5	0	-0,5						
1x	0	0	0		1	+0,5	+0,5	0	-					
2x	-0,5	-0,5	0											
4x	0	+0,5	+0,5	+	2	0	0	0						
5x	0	+1,5	+1,5											
6x	-0,5	+1,0	+0,5	+	3	+0,5	0	-0,5	-					
7x	-1,0	0	-1,0	-										
x1	0	-1,0	+1,0	+										
x2	0	+1,0	+1,0	+										
x3	-0,5	0	-0,5											
x4	-1,0	+1,0	0											
Итого			+11,5	5+				-3,0	2-				+6,0	2+

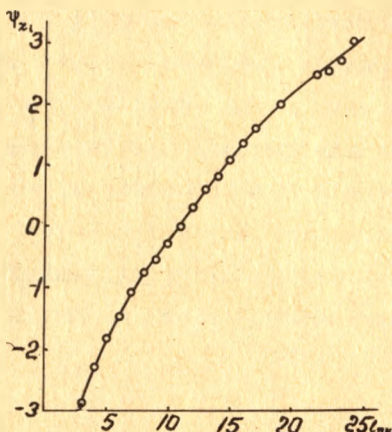
суммируются все значения Δ и все знаки. Из этой таблицы видно, что распределение длительностей по инициалам и тону отдает явное предпочтение длительностям целых слогов. Что же касается распределения по медиалам и финалям, то здесь наблюдается несколько иная картина. Но тут следует отметить, что величина Δ никогда не превышает 0,5, а отрицательный знак нам дают медиали, чье распределение вообще носит весьма случайный характер. Это явление можно объяснить тем, что и медиали и финали вместе выступают в качестве члена структурного противопоставления по отношению к инициалам. Но даже отбросив это соображение и подходя к вопросу совершенно формально, при сложении для всех трех частей таблицы мы получим +14,5 и 5+, что является убедительным свидетельством в пользу целых слогов.

Весьма интересно отметить, что в результате анализа распределения отклонений средних длительностей вполне естественно прийти к выводу, что длительность является в той же степени компонентной тона, как и фонетического состава. Действительно, длительности слогов в зависимости от тонов дают отклонения того же порядка, что и отклонения длительностей слогов в зависимости от медиалей и финалей. Это наблюдение естественнее всего было бы объяснить на основании предположения, выдвинутого в [8] (стр.55) о том, что "в связной речи часть признаков от тона отторгается; они манифестируют там уже не тон, а интонацию", однако это предположение следует еще проверить. Здесь лишь следует отметить, что резкое отклонение по длительности дают финали х3 и х4, а также нулевой тон.

3.4.3. В заключение рассмотрения вопроса о длительно-

стях следует остановиться на их распределении в чистом виде, то есть его форме и контрастности. Общее распределение длительностей для наибольшего по объему текста 10 было проверено методом спрямления нормальной кривой; данные этой проверки приведены на рис.9.

Рис.9



Этот рисунок свидетельствует о том, что распределение длительностей не имеет столь сложной формы, как распределение значений m , и отличается некоторой положительной асимметрией, которая вполне вероятно объясняется результатом смещения нескольких нормальных распределений. Гистограммы распределения длительностей для всех текстов с инициалами и без инициалей здесь приведены не будут; для примера приведем здесь только распределение длительностей с инициалами для текстов 10 и 19 (рис.10 и 11). Из всех рисунков было бы видно, что в большинстве случаев более контрастную картину дают распределения слогов с инициалами (это еще один аргумент в их пользу). Почти во всех случаях можно предполагать наличие трех контрастных областей, свидетельствующих о трех степенях сегментального ударения (об этом см. [5], [6]), но резко они выделяются далеко не во всех случаях.

Мне кажется целесообразным именно в этом месте на конкретном примере продемонстрировать основные теоретические проблемы, связанные с принципом контрастности.

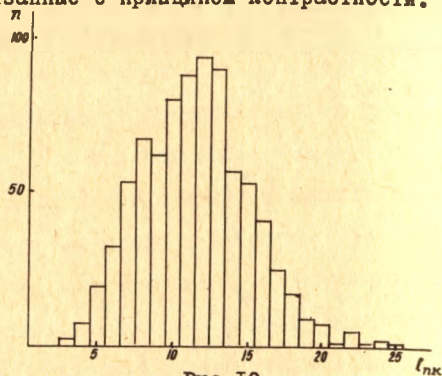


Рис. IО

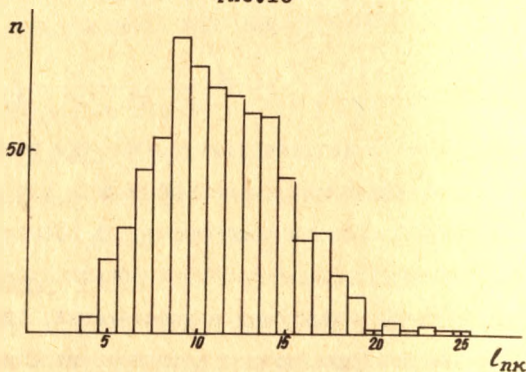


Рис. II

Принцип контрастности был сформулирован в [6] как обязательность противопоставления в области значений означающего при наличии противопоставления означаемых. В данном случае наличие трех контрастных областей подразумевает наличие трех степеней ударенности. Однако ввиду различной длительности слогов, связанной с их тоном и фонетическим

составом, показанной в § 3.4.2, эти контрастные области оказываются смещенными по отношению друг к другу. Это обстоятельство можно охарактеризовать как принцип относительности. В результате действия принципа относительности нарушается степень контрастности исходного распределения, показанного на вышеприведенных рисунках. Насколько возрастает контрастность распределения при рассмотрении слогов одного тона видно из рис. I2 и I3, где приведены данные о распределении длительности слогов нулевого тона для текстов I0 и I9 соответственно. Это распределение гораздо контрастнее распределения на рис. I0 и I1 и вполне определенно говорит о наличии трех контрастных областей. Но ведь и среди слогов нулевого тона

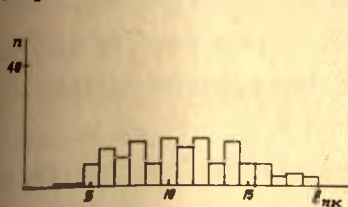


Рис. I2

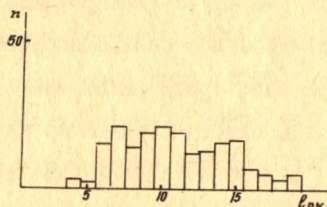


Рис. I3

были, например, слоги с финалями x_3 и x_4 . Следовательно, если бы мы могли их исключить, контрастность была бы еще больше, но при этом мы бы рисковали получить выборку недостаточного объема. Правомерно было бы получать данные для каждого слога определенного фонетического состава, но тогда почти во всех выборках будет меньше десяти наблюдений и ни о какой контрастности не может идти и речи. Это своего рода принцип лингвистической неопределенности: произведение степени контрастности на процент ошибки является константой.

Однако это утверждение не относится к выделению контрастных единиц шкалы. Контрастность этого рода предполагается заложенной в самой природе лингвистического материала, и на степень точности ее определения и выделения единиц ее шкалы влияет только степень точности эксперимента. На выделение КЧ сильно влияет перемещение начальной точки отсчета (об этом см. [5] и [6]); при определении КД на основании анализа распределения длительностей слогов мы сталкиваемся с несколько иной проблемой. Здесь, как мне кажется, недостаточной является точность определения границ слога, которая не позволяет нам определить его длительность точнее, чем до I ПК. Поэтому за исключением трех контрастных областей иную дискретность в распределениях установить трудно. Из этого можно сделать предварительный вывод, что единицами КД являются единицы одного порядка с ПК.

Теперь остается лишь провести окончательное разграничение между контрастными значениями означающего, имеющими соответствия для означаемых и не имеющими таковых. КД не имеет никакого соответствия для означаемых, определенное соответствие имеет лишь определенная комбинация КД; КД — это просто единица измерения, только не частоты, а длительности. Но КП, например, уже имеет совершенно определенное соответствие для означаемого — тип сегмента. Нетрудно видеть, что КП измеряются в единицах КД (вопрос о том, те же это единицы, что и для измерения длительности слогов, в данной связи не имеет значения). Таким образом мы окончательно формулируем положение о контрастных единицах измерения и контрастном выражении означаемого. Контрастные единицы измерения должны существо-

вать вне всякой связи с означаемыми, они являются минимальными единицами, атомами, из которых строится вся система языка. Эту иерархию можно показать на следующем примере: утверждение "длительность слога, имеющего такую-то степень сегментального ударения, составляет от 7 до 12 КД, что находится в соотношении контрастности к длительности слога, имеющего следующую степень ударения, составляющую от 18 до 22 КД" вполне может быть истинным.

4. ЗАКЛЮЧЕНИЕ

4.1. Необходимость машинной обработки.

Как уже говорилось единственным выходом из заколдованного круга является задание всех возможных параметров с целью получения наиболее контрастного распределения. Иначе это можно формулировать как перебор всех логических возможностей. До сих пор успех любого исследования определяли интуитивным выбором исследователем некоей логической возможности — первоначальной гипотезы, на которой строилось исследование. Если такая гипотеза была выбрана правильно — исследование давало положительные результаты, если нет — то ценность исследования была практически равна нулю, так как число логических возможностей достаточно велико и утверждение "надо делать не так" несет очень мало информации. Но даже удовлетворительные результаты исследования не свидетельствуют о том, что первоначальная гипотеза является единственно правильной, другая гипотеза может дать еще лучшие результаты. Даже в исследовании,

весьма ограниченном по объему, человек не может задать и тем более проверить все логические возможности. Это может выполнить лишь машина. Из этого и вытекает необходимость машинной обработки: точность и объективность исследования требует проверки всех логических возможностей, что может сделать лишь машина. Функции же человека в этом случае сводятся к тому, чтобы, во-первых, задать машине эти логические возможности (в настоящей работе такими логическими возможностями были различные значения параметров x , y , k и "величина макросегментальной паузы"), во-вторых, задать машине исходные данные с определенной, достаточно большой степенью точности, в третьих, разработать алгоритм исследования и перевести его на язык машины. Затем должны быть выработаны критерии принятия всей гипотезы в целом (в данном случае подтверждения того, что восприятие интонационной структуры речи в определенной степени соответствует модели, приведенной в § I.1), а также критерии принятия определенных значений параметров и прекращения дальнейших экспериментов.

Как видно из вышеизложенного, основная часть работы производится машиной, а главная задача исследователя (лингвиста) на первом этапе получить данные в таком виде, чтобы машина могла их обработать и предусмотреть логические возможности.

4.2. Практическое значение подобного рода исследования.

Целью настоящего исследования является отыскание путей автоматического ввода речи в машину, и его задачей, грубо говоря, является научить машину в нужных местах ставить зна-

ки препинания и обозначения тонов над соответствующими слогами. На самом же деле задача исследования несколько шире.

Ни у кого не вызывает сомнения, что в интонации содержится гораздо больше информации, чем в знаках препинания. С другой стороны, при автоматическом распознавании удавалось, например, распознать не более 80% фонем [7]. Можно предположить, что избыток информации в интонации поможет распознавать другие единицы фонетической структуры.

Этот вопрос можно изложить следующим образом. Избыточность информации в письменной форме языка позволила сократить до минимума набор и количество знаков препинания в тексте. Действительно, избыточность письменного языка очень велика. В то же время избыточность в устной форме языка вообще отсутствует или даже имеет отрицательное значение — мы пока еще не можем распознавать все фонемы. Таким образом, следует сделать вывод, что соотношение знаков препинания и букв весьма отличается от соотношения интонации и фонем.

Сказанное можно подытожить следующим утверждением: избыточность в одной части структуры компенсирует недостаточность (отрицательную избыточность) в другой ее части. Общей тенденцией развития языка следует признать стремление избыточности к нулю. Большая величина избыточности в его письменной форме говорит лишь о том, что письменная форма не является, строго говоря, одной из форм существования языка, а скорее определенным нарушением его структуры.

4.3. Значение фонетических исследований для формализации лингвистики

Формализация может начинаться лишь тогда, когда существуют дискретные единицы. Так, например, формализация фонологии началась лишь после введения дифференциальных признаков. Поэтому в основном формализуется лишь письменная форма языка, здесь уже дан готовый набор дискретных единиц: букв, пропусков, знаков препинания. Часто, также исследуются буквы при помощи дифференциальных признаков, но так как каждая буква однозначно записывается комбинацией дифференциальных признаков, исследования букв в чистом виде и в дифференциальных признаках практически однозначны и все они одинаково восходят к письменной форме языка.

Мне кажется, что изучение устной формы языка имеет большее значение по следующим причинам:

(1) письменную форму языка можно рассматривать как некоторое нарушение структуры языка;

(2) устная форма является основной формой существования языка;

(3) это единственная возможность полной формализации языка.

Формализация может быть полностью оправдана тогда, когда она начинается с самого начала, а не с середины, что происходит при формализации письменной формы языка. Действительно, при создании письменной формы языка буквы не выделялись путем чисто формальной процедуры. Формализация языка может производиться с двух концов — или со стороны семантики, или со стороны акустики. Для формализации со стороны се-

интики необходимо формальное определение всех видов значения, что, вероятно, не будет сделано еще долгое время. Остается лишь акустическая сторона. Здесь мы имеем физические величины, определяемые (или могущие определяться) приборами точно и однозначно, и именно отсюда должна начинаться формализация.

Представляется, что лингвистика станет ^(точной) наукой тогда, когда будут созданы две машины. Первая из них, слушая речь, записывала бы ее в знаках письменности данного языка (или метаязыка; последнее проще). Такая машина нужна, главным образом, для контроля и для увязки этих двух форм существования языка. Вторая машина должна, пользуясь результатами деятельности первой машины, выдавать идентичный текст на любом другом языке (или метаязыке, что опять же проще). Вся лингвистика как наука должна заключаться в устройстве этих двух машин. Основной вопрос заключается в том, на каком уровне формализация языка столкнется с семантикой. Моя работа является первым шагом определения устройства одного небольшого блока первой машины, сделанной в надежде, что ее функционирование не зависит от работы второй машины.

Трудно судить о связи этого блока с другими блоками машины. Если предположить, что процесс восприятия речи сводится к снятию спектров через время X , то, кроме печатания знаков препинания, этот блок будет иметь значение для синхронизации ее работы.

4.4. Преимущество акустических характеристик

Большинство современных исследований по автоматическому распознаванию речи строится на регистрации артикуляторных

характеристик (см.соответствующий обзор в [7] и самую эту работу). Мне представилось более целесообразным исходить из акустических характеристик по двум причинам:

(1) исходя из интересов формализации гораздо удобнее пользоваться акустическими характеристиками, так как артикуляторные в свою очередь требуют формализации ;

(2) как видно из неоднократно цитированной работы [7], ввод артикуляторных характеристик требует большого числа каналов, что чрезвычайно затрудняет работу машины, делая ее подчас вообще невозможной. Ввод же акустических признаков требует максимум двух каналов - один для интенсивности и один для частоты. Кроме того, оборудование для записи артикуляторных движений в значительной степени затрудняет процесс речи и делает ее неестественной. Не следует забывать и то обстоятельство, что преимущественному изучению акустических характеристик способствует уже разработанная в настоящее время теория их автоматического определения.

ЛИТЕРАТУРА

1. 常用汉字拼音表。北京, 1958.
2. 北京话轻声词汇。上海, 1957.
3. Mathews' Chinese-English Dictionary. Shanghai, 1931.
4. Т.П.Задоевко. Краткий очерк системы тонов современного китайского языка. Сб. "Вопросы китайской филологии", М., 1963.
5. А.М.Карапетьянц. Тон и интонация современного пекинского диалекта. М., 1962 (Курсовая работа).
6. А.М.Карапетьянц. Изучение тона и интонации пекинского диалекта. Сб. "Исследования по фонологии восточных языков" (в печати).
7. Речь. Артикуляция и восприятие. М.-Л., 1965.
8. М.К.Румянцев. Интонация предикативных и атрибутивных синтагм. Сб. "Спорные вопросы строя китайского языка". М., 1965, стр.19-61.
9. Е.Скучик. Теория акустики. Т.II. М., 1959.
10. В.Д.Урбах. Биометрические методы. М., 1964.

A FORMAL APPROACH TO STUDYING MANDARIN INTONATION

An algorithm, based upon the determination of contrasting units of intonational scale, is submitted for description of modern pekinese dialect. Such units are contrasting frequencies (CF) and contrasting durations (CD), special attention was paid to the reliability of CF estimation. The results of computer research of intonograms of 7 texts were considered, consisting of data on distribution of duration of pauses and syllables (as well as parts of syllables), and frequencies in the syllables. Three types of pauses were established: macrosegmental, mesosegmental and microsegmental. The measurements of frequencies under different conditions indicated that the number of CF for different texts ranges from 3 to 6, the time between two successive pitch perceptions being about $3/50$ sec, and the duration of macrosegmental pause irrelevant to the time organisation of the spoken chain. The distribution of durations confirmed that only the durations of whole syllables enter into the system of speech organisation. The syllables pronounced with the 2nd tone have the longest duration in the spoken chain, then come syllables with the 1st, 3rd and 4th tone respectively. The syllables of the neutral tone and the syllables with finals an and ang have striking deviation from the medial duration (the first type syllables are much shorter, while the second type syllables are much longer). The duration of a syllable in the spoken chain depends as much on its tone, as on its phonetical composition. The preliminary estimation of the smallest unit of time organisation of the spoken chain is near to $1/50$ sec.

КАПЛУН М.И.
ЛЭФ ИВЯ при МГУ

ИНТОНАЦИОННАЯ ДИФФЕРЕНЦИАЦИЯ ПРЕДИКАТИВНОЙ И АТТРИБУТИВНОЙ (ОПРЕДЕЛИТЕЛЬНОЙ) СВЯЗИ В СОВРЕМЕННОМ РУССКОМ ЯЗЫКЕ

Проблема предикативности является одной из актуальнейших проблем в языкознании. Объясняется это тем, что понятие предикативности теснейшим образом связано с понятиями "предложение" и "словосочетание", являющимися основными категориями синтаксиса, с определением структурно-семантических границ данных категорий.

Одним из основных способов выражения предикативности является интонация.

В данной статье излагаются результаты экспериментального исследования, предпринятого нами с целью выяснения механизма интонационной дифференциации предикативной и атрибутивной (определяющей) связи.

Анализируется лишь одна акустическая характеристика — частота колебаний основного тона, ее функция в рассматриваемой дифференциации.

Запись предикативных и определяющих синтагм¹⁾ была проведена вне контекста, так как необходимо было получить идеальные модели противопоставления. В контексте же вполне

1) Термин синтагма употребляется нами условно как наименование исследуемых синтаксических единиц. Фактически предикативное сочетание двух элементов представляет собой две синтагмы.

возможна интонационная нейтрализация, что для рассматриваемого эксперимента весьма нежелательно. Поэтому частота колебаний основного тона анализировалась лишь в тех синтагмах, которые интонационно достаточно четко различались, что было объективно подтверждено аудитивным анализом. Аудиторам были предложены интонации синтагм в последовательности: ОС – ПС.¹⁾ Пять аудиторов безошибочно определили по интонации тип связи во всех шестидесяти реализациях. Данные акустического анализа необходимы при выяснении роли частоты колебаний основного тона в данном противопоставлении.

Исследование частоты колебаний основного тона проводилось путем сопоставления интонации предикативных и определенных синтагм. В эксперимент включались предикативные и определительные синтагмы типа (косой-заяц – ПС) и (косой заяц – ОС). В устной реализации в синтагмах данного типа предикативная связь была выражена чисто интонационно; первый элемент синтагмы – субстантивированное прилагательное выступало в роли подлежащего (субъект высказывания), второй элемент функционировал как сказуемое (предикат высказывания).

Для исследования частоты колебаний основного тона в данном типе синтагм были подобраны следующие пары ПС и ОС:

1. (грозный – царь – ПС) и (грозный царь – ОС). Первый элемент синтагм является двусловным словом с ударением на первом слоге (ударный и заударный слоги). Второй элемент синтагм односложное слово.

1) ПС – предикативная синтагма, ОС – определительная синтагма.

2.(седой - старик - ПС) и (седой старик - ОС) ; первый элемент - двусложное слово с ударением на втором слоге (предударный и ударный слоги). Второй элемент двусложное слово с ударением на втором слоге.

3.(голодная - степь - ПС) и (голодная степь - ОС) ; первый элемент - четырехсложное слово с ударением на втором слоге (предударный, ударный, заударные слоги). Второй элемент - односложное слово.

Всего четырьмя дикторами было произнесено 15 пар синтагм, что в общей сложности дало 60 реализаций. Предикативные синтагмы были произнесены дикторами как относительно законченные двусоставные предложения с обычным фразовым ударением в форме простой констатации факта. Определительные синтагмы выступали как обычные номинативные образования.

Цель эксперимента заключалась в том, чтобы выяснить акустические параметры основного тона в предикативных и определительных синтагмах, соотнести эти акустические параметры с аудитивным восприятием и решить вопрос о их релевантной или иррелевантной функции при интонационной дифференциации рассматриваемых типов синтагм.

Как в предикативных, так и в определительных синтагмах (косой - заяц - ПС) и (косой заяц - ОС) второй элемент синтагм у всех четырех дикторов всегда реализовывался в более низкой полосе частот, нежели первый элемент, причем, частота основного тона второго элемента падает к концу звучания, что связано, видимо, с интонационным признаком утверждения, который является характерным для данного типа синтагм.

Данные о частотных характеристиках элементов у различных дикторов сведены в таблицу I.

Таблица I

Синтагма (седой - старик - ПС)	Частота колебаний основного тона, Гц			
	Первый элемент		Второй элемент	
	Начало	Конец	Начало	Конец
Диктор I	175	200	150	100
Диктор 2	110	175	110	100
Диктор 3	210	340	350	175
Диктор 4	250	375	250	225
(седой старик - ОС)				
Диктор I	175	210	175	100
Диктор 2	125	150	150	110
Диктор 3	240	350	350	225
Диктор 4	250	375	300	225

Перепад частот по конечным точкам звучания элементов ПС в музыкальных тонах равен:

Диктор I - октава

Диктор 2 - септима

Диктор 3 - б.секста

Диктор 4 - септима

В определительной синтагме (седой старик) наблюдается в общем менее резкий перепад частот по конечным точкам звучания элементов, за исключением данных диктора I.

Диктор I - м.сек. через октаву

Диктор 2 - б.терция

Диктор 3 - б.секста

Диктор 4 - м.секста

В предикативной синтагме (седой-старик) произнесенной диктором I, первый элемент реализовался в полосе частот 150-300 гц (частотный диапазон элемента равен октаве), а в определительной (седой старик) - в полосе частот 160-250 гц (частотный диапазон элемента равен м.сексте). Различие между частотными диапазонами первых элементов ПС и ОС равно четырем полутонам.

2-й эл.ПС: 190 - 100гц (септима)

2-й эл.ОС: 175- 100гц (м.септима)

Разница же между частотными диапазонами вторых элементов синтагм равна одному полутону.

Предударный слог первого элемента в предикативной синтагме (седой-старик) и в определительной синтагме (седой старик) реализовался приблизительно в одинаковой полосе частот: 175 - 150 гц в ПС и 175 - 160 гц - в ОС.

Движение частоты колебаний основного тона в ударном слоге первых элементов синтагм является фонологически значимым для интонационной дифференциации рассматриваемых типов связи.

Ударный слог первого элемента ПС у диктора I характеризуется восходяще-нисходящим движением основного тона. Начинается ударный слог с частоты в 225 гц. В дальнейшем в течение 100 мсек звучания частота колебаний основного тона повышается до 300 гц. В конце же слога тон падает до 200 гц. Восходящий интервал равен кварте (100 мсек звучания), а нисходящий — квинте (80 мсек звучания).

В определительной синтагме в ударном слоге также наблюдается восходяще-нисходящее движение частоты основного тона. Однако восходящий интервал равен только б.терции, а нисходящий — м.терции.

Предударный слог второго элемента предикативных и определительных синтагм характеризуется более высокой частотой основного тона по сравнению с ударным слогом.

Предударный слог второго элемента предикативной синтагмы начинается со 150 гц, и за 80 мсек звучания частота колебаний его тона повышается до 190 гц, что составляет б.секунду.

В ударном слоге происходит довольно резкое понижение частоты основного тона: на септиму за 200 мсек.

Для предударного слога второго элемента определительной синтагмы характерно понижение частоты основного тона (с интервалом в б.терцию), которое продолжается и на ударном слоге (б.терция), хотя в середине слога происходит не-

большое повышение, равное двум полутонам. В произнесении диктора I частотный диапазон предикативной синтагмы (седой-старик) равен квинте через октаву, а определительной синтагмы (седой старик) – терции через октаву.

Таким образом, у диктора I частотный диапазон предикативной синтагмы в целом и ее элементов значительно больший, чем определительной. Перепад частот в конце первого элемента и начале второго и контраст частот между высшими точками элементов в предикативной синтагме также больший, чем в определительной. Частотный контраст между элементами фиксируется как в предикативной синтагме, так и в определительной, первый элемент выше, второй – ниже, однако для предикации характерен более резкий частотный контраст, чем для номинации.

У диктора 2 в предикативной синтагме (седой-старик) и определительной (седой старик) также наблюдается частотный контраст между элементами, как и у диктора I, не в смысле совпадения частот элементов в данных типах связи, а в соблюдении общего принципа, когда частота колебаний основного тона в конце первого элемента большая, чем в начале второго.

Общий частотный диапазон предикативной синтагмы-септима, определительной – квинта. Частотный диапазон элементов предикативной синтагмы также больший, чем определительной.

ПС (седой-старик)

ОС (седой старик)

1-й эл. – м.септима

1-й эл. – м.терция

2-й эл. – ув.кварта

2-й эл. – б.терция

Слоговой гласный предударного слога первого элемента

ПС и ОС очень незначителен по времени, всего 40 мсек, и движение частоты основного тона в обеих синтагмах этого слога

характеризуется близкими по своему значению данными ПС - 110 гц, ОС 125 гц.

Если у диктора 1 в ударном слоге первого элемента ПС и ОС наблюдалось ярко выраженное восходяще-нисходящее движение основного тона, то у диктора 2 восходяще-нисходящую линию можно наблюдать только в предикативной синтагме и то с не-большим перепадом частот - 25 гц. В определительной синтагме ударный слог начинается со 140 гц, и этот уровень сохраняется до середины слога, затем происходит повышение тона до 150 гц, а к концу слога тон опять понижается до 140 гц.

В предикативной синтагме ударный слог начинается частотой колебаний в 150 гц, затем на коротком отрезке звучания (40 мсек) идет незначительное понижение тона (до 140 гц), после чего тон повышается до начального уровня (150 гц), где происходит его фиксация (ровное движение в течение 40 мсек). За периодом фиксации вновь происходит повышение тона (до 175 гц) с последующим ровным участком его движения (40 мсек) и спадом на 25 гц, а в последние 40 мсек звучания тон возвращается к своей высшей точке (175 гц).

Обращает внимание то, что у диктора 1 примерно с середины ударного слога начиналось равномерное понижение частоты основного тона, которое к концу звучания слога достигало интервала равного квинте, в то время как у диктора 2 конец звучания ударного слога находится на том же уровне, что и высшая точка слога. Заметим также, что повышение частоты основного тона в конце звучания ударного слога первого элемента ПС характерно и для других дикторов.

Движение тона в предударном и ударном слогах второго

элемента ОС нисходящее, двуступенчатое. После каждого падения тона наблюдается небольшой период фиксации (40 м сек).

В предикативной синтагме противоположная картина: для предударного слога в целом характерно восходяще-нисходящее движение основного тона. Начинается он с частоты в 110 гц и повышается до 140 гц (б.терция). Затем, в середине слога (на протяжении 60 м сек) движение основного тона ровное, на уровне 140 гц, а к концу звучания слога тон понижается до 125 гц (б.секунда).

Ударный слог второго элемента как в предикативной, так и в определительной синтагмах характеризуется понижением частоты основного тона: в ПС на ув.кварту, в ОС на б.терцию.

Анализ интонограмм предикативной синтагмы (седой-старик) и определительной (седой старик) показывает, что у диктора 3 в конфигурации изменений основного тона ОС и ПС больше общего, нежели у предыдущих дикторов.

Так предударный слог первого элемента ПС и ОС характеризуется восходяще-нисходящим движением основного тона. Различно лишь то, что этот контур реализуется в разных полосах частот. В предикативной синтагме предударный слог реализовался в полосе частот 200 - 240 гц (б.секунда). В определительной 240-260 гц (м.секунда). В ударном слоге первого элемента ПС происходит повышение частоты колебаний основного тона от 250 гц в начале слога до 340 в конце (кварты).

Повышение частоты основного тона в ударном слоге первого элемента определительной синтагмы происходит не так, как в предикативной. Если в предикативной синтагме повышение произошло за отрезок времени в 120 м сек, то в определительной только за 100 м сек.

Изменения в движении основного тона второго элемента исследуемых синтагм очень близки по общей конфигурации, хотя по времени звучания второй элемент предикативной синтагмы на 80 мсек короче определительной, и реализовались они в разных полосах частот: ПС 175-250 гц (ув. кварта)

ОС 200-350 гц (м. септима)

Предикативная синтагма (седой - старик) реализовалась в полосе частот 175-350 гц (октава), а определительная 200-375 гц (септима).

ПС

ОС

1-й эл. - м. секунда

1-й эл. - м. секста

2-й эл. - ув. кварта

2-й эл. - м. септима

В определительной синтагме (седой старик) у диктора 2 частота колебаний основного тона второго элемента ниже, чем первого, а у диктора 4 в данной синтагме схема распределения элементов по частоте колебаний основного тона нарушена. Наивысшая точка 1-го элемента - 375 гц, а второго - 400 гц. В предикативной синтагме схема распределения элементов синтагмы по частотным диапазонам остается прежней, т.е. первый элемент выше, второй ниже.

Для предударного слога первого элемента определительной синтагмы характерно восходяще-нисходящее движение основного тона с перепадом частот в 50 гц. В начале предударного слога первого элемента предикативной синтагмы наблюдается повышение частоты основного тона с 250 до 275 гц, а в середине и конце слога движение основного тона остается на этом же уровне.

В отличие от дикторов 1 и 2, у диктора 4 ударный слог первого элемента предикативной синтагмы начинается значительно

но ниже, чем предударный слог. (Ср. у диктора 1 перепад частот между предударным (ниже) и ударным (выше) слогами составил 6.терцию, у диктора 2 - кварту. У диктора 4 обратный перепад частот составил кварту).

В определительной синтагме у диктора 4 так же, как у дикторов 1 и 2 ударный слог первого элемента выше предударного. Для начала и середины ударного слога первого элемента ПС и ОС характерна почти идентичная конфигурация изменений основного тона, различие заключается лишь в том, что эта конфигурация реализуется в неодинаковой полосе частот в (ПС) 200-250 гц, а в (ОС) 290-360 гц. Частотный диапазон предикативной синтагмы 200-375 гц (6.септима), определительной - 225-400 гц (м.септима). Частотный диапазон элементов ПС и ОС:

ПС	ОС
1-й эл. - 6.септима	1-й эл. - квинта
2-й эл. - квинта	2-й эл. - м.септима

В ударном слоге первого элемента ПС повышение тона представляет собой непрерывное нарастание частот от некоторой точки в первой половине слога до его окончания.

В ударном слоге ОС повышение тона менее значительное (в ОС 275-360 гц, в ПС 200-375 гц) и распространяется оно не на все звучание слога, в конце звучания фиксируется участок почти ровного движения тона (60 мсек). Выше было отмечено, что начало второго элемента (60 мсек) ОС (предударный слог) выше по частоте колебаний основного тона конца первого элемента (ударный слог). Таким образом, у диктора 4 нарушается тональное взаимодействие между элементами син-

тагм, наблюдаемое у других дикторов, когда первый элемент был выше, второй – ниже.¹⁾

В начале предупредительного слога второго элемента ОС за 20 мсек происходит резкое повышение частоты основного тона с 300 до 400 гц (кварта), а в конце слога за отрезок времени равный также 20 мсек происходит понижение тона с 390 до 310 гц (б.терция). Движение основного тона в предупредительном слоге второго элемента предикативной синтагмы у диктора 4 почти идентично данным дикторов I и 2.

Для ударного слога второго элемента предикативной синтагмы характерно нисходяще-восходящее движение основного тона, а для определительной наоборот – восходяще-нисходящее движение.

В предикативных и определительных синтагмах для первого элемента, состоящего из двух слогов, последний из которых является ударным, характерно, в основном, восходящее движение основного тона, хотя у диктора I наблюдалось и восходяще-нисходящее движение. Однако, ни у одного из дикторов частота колебаний основного тона в конце ударного слога не поднималась выше частоты середины этого слога. Например, у дикторов I и 3 в первом элементе предикативной синтагмы ударный слог заканчивается частотой в 200 и 340 гц; высшая точка середины слога 300 и 350 гц.

1) Является ли такое нарушение вариантом по отношению к некоторой норме или это результат какого-то эмоционального наложения сказать пока трудно. Последнее констатировал один из аудиторов при слуховом анализе интонационных реализаций

В предикативной и определительной синтагмах наблюдается контраст частот между концом первого элемента и началом второго: конец первого элемента выше, начало второго - ниже. Между предикативно сочетающимися элементами отмеченный контраст частот значительно больше, чем в определительной синтагме, что, по-видимому, должно являться релевантным признаком предикации.

Диктор 1	Диктор 2
ОС - 6.секунда	ОС - 140-140 гц = 0
ПС - кварта	ПС - квинта
Диктор 3	Диктор 4
ОС - 350-350 гц = 0	ОС - м.секунда
ПС - кварта	ПС - квинта

Вторые элементы синтагм по общей конфигурации изменений основного тона близки друг другу у всех дикторов. В определительной синтагме для второго элемента характерно нисходящее движение основного тона, а в предикативной - восходяще-нисходящее движение, причем, обычно, нисходящий интервал значительно больше восходящего (ср. диктор 3, восходящий интервал - м.секунда, нисходящий - квинта).

Исследуя движение основного тона в синтагмах типа (седой - старик) ПС и (седой старик)ОС необходимо сравнить ударные слоги первых элементов синтагм, а также ударные слоги первых и вторых элементов внутри ПС и ОС.

В таблице 2 сопоставлены данные первого элемента ПС и первого элемента ОС по времени звучания и частотным диапозонам.

таблица 2

Синтагма	Первый элемент ПС			Первый элемент ОС		
	Время звуч. уд. слога в м сек.	Время звуч. слогооб-раз. гласн.	Част. диал. в гц и интервал	Время звуч. уд. сл. в мсек.	Время звуч. слогооб-раз. гласного интервал	Частот. диапазон в гц и интервал
Диктор 1	260	140	200-300 квинта	220	100	200-250 б. терция
Диктор 2	220	100	140-175 б. терция	180	80	140-150 м. секунда
Диктор 3	240	140	250-350 ув. кварта	200	100	250-375 квинта
Диктор 4	260	140	200-375 септима	200	100	275-375 кварта

Если сравнить ударные слоги элементов внутри синтагм, то по наивысшим точкам звучания ударный слог первого элемента по частоте колебаний основного тона будет всегда выше ударного слога второго элемента.

Диктор I

ПС седой-старик

ОС седой старик

уд.сл. I эл. - 300 гц	} б.тер- ция	уд.сл. I эл. - 250 гц	} октава
уд.сл. 2 эл. - 125 гц		уд.сл. 2 эл. - 125 гц	

через октаву

Диктор 2

ПС седой - старик

ОС седой старик

уд.сл. I эл. - 175 гц	} м.секста	уд.сл. I эл. - 150 гц	} м.тер- ция
уд.сл. 2 эл. - 110 гц		уд.сл. 2 эл. - 125 гц	

Диктор 3

ПС седой-старик

ОС седой старик

уд.сл. I эл. - 375 гц	} б.секста	уд.сл. I эл. - 375 гц	} квинта
уд.сл. 2 эл. - 275 гц		уд.сл. 2 эл. - 250 гц	

Диктор 4

ПС седой-старик

ОС седой старик

уд.сл. I эл. - 350 гц	} кварта	уд.сл. I эл. - 375 гц	} б.тер- ция
уд.сл. 2 эл. - 260 гц		уд.сл. I эл. - 300 гц	

Движение основного тона в синтагмах типа (седой-старик) ПС и (седой старик) ОС образует контраст частот между предударным и ударным слогами первого элемента. Сравнение слогов по перепаду частот между концом предударного и началом ударного слогов дает следующие результаты:

Диктор 1

ПС		ОС
предуд.сл. - 150 гц	} квинта	предудар.сл. - 160 гц
ударн.сл. - 225 гц		удар.сл. - 200 гц
		} б.тер- ция

Диктор 2

ПС		ОС
предуд. сл. - 110 гц	} кварта	предуд. сл. - 125 гц
ударный сл. - 150 гц		ударный сл. - 140 гц
		} б.секун- да

Диктор 3

ПС		ОС
предуд.сл. - 200 гц	} б.терция	предуд.сл. - 240 гц
ударн.сл. - 250 гц		удар.сл. - 250 гц
		} м.сек.

Диктор 4

У диктора 4 в ПС наблюдается обратный контраст частот между слогами первого элемента.¹⁾

ПС		ОС
предуд.сл. - 275 гц	} ув.кварта	предуд.сл. - 250 гц
ударн.сл. - 200 гц		удар.сл. - 290 гц
		} б.секунда

У всех дикторов, за исключением диктора 4, контраст частот между предударным и ударным слогами первого элемента в предикативной синтагме больший, чем в определительной.

1) Аналогичное распределение частот между слогами в двух-сложном слове, где ударным является второй слог, мы находим в работе Л.В.Златоустовой "Фонетическая структура слова в потоке речи". Казань, 1962, стр.7.

Контраст частот между предударным и ударным слогами первого элемента ПС был большим и тогда, когда сопоставлялись высшие точки частоты предударного и ударного слогов первого элемента синтагм.

Диктор 1

ПС

ОС

предуд.сл. - 175 гц	} б.секста	предударн.сл. - 175 гц	} ув. кварта
ударн.сл. - 300 гц		ударн.сл. - 250 гц	

Диктор 2

ПС

ОС

предуд.сл. - 110 гц	} м.секста	предуд.сл. - 125 гц	} б.се-кунда
ударн.сл. - 175 гц		ударн.сл. - 150 гц	

Диктор 3

ПС

ОС

предуд.сл. - 240 гц	} квинта	предуд.сл. - 260 гц	} ум. квинта
ударн.сл. - 350 гц		ударн.сл. - 275 гц	

Диктор 4

ПС

ОС

предуд.сл. - 275 гц	} кварта	предуд.сл. - 300 гц	} б.тер-ция
ударн.сл. - 375 гц		ударн.сл. - 375 гц	

В предикативной синтагме (седой-старик) и определительной (седой старик) второй элемент реализовался в большинстве случаев в более низкой полосе частот, нежели первый элемент, что объяснялось отчасти тем, что первый элемент синтагмы заканчивался ударным слогом, частота колебаний основного тона которого была значительно выше по сравнению с предударным слогом второго элемента синтагмы.

Сравнивая ПС(седой - старик)и ПС(грозный - царь), можно заметить, что ПС(грозный - царь)имеет ряд особенностей в движении основного тона. В предикативной синтагме (грозный-царь) нет того перепада частот между элементами синтагмы, который наблюдался в предыдущей синтагме,¹⁾ что, по-видимому, объясняется тем, что в данной синтагме ударный слог первого элемента начинает синтагму, реализуясь, в более высокой полосе частот, чем заударный слог, находящийся в конце элемента. Второй элемент синтагмы реализовался приблизительно в той же полосе частот, что и заударный слог первого элемента.

Если в предикативной синтагме(седой - старик)наблюдался прямой контраст частот между конечными точками, то в предикативной синтагме(грозный - царь)данный контраст наблюдается не всегда. Возможны бесконтрастные реализации, либо же реализации с обратным контрастом.

ПС (грозный - царь)

Диктор 1		Диктор 2	
I эл. - 100 гц	} 6.секунда	I эл. - 100 гц	} контраста не наблюдается
2 эл. - 125 гц		2 эл. - 100 гц	
Диктор 3		Диктор 4	
I эл. - 150 гц	} 6.терция	I эл. - 375 гц	} м.терция через октаву
2 эл. - 190 гц		2 эл. - 160 гц	

1) Имеется в виду контраст частот между элементами, образованный конечной точкой звучания первого элемента и начальной точкой второго элемента.

В определительной синтагме (грозный царь) прямой перепад частот между концом первого элемента синтагмы и началом второго наблюдается у всех четырех дикторов.

Диктор 1

1 эл. - 150 гц } м.секунда
2 эл. - 140 гц } да

Диктор 2

1 эл. - 150 гц } м.секунда
1 эл. - 140 гц }

Диктор 3

1 эл. - 310 гц } квинта
2 эл. - 210 гц }

Диктор 4

1 эл. - 400 гц } октава
2 эл. - 200 гц }

Если сравнить высшие точки в звучании элементов синтагм, т.е. ударные слоги в том и в другом элементе, то контраст частот между элементами будет как в предикативной, так и в определительной синтагмах.

Диктор 1

ПС

1 эл. - 275 гц }
2 эл. - 150 гц } септима

ОС

1 эл. - 210 гц }
2 эл. - 140 гц } квинта

Диктор 2

ПС

1 эл. - 190 гц }
2 эл. - 125 гц } м.секста

ОС

1 эл. - 175 гц }
2 эл. - 140 гц } б.секунда

Диктор 3

ПС

1 эл. - 390 гц }
2 эл. - 250 гц } м.секста

ОС

1 эл. - 350 гц }
2 эл. - 210 гц } б.секста

Диктор 4

ПС		ОС	
I эл. - 425 гц	} октава	I эл. - 425 гц	} септима
2 эл. - 210 гц		2 эл. - 225 гц	

У трех дикторов контрасты частот в предикативной синтагме больше, чем в определительной. Исключение составляют данные диктора 3, у которого контраст частот между высшими точками элементов определительной синтагмы оказался на один полутон больше, чем в предикативной. (У диктора I на восемь полутонов, у диктора 2 на шесть, у диктора 4 на два полутона)

Предикативная синтагма (грозный царь) в произнесении диктора I реализовалась в полосе частот 100-275 гц (кварта через октаву), а определительная (грозный царь) - 110-210 гц (септима).

ПС	ОС
I-й эл. - 100-275 гц (кварта через октаву)	I-й эл. - 110-210 гц (септима)
2-й эл. - 110-150 гц (кварта)	2-й эл. - 110-140 гц (б.терция)

Таким образом, частотный диапазон предикативной синтагмы в целом и ее элементов значительно больше, чем определительной.

Ударный слог первого элемента предикативной и определительной синтагм характеризуется восходящим движением основного тона, но в конце слога наблюдается небольшое понижение, большее - в определительной, меньшее - в предикативной синтагме.

Начинается ударный слог первого элемента предикативной синтагмы с частотой в 140 гц и после небольшого начального понижения, равного двум полутонам, происходит резкое повышение частоты основного тона с 120 до 275 гц; к концу слога тон понижается до 260 гц.

В ударном слоге первого элемента определительной синтагмы движение основного тона почти аналогично предикативной, хотя в определительной синтагме увеличение частоты основного тона меньше (с 110 до 210 гц), а в конце слога нисходящее движение значительно больше, чем в предикативной. (Ср. в ПС нисходящий интервал - м.секунда, а в ОС - ув.кварта).

Заударному слогу первого элемента предикативной и определительной синтагм присуще нисходящее движение основного тона, в ОС - с 160-150 гц, в ПС - с 150 гц до 100 гц, однако в предикативной синтагме заударный слог реализовался в более низкой полосе частот, чем в определительной. Резкое падение тона на заударном слоге первого элемента предикативной синтагмы, по сравнению с определительной, происходит скорее всего оттого, что ударный слог ПС реализовался в более высокой полосе частот, чем в ОС.

Вторые элементы синтагм по общему рисунку изменений основного тона очень близки и реализовались приблизительно в одинаковой полосе частот: ПС - 110-150 гц (кварта), ОС - 110-140 гц (б.терция).

В движении основного тона в ПС(Грозный - царь) и ОС(Грозный царь) наблюдаются общие закономерности, присущие реализациям различных дикторов, в данном случае дикторов I и 2, но не в плане совпадения частот, а в распределении частот-

ных контрастов между элементами синтагм и между слогами в элементах.

Частотный диапазон предикативной синтагмы (грозный - царь) - диктор 2 - септима, определительной (грозный царь) - м.септима (как и у диктора I частотный диапазон ПС больше, чем в ОС). Элементы синтагмы реализовались в полосе частот:

ПС	ОС
1-й эл. - 100-190 гц (септима)	1 -й эл.-125-175 гц (ув.кварта)
2-й эл. - 100-125 гц (б.терция)	2-й эл.-100-140 гц (ув.кварта)

(В отличие от диктора I у диктора 2 только частотный диапазон второго элемента ОС оказался на два полутона больше второго элемента ПС).

Ударный слог первого элемента синтагмы характеризуется почти идентичной общей конфигурацией изменений основного тона, различие состоит в том, что частотный диапазон ударного слога предикативной синтагмы на четыре полутона больше, чем в определительной. Как и у диктора I в конце ударного слога наблюдается падение частоты основного тона в ПС с 190 до 160 гц (м.терция), в ОС с 175 до 150 гц (б.терция).

В заударном слоге первого элемента предикативной синтагмы, который реализовался в более низкой полосе частот, по сравнению с ударным слогом, происходит понижение основного тона с 125 гц до 100 гц (б.терция).

Заударный слог первого элемента определительной синтагмы реализовался в полосе частот 140-150 гц (м.секунда).

(Ср. с заударным слогом ПС - 150-100 гц); для него также характерно некоторое начальное повышение частоты, в середине же и в конце - ровное движение тона (150 гц).

Для второго элемента определительной синтагмы характерно нисходяще-восходящее движение основного тона в начале и середине и резкое понижение в конце слова, равное кварте. Нисходящее движение наблюдается также и в конце второго элемента предикативной синтагмы, но на три полутона ниже, чем в определительной. Звучание вторых элементов ПС и ОС заканчивается частотой колебаний в 100 гц.

Предикативная синтагма (грозный - царь) и определительная (грозный царь) у диктора 3 по своим частотным характеристикам в принципе мало чем отличается от реализаций предыдущих двух дикторов.

Предикативная синтагма реализовалась в полосе частот 140-390 гц (ув.кварта через октаву), определительная 190-350 гц (септима).

ПС

1-й эл. - 140-390 гц (ув.кварта через октаву);

2-й эл. - 160-250 гц (квинта)

ОС

1-й эл. - 200-350 гц (м.септима);

2-й эл. - 190-210 гц (б.секунда).

Частотный диапазон предикативной синтагмы в целом и ее элементов у диктора 3 как и у двух предыдущих дикторов больший, чем в определительной.

Для ударного слога первого элемента предикативной и определительной синтагм характерно равномерное повышение частоты основного тона в ПС с 225 до 390 гц, в ОС с 200 до 350 гц. В конце слога, как и у дикторов 1 и 2, в предикативной синтагме происходит небольшое понижение тона с 390 до 375 гц (м.секунда).

Движение основного тона в начале и в середине заударного слога является продолжением движения основного тона ударного слога, который заканчивается на уровне 350 гц. В конце заударного слога происходит понижение основного тона (на б.терцию).

Заударный слог первого элемента предикативной синтагмы реализовался в более низкой полосе частот по сравнению с определительной. (Ср. ОС - 310-350 гц; ПС - 140-200 гц). Этот слог характеризуется нисходящим движением основного тона, интервал которого в середине и конце слога равен кварте.

Вторые элементы синтагм несколько различаются по общей конфигурации изменений основного тона. Если для второго элемента определительной синтагмы характерно ровное движение основного тона с изменениями вверх и вниз не более, чем на 10 гц, то в начале и в конце второго элемента предикативной синтагмы происходит понижение основного тона с 250 до 160 гц, а в конце повышение на б.терцию.

Акустические характеристики предикативной синтагмы (грозный-царь) и определительной (грозный царь) у диктора 4 отличаются от характеристик предыдущих дикторов. Основное различие состоит в том, что у дикторов 1, 2 и 3 ударные слоги первых элементов синтагм обычно реализовались в более высокой полосе частот, чем заударный слог. У диктора 4 обратная картина: ударный слог первого элемента ПС реализовался в полосе частот 190-225 гц, а заударный 250-425 гц. Это можно квалифицировать как построение с обратным слоговым контрастом. Однако при аудитивном анализе пять аудиторов зафиксировали в данной синтагме предикативную связь. По-видимому,

предикатива создается здесь за счет большего времени звучания первого элемента ПС чем ОС, особенно в гласной части. Например, время звучания первого элемента ПС (грозный - царь) = 340 мсек; первого элемента ОС (грозный царь) - 500 мсек, причем время звучания ударного "о" в ПС = 180 мсек, в ОС = 140 мсек; в ПС (седой - старик) звучание первого элемента = 420 мсек (звук "о" в ударном слоге = 140 мсек). В ОС "седой старик" звучание первого элемента = 340 мсек (звук "о" = 100 мсек).

Аналогичные данные получены в ПС (голодная - степь) и ОС (голодная степь).

Предикативная синтагма (грозный-царь) у диктора 4 реализовалась в полосе частот 160-425 гц; интервал 6.кварта через октаву.

ПС	ОС
1-й эл. - 190-425 гц (6.секунда через октаву)	1-й эл. - 250-425 гц (6.секста)
2-й эл. - 160-210 гц (кварта)	2-й эл. - 175-225 гц (6.терция)

Ударный слог первого элемента предикативной синтагмы у диктора 4 характеризуется нисходящим движением основного тона (Ср.с данными дикторов 1,2,3). Реализовался он в полосе частот 225-190 гц, а ударному слогу первого элемента определенной синтагмы присуще восходящее движение основного тона. Этот слог реализовался в более высокой полосе частот по сравнению с тем же слогом в предикативной синтагме (250-325 гц). У предыдущих дикторов ударный слог первого элемента ПС реализовался в более высокой полосе частот, чем ОС.

Заударный слог первого элемента предикативной синтагмы начинается с частоты в 260 гц (самая высокая точка ударного слога 225 гц). Для заударного слога характерно восходящее движение основного тона с 260 до 425 гц. В конце происходит некоторое понижение тона.

Если сравнить любые точки ударного и заударного слогов в ПС, т.е. высшие точки, перепад частот между концом ударного и началом заударного слогов, соотнести их частотные диапазоны – все равно у диктора 4 получится обратный слоговой контраст частот: заударный слог выше, ударный ниже.

Второй элемент предикативной синтагмы реализовался в более низкой полосе частот, чем второй элемент определительной синтагмы (ПС- 160-210 гц, ОС 175-225 гц).

В движении основного тона первого элемента предикативной синтагмы (седой – старик) и определительной (седой старик), фиксируется контраст частот между предударным и ударным слогами первого элемента как в ПС, так и в ОС, только в ПС частотный контраст больший.

В определительной синтагме (грозный царь) перепад частот в конце ударного и начале заударного слогов первого элемента весьма незначителен, иногда его совсем нет; не наблюдается также обратный перепад частот, в то время как в предикативной синтагме (грозный – царь) такой перепад частот отмечается у трех дикторов из четырех.

Диктор I

ПС		ОС
ударн.сл. – 260 гц	м.септима	ударн.сл. – 150 гц (м.секунда)
заударн.сл. – 150 гц		заударн.сл. – 160 гц (обратный контраст частот)

Диктор 2

ПС		ОС
ударн.сл. - 160 гц	ум.квинта	ударн.сл. - 150 гц м.секунда
заударн.сл. - 110 гц		заударн.сл. - 140 гц

Диктор 3

ПС		ОС
ударн.сл. - 375 гц	септима	ударн.сл. - 350 гц } 0
заударн.сл. - 200 гц		заударн.сл. - 350 гц

Сравнивая высшие точки частот ударного и заударного слогов первого элемента предикативной синтагмы также обнаруживаем значительный частотный контраст, в то время как в определительной синтагме этот контраст фиксируется не всегда, а если таковой и отмечается, то значительно меньший, чем в предикативной.

Диктор I

ПС		ОС
ударн.сл. - 275 гц	септима	ударн.сл. - 210 гц } 0
заударн.сл. - 150 гц		заударн.сл. - 175 гц

Диктор 2

ПС		ОС
ударн.сл. - 190 гц	квинта	ударн.сл. - 175 гц } м.терция
заударн.сл. - 125 гц		заударн.сл. - 150 гц

Диктор 3

ПС		ОС
ударн.сл. - 390 гц	септима	ударн.сл. - 350 гц } 0
заударн.сл. - 200 гц		заударн.сл. - 350 гц

В предикативной синтагме (голодная-степь) и определительной (голодная степь) первый элемент характеризуется восходяще-нисходящим движением основного тона, что, видимо, объясняется тем, что частота основного тона в ударном слоге больше, чем в предударном и заударном слогах элемента. Конец первого и начало второго элемента этой синтагмы образуют обратный перепад частот, т.е. конец первого элемента ниже начала второго. (Ср. с ПС седой - старик). Обратный перепад частот между элементами можно, по-видимому, объяснить резким падением частоты основного тона на заударном слоге после увеличения частоты на ударном слоге первого элемента. Обратный перепад частот в определительной синтагме (голодная степь) наблюдается только у диктора I.

Диктор I

ПС		ОС
I-й эл. - 100 гц	квинта	I-й эл. - 100 гц
2-й эл. - 150 гц		2-й эл. - 150 гц
		квинта

Диктор 2

ПС		ОС
I-й эл. - 100 гц	м.секунда	I-й эл. - 140 гц
2-й эл. - 110 гц		2-й эл. - 125 гц
		б.секунда

Диктор 3

ПС		ОС
I-й эл. - 190 гц	б.секунда	I-й эл. - 350 гц
2-й эл. - 210 гц		2-й эл. - 260 гц
		кварта

Диктор 4

ПС

ОС

1-й эл. - 400 гц
 2-й эл. - 240 гц

б.секста

1-й эл. - 360 гц
 2-й эл. - 200 гц

септима

Нетрудно заметить, что во всех рассмотренных выше предикативных и определительных синтагмах в произнесении диктора 4 частота основного тона в конце первого элемента оканчивалась значительно выше, чем в начале и середине, независимо от того, каким слогом заканчивается первый элемент — ударным или заударным.

Очевидно, этим объясняется перепад частот между элементами предикативной синтагмы (**голодная — степь**), когда первый элемент выше, второй — ниже.

В предикативной синтагме (**голодная — степь**) и определительной (**голодная степь**) как в синтагмах типа (**седой — старик**) и (**грозный — царь**) наблюдается контраст частот между высшими точками элементов синтагм.

Диктор I

ПС

ОС

1-й эл. - 300 гц
 2-й эл. - 150 гц

октава

1-й эл. - 225 гц
 2-й эл. - 150 гц

квинта

Диктор 2

ПС

ОС

1-й эл. - 175 гц
 2-й эл. - 110 гц

м.секста

1-й эл. - 150 гц
 2-й эл. - 125 гц

б.секунда

Диктор 3

ПС		ОС
1-й эл. - 390 гц	септима	1-й эл. - 350 гц
2-й эл. - 210 гц		2-й эл. - 260 гц
		кварта

Диктор 4

ПС		ОС
1-й эл. - 400 гц	м.секста	1-й эл. - 400 гц
2-й эл. - 250 гц		2-й эл. - 250 гц
		м.секста

(Только у диктора 4 частотный контраст в предикативной и определительной синтагмах одинаковый, у остальных трех дикторов частотный контраст в ПС больше, чем в ОС)

Так же, как в предыдущих синтагмах, частотный диапазон предикативной синтагмы больше, чем определительной. (См. таб. 3-4)

Таблица 3

Синтагма	Диктор	Частотный диапазон (интервал)
(голодная - степь) ПС	1	квинта через октаву
	2	м.септима
	3	кварта через октаву
	4	м.секунда через октаву
(голодная степь) ОС	1	б.секунда через октаву
	2	квинта
	3	б.секста
	4	октава

Частотный диапазон первого элемента ПС также больше, чем в ОС.

Синтагма	Частотный диапазон I-го элемента в Гц	Интервал	Частотный диапазон 2-го элемента	Интервал
(голодная - степь) (ПС)	дик. I : 100-300	квинта через октаву	125 - 150	б. терция
	дик. 2 : 100-175	м. септима	100 - 110	м. секунда
	дик. 3 : 150-390	кварта через октаву	200 - 210	м. секунда
	дик. 4 : 190-400	м. сек через октаву	200 - 250	б. терция
(голодная степь) (ОС)	дик. I : 100-225	б. сек. через октаву	125 - 150	б. терция
	дик. 2 : 140-150	м. секунда	100 - 125	б. терция
	дик. 3 : 210-350	б. секста	210 - 260	б. терция
	дик. 4 : 225-400	м. септима	200 - 250	б. терция

Если в предыдущих синтагмах частотный диапазон второго элемента ПС чаще всего больше, чем в ОС, то в данной синтагме одна часть реализаций характеризуется равными диапазонами, другая - большими в ОС.

В предикативной синтагме (голодная - степь) в отличие от определительной (голодная степь) всегда наблюдается перепад частот в конце (**ударного (выше)**) и начале заударного (**ниже**) слогов первого элемента.

Диктор 1

ПС		ОС
ударн.сл. - 300 гц	октава	ударн.сл. - 210 гц
заударн.сл. - 150 гц		заударн.сл. - 150 гц
		кварта

ПС	Диктор 2	ОС
ударн.сл. - 175 гц	м.терция	ударн.сл. - 150 гц
заударн.сл. - 150 гц		заударн.сл. - 140 гц
		м.секунда

Диктор 3

ПС		ОС
ударн.сл. - 350 гц	кварта	ударн.сл. - 300 гц
заударн.сл. - 260 гц		заударн.сл. - 325 гц
		м.секунда
		обратный перепад частот

Диктор 4

ударн.сл. - 190 гц	м.секунда	ударн.сл. - 260 гц	б.терция
заударн.сл. - 200 гц	обратный перепад частот	заударн.сл. - 325 гц	обратный перепад частот

Эти данные говорят о том, что в тех случаях, когда и предикативная, и определительная синтагмы имеют прямой перепад

частот между элементами, то в предикативной синтагме он **значительнее**, чем в определительной.

Если сравнить высшие точки частот ударного и заударного слогов первого элемента синтагм, то в предикативной синтагме высшая точка частоты ударного слога у трех дикторов выше, чем в заударном слоге, в (ОС) такой закономерности не наблюдалось.

Диктор 1

ПС		ОС	
ударн.сл. - 300 гц	} октава	ударн.сл. - 225 гц	} м.терция
заударн.сл. - 150 гц		заударн.сл. - 190 гц	

Диктор 2

ПС		ОС	
ударн.сл. - 175 гц	} м.терция	ударн.сл. - 150 гц	} 0
заударн.сл. - 150 гц		заударн.сл. - 150 гц	

Диктор 3

ПС		ОС	
ударн.сл. - 390 гц	} квинта	ударн.сл. - 325 гц	} м.секунда
заударн.сл. - 260 гц		заударн.сл. - 350 гц	
			} обратный контраст частот

Диктор 4

ПС		ОС	
ударн.сл. - 275 гц	} квинта	ударн.сл. - 260 гц	} м.секста
заударн.сл. - 400 гц		заударн.сл. - 400 гц	

/Обратный контраст частот/

В синтагмах (**голодная-степь**) ПС и (**голодная степь**) ОС как и в предыдущих синтагмах общий контур изменений основного тона идентичен у дикторов 1,2,3.

Для предударного слога первого элемента ПС и ОС характерно ровное движение основного тона, иногда с небольшим повышением в конце слога равным 10-15 гц.

Для ударного слога первого элемента ПС и ОС характерно восходящее движение основного тона, причем, в конце этого слога находится высшая точка частоты не только слога, но и всей синтагмы. Ударный слог первого элемента ПС обычно реализуется в более высокой полосе частот, чем в ОС. Изменения основного тона в заударном слоге первого элемента не у всех дикторов одинаковы.

Заударный слог первого элемента ПС (дикторы 1,2,3) характеризуется нисходящим движением основного тона, а в ОС у диктора 2 в этом слоге ровное движение основного тона, с небольшим повышением и последующим понижением в середине слога. У диктора 3 заударный слог реализовался с небольшим повышением частоты основного тона (с 325 до 350 гц):

Необходимо отметить, что заударный слог первого элемента предикативной синтагмы у дикторов 1,2,3 реализовался в более низкой полосе частот, чем в определительной.

Анализ частоты основного тона позволил сделать следующие выводы:

1. Частотный интервал в ударном слоге первого элемента ПС больше, чем в ударном слоге первого элемента ОС.

2. Частотный диапазон предикативной синтагмы в целом, и ее элементов почти всегда больше, чем определительной.

3. Перепад частот между элементами в целом, а также частотный контраст между высшими точками элементов в предикативной синтагме чаще всего больше, чем в определительной.

4. В определительных синтагмах не всегда наблюдается контраст частот в первом элементе между предударным и ударными слогами (в синтагмах типа седой старик) и между ударным и заударным слогами (в синтагмах типа грозный царь).

Во время аудитивного анализа, предшествовавшего акустическому (см. стр. 276) аудиторам было предложено прослушать и определить тип связи в 15 парах синтагм типа (косой-заяц) ПС и (косой заяц) ОС в произнесении четырех дикторов. Все аудиторы в 60 реализациях правильно определили тип связи, тем самым четко дифференцируя акустические характеристики предикации и номинации. Это было исходным пунктом для постановки вопроса об акустической основе данной дифференциации. Подробный акустический анализ, данный нами выше, преследовал цель обнаружить в акустической структуре предикативных и определительных синтагм те признаки, которые составляют эту основу. В связи с этим выдвигается проблема локализации признаков, обеспечивающих данную дифференциацию. Локализуются ли они на протяжении всего звучания синтагмы или же сосредоточены в каких-то ее точках (отрезках звучания). Решению этой проблемы было посвящено несколько экспериментов на восприятие.

Эксперименты на восприятие

Сущность первого эксперимента на восприятие заключалась в том, что аудиторам предлагалось оценить не синтагмы в целом, а только их элементы. При аудировании вторых элементов перед аудиторами была поставлена задача определить, имеются ли какие-либо различия в интонационных реализациях вторых

элементов ПС и ОС. (15 пар синтагм диктор I, 15 пар - дикт.2).

Аудитор I в 12 парах из 30 поставил знак равенства между интонациями вторых элементов ПС и ОС. Другими словами, интонации этих элементов для данного аудитора неотличимы. Четыре пары вторых элементов у диктора I и две пары у диктора 2 воспринимались им как "очень похожие".

Аудитор 2 из такого же количества реализаций в 8 парах не нашел различий в интонационном оформлении вторых элементов синтагм. Аудитор 3 отметил полное тождество интонаций вторых элементов в 9 парах, аудитор 4 - в 13 и аудитор 5 - в 15 парах.

В тех случаях, когда аудиторы замечали какие-либо акустические различия в реализации вторых элементов синтагм, то эти различия не связывались с какими-нибудь значениями. А это может свидетельствовать о том, что интонации вторых элементов синтагм сами по себе не несут информации о типе связи. Это подтверждается и анализом акустических характеристик этих элементов. В них не отмечается той динамики в движении тона, какую мы наблюдаем в первых элементах. Этот вывод вытекает и из другого эксперимента, а именно: пересадки вторых элементов ПС к первым элементам ОС и наоборот. Аудирование синтагм с подклеенным ("инородным") вторым элементом показало, что аудиторы почти безошибочно определяют тип связи в смонтированных построениях, не замечая при этом никакой интонационной неестественности.

При прослушивании первых элементов синтагм, без вторых элементов, аудиторы отмечали, что для первого элемента ОС характерно ровное движение тона. К первому элементу ОС

аудиторы легко присоединяли второй элемент – существительное (определяемое слово), восстанавливая синтагму с определенной связью. Например, аудитор 4

большой (успех),	аудитор 5	голодная (овца)
грозный (муж)		старый (волк)
горькая (правда)		большой (успех)
седой (старик) и т.д.		грозный (муж) и т.д.

Следовательно, интонация первого элемента ОС в большинстве случаев оказывалась информативной. Причем, информация была однозначной.

Прислушивая же первые элементы ПС, аудиторы воспринимали их как вопросы, незавершенные высказывания, переспросы. Таким образом, по звучанию первого элемента ПС аудиторы не могли установить того, как он связан со вторым элементом, как это было в ОС. Интонационная информация интерпретировалась неоднозначно. Здесь для того, чтобы аудиторы могли правильно определить тип связи, необходимо подставить второй элемент синтагмы. Эта подстановка устраняет констатируемую выше многозначность. Происходит это потому, что акустические характеристики элементов становятся соотносительными, образуется некоторый интонационный контраст (слово) между элементами предикативных синтагм.

Как показывают акустический и аудитивный анализ, основная информация о исследуемых типах связи локализована в первом элементе ОС и в первом элементе ПС + контраст частот (слово) между элементами синтагмы. Сравнение акустических характеристик слогов первых элементов ПС и ОС показало, что

в ударных слогах фиксируются значительные изменения их основных характеристик: увеличивается время звучания и частотный интервал, а также интенсивность; изменяется частотный регистр. Это привело нас к мысли, что решающая информация о типах связи локализована в ударных слогах. Но некоторые изменения акустических характеристик происходят и в безударных слогах. Например, в ПС (седой - старик дик. I) начало предупредного слога первого элемента характеризуется ровным движением тона на отрезке звучания в 40 мсек, затем в середине слога тон падает на 6. секунду, а в конце слога - ровный участок звучания в 40 мсек. В начале предупредного слога первого элемента ОС (седой старик) наблюдается падение тона на м. секунду, а в середине и конце - ровное его движение (40 мсек). По частоте колебаний основного тона рассматриваемые предупредные слоги первого элемента ПС и ОС отличаются друг от друга. Разнятся эти слоги и по другим акустическим параметрам: время звучания слогаобразующей гласной в ПС - 80 мсек, в ОС - 60 мсек; пик интенсивности слога в ПС - 12 мм, в ОС - 10 мм. В связи с этим возникает вопрос о функциональной значимости отмечаемых изменений акустических характеристик в безударных слогах; необходимо выяснить: может ли осуществиться интонационная дифференциация только за счет изменений акустических характеристик в ударном слоге? Чтобы ответить на этот вопрос, необходимо уравнивать акустические характеристики безударных слогов ОС и ПС. Это может быть достигнуто пересадкой¹⁾ безударного слога опреде-

1) Методом пересадки звуков пользовался Н.И. Дукельский. Принципы сегментации речевого потока. Изд. АН СССР, 1962, стр. 5. М.К. Румянцев осуществил пересадку части звучания предложения. Сб. "Интонация и звуковой состав". Изд. МГУ, 1965, стр. 94.

лительной синтагмы в предикативную. Такой эксперимент и был осуществлен нами по следующей схеме: в синтагмах типа (седой-старик) ПС и (седой старик) ОС к ударному слогу первого элемента ПС подклеивался предударный слог первого элемента ОС, а к ударному слогу первого элемента ОС подклеивался предударный слог первого элемента ПС; в синтагмах типа (грозный-царь) ПС и (грозный царь) ОС к ударному слогу первого элемента ПС подклеивался заударный слог первого элемента ОС и наоборот; в синтагмах типа (голодная-степь) ПС и (голодная степь) ОС к ударному слогу первого элемента ПС подклеивался предударный и заударный слоги первого элемента ОС, а к ударному слогу подклеивались предударный и заударный слоги первого элемента ПС. В смонтированных таким образом интонационных структурах вся информация об исследуемых типах связи была сосредоточена в ударных слогах первого элемента синтагм, и этой информации, как показал аудитивный анализ, оказалось достаточно для верного их восприятия. Во всех II парах синтагм аудиторами были правильно опознаны типы связи. Правильному опознаванию определительной связи не мешали не совсем нейтральные акустические признаки безударных слогов, перенесенных в ОС из предикативных синтагм. Следовательно, выявленные акустические признаки, локализованные в ударных слогах первых элементов ПС и ОС, являются фонологически значимыми для дифференциации предикации и номинации. Акустические же характеристики, сосредоточенные в предударных и заударных слогах первых элементов ПС и ОС, несут избыточную информацию.

Kaplun M.I.

L.E.Ph., I.O.L. M.S.U.

INTONATIONAL DIFFERENTIATION OF PREDICATIVE AND ATTRIBUTIVE (QUALIFYING) BONDS IN MODERN RUSSIAN

The author gives the results of an experimental study of the mechanism of the intonational differentiation of predicative and attributive (qualifying) bonds.

In predicative and attributive syntagms of the type "косой-заяц" - Ц vs "косой заяц" - OC only one acoustic feature, fundamental frequency, and its function in the differentiation was analysed.

The object of the detailed acoustic analysis was to detect certain indications in the acoustic structure of predicative and attributive syntagms which could serve as the basis for their differentiation.

In this connection it was necessary to solve the problem of the location of differentiating indications, i.e. establish whether they are located throughout the sounding of the syntagms or are concentrated at certain points (sounding sections). To solve the task several experiments on perception were undertaken in addition to instrumental analysis.

ОПЫТ ЭКСПЕРИМЕНТАЛЬНО-ФОНЕТИЧЕСКОГО ИССЛЕДОВАНИЯ
СЛОВЕСНОГО УДАРЕНИЯ В ИНДОНЕЗИЙСКОМ ЯЗЫКЕ

Акустическая природа словесного ударения в индонезийском языке почти совсем не изучена. Отдельные высказывания о характере ударения в этом языке встречаются в научных работах по индонезийским языкам, учебниках и грамматиках, в работах по общей фонетике. Так, Р.Брандстеттер писал, что "в индонезийских языках акцентуируемый слог может отличаться от неакцентуируемого большей громкостью, более высоким тоном или большей длительностью".¹⁾ С.Т.Алишахбана, пытаясь определить общий характер индонезийского ударения, пришел к выводу, что в этом языке существуют элементы силового, долготного и музыкального ударения, которые часто комбинируются друг с другом. "... однако, в целом, силовое ударение не играет большой роли в индонезийском языке, более важно долготное ударение".²⁾ Проф.Бурниг Салех в своих лекциях по джакартскому диалекту, прочитанных в Институте восточных языков при МГУ, отмечал, что в джакартском диалекте, как и литературном языке, словесное ударение не силовое, а темпоральное, в отличие от фразового. Фразовое ударение во всех индонезийских языках, по его мнению, играет более важную роль, чем словесное.

1) R.Brandstetter, An introduction to Indonesian linguistics, London, 1916, p.340.

2) S.T.Alisjahbana, Tata bahasa baru bahasa Indonesia, djl. I, Djkt. 1958, h.24.

В свое время Е.Д.Поливанов, говоря о так называемом музыкальном ударе, замечал, что музыкальное (или тоническое) ударение известно и в малайском языке.³⁾ А.С.Теселкин и Н.Ф.Алиева полагают, что в индонезийском языке количественное ударение, ссылаясь при этом на наиболее распространенную точку зрения. При этом обращается внимание и на то, что "хотя имеются также элементы и силового, и музыкального удара, долгота звучания тем не менее служит основной, преобладающей над остальными чертой акустической характеристики ударения".⁴⁾

Можно, наконец, привести и такое мнение: "В индонезийском языке вообще нет словесного удара. ... на наш взгляд, более важным, чем ударение является мелодика слова или предложения".⁵⁾

Методика исследования

Для исследования акустических характеристик словесного удара в индонезийском языке была составлена специальная программа и проведены три эксперимента. В первом эксперименте с помощью интонографа конструкции ЛЭФИПР записывались отдельные двусложные слова с одинаковыми гласными. Во втором эксперименте записывались двусложные и трехсложные пары слов в двух вариантах их реализации: с ударением на первом или

3) Е.Д.Поливанов, Введение в языковедение (для востоковедных вузов), Л. 1928, стр.117.

4) А.С.Теселкин, Н.Ф.Алиева, Индонезийский язык, ИВЛ, 1960, стр.12.

5) Soedarmo, *Tekanan kata dan perscalannja*, Medan Bahasa, II, 1957.

втором слоге (в трехсложных словах) и с ударением на последнем слоге. Таким образом, во втором варианте осуществлялся перенос ударения с одного слога на другой. В третьем эксперименте записывались отдельные предложения, в которых были использованы слова, включенные в два первых эксперимента.⁶⁾

Во втором эксперименте первый и второй варианты реализации часто соотносились между собой как реализации с естественной и искусственной акцентуацией. При этом, по мнению дикторов, в некоторых парах искусственная акцентуация, если и не искажала смысловую оболочку слова, то во всяком случае нарушала его естественную акцентную структуру.

Интонограммы экспериментальных слов и предложений были подвергнуты расшифровке и анализу. Движение основного тона и время звучания исследуемых слогов были нанесены на графики, где частота основного тона (в герцах) отложена по оси ординат в логарифмическом масштабе, время звучания (в миллисекундах) по оси абсцисс в линейном масштабе.

В связи с тем, что акустические характеристики ударения наиболее ярко проявляются в слоговом гласном, в дальнейшем при сравнении ударных и неударных слогов учитываются, главным образом, акустические параметры слогового гласного.

Ниже рассматриваются акустические характеристики ударных и неударных слогов в последовательности: интенсивность

6) Запись проводилась в фонетической лаборатории ИВЯ при МГУ (руководитель доц. М. К. Румянцев). В качестве дикторов были приглашены преподаватели индонезийского языка проф. Интойо (уроженец Средней Явы) и доц. А. Т. Эффенди (уроженец Северной Суматры). Записи на интонографе предшествовала запись на ферромагнитную пленку. Перед записью дикторы знакомились с экспериментальным материалом, который затем неоднократно прослушивался с целью определить место ударения в слове.

(динамический компонент), время звучания (темпоральный компонент), движение основного тона (тоновый компонент).

Интенсивность (динамический компонент)

При анализе интонограмм двусложных слов с одинаковыми гласными было отмечено регулярное увеличение слоговой интенсивности ударного гласного. Примеры такого увеличения приводятся в таблице I.

Таблица I

Слово	Амплитуды слоговой интенсивности	
	Ударный слог	Безударный слог
buku	7,5 мм	5 мм
tinggi	6 мм	3 мм
ramah	10,5 мм	7 мм

В двусложных и трехсложных словах при переносе ударения с одного слога на другой также фиксировалось увеличение интенсивности ударного гласного. Рассмотрим некоторые примеры.

Пример I. Слово pasir . Состоит из двух слогов, первый из которых (pa) открытый, второй (sir) закрытый. Гласная а сама по себе более громкая, чем ¹ . При ударении на втором слоге (рис.1), уровень амплитуд первого слога достигает 11 мм, второго - 13 мм.

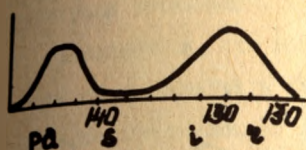


Рис. 1

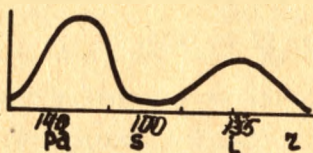


Рис. 2

На рис. 2 изображено распределение слоговой интенсивности при переносе ударения на первый слог. Уровень его интенсивности возрастает до 18 мм, во втором слоге амплитуды интенсивности уменьшаются до 9 мм. Наибольшая интенсивность отмечается примерно в середине звучания гласной, к концу слога интенсивность падает.

Пример 2. Слово *шлшш*. При ударении на первом слоге (ш) уровень амплитуд его слоговой интенсивности достигает 14 мм, во втором слоге - 13 мм (рис. 3).

Наибольшая интенсивность также отмечена примерно в середине слога.

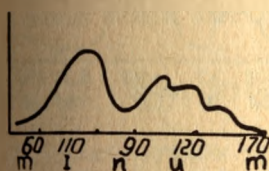


Рис. 3

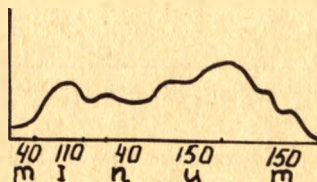


Рис. 4

При переносе ударения на второй слог (рис. 4) его интенсивность увеличивается незначительно - до 13,5 мм, в первом же слоге амплитуды интенсивности сокращаются до 9,5 мм.

Пример 3. Слово *djatuh*. При ударе на первом слоге (*dja*) уровень интенсивности этого слога достигает 14 мм, второго 12 мм (рис. 5)

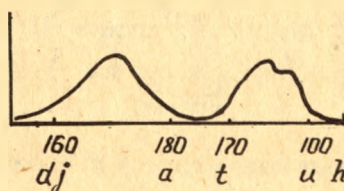


Рис. 5.

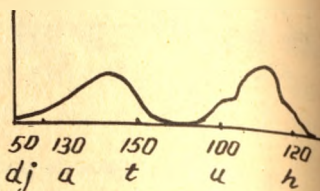


Рис. 6.

При переносе удара на второй слог (*tuh*) интенсивность этого слога увеличивается до 13 мм, интенсивность первого слога уменьшается до 12 мм (рис. 6).

Пример 4. Слово *karangan*. При ударе на третьем слоге (*an*) интенсивность этого слога достигает 11 мм, во втором слоге (*rang*) также 11 мм, в первом (*ka*) 10 мм (рис. 7).

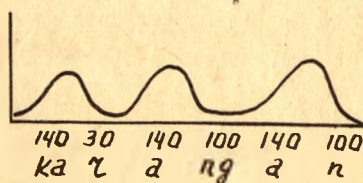


Рис. 7.

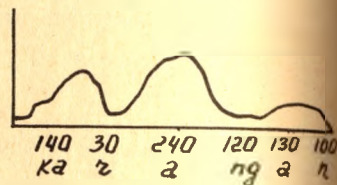


Рис. 8.

При переносе удара на второй слог (*rang*) интенсивность этого слога возрастает до 12,5 мм, интенсивность ударного слога (*ka*) - 11 мм, заударного (*an*) - 6 мм (рис. 8).

Аудитивное восприятие экспериментальных слов и анализ интонограмм показывают, что слог, воспринимаемый как ударный, произносится, как правило, с большей силой артикуляции, которой соответствуют и большие амплитуды слоговой интенсивности. Все это позволяет считать интенсивность (динамический компонент) одним из акустических признаков, обеспечивающих восприятие ударного слога в слове.

В то же время следует отметить, что при анализе интонограмм слов с естественной акцентуацией наблюдались случаи, когда амплитуды слоговой интенсивности менее громкого гласного (например *i*) в ударном слоге не превышали амплитуд интенсивности более громкого гласного (например, *a*) в безударном слоге, а лишь совпадали с ними или даже были меньше их. Этот факт еще раз подтверждает мысль о том, что основу функциональных противопоставлений составляют не абсолютные, а относительные значения акустических характеристик.

Относительное увеличение интенсивности в ударном слоге фиксировалось и при записи слов с одинаковыми гласными в позиционно разных слогах.

Время звучания (темпоральный компонент)

Время звучания ударного слога в изолированных словах возрастает, иногда довольно значительно, и в то же время оно обусловлено характером слога. Рассмотрим некоторые примеры.

Пример I. В слове *datang* при ударе на первом слоге (*da*) время звучания гласного *a*—180 мсек в безударном — 100 мсек. При переносе удара на второй слог *tang* ,

время звучания гласного в этом слоге возрастает до 120 мсек, время звучания гласного в первом слоге сокращается до 140-160 мсек.

Пример 2. В слове *djatuh* при ударе на первом слоге (*dja*) время звучания гласного *a* в этом слоге - 180 мсек, гласного *u* во втором слоге - 100 мсек. При переносе удара на второй слог (*tuh*) время звучания гласного *a* становится 150 мсек, гласного *u* - 120 мсек.

Пример 3. В слове *lulus* при ударе на первом слоге (*lu*) время звучания ударного гласного *u* - 170 мсек. Время звучания этого же гласного в безударном слоге - 120 мсек. При переносе удара на второй слог (*lus*) длительность ударного гласного возрастает до 150 мсек, длительность безударного гласного сокращается до 150 мсек.

Пример 4. В слове *karangan* при ударе на втором слоге (*rang*) время звучания ударного гласного - *a* - 210 мсек, время звучания того же гласного в слоге - *ka* - 140 мсек, в слоге *an* - 120 мсек. При переносе удара на последний слог, длительность ударного гласного становится 145 мсек, длительность того же гласного в слоге - *ka* - 100 мсек, в слоге *rang* - 140 мсек.

Таким образом, время звучания (темпоральный компонент) также является одним из составляющих удара в индонезийском языке. Однако этот компонент в какой-то степени обусловлен характером слога, так как длительность ударного гласного в закрытом слоге, при прочих равных условиях, меньше или равна длительности безударной гласной в открытом слоге.

частота колебаний основного тона (тоновый компонент)

В двусложных словах с одинаковыми гласными отмечалось регулярное повышение частоты колебаний основного тона в ударном гласном.

Пример I. Слово *gigh*. При ударе на первом слоге максимальный уровень частоты колебаний основного тона ударного гласного равнялся 175 гц, неударного 125 гц (рис.9).

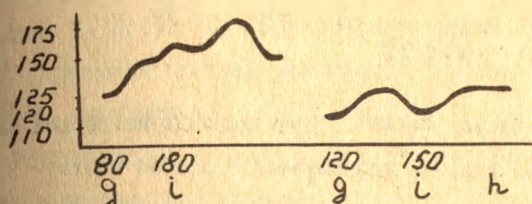


Рис. 9

Увеличение частоты колебаний основного тона начинается в начале слога и достигает максимального уровня в середине слога. Контур движения частоты основного тона ударного гласного восходяще-нисходящий (125-175-150 гц), безударного гласного вначале восходящий (110-125 гц), затем нисходяще-восходящий (125-110-125 гц). Частотный интервал между ударным и безударным слогом (по наивысшим частотам) составляет примерно кварту.

При переносе ударения на второй слог, частота колебаний основного тона ударного гласного возрастает до 175 гц; у предударного гласного наивысший уровень характеризуется частотой колебаний в 150 гц. Контур движения тона в ударном

слоге восходящий (150-175 гц), в безударном - также восходящий с незначительным падением частоты в конце (125-150-140гц). Частотный интервал между безударным и ударным слогом равен 6.секунде (рис.10).

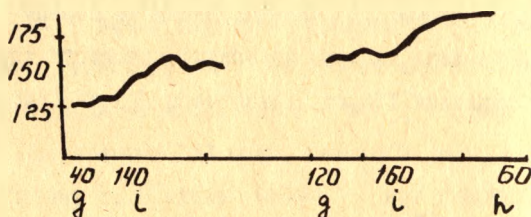


Рис.10

Пример 2. Слово *Julius*. При ударении на первом слоге (*ly*), ударный гласный реализуется в полосе частот 160-200 гц, безударный с частотой 125 гц. Повышение тона до 200 гц в ударном слоге отмечено в середине слога. Контур движения тона в ударном слоге восходяще-нисходящий (160-200-175 гц), в безударном слоге-ровный (125 гц). Частотный интервал между ударным и безударным слогом равен 6.сексте (рис.11).

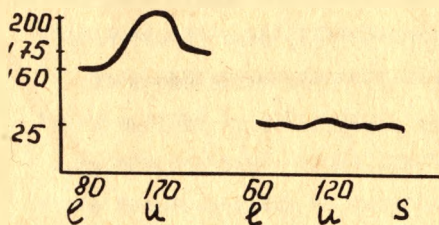


Рис.11

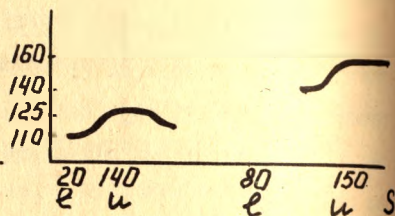


Рис.12

При переносе ударения на второй слог (*lyus*) отмечается повышение основного тона этого слога до 160 гц, в

предударном понижении до 125 гц. Повышение основного тона, отмечаемое в начале звучания, сохраняется до конца слога. Контур движения тона в ударном слоге восходящий (140-160 гц), в предударном - восходяще-нисходящий (110-125-120 гц). Частотный интервал между безударным и ударным слогами - б.терция (рис.12).

При записи слов с разными гласными и трехсложных слов также отмечалось регулярное повышение основного тона ударного гласного.

Пример 3. Слово *djatuh*. При ударе на первом слоге (*dja*) ударный гласный реализуется в полосе частот 150-160 гц, безударный - 125-110 гц.

Контур движения основного тона ударного слога характеризуется некоторым восхождением частот во второй половине своего звучания. В самом конце звучания фиксируется небольшое падение тона. Контур движения тона у заударного слога - нисходящий (125-110 гц). Частотный интервал между безударным и ударным слогами - б.терция (рис.13).

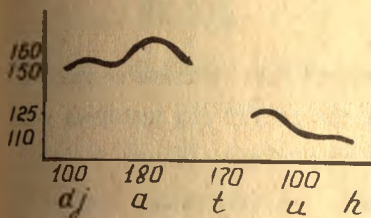


Рис. 13

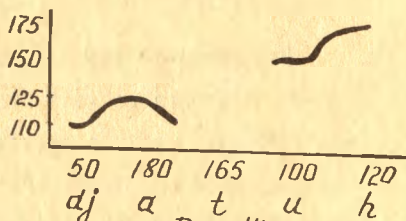


Рис. 14

При переносе ударения на второй слог (*tuh*) частота колебаний основного тона в гласном - *u* увеличивается до 175 гц. Меняется контур движения основного тона в этом

слоге: из нисходящего он становится восходящим (150-175 гц), в безударном слоге (dja) в основной части звучания тон ровный, завал частот в конце слога несколько больший, чем в предыдущей реализации, и весь слог dja реализуется в этом случае в более низкой полосе частот (110-125-110 гц). Частотный интервал между безударным и ударным слогами равен кварте (рис.14).

Пример 4. Слово *tipum*. При ударе на первом слоге (*ti*) ударный гласный реализуется в полосе частот 150-175 гц, заударной - 140-150 гц. Контур движения основного тона ударного гласного восходящий (150-175 гц), заударного нисходяще-восходящий (150-140-150 гц). Частотный интервал между безударным и ударными слогами - 6.секунда (рис.15).



Рис. 15

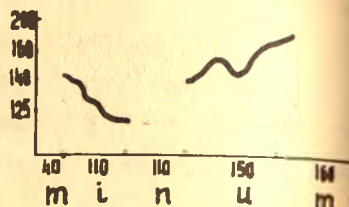


Рис. 16

При переносе ударения на второй слог частота колебаний (*num*) основного тона повышается до 175 гц, меняется контур движения тона: из нисходяще-восходящего (почти ровного) он превращается в явно восходящий, а в безударном слоге *ti* восходящее движение тона сменяется нисходящим, и весь слог реализуется в более низкой полосе частот (140-125 гц). Частотный интервал между безударным и ударными слогами - 6.терция (рис.16).

Пример 5. Слово *karangan*. При ударе на втором слоге (*rang*) частота колебаний основного тона ударной гласной 150-160 гц, предударной 140-150 гц, заударной - 110-125 гц. Контур движения основного тона ударного гласного восходяще-нисходящий (150-160-150 гц), предударного нисходяще-восходящий (150-140-150 гц), заударного нисходящий (125-110 гц). Частотный интервал между предударным и ударным слогом положительный, равен м.секунде, между ударным и безударным слогом - отрицательный, равен 6.терции (рис.17).

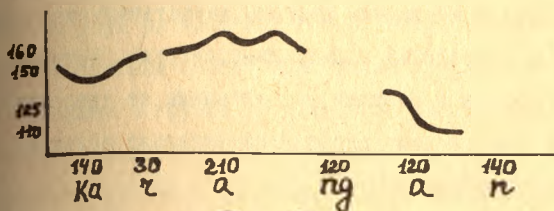


Рис 17

При переносе удара с одного слога на другой отмечен ряд случаев, когда частота колебаний основного тона заданного ударного слога оказывалась не выше, чем у безударного, хотя и повышалась по сравнению с безударной реализацией того же слога. Например, в слове *datang* ударный слог (*da*) реализуется в полосе частот 140-160 гц (140-160-150 гц), а безударный слог (*tang*) в полосе 110-125 гц (125-110-125). При переносе удара на слог *tang* в слове *da* фиксируется последовательность частот 140-150-150 гц, а в слове *tang* - 150-140-125 гц.

Суммируя результаты анализа всего экспериментального материала, можно сделать следующие выводы, касающиеся движе-

ния основного тона в ударных и безударных слогах:

а) при ударе, частота колебаний основного тона ударного гласного, как правило, повышается независимо от качества гласного и характера слога. Следовательно, высота основного тона — тоновый компонент является одним из составляющих индонезийского ударения;

б) при изолированном произнесении слов контур движения основного тона в ударных слогах чаще всего восходящий или восходяще-нисходящий, в заударных слогах — нисходящий;

в) при естественной акцентуации слова частотный интервал между ударным и неударным слогами характеризуется значениями, лежащими в пределах м.секунды-б.секунды. Интервал между ударным слогом и слогами, отделенными от ударного другим слогом (в многосложных словах), значительно больший (в пределах б.терции-кварты);

г) в индонезийском языке в многосложных словах, помимо ударных и безударных слогов, существуют и полуударные, акустические параметры которых меньше, чем у ударных, но больше, чем у безударных. Полуударные слоги обычно непосредственно предшествуют ударному слогу.

Акустические признаки словесного ударения в индонезийском языке складываются из трех компонентов: динамического, темпорального и тонового, которые реализуются как относительное увеличение амплитуд слоговой интенсивности, относительное увеличение времени звучания и относительное повышение частоты колебаний основного тона ударного гласного.

РЕАЛИЗАЦИЯ СЛОВЕСНОГО УДАРЕНИЯ ВО ФРАЗЕ

Запись простых и некоторых видов сложных предложений (фраз) осуществлялась для того, чтобы получить общую картину их интонационного оформления, выявить интонационные участки с повышением и понижением тона, проверить действие акустических признаков словесного ударения на этих участках.

Исследование акустических признаков тех или иных интонационных дифференциаций не входило в нашу задачу⁷⁾.

В наших экспериментальных фразах было от одной до пяти ритмических групп, причем наиболее часто встречались фразы, состоящие из двух ритмических групп. Это, главным образом, простые двусоставные повествовательные предложения. В таких предложениях в группе подлежащего, как правило, отмечалось общее повышение тона, в группе сказуемого — общее понижение. Однако в пределах группы подлежащего при вном актуальном членении отмечались случаи понижения тона к началу каждого нового члена синтагмы и повышения к концу синтагмы.

Рассмотрим несколько типичных примеров из нашего материала и проследим, как интонация той или иной ритмической группы предложения накладывается на словесное ударение, как

7) Исследование интонации индонезийского предложения по существу только начинается. См. O.Essen, *Die sprachmelodische Gestaltung der Aussage und Frage in der Bahasa Indonesia, Oriens Extremus*, XI, 1959;

Е.С.Белкина, Интонация основных коммуникативных типов предложений в индонезийском языке, Сб. Интонация и звуковой состав, МГУ, 1965.

взаимодействуют акустические признаки словесного ударения и интонации.

Акустические параметры того или иного слова в изолированной позиции условно назовем параметрами 1-й реализации, параметры того же слова в предложении – параметрами 2-й реализации.

Пример I. *Dekat sungai itu' ada danau*⁸⁾ "Возле этой реки есть озеро". Предложение состоит из двух ритмических групп.

Общее движение тона в предложении графически представлено на рис.18 (см.приложение).

В первом слове предложения отмечается некоторое повышение тона, затем следует понижение, захватывающее всю первую половину фразы вплоть до ее последнего слога *tu*. В слоге *tu* фиксируется резкое повышение тона. Вторая ритмическая группа, в целом, реализуется, примерно, в той же полосе частот, что и первая, если исключить последний слог первой ритмической группы. Во второй ритмической группе от ее первого слога до слога *da* включительно наблюдается умеренное восхождение тона, которое заканчивается весьма заметным падением на последнем слове *danau*.

В первой ритмической группе значительное понижение тона в слоге " i " с последующим резким повышением в слоге " tu " (частотный интервал – квинта) может вызываться особенностями реализации такой интонационной дифференциации как незавершенность высказывания, когда падение частот как бы подчеркивает начало их резкого нарастания и служит его фоном.

8) Штрихом обозначается граница ритмической группы.

Некоторое повышение частоты колебаний основного тона на слоге "da" в слове *danau* с резким последующим понижением на последнем слоге вызывается, очевидно, реализацией завершенности высказывания.⁹⁾

Слово *sungai*, имевшее в I-й реализации ударение на первом слоге "su", в предложении *Dekat sungai itu ada danau* локализуется на интонационном участке с некоторым понижением частоты, при этом аудитивно констатируется переход ударения на второй слог *ngai*. Основные акустические параметры этого слова в обеих реализациях представлены в таблице 2.

Таблица 2

<i>sungai</i>	I-я реализация		2-я реализация	
	<i>su</i>	<i>ngai</i>	<i>su</i>	<i>ngai</i>
Слоги				
Интенсивность в мм	12	9	8	9
Время звучания в мсек	120	240	80	180
Частота колебаний в Гц	175-160-175	140-125	150	140-125

9) "Интонационная дифференциация завершенности и незавершенности высказывания осуществляется акустическими средствами, локализованными в одной (фонологической) зоне звучания предложения - в его последнем слове".

См. М. К. Румянцев, Интонационное выражение завершенности и незавершенности высказывания, Сб. Интонация и звуковой состав, МГУ, 1965.

В I-й реализации ударный слог превалирует над безударным по интенсивности. Частота колебаний основного тона ударного слога выше, чем у безударного. Частотный интервал между этими слогами (по максимальным значениям) — 6.секунда. Контур движения тона в ударном слоге можно охарактеризовать как ровный, в безударном — нисходящий (интервал понижения — 6.секунда). Длительность ударного слога меньшая, чем безударного. Отношение времени звучания ударного слога к безударному равно 0,5. Это может быть вызвано большим количеством слоговых элементов в безударном слоге, а также их качеством.

Во второй реализации этого слова интенсивность первого слога значительно уменьшается. Это приводит к тому, что второй слог — *ngai* — оказывается выделенным по интенсивности, хотя в количественном отношении его интенсивность остается такой же, как и в первой реализации.

На интонационных участках, предшествующих фонологической зоне, происходит временная компрессия элементов. Сжатие во времени звучания особенно заметно в начале предложения. Ударный слог так же, как и в первой реализации, где он был безударным, несколько длительнее, чем слог *su*. Отношение времени звучания слога *su* к слогу *ngai* несколько уменьшается. В I-й реализации оно равно 0,5, во второй — 0,44. Это свидетельствует о том, что доля времени слога, ставшего ударным, возрастает.

В предложении слово *bungai* реализуется примерно в том же диапазоне частот, что и в изолированной позиции. Безударный слог выше по частоте колебаний, чем ударный. Частотный интервал между ними — м.секунда. Контур движения основ-

ного тона в безударном слоге ровный, в ударном нисходящий (интервал падения частот - 6.секунда). Деформация контура во второй реализации вызвана общим положением тона на этом участке.

Слово *danau* с ударением на первом слоге *da* в предложении оказывается на интонационном участке с некоторым повышением тона на первом слоге и последующим резким падением частоты на втором слоге - *nau*. В предложении перехода ударения на другой слог не происходит. Акустические параметры слова *danau* в обеих реализациях см. в таблице 3.

Таблица 3

<i>danau</i>	1-я реализация		2-я реализация	
	<i>da</i>	<i>nau</i>	<i>da</i>	<i>nau</i>
Слоги				
Интенсивность в мм	I3	I0	I3	I0
Время звучания в мсек	220	300	I20	280
Частота колебаний в гц	I25-I60	I40-I25	I60	I25-I00

В 1-й реализации ударный слог *da* превалирует по интенсивности над безударным и реализуется в более высоком частотном диапазоне. Частотный интервал между этими слогами - 6.секунда. Контур движения основного тона в слоге *da* восходящий (интервал повышения - 6.терция), в слоге *nau* - нисходящий (интервал понижения - 6.секунда). Длительность

ударного слога меньшая, чем у безударного. Отношение времени звучания ударного слога к безударному ($da : \text{пав}$) примерно 0,7.

Во второй реализации интенсивность ударного слога тоже больше, чем у безударного. Однако безударный слог, как и в I-й реализации длилнее ударного, хотя доля времени ударного слога уменьшается. Отношение от 0,7 сокращается до 0,42, что, видимо, связано с конечной позицией данного слога в предложении. Частота колебаний основного тона ударного слога выше, чем у безударного, частотный интервал между этими слогами - б.терция. Контур движения основного тона ударного слога - ровный (в I-й реализации - восходящий), безударного - нисходящий, интервал падения - м.секунда.

Деформация контура в ударном слоге объясняется общим понижением интонации на этом участке. Некоторое повышение частот перед их резким падением лишь сглаживает контур в пределах того же частотного диапазона, но не превращает его из восходящего в нисходящий. Значительное понижение тона к концу предложения усиливает нисходящий контур слога пав .

Сопоставление интонационных характеристик других ударных слогов рассматриваемого предложения показывает, что ударный слог превалирует над безударным по интенсивности, хотя его длительность и частота колебаний основного тона коррелируют с динамикой интонации того или иного участка.

Наибольшие значения акустических параметров отмечены в слоге tu (itu) в зоне незавершенности высказывания; частотный интервал между слогами i и tu равен октаве.

Пример 2. *Barak anak itu ada buku banjak* "у отца этого ребенка много книг". В рассматриваемой реализации предложение представляет собой две ритмические группы.

Общее движение частоты колебаний основного тона в этом предложении аналогично тому, которое мы наблюдали в первом примере (см. рис. 19).

Повышение частоты колебаний и в этом предложении отмечается в его начале. Понижение частот перед их резким повышением в зоне незавершенности (слог *tu*) отмечается к концу слова *anak* - началу слова *itu* . Во второй ритмической группе некоторое повышение частоты с последующим понижением отмечено в первом слове *ada* . Некоторое повышение частоты перед резким падением в зоне завершенности фиксируется в слове *bu* (слово *buku*).

В данном предложении слово *barak* с ударением на том же слоге (*pak*), что и в изолированной реализации, локализовано на участке повышения частот. Сопоставляемые параметры реализации этого слова в изолированной позиции и на данном участке фразовой интонации приводятся в таблице 4.

Таблица 4

barak	1-я реализация		2-я реализация	
	ba	pak	ba	pak
Слоги				
Интенсивность в мм	9,5	10	12	12
Время звучания в мсек	260	125	130	130
Частота колебания в гц	150-175-160	200-175-200	160-200	200-190-200

Ударный слог - **рак** - в первой реализации превалирует по интенсивности над безударным, во второй - равен ему. Вероятно, интенсивность в большей степени связана здесь с интонационными характеристиками на данном участке фразы.

Время звучания безударного открытого слога **да** во второй реализации подвергается значительной компрессии, длительности ударного слога остается практически неизменной. Соотношение длительности ударного и безударного слогов изменяется в сторону значительного увеличения доли ударного слога по сравнению с первой реализацией (в I-й реализации отношение - 0,43, во второй - 1).

Ударный слог реализуется в более высокой полосе частот, чем безударный, хотя по сравнению с изолированной позицией в безударном слоге фиксируется повышение частоты, связанное, видимо, с общей динамикой интонации на этом участке. В изолированном произнесении частотный интервал между ударным и безударным слогом (по пиковым значениям) - 6.секунда. Первая (большая) половина ударного слога во второй реализации произносится на уровне конечной точки безударного слога с некоторым завалом частот во второй половине слога. Контур движения частоты основного тона ударного гласного в первой реализации нисходяще-восходящий, во второй-ровный. В безударном слоге в первой реализации-восходящий с незначительным понижением, во второй - восходящий (интервал восхождения-кварта).

В первом слоге тоновой контур усиливается за счет интонационной тенденции повышения тона на этом участке, во втором - выравнивается.

Слово **anak** с ударением на первом слоге оказывается в рассматриваемом предложении на интонационном участке с некоторым общим понижением тона, причем фиксируется переход ударения на второй слог. Акустические параметры этого слога в обеих реализациях приводятся в таблице 5.

Таблица

anak	I-я реализация		2-я реализация	
	a	nak	a	nak
Слоги				
Интенсивность в мм	7	6	8	7
Время звучания в мсек	300	120	80	110
Частота колебаний в Гц	140-160	175-140	150-175- 150	175-140

В первой реализации ударный слог **a** превалирует по интенсивности над безударным **nak**. Во второй реализации интенсивность слога **a** также оказывается большей, чем в слоге **nak**, который становится ударным. Это может быть связано с общим понижением интонации на этом участке.

Время звучания слога **a** значительно сокращается, во втором слоге - **nak** - сокращается незначительно. Соотношение по длительности ударного и безударного слога в первой реализации - 0,4, во второй - 0,72. Весьма вероятно, что признак интенсивности как бы "исчезнувший" во второй реализации, компенсируется большим временем звучания.

Контур движения основного тона ударного слога в первой реализации восходящий (интервал восхождения большая секунда),

в безударном слоге – нисходящий (интервал–большая терция). Частотный интервал между слогами–малая секунда. Во второй реализации в ударном слоге нисходящий контур (интервал–б.терция), в безударном – восходяще-нисходящий. Деформация контура и здесь объясняется интонационными причинами.

Слово *byku* с ударением на первом слоге в предложении занимает позицию на участке некоторого повышения частот перед их падением в фонологической зоне завершенности высказывания.

Интонационные характеристики слова *byku* в обеих реализациях приводятся в таблице 6.

Таблица 6

<i>byku</i>	I-я реализация		2-я реализация	
	<i>by</i>	<i>ku</i>	<i>by</i>	<i>ku</i>
Слоги				
Интенсивность в мм	7,5	5	4	3
Время звучания в мсек	240	250	100	120
Частота колебаний в гц	150–175	175–140	160–175	175–190

В обеих реализациях ударный слог – *by* превалирует по интенсивности над безударным.

Длительность безударного слога – *ku* в обеих реализациях большая, чем у ударного (соотношение в I-й реализации – 0,9, во второй – 0,83). При этом во второй реализации при значительном сокращении времени звучания, доля ударного слога даже несколько уменьшается. Это, вероятно, связано с

некоторым повышением частного диапазона на этом участке.

Контур движения ударного слога в обеих реализациях восходящий (интервал восхождения соответственно м. и б. секунды), основной тон реализуется, примерно, в одинаковой полосе частот. Контур движения основного тона в безударном слоге в 1-й реализации нисходящий (интервал-м.секунда), во второй-восходящий (интервал-м.секунда). Изменение контура и в рассматриваемом случае вызывается повышением тона на этом интонационном участке.

В словах *ada* и *banjak* слог, воспринимаемый аудитором как ударный, обладает большей интенсивностью, чем безударный, однако длительность и частотные регистры этих слогов коррелируют, видимо, с интонацией на соответствующем участке.

Наивысшие частотные и временные параметры отмечены в слоге *tu(itu)*, реализуемом в фонологической зоне незавершенности высказывания (220 мсек, 190-225 гц), наименьшие - в слоге *njak (banjak)*, реализуемом в фонологической зоне завершения высказывания (80 мсек, 125-100 гц).

Пример 3. *Pagar itu' tinggi sekali* "Этот забор очень высокий". В рассматриваемой реализации предложение представляет собой две ритмические группы: первая группа - группа подлежащего (данное), вторая - группа сказуемого (новое).

Общее движение основного тона в предложении представлено на рис. 20.

В начале предложения частотный рисунок относительно ровный. Некоторое понижение частот перед их повышением в зоне незавершенности отмечается в слоге *i*. Вторая ритмиче-

ская группа характеризуется общим понижением тона, хотя начальный ее участок реализуется в довольно высокой полосе частот (225-200 гц).

Слово *pa:gar* с ударением на втором слоге начинается предложение и реализуется с относительно ровным тоном. Акустические характеристики слова *pa:gar* в обеих реализациях см. в таблице 7.

Таблица 7

pa:gar	I-я реализация		2-я реализация	
	pa	gar	pa	gar
Слоги				
Интенсивность в мм	7	II	IO	I2,5
Время звучания в мсек	290	180	140	120
Частота колебаний в гц	160-140	175-150	190-175- 190	190

В обеих реализациях ударный слог превалирует над безударным по интенсивности.

Время звучания ударного слога в обеих реализациях меньше, чем время звучания безударного слога (соотношение в I-й реализации - 0,62, во второй - 0,85). Доля ударного слога в общей длительности слогов во 2-й реализации возрастает.

В I-й реализации частота основного тона ударного гласного выше, чем у безударного (частотный интервал между ними - м.секунда). Во 2-й реализации оба слога реализуются примерно в одинаковом частотном диапазоне.

В I-й реализации контур движения основного тона в ударном слоге нисходящий (интервал - 6.секунда), во второй - ровный.

В безударном слоге в I-й реализации слоговой контур нисходящий (частотный интервал - 6.секунда), во второй - нисходяще-восходящий. Интонационные факторы на этом участке выравнивают контур в обоих слогах.

Слово *tinggi* с ударением на первом слоге располагается в начале второй ритмической группы с общим понижением тона. Акустические параметры этого слова в обеих реализациях см. в таблице 8.

Таблица 8

tinggi	I-я реализация		2-я реализация	
	ting	gi	ting	gi
Интенсивность в мм	6	3	5	3,5
Время звучания в мсек	90	200	100	160
Частота колебаний в гц	190-175	190-125	210-225- 210	200-190

В обеих реализациях фиксируется большая интенсивность в ударном слоге *ting*. Длительность ударного слога меньшая, чем у безударного (соотношение в I-й реализации - 0,45, во второй - 0,62). Доля ударного слога в общей длительности во второй реализации несколько возрастает.

Контур движения основного тона в ударном слоге в I-й реализации нисходящий (частотный интервал - м.секунда), во второй - восходяще-нисходящий. Слоговой контур в безударном слоге в I-й реализации нисходящий (частотный интервал - малая секста), во второй - также нисходящий (частотный интервал - малая секунда), но в более высоком частотном регистре. Выпрямление контура в ударном слоге вызывается общим повышением тона на этом участке.

Наибольшее время звучания и частота колебаний основного тона отмечены в зоне незавершенности высказывания (слово *itu*), в слоге *tu* время звучания 240 мсек, частота колебаний 225-200 гц.

В результате анализа интонограмм всех записанных нами предложений представляется возможным сделать следующие выводы:

Акустические признаки интонации накладываются на акустические признаки словесного ударения, образуя сложный контур движения основного тона, разную слоговую интенсивность и т.д.

В связной речи темпоральный и тоновый компоненты чаще всего уходят на реализацию интонационных дифференциаций, хотя в нашем материале доля ударного слога в общей длительности звучания слова во многих случаях возрастала. Контур движения основного тона в ударных и безударных слогах зависит от динамики интонации на том или ином участке и под ее влиянием может деформироваться. Динамический компонент, видимо, наиболее прочно связан со словесным ударением. В нашем материале случаи его "исчезновения" под влиянием интонации единичны.

В тех случаях, когда интенсивность используется для реализации какой-либо интонационной дифференциации, этот признак компенсируется другим - темпоральным. Для восприятия словесного ударения, видимо, достаточно одного динамического признака, однако слоги более длительные и с большей частотой колебаний основного тона (иными словами слоги, реализуемые на участках с повышением интонации или в фонологической зоне незавершенности) также могут восприниматься на слух как ударные. Это явление создает впечатление двойного ударения в слове или равноударности слогов. Контур движения основного тона и частотный интервал в слоге не входят в число признаков словесного ударения, так как коррелируют с интонационным рисунком того или иного участка предложения.

Приложение

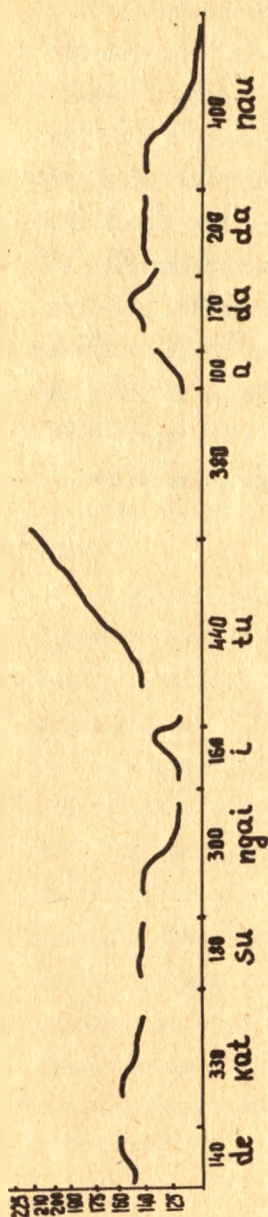


Рис. 18

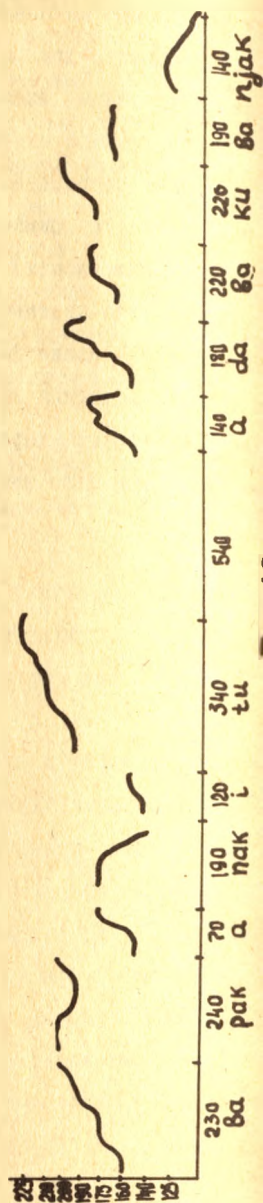


Рис. 19

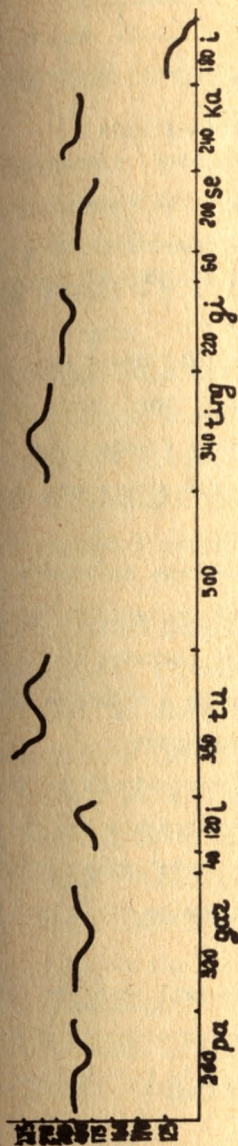


Рис. 20

Pavlenko A.P.

L.E.Ph., I.O.L. M.S.U.

AN EXPERIMENTAL-PHONETIC STUDY OF WORD-STRESS
IN THE INDONESIAN LANGUAGE

The article gives the results of a study of the acoustic parameters of word-stress in Indonesian.

The speech of two informants from different regions of Indonesia (Java and Sumatra) was studied. Word pairs with a successive stress shift from one syllable to the other, and simple twomember sentences containing the same words were recorded on an intonographer.

The experiment showed that word-stress in Indonesian is realised through a relative increase of syllabic intensity (dynamic component), through a relative increase of the syllable duration (temporal component) and through a relative increase in frequency of the main tone (tone component).

The most constant indication of word-stress is the dynamic component. The temporal and tone components are not infrequently used for realisation of other phonological differentiations or correlate with the general dynamics of intonation at some part of the sentence.

In those cases (rare in our material) when the dynamic component does not happen to be connected with the rhythmic structure of the word, other components, e.g. temporal, are used instead.

В.И.ПЕТРЯНКИНА
ЛЭФ ИВЯ ПРИ МГУ

ДОЛГОТНАЯ ХАРАКТЕРИСТИКА ГЛАСНЫХ ЯЗЫКА СУАХИЛИ

До настоящего времени длительность гласных языка суахили экспериментально не изучалась.

Между тем выяснение времени звучания гласных в суахили важно для решения многих фонетических и фонологических проблем данного языка.

Проведенный нами анализ длительности гласных позволил:

(1) определить время звучания того или иного гласного звука в ударных и безударных позициях, что оказалось весьма полезным при выяснении характера редукции гласных ;

(2) установить некоторую иерархию гласных по времени звучания ;

(3) проследить характер изменения долготы ударных и безударных гласных в разных фонетических условиях ;

(4) судить об изменении долготы гласных в связи с изменением стилей речи (разговорный и полный) ;

(5) характеризовать назализацию гласных ;

(6) дать фонематическую трактовку звуко сочетаний.

Для анализа длительности гласных на интонографе конструкции ЛЭФИПР было записано 252 суахилийских слова. Каждое слово записывалось по несколько раз в двух темпах (быстром и медленном).

Так как запись на магнитофонную ленту производилась не в студийных условиях, время звучания исследуемых звуков

по интонаграммам оказалось возможным определить с несколько большей, чем обычно погрешностью.

При подготовке к записи экспериментальный материал - слоги, слова были расположены строчками (по три слова или слога в каждой строчке).

Строчки располагались одна над другой, при этом исследуемые звуки находились в определенных фонетических условиях в составе сопоставляемых по вертикали слов.

Подобная композиция материала была продиктована желанием уравнивать интонационные условия произнесения слогов и слов.

Надежность выводов зависит от правильного сегментирования речевого потока, то есть от определения границ звуков при анализе записей. Нами за основу были приняты положения Л.В.Щербы.¹⁾

Однако в связи с наличием в суахили придыхательных смычных согласных нами измерялась не артикуляторная, а сонорная длительность гласных, т.е. с момента начала голосовых вибраций.

Долгота звуков измерялась в миллисекундах (мсек) и была исследована:

1. В двусложных словах с открытыми слогами (ударные и безударные):

а) перед смычными глухими, смычными звонкими, щелевыми глухими, щелевыми звонкими, носовыми и плавными согласными; после указанных согласных;

1) Л.В.Щерба. Русские гласные в качественном и количественном отношении. СПб, 1912, стр.106-125.

б) длительность конечных гласных после указанных групп согласных ;

в) длительность начальных гласных двусложных слов.

2. В трехсложных и многосложных словах.

ДОЛГОТА ГЛАСНЫХ В ДВУСЛОЖНЫХ СЛОВАХ

Наиболее четкое различие в длительности гласных языка суахили прослеживается при их положении перед разного рода согласными.

В связи с этим была исследована длительность ударных и безударных гласных перед смычными (глухими и звонкими), щелевыми (глухими и звонкими), носовыми и плавными согласными. В таблице I показано среднее время звучания гласных перед последующими согласными.¹⁾

Таблица I¹⁾

Ударный гласный					
перед смычным		перед щелевым		перед плавным	перед носовым
глухой	звонкий	глухой	звонкий		
160	170	173	180	перед 1 200 перед г 210	230

На основе данных, приведенных в таблице I, представляется возможным установить следующую иерархию длительности ударного гласного: наименьшее звучание - перед смычным

¹⁾ Средние данные получены в результате анализа 25-30 реализаций слогов и слов.

глухим, затем наблюдается увеличение длительности в позициях перед смичным звонким, перед щелевым глухим, щелевым звонким, плавным (особенно [r]) и носовым.

Как известно, в безударном положении звуки менее устойчивы в своей длительности, чем в ударном.

Некоторые лингвисты указывали на различную длительность гласных разного качества при прочих равных условиях.

Так, В.А.Богородицкий писал о том, что "... звучание узких гласных несколько длительнее по сравнению с средне-широкими".¹⁾

Л.В.Щерба, наоборот, утверждал, что узкие гласные короче широких.

Долгота ударных гласных в языке суахили исследовалась нами в словах, в которых гласные находились в одинаковых фонетических условиях. Данные эксперимента сведены в таблицу 2.

Таблица 2

Слова	Гласный	Длительность гласных
tambo	a	160
tombo	o	158
tembo	e	110
tumbo	u	90
timbo	i	73

1) В.А.Богородицкий. Фонетика русского языка в свете экспериментальных данных. К., 1930, стр.7.

Таким образом, мы имеем основание предполагать, что в языке суахили длительность узких гласных меньше, чем длительность широких.

Исследование гласных в середине двусложных слов дало возможность судить об определенной закономерности изменения длительности этих гласных.

При исследовании конечных гласных слов целесообразным явилось их изучение в положении после тех же групп согласных.

Долгота конечного неударного гласного во многом зависит от стиля и темпа речи, однако и здесь его длительность имеет такую же зависимость от согласных, как и в середине слова.

Небезынтересным оказалось выявление длительности ударных и конечных безударных гласных в разговорном и полном стилях речи и их сопоставление (см.табл.3).

Таблица 3

Слова	Стили речи			
	разговорный		полный	
	ударный гласный	безударный гласный	ударный гласный	безударный гласный
dada	220	115	228	140
baba	216	120	230	152

При изменении стиля речи (но при одном темпе) длительность ударного гласного изменяется незначительно, конечного неударного — меняется намного заметнее.

Конечные безударные гласные короче ударных: в разговорном стиле речи длительность безударных гласных почти в два раза меньше длительности ударных, в полном стиле — приблизительно в полтора раза.

В языке суахили в тех случаях, когда слово начинается гласным без предшествующего согласного, длительность гласного ударного абсолютного начала несколько больше, нежели длительность ударного гласного начального слога, начинающегося с согласного (см. табл. 4).

Таблица 4

Слова	Долгота гласного в абсолютном начале слова	Слова	Долгота гласного начального слога, начинающегося с согласного
imba	184	mimi	96
ona	194	hodi	148

В языке суахили значительный отрезок длительности гласного звука может быть подвержен назализации.

Наиболее значительной оказалась назализация гласных в конце слов. Л. В. Щерба назализацию конечных звуков объяснял "общей вялостью артикулирующих органов, в том числе и мягкого неба, наступающей к концу слова".¹⁾

1) Л. В. Щерба. Русские гласные в качественном и количественном отношении С. — П., 1912, стр. 91.

Особо сильную назализацию в языке суахили можно наблюдать у гласных, стоящих в конце слова после носовых согласных, а также их соединений: mb, nd, ndz, ng, mv, nz, особенно mb .1)

В этих случаях назализируется весь отрезок длительности конечного гласного.

При исследовании ударных гласных, стоящих перед сочетаниями типа "носовой плюс согласный", оказалось, что эти гласные сильно назализируются, причем назализация захватывает примерно 3/4 длительности гласного. Особо подвержены назализации суахилийские гласные [u] и [ɔ].

Можно полагать, что долгота гласных, стоящих перед сочетаниями носовых с согласными, сокращается, она как бы отнимается носовым элементом консонантного сочетания.

Ранее это отмечалось также A.N. Tucker и E.O. Ashton на, которые указывали при этом случаи сильной назализации гласных.

ДОЛГОТА ГЛАСНЫХ В ТРЕХСЛОЖНЫХ И МНОГОСЛОЖНЫХ СЛОВАХ

Кроме отмеченных случаев изменения длительности гласных, удалось определить, что долгота гласных зависит от места ударения, а также от количества слогов в слове.

1) См. об этом A.N. Tucker, E.O. Ashton, Swahili phonetics. London, 1942, p. 84, где они отмечали, что "Vowels in syllables beginning with a nasal consonants (especially m) are often nasalized."

В языке суахили гласные в ударных и безударных слогах многосложных слов имеют различную длительность: гласные ударного слога звучат дольше (как и в двусложных словах), например:

o | n | a | n | a
100 | 205 | 120

w | a | l | i | k | u | j | a
130 | 84 | 144 | 60

Однако длительность гласного ударного слога в многосложных словах меньше длительности гласного ударного слога двусложных слов:

k | o | mbe
166

v i k | o | mbe
138

ú | so
160

w a l i k | u | i a
144

p | é | cha
130

pek | é | cha
116

Длительность конечного гласного слова меньше длительности конечного гласного первого слога:

w | a | l | i | k | u | j | a
130 | 84 | 144 | 60

th | e | l | a | th | i | n | i
94 | 66 | 138 | 84

Примеры показывают, что в суахилийских многосложных словах гласный второго предударного слога дольше гласного первого предударного и заударного слогов.

ДЛИТЕЛЬНОСТЬ СОЧЕТАНИЙ ГЛАСНЫХ

В языке суахили очень часты сочетания гласных звуков:

leo, ndio, kua, kaa, juu и т.д.

Кроме того, возможны слова, состоящие из одних гласных:

- ua "убивать" - -oa "жениться"
aau "он убивает" и т.д.

Р. Берт 1) отмечает следующие сочетания гласных как наиболее частые

a - e	e - a	a - u
a - o	o - a	i - o
e - i		
o - e	i - e	i - u
u - e	e - o	
	e - u	

В связи с этим возникают проблемы определения соотношения длительности сочетающихся звуков и фонематической трактовки данных сочетаний: составляют ли они сочетание двух самостоятельных фонем или же являются одной фонетически сложной фонемой.

1) F. Burt, Swahili grammar and vocabulary. London, 1910, p.9.

Осциллографический анализ подводит нас к той мысли, что в языке суахили сочетания гласных типа *ao, ai, oi* и т.д. не являются единым слогообразующим элементом, они распадаются на две самостоятельные единицы. Об этом свидетельствует, в частности, двувершинное распределение интенсивности в этих сочетаниях.

Немаловажное значение имеет также исследование того, как возникают данные сочетания. Чаще всего это связано с опусканием согласных: *l/r* перед гласными, что можно наблюдать в таких суахилийских словах как *lia, lea, kua* в то время, как в других языках группы банту (например, кикую) имеем *gira, gega, kũga* и т.д.

Это, с нашей точки зрения, может говорить в пользу делимости на слоги данных сочетаний и об их бифонематичности.

При сопоставлении долготы каждого гласного звука исследуемых сочетаний оказалось, что первый ударный гласный длиннее второго, конечного:

<i>ma l i m</i>	!	<i>a</i>	!	<i>o</i>
	!	296	!	190
<i>a m b</i>	!	<i>a</i>	!	<i>o</i>
	!	240	!	150
<i>n d</i>	!	<i>i</i>	!	<i>o</i>
	!	210	!	138
<i>l</i>	!	<i>o</i>	!	<i>o</i>
	!	254	!	200

Второй гласный, стоящий после гласного широкого образования, длительнее гласного, стоящего после гласного узкого образования, т.е. сочетания "ao", "eo", звучат дольше, чем сочетания "io", "uo".

Особо следует остановиться на сочетании двух одинаковых гласных с большим временем звучания в словах типа каа, жуу, ндоо.

В суахили длинные гласные встречаются и на стыке морфем (ninaandika : первое "a" - конец префикса настоящего времени, второе "a" - начало глагольной основы) и в пределах одной морфемы: каа, жуу.

Бифонематичность долгого гласного в первом случае не подлежит сомнению.

Учитывая положения Трубецкого¹⁾, Щербы и Зиндера²⁾ о соотношении длительности монофонематичных и бифонематичных долгих гласных, нами определена длительность долгих гласных и сопоставлена с длительностью других, кратких гласных, находящихся в тех же фонематических условиях.

Получены следующие результаты (при медленном темпе речи, см.табл.5):

1) N. Trubetzkoy. Zur allgemeinen Theorie der phonologischen Vokal-systeme (TCLP, t.1, 1929) 26, 50-53.

2) Л.Л.Зиндер. Общая фонетика. Л., 1960, стр.194.

Таблица 5

Сочетание	Длительность	Долгий гласный	Длительность
ao	340	aa	360
eo	298	oo	300
ua	280	uu	272

Длительность долгого гласного приближается к длительности сочетания двух разных кратких гласных; длительность долгого гласного не меньше длительности сочетания двух разных кратких гласных.

Это дает основание говорить о бифонематичности долгого гласного.

Кроме того, распределение интенсивности в исследуемых долгих гласных двухвершинно, что говорит о сочетании двух фонем.

Исследуемые сочетания звуков делятся на два слога:

ka	-	a	
ju	-	u	
ndo	-	o ^{I)}	и т.д.

При сопоставлении длительности ударного и неударного гласного оказалось, что наибольшая длительность присуща ударному гласному.

Это было замечено и другими исследователями.

I) F. Burt. *Swahili grammar and vocabulary*. London, 1910, p. 3.
D.V. Perrot. *Teach yourself books. Swahili*. London, 1964, p. 6.

Petryankina V.I.

L.E.Ph, I.O.L. M.S.U.

LENGTH FEATURE OF VOWELS IN SWAHILI

The article gives the results of an experimental study of duration of vowels. The study made it possible to determine the duration of vowels in Swahili in stressed and unstressed positions, to establish the vowel hierarchy with respect to this sounding particular feature, to trace the nature of change in duration of stressed and unstressed vowels under various phonetic conditions for various speech styles (colloquial and literary), to give characteristics ^{of} vowel nasalisation, and to give a phonematic treatment of sound combinations. The vowel duration was measured in bi-syllabic words with open syllables (stressed and unstressed), before the consonants of various quality and after them.

A study was made of the duration of initial and final vowels of bi-syllabic, three-syllabic and multi-syllabic words.

В.И.ПЕТРЯНКИНА
ЛЭФ ИВЯ ПРИ МГУ

ВРЕМЯ ЗВУЧАНИЯ СОГЛАСНЫХ ЯЗЫКА
СУАХИЛИ

Выяснение времени звучания согласных языка суахили показало, что наиболее долгими являются носовые согласные. Все другие согласные по времени звучания располагаются в последовательности: глухие щелевые, глухие смычные, звонкие смычные, звонкие щелевые и наиболее короткие - плавные, особенно [r] (см.таблицу I).

Таблица I

Время звучания начальных согласных

Смычные		Щелевые		Носовые	Плавные
глухие	звонкие	глухие	звонкие		
173	150	184	132	200	1 - 130 r - 100
Согласный в интервокальном положении					
160	134	176		192	1 - 110 r - 83

Определяя время звучания согласных звуков языка суахили, необходимо особо остановиться на длительности глухих смычных придыхательных [p^h], [t^h], [k^h] и определить их место в общей системе долгот согласных языка суахили.

При анализе осциллограммы нами учитывалась не артикуляторная, а сонорная длительность гласного звука, т.е. начало гласного определялось не началом взрыва (при сочетании глухого смычного согласного с гласным), а началом голосовых вибраций.

Л.Р.Зиндер справедливо говорит о том, что время смычки согласного необходимо относить к его общей длительности.

"Акустический эффект, производимый смычными, не сводится к моменту взрыва: звонкие смычные слышны и во время смычки, а в глухих отсутствие звучания до взрыва тоже следует рассматривать не как нуль, а как некую величину, хотя и отрицательную ... в реальном употреблении в речи "пауза" во время смычки имеет такое же значение, как и всякий другой звук речи".¹⁾

Для суахили характерны смычные, в которых хорошо слышно придыхание, следуемое за взрывом. "В придыхательных затворных, - писал А.И.Томсон, - после взрыва следует сильное выдыхание, производящее более или менее сильный шум в самой полости рта".²⁾

Придыхательность согласных увеличивается по мере увеличения длительности выдоха.

"При большой воздушности выдох продолжается некоторое время после взрыва, вызывая шум в полости рта, возбуждая ее собственные тоны.

1) Л.Р.Зиндер. Общая фонетика. Л., 1962, стр.125.

2) А.И.Томсон. Общее языковедение. Одесса, 1910, стр.190.

Характер последних обычно зависит от последующего гласного, так как язык во время придыхания переходит к тому положению, которое необходимо для артикуляции этого гласного. При слабой воздушности выдох после взрыва очень краток, и согласный не имеет придыхания".¹⁾

На осциллограммах смычка представлена нулевой линией, а взрыв - рисунком высокочастотного шума, как у глухих щелевых.

Анализ осциллограмм слов с придыхательными согласными [p^h], [t^h], [k^h] и их сопоставление с непридыхательными [p], [t], [k] показали что:

1. В языке суахили в одних и тех же фонетических условиях возможно употребление обеих групп смычных согласных (придыхательных и непридыхательных). "Когда воздушность согласного связана с определенным фонетическим положением, она, разумеется, не может служить различительным фонематическим признаком, но, очевидно, способствует разграничению позиционных оттенков фонем".²⁾

В суахили смычные согласные различной длительности могут появляться в одинаковых фонетических условиях, что дает возможность говорить о фонематической значимости придыхательных согласных

2. Придыхательные согласные имеют гораздо большую длительность, чем непридыхательные.

1) Л.Р.Зиндер. Общая фонетика. Л., 1962, стр.116.

2) Л.Р.Зиндер. Уч.зап.Лен-го у-та. Вопросы фонетики, Л., 1964, стр.12.

3. Наибольшей длительностью характеризуется заднеязычный [к^h].

4. При быстром произнесении длительность придыхательных согласных сокращается примерно на 18% общей длительности звука.

Такая редукция "заключается, прежде всего, в нарушении резкости взрыва, связанном с ослаблением смычки, что легко может быть объяснено сокращением длительности согласных в быстром произношении".¹⁾

5. Длительность придыхательных смычных в большинстве случаев превосходит длительность целевых смычных согласных, но не превышает длительности носовых.

ДЛИТЕЛЬНОСТЬ МЯГКОГО СОГЛАСНОГО

В суахили мягким согласным, который можно сопоставить с соответствующим твердым, является согласный [ɰ] (ny)

Исследование показало, что мягкий [ɰ] дольше твердого [n] :

o! n !a n a	-	p o! ny !o k o
! 170!		! 190 !

! n !a n i	-	! ny !a ma
! 165 !		! 186 !

Мягкий носовой [ɰ] по длительности можно поставить перед носовыми [m] и [n] .

1) Л.Р.Зиндер. Общая фонетика. Л., 1962, стр.117.

ДЛИТЕЛЬНОСТЬ ЗВОНКИХ И ГЛУХИХ СОГЛАСНЫХ

В языке суахили длительность звонких согласных меньше длительности соответствующих глухих.

Длительность звонких смычных больше длительности звонких щелевых.

Следовательно, учитывая данные по времени звучания всех групп согласных, можно расположить их в следующем порядке: мягкий носовой, носовые, придыхательные глухие смычные, глухие щелевые, глухие смычные, звонкие смычные, звонкие щелевые, плавные.

Длительность согласных ударных слогов суахилийских слов больше длительности согласных неударных слогов:

m	!	i	!	m	!	i
I92	!		!	I40	!	

d	!	a	!	d	!	a
20	!		!	I2	!	

th	!	e	!	la	!	th	!	i	!	n	!	i
40	!				!	I72	!				!	

o	!	n	!	a	!	n	!	a
I64	!				!	I10	!	

o	!	n	!	a
I50	!			

Согласный звук в сочетаниях с другими согласными сокращается по времени звучания по сравнению с согласными интервокального положения:

n	!	d	!	o	!	v	!	u
I40	!							

o	!	n	!	a	!	n	!	a
I64	!							

Большинство сочетаний согласных в суахили составляют "носовой [м], [н] плюс смычный".

Длительность консонантных сочетаний с [м] оказалась больше длительности сочетаний с [н].

Сопоставление длительности каждого компонента внутри сочетаний выражается отношением 10:1, т.е. длительность носового согласного звука приблизительно в 10 раз больше длительности второго согласного:

<u>! ! m ! b ! a</u>	<u>! a ! m ! b ! a</u>	<u>n y u ! m ! b ! a</u>
! 120 ! 12 !	! 146 ! 16 !	! 130 ! 14 !
<u>n ! d ! o v u</u>	<u>f u ! n ! d ! i</u>	
! 140 ! 14 !	! 116 ! 12 !	

Консонантные сочетания из двух одинаковых звуков для суахили нетипичны и встречаются редко (в заимствованных словах или диалектах).

По своей длительности они приближаются к сочетаниям двух других согласных; преобладающим является первый компонент:

<u>! a ! l ! l ! a h</u>
! 188 ! 50 !

Для суахили характерно наличие слоговых согласных [м] и [н].

В связи с этим небезынтересно поставить вопрос о длительности этих согласных.

Выполняя функцию слогообразующих, согласные [м] и [н] обладают длительностью намного большей, чем соответствующие

неслогообразующие:

$\frac{\text{!} \text{ m } \text{!}}{\text{!} 214 \text{!}}$ tu , $\frac{\text{!} \text{ m } \text{!}}{300 \text{!}}$ bwa , $\frac{\text{!} \text{ n } \text{!}}{283 \text{!}}$ ne - $\frac{\text{!} \text{ m } \text{!}}{\text{!} 140 \text{!}}$
 $\frac{\text{!} \text{ m } \text{!}}{200 \text{!}}$ a m b o .

Если встречаются два одинаковых согласных, один из которых слоговой, то каждый из них характеризуется различными долготами:

$\frac{\text{!} \text{ n } \text{!}}{283 \text{!}}$ n $\frac{\text{!} \text{ n } \text{!}}{80 \text{!}}$.

Длительность слогообразующих согласных приближается к длительности слогообразующих гласных, а иногда и превосходит ее:

$\frac{\text{!} \text{ m } \text{!}}{214 \text{!}}$ tu - $\frac{\text{!} \text{ o } \text{!}}{194 \text{!}}$ na . $\frac{\text{!} \text{ d } \text{!}}{\text{!} 228 \text{!}}$ a da .

Petryankina V.I.

L.E.Ph., I.O.L. M.S.U.

DURATION OF CONSONANTS IN SWAHILI

The article gives the results of an experimental study of duration of consonants. The consonant hierarchy is established with respect to their duration. The author analyses the duration of the voiceless plosives (pronounced with aspiration) /p^h/, /t^h/, /k^h/, gives the length characteristics of the soft consonant /ɸ/ comparing it to the corresponding dark /n/ and /m/. Data are also given concerning the duration of voiceless and voiced consonants, consonant clusters and the syllabic consonants /m/ and /n/.

М.К.РУМЯНЦЕВ
ЛЭФ ИВЯ при МГУ

ВРЕМЯ ЗВУЧАНИЯ ЭЛЕМЕНТОВ СЛОГА (на материале китайского языка)

Время звучания элементов слога определяется двумя факторами: общеслововым временем и количеством структурных элементов слога. Существуют определенные пропорции в распределении времени по элементам слога. Эти пропорции не являются, по-видимому, межязыковыми. В какой-то степени они определяют национальное своеобразие слога, а также отличают одну реализацию слога от другой в каждом данном языке.

Структура слога, его внутренние механизмы экспериментально-фонетически изучены еще очень мало, хотя сведения, которые мы могли бы здесь получить, очень важны как в теоретическом, так и прикладном плане. Попытки синтеза слогов и слов из отдельных звуков речи выявили важнейшую роль переходных процессов внутри слога и слова.¹ Двух, трех и четырехэлементный слог — это не просто следование звуков речи друг за другом, а их взаимная приспособляемость — естественный переход одного звука в другой.² Эта приспособляемость и создает единство слога, впечатление того, что воспринимаемые нами звуки, не самостоятельные единицы, а входят в единый блок, как его составные элементы. Переход одного звука в другой создается последовательной сменой артикуляций, и эта смена

1. P. Menzerath, *Der Diphthong*, Bonn und Berlin, 1941.

C.M. Harris, *A study of the building blocks in speech* JASA, 1953, vol. 25, N 5.

2. А.Абеле, К вопросу о слоге, *slavia*, 1924-25 г., т.3, № 1, стр.19.

артикуляций при малой протяженности звучания (внутри слога) — процесс более тонкий, нежели смена артикуляций между слогами в слове. Смена артикуляций (приспособляемость одного звука к другому), временные пропорции элементов слога и цельность слога — все это между собой связано: одно обуславливает другое.

Совершенно не изучен вопрос о том, что происходит с временными пропорциями элементов слога в тех случаях, когда увеличивается или уменьшается его время звучания. Сохраняются ли эти пропорции и каковы пределы их варьирования? Когда, например, речь идет о медиалах и конечнослоговых элементах, существенным является не только то, до каких пределов эти элементы могут сокращаться, а и то, до каких пределов они могут распространяться, так как в каждой данной реализации должна существовать граница их распространения. В противном случае слог потеряет свое национальное своеобразие, будет неестественным либо превратится в сочетание двух слогов. Пределы же сокращения звучания этих элементов обуславливаются, прежде всего, нашими возможностями их восприятия. Если элемент сократится настолько, что не будет воспринят, слог из двух или трех-элементного превратится в одно- или двухэлементный. Кроме того, и это основное: сокращение звучания медиалей и конечных элементов находится в зависимости от времени звучания основной гласной; при большом времени звучания основной гласной не может быть очень малого звучания медиалей и, наоборот. Должна быть соблюдена определенная пропорция. В противном случае, слог утратит свои характерные черты, если не разрушится совсем.

Аудитивная оценка слогов варьированных по временным пропорциям их элементов, позволяет предположить, что переходные процессы и временные пропорции элементов слогов функционально не дублируют друг друга, хотя два этих фактора и тесно связаны между собой. Так, например, сами по себе, без переходных участков звучания правильные временные пропорции между элементами не создают у слушающего впечатления цельности фонетического комплекса. Справедливо и другое: нарушение временных пропорций между элементами (при сохранении переходных участков звучания) приводит слог к таким внутренним сдвигам, которые предопределяют его неестественные реализации либо вообще лишают право называться слогом данного языка.

В создании цельности слога, как известно, участвует еще и третий фактор — распределение интенсивности по элементам слога, но это, безусловно, предмет специального исследования.

Выяснение временных соотношений элементов слога приобрело в настоящее время особое значение в связи с проблемой автоматического (машинного) распознавания речи. Может ли слог быть единицей распознавания? Это во многом зависит от того, насколько он вариабелен в речи, сколь устойчивы его акустические характеристики. Чтобы ответить на этот вопрос нужны достоверные данные о реализациях акустических характеристик слога в речи.

В наших экспериментах выяснение временных пропорций элементов слогов проводилось: I) на более или менее нейтральных реализациях слогов, вне четко ощутимых интонационных влияний или в какой-то одной интонационной схеме. Экспери-

ментальный материал при этом составляли отдельные слоги, произнесенные в определенном порядке по заранее разработанной инструкции; 2) на реализациях слогов в связной речи различного темпа. В этом случае анализировались временные пропорции элементов слогов при большом варьировании общего времени звучания слога. Нейтральные реализации в нашем материале представлены как интонограммами, так и спектрограммами. Связная речь - только интонограммами.

ВРЕМЯ ЗВУЧАНИЯ МЕДИАЛИ СЛОГА

М е д и а л ь - i . При исчислении времени звучания медиали - i было проанализировано свыше 900 спектрограмм и интонограмм различных слогов. При этом выяснилось, что временные пропорции медиали - i к основной гласной в значительной степени варьируют в зависимости от с т р у к т у р ы с л о г а: от числа его элементов и наличия или отсутствия инициали в составе слога. В нашем экспериментальном материале двухэлементные безинициальные слоги (iā) характеризуются меньшим отношением времени звучания медиали - i к основной гласной, чем слоги трехэлементные (iāo). Среднее отношение времени звучания медиали - i к гласной -a в слоге iā (из 100 реализаций) оказалось равным 0,377 ($\bar{C} = 0,0404$), в слоге iāo - 0,492 ($\bar{C} = 0,0836$). Обусловлено это тем, что в трехэлементном слоге iāo часть общеслогового времени забирает конечный элемент слога - o. Основная гласная (-a) при этом укорачивается. Медиаль же - i в слоге iāo либо остается равной медиали - i в слоге iā, либо укорачивается, но не на такое время, кото-

рое занял конечный элемент (-o), а на меньшее. В результате такого перераспределения времени и происходит некоторая динамика пропорций: отношение времени звучания медиали к основной гласной увеличивается.

Существует два варианта реализации медиали - i : реализация при нуле инициали (медиаль сама начинает слог) и реализация после инициали. В двух и трехэлементных слогах без инициали (типа iā, iāo, iān, iāng) медиаль -i, как правило, более продолжительна, чем в слогах с инициалью (типа xiā, xiāo, jiā, qiā, tiāo). Объясняется это тем, что слоги с последовательностью элементов: "инициаль + медиаль - i + основная гласная" представляют собой фонетические структуры со слитной реализацией комплекса "инициаль + медиаль". Медиаль в этой реализации прочно спаяна с инициалью. Акустически в таких случаях легче произвести сечение не по традиционному шву: "инициаль / / + медиаль + основная гласная + конечный элемент" (например, x//iā, d//iān), а после медиали: "инициаль + медиаль // + основная гласная + конечный элемент" (xi//ā, di//ān).

Так обстоит дело даже при непалатализирующихся инициалах.¹⁾ При палатализации же слияние инициали с медиалью - i еще более прочное.²⁾ Акустически медиаль входит в реализацию пала-

1) К инициалам, непалатализирующимся перед гласным - i (слоговым и неслоговым) относятся губные p, b, m переднеязычные t, d и сонорные n, l . См. А.А. Драгунов и Е.Н. Драгунова, Структура слога в китайском национальном языке. Советское Востоковедение № 1, 1955, изд. АН СССР, стр. 62.

2) Палатализируются перед - i свистящие согласные ts', dz', s . См. А.А. Драгунов и Е.Н. Драгунова. Там же.

палатализованной инициали как обязательный ее компонент.³⁾ Основная гласная может следовать непосредственно за палатализованной согласной: $x\bar{a}$, $j\bar{a}$, $q\bar{a}$. Время звучания медиали - i в таких слогах может быть минимальным - 30-40 мсек. И действительно, наименьшие значения во временных отношениях (0,2 и меньше) зафиксированы нами в слогах с палатализованными инициалами, хотя в целом слоги с палатализованными и непалатализованными инициалами характеризуются в наших экспериментах одними и теми же средними значениями отношений. Так, в слове ($t\bar{i}ao$) с непалатализуемой придыхательной инициальной t среднее отношение времени звучания медиали - i к гласной - a (52 реализации) составило 0,208 ($\sigma = 0,0334$).

В безинициальном слове того же гласного состава ($i\bar{a}o$) среднее отношение (из 58 реализаций) равно 0,492 ($\sigma = 0,0836$).

Фактором, определяющим малую величину отношения в слогах типа $t\bar{i}ao$, является малое время звучания самой медиали. Количественная структура слога представляет собой другой фактор, поддерживающий противоположную тенденцию - увеличения малых

3) Мягкость начального согласного в русском слове также, как известно, "связывается со специфическим i - образным звучанием начального гласного перехода". См. Л.В.Бондарко, Структура слога и характеристики фонем. Вопросы языкознания, № 1, 1967, стр.40, а также Л.Р.Зиндер, Л.В.Бондарко, Л.А.Вербицкая, Акустическая характеристика различия твердых и мягких согласных в русском языке, Сб. "Вопросы фонетики", 1964 (Уч.зап.ЛГУ), 325, серия филолог. наук, 69.

значений отношения медиали к гласной за счет сокращения времени звучания последней. Ведь в финали слога *tiāo* - три элемента и, следовательно, на основную гласную -а приходится меньшая временная норма, чем в слогах с двухэлементной финалью, таких как *xiā*, *qiā*, *jiā* и т.д. Этот второй фактор (особенно в связной речи) нередко оказывается весьма действенным. Основная гласная в трехэлементной финали реализуется в крайне сжатое время. Медиаль - *i* при этом также сокращается, но не пропорционально сокращению гласной, в связи с чем временные отношения этих элементов сдвигаются к большим значениям.

Наиболее значительную порцию общего времени трехэлементной финали забирают конечные - *n* и -*ng* (см. об этом стр. 384).

В наших общих материалах, представленных четырехэлементными слогами разного состава и разных тонов, среднее отношение времени звучания медиали - *i* к гласной -а составляет 0,495 (из 100 реализаций).

Варьирование отношений в этих слогах осуществлялось в пределах от 0,2 до 1. Крайние значения шкалы варьирования временных отношений зафиксированы лишь в отдельных реализациях. Отношение 0,2 - в одной реализации (1%), 1 - в трех реализациях (3%); отношения от 0,21 до 0,3 - в 6%, от 0,31 до 0,35 - в 12%, от 0,36 до 0,5 - в 44%, от 0,51 до 0,7 - в 19%, от 0,71 до 0,8 - в 5%, от 0,81 до 0,9 - в 3%.

ТОН И ВРЕМЕННЫЕ ОТНОШЕНИЯ МЕЖДУ ЭЛЕМЕНТАМИ СЛОГА

При анализе временных отношений элементов слога весьма существенной является проблема функциональной значимости вы-

являемых пропорций в типологическом плане. Что определяют эти пропорции? Слог как таковой – в любом языке или слог данной языковой системы? Отличаются ли, например, по временным отношениям элементов китайские слоги $i\bar{a}$, $i\acute{a}$, $i\check{a}$, $i\grave{a}$ и немецкий или русский слоги $\bar{a}$, $\acute{a}$? Ответить на этот вопрос можно будет только тогда, когда будут выявлены временные пропорции между элементами слогов в языках различных систем. В пределах нашего материала мы пытались установить, изменяются ли временные пропорции между элементами слога в зависимости от тона. Если такая зависимость существует, то вопрос об универсальности временных пропорций вряд ли может быть решен положительно, так как слоги тональных языков будут отличаться своим временным распределением от слогов языков нетональных. Да и сами различия внутри одной языковой системы, если они связаны с тоном, будут свидетельствовать не только о физиологических, но и об их лингвистических функциях.

Выяснение зависимости временных пропорций элементов слога от тона осуществлялось на двухэлементном слоге, состоящим из медиали $-i$ и слогаобразующей гласной $-a$ в разных тонах. Сравнивались реализации: $i\bar{a}$, $i\acute{a}$, $i\check{a}$, $i\grave{a}$.

При записи на магнитофонную пленку четыре тона были произнесены одним диктором. Каждый тон по 100 раз.^I Интонационные условия произнесения уравнивались единой синтаксической конструкцией фразы, с цитируемым слогом в качестве

I. Все 400 произнесений были осуществлены сразу – за четыре сеанса. За один сеанс диктор наговаривал 100 фраз. Перерыв между сеансами продолжался 10–15 минут.

объекта исследования. Например *iā zì shì Lǎo Lǐ xiě de* "Иероглиф *iā* написан почтенным Ли".²

С магнитофонной пленки сигналы подавались на 48 канальный спектроанализатор конструкции ЛЭФИПР.

Спектрограммы звуков, зафиксированные на 35 мм кинопленке, представляют собой последовательности кадров — спектральных срезов, произведенных через каждые 15 мсек звучания.

По спектральным характеристикам срезов визуально прослеживалось, что происходит с данным звуком в исследуемом слого через каждые 15 мсек его звучания: остается ли он на данном срезе по своему качеству тем же звуком, который был представлен предшествующим срезом, например, звуком — *i* или изменяется, становится более открытым и в последующих срезах переходит в звук — *a*. Грубо говоря, задача сводилась к тому, чтобы по спектрограммам определить участки звучания — *i* и — *a* в слого *iā*.

В начале данного раздела было упомянуто, что перцептивное ощущение цельности слога, а не простой последовательности звуков создается взаимным проникновением его элементов друг в друга. При этом взаимопроникновении важную роль играют и исследуемые нами временные пропорции элементов слога.

-
2. Этот прием уравнивания условий произнесения слогов, несколько его видоизменив, мы заимствовали у Р.Л.Бротзмана. См. R.L.Brotzman. Vowel Formant Values. The Ohio State University. Project on linguistic analysis. Report N 6. 1963.

В звучании любого слога, состоящего более чем из одного элемента, можно определить временные отрезки, на которых осуществляется взаимопроникновение элементов, или иначе, идут "переходные" процессы.

Современные исследования на материале различных языков вскрывают большую и разноплановую информативность этих процессов. Естественно, что взаимопроникновение элементов происходит на тех участках звучания, которыми они соприкасаются друг с другом.

Отрезки времени, на которых осуществляются переходные процессы, принадлежат не одному какому-нибудь элементу, а сразу двум, поэтому и границы этого соприкосновения, как правило, нечетки. Все это очень затрудняет процедуру измерения времени звучания элементов слога, но не делает ее совсем невозможной. Ошибки возможны как при определении времени звучания переходных участков, так и при их условном сегментировании, т.е. решении отнести те или иные отрезки переходного звучания к предшествующему или последующему элементу слога. Но эти ошибки вряд ли могут существенно помешать выяснению общих временных пропорций элементов слога.

Медиаль — $i\bar{a}$ в слове $i\bar{a}$ вся (от начала до конца) представляет собой переходный процесс: от закрытого (узкого) иотированного элемента к образованию открытому $[e\bar{a}]$ — образному, переходящему в —а. Стационарных участков в звучании медиали — i почти нет. Но в наших целях трудно интерпретируемым является не все звучание медиали — i , а лишь тот ее участок, который непосредственно переходит в —а. Обычно это один — два спектральных среза, т.е. 15–30 мсек

звучания. В наших измерениях время непосредственно-переходного участка, т.е. того спектрального среза, за которым следовало явное -а, делилось пополам, как принадлежащее не только медиали - i, но и основной гласной -а. Иными словами, сегментируя переходные участки непосредственного стыка медиали - i и гласной -а, можно ошибиться в интерпретации одного спектрального среза: отнести его к звучанию предшествующего элемента слога, а не последующего или, наоборот. Ошибка в интерпретации двух срезов менее вероятна.

Анализ времени звучания элементов слогов $i\bar{a}$, $i\acute{a}$, $i\check{a}$, $i\grave{a}$ по спектрограммам показал, что принадлежность к тону приводит к некоторым сдвигам временных пропорций медиали и основной гласной. И этот сдвиг обусловлен тем, что тоны различаются своим временем звучания. Основу же временных различий тонов составляет закономерное варьирование времени звучания основной гласной. Медиаль при этом может быть более или менее постоянной. Связь между временем звучания тона и отношением времени звучания медиали к основной гласной обратна пропорциональная: чем продолжительнее тон (основная гласная), тем меньшее отношение. Четвертый тон (самый короткий) характеризуется самым большим отношением медиали к основной гласной.

В нашем экспериментальном материале среднее отношение времени звучания медиали - i к основной гласной оказалось равным: первый тон - 0,377 (100 реализаций) ($\bar{C} = 0,0404$),
 второй тон (100 реализаций) - 0,351 ($\acute{C} = 0,0530$),
 третий тон (100 реализаций) - 0,439 ($\check{C} = 0,0631$),
 четвертый тон (100 реализаций) - 0,485 ($\grave{C} = 0,0627$).

Третий тон в своей полной реализации (с конечной восходящей частью) - самый продолжительный, и отношение времени звучания его медиали к звучанию основной гласной должно быть меньшим, чем у второго тона. В наших данных это отношение больше, чем у второго и первого тонов. Объясняется это тем, что все реализации третьего тона в описываемом эксперименте не были полными. В синтаксической конструкции *lǎo lǐ xiě de* третий тон слога *lǎ* перед четвертым тоном слога *xiě* по правилам сандхи давал комбинаторный вариант третьего тона - реализацию без конечной восходящей части. При этом время звучания его основной гласной было меньшим, чем у второго и первого тонов.

Полные реализации третьего тона в других наших экспериментах характеризовались средним отношением времени медиали к основной гласной равным 0,286.

Исчисление средних значений является весьма показательным при сравнении величин подверженных тому или иному варьированию. Применительно к данному эксперименту средние значения (средние отношения) дают возможность сравнить между собой группы варьированных реализаций медиали и основной гласной: группу реализаций первого тона со вторым, третьим и четвертым. Это, безусловно, важно, если цель исследования заключается в том, чтобы выяснить, зависят ли временные пропорции медиали и основной гласной от тона.

Лингвистически не менее важной представляется и другая задача: выявить пределы варьирования временных отношений, узнать, при каких временных отношениях медиали и основной гласной слог остается равным самому себе - воспринимается

как данный слог китайского языка, и какие соотношения ему противопоставлены. При такой постановке вопроса показательными оказываются уже не средние отношения, а отношения конкретных реализаций.

Диапазон варьирования временных отношений медиали и основной гласной в нашем материале лежит в пределах:

первый тон 0,288 - 0,522

второй тон 0,267 - 0,476

третий тон 0,296 - 0,577

четвертый тон 0,371 - 0,645

Временные отношения между медиалью и основной гласной варьируют лишь в определенных пределах. И особенно существенно здесь то, что в связной речи наблюдаются те же временные пропорции, что и в изолированно произнесенных слогах. Время звучания слога в целом может быть большим или малым, временные пропорции медиали и основной гласной не выходят при этом за пределы некоторой области допустимого варьирования.

В нашем эксперименте представлена следующая временная шкала слоговых реализаций.

Первый тон от 344 до 179 мсек звучания.

Второй тон от 615 до 254 мсек.

Третий тон от 359 до 194 мсек.

Четвертый тон от 299 до 209 мсек.

При значительном уменьшении времени звучания слога отношение времени звучания медиали к основной гласной сдвигается к большим значениям допустимого варьирования. Происходит это потому, что медиаль представляет собой эле-

мент менее вариабельный, чем основная гласная. Более или менее безболезненно для восприятия этот элемент сокращается лишь до известного предела. В безинициальных слогах типа $i\bar{a}$, $i\bar{a}o$, $i\bar{a}n$, $i\bar{a}ng$ таким пределом являются ≈ 50 мсек звучания. При большом сокращении времени звучания основной гласной, например, до 120 мсек даже 50 мсек звучания медиали составят отношение 0,416. Медиаль в 60 мсек займет уже половину звучания гласной.

В описываемом эксперименте наблюдается большая устойчивость в звучании медиали $-i$. При значительных различиях во времени звучания основной гласной медиаль остается постоянной, либо варьирует весьма незначительно. Время звучания медиали составляет ≈ 80 мсек (в 324 реализациях из 450) при звучании основной гласной в пределах 307 - 142 мсек. При времени звучания гласной в 187 мсек и менее медиаль сокращалась иногда до ≈ 67 мсек. Увеличение времени звучания медиали до ≈ 100 мсек фиксировалось в реализациях со временем звучания основной гласной 187 мсек и более. Из 450 реализаций лишь в четырех случаях медиаль $-i$ звучала свыше 100 мсек: $i\bar{a}$ - II2 + 232, $i\bar{a}$ - I27 + 262, $i\bar{a}$ - I27 + 202, $i\bar{a}$ - II2 + 262.

Если среднее время звучания медиали $-i$ в безинициальных слогах принять за 100 мсек, то крайними границами ее оптимальных реализаций будет время звучания $\approx$ в 50 и 150 мсек.

То, что сокращение времени звучания медиали $-i$ до 50 мсек не влияет существенным образом на восприятие слогов $i\bar{a}$, $i\bar{a}$, $i\bar{a}$, $i\bar{a}$ подтверждается экспериментом стирания на-

чального участка медиали. В результате стирания начальных участков звучания медиали - i были получены реализации с различной продолжительностью ее звучания: 15, 22, 37, 52, 67 мсек.

Реализации, в которых была стерта вся медиаль, за исключением последних 15 мсек, представляющих собой непосредственный переход к гласной -а, воспринимались аудиторам как слог а с небольшим э-образным призвуком. При 22 миллисекундной медиали слог уже воспринимается как ia, но удаление предшествующего участка звучания - i еще явно чувствуется. Слоги с медиалью в 37 мсек квалифицировались аудиторам как почти естественные, еле отличимые от обычных реализаций. При времени звучания медиали - i в 52 мсек слог воспринимались как естественные.

При быстром темпе речи такое время медиали - i (52 мсек) зафиксировано нами во многих реализациях. Длительность гласной -а при этом была минимальной: 127 - 142 мсек.

Выше мы отметили, что другую границу оптимальных реализаций медиали - i составляет время равное $\approx$ 150 мсек. Такое и даже большее время звучания медиали - i зафиксировано нами у некоторых дикторов в изолированных произнесенных слогах. При этом наблюдается одна особенность: первая половина звучания медиали - i только намечена слабым сонантообразным спектром, не имеющим характерной для - i двухформантной структуры. Такая структура в рассматриваемых реализациях присуща только второй половине медиального звучания, т.е. её [εæ] - образной части. Занимает она от 50 до 100 мсек.

Эксперименты со стиранием начальных участков звучания медиали - i показывают, что эти участки не несут существенной для восприятия информации.¹

Сонантообразное начало присутствует почти во всех реализациях медиали - i, но обычно оно не занимает более 15-30 мсек звучания. В изолированно произнесенных слогах это время иногда более чем удваивается.

М е д и а л ь - и. В реализации медиали - и наблюдаются два участка: участок и-образного звучания, непосредственно примыкающий к инициали и вызывающий лабиализацию последней, и участок перехода от -и к -а. Все звучание медиали -и занимает обычно 40-80 мсек. Причем 40 мсек - это звучание близкое к пороговому. Дальнейшее сокращение времени существенным образом сказывается на восприятии не только самой медиали, но и всего слога, так как сокращение медиального времени предполагает жесткую временную норму для всего слога.

60-80 мсек звучания медиали - и фиксируется в изолированно произнесенных слогах, т.е. при большой общеслоговой временной норме. Реализация же этой медиали за время большее, чем 80-100 мсек повлечет за собой еще большее увеличение общеслогового времени. В слогах трехэлементного состава (типа

1. Аналогичные данные были получены ранее Н.И.Дукельским на материале русского языка. В его характеристиках начальных сонантов мы находим такое замечание: "Зачеркивание начала речевого потока, соотносимого с j- в сочетаниях j+vc оказалось возможным в пределах ~ 70% общей длительности согласного. Часть j-, составляющая ~ 0,5 его общей длительности, распознается хорошо". Н.И.Дукельский, Принципы сегментации речевого потока. Изд. АН СССР, Москва, Ленинград, 1962, стр.36.

huā) отношение времени звучания медиали -и к гласной меньше, чем в слогах четырехэлементных (типа huān). Это, по-видимому, можно считать общим правилом временных пропорций для всех медиалей и конечных элементов слога: увеличение элементов слога (финали) сдвигает отношение к большим значениям. В нашем материале среднее отношение времени звучания медиали -и к гласной в слогах типа huā равно 0,253, в слогах типа huān - 0,439.

ВРЕМЯ ЗВУЧАНИЯ КОНЕЧНОСЛГОВЫХ НАЗАЛЬНЫХ -n И -ng

В исследовании Р.Л.Бротцмана Vowel Formant Values ¹⁾ приводятся предварительные данные о соотношении времени звучания назальных окончаний слога -n, -ng и основной гласной. Автор публикации специально оговаривает, что материал, которым он пользовался, был ограниченным (limited number of samples) и проанализированы были реализации только одного диктора: по 28 реализаций на каждый тон при назальном -n и по 21 реализации при -ng .

Определяя отношение времени звучания назальных конечных элементов -n и -ng к основной гласной, автор пришел к выводу, что "при велярном назальном отношении всегда больше, чем при альвеолярном; при велярном назальном отношении часто приблизительно вдвое больше, чем при альвеолярном".²⁾

1) R.L. Brotzman. Vowel Formant Values. The Ohio State University. Project on Linguistic Analysis. Report N 6. Research on Mandarin phonology. 1963.

2) Указан. сочин. стр. 10.

В общей вводной статье W. S. Y. Wang этот вывод формулирует следующим образом: "Когда за гласным следует альвеолярный носовой согласный, то гласный по своей долготе примерно равен этому альвеолярному согласному. Но если последующий согласный велярен, то гласный может составлять лишь одну четверть его долготы".¹⁾

Наши данные не позволяют нам полностью присоединиться к выводу Р.Л.Бротцмана.

Мы проанализировали время звучания назальных -n и -ng на двух экспериментальных материалах. Один материал представлял собой 315 реализаций альвеолярного -n и 308 реализаций велярного -ng в слогах различной количественной структуры. Как слогообразующие выступали гласные a, i, u, ə. Время звучания альвеолярного -n, велярного -ng и соответствующих гласных исследовалось в каждом тоне отдельно: первый тон альвеолярного -n - 100 реализаций, второй тон - 75, третий тон - 64, четвертый тон - 76, первый тон велярного -ng - 66 реализаций, второй тон - 91, третий тон - 81, четвертый тон - 70 реализаций. В этот материал включались реализации четырех дикторов. В него вошли как изолированно произнесенные слоги, так и, частично, слоги из связной речи.

При исчислении времени звучания назальных -n и -ng по спектрограммам и интонограммам мы не относили к ним конечный назализованный участок звучания гласного, непосредственно примыкающий к переходу от гласного к назальным конечным. Время же самого перехода от гласного к конечному назальному (30-40 мсек) делилось пополам: первая его поло-

1. Указан. сочин., стр. 6.

вина квалифицировалась как время гласного, вторая половина - согласного.

Статистическая обработка данных таких измерений показала, что в среднем отношение времени звучания велярного - *ng* к основной гласной примерно такое же, как и у альвеолярного - *n*. Оно варьирует в пределах от $\approx 0,5$ до 2 со средним значением ≈ 1 .

Некоторое варьирование временных отношений в данном материале обусловлено различием слогообразующих гласных. Так, например, отношение времени звучания назальных окончаний ко времени звучания гласной -*a* несколько меньше, чем к гласной -*e* (ср. слоги *fān* и *fēn*, *fāng* и *fēng*). Но этот фактор не может существенным образом повлиять на эксперимент, поскольку он равно действует внутри каждого тона с альвеолярным - *n* и велярным - *ng*.

Второй экспериментальный материал был составлен из слогов одной и той же количественной структуры при одной слогообразующей гласной -*a*: 482 реализации слогов *tān*, *tán*, *tān*, *tàn* и 480 реализаций слогов *tāng*, *táng*, *tāng*, *tàng*. Слог *tan*: первый тон - 122 реализации, второй тон - 130, третий тон - 115, четвертый тон - 115. Слог *tang*: первый тон - 117 реализаций, второй тон - 127, третий тон - 121, четвертый тон - 115.

Слоги были произнесены двумя дикторами по схемам, предусматривающим некоторый стереотип реализации.

В материале одного диктора слоги располагались строками: каждая строка - пять слогов, произносимых на одном выдохе, например, *tān*, *tán*, *tān*, *tān*, *tān*. Затем другая та-

там же строка и т.д. Другому диктору было предложено использовать схему Р.Л.Бротцмана (см.стр.376). Время звучания элементов слогов в этом материале исчислялось по интограммам. Принцип измерения был таким же, как и в предшествующем эксперименте: время переходного участка от -а к -п или -ng делилось пополам. Правильность сегментирования слога - определение участков звучания -а, перехода к назальным окончаниям, самих назальных - проверялась в данном эксперименте специальной операцией: удалением части звучания назальных окончаний.

С конца звучания назальных элементов удалялись отрезки различной продолжительности: $1/3$ всего звучания и меньше, $1/4$, $1/2$, $3/4$, весь назальный элемент, оставлялся только переходный участок или удалялся и он. Последующее прослушивание оставшейся части слога аудиторами, сравнение интограмм полных и урезанных реализаций давало возможность с большей точностью произвести сегментирование слога.

При такой обработке материала также не было обнаружено того разительного расхождения во времени звучания -п и -ng, о котором сообщает Р.Л.Бротцман. Средние значения отношения времени звучания -п и -ng к гласной -а приведены ниже.

Дик. I	tān - 1,430(=0,282)	tāng - 1,982(=0,451)
	tān - 0,874(=0,146)	tāng - 1,198(=0,1695)
	tān - 0,732(=0,1624)	tāng - 0,929(=0,1775)
	tān - 0,637(=0,0963)	tāng - 0,682(=0,0816)

Дик. 2	$\bar{t}\bar{a}n - 1,423(=0,1117)$	$\bar{t}\bar{a}ng - 1,306(=0,1535)$
	$\bar{t}\bar{a}n - 1,003(=0,0985)$	$\bar{t}\bar{a}ng - 1,023(=0,1068)$
	$\bar{t}\bar{a}n - 0,870(=0,0721)$	$\bar{t}\bar{a}ng - 1,105(=0,1233)$
	$\bar{t}\bar{a}n - 0,923(=0,0847)$	$\bar{t}\bar{a}ng - 1,367(=0,1722)$

Р.Л.Бротцман в цитированном выше исследовании пишет: "Нет заметного влияния тона на коэффициент временных отношений конечного назального согласного к гласному. Поэтому представляется, что хотя тон и оказывает влияние на время звучания гласного, но такому же влиянию подвержены и назальные окончания, в связи с чем коэффициент отношения сохраняется".^I

По нашим данным только в четвертом тоне происходит более или менее пропорциональное влияние тона на гласную и конечнослоговой назальный элемент. Что же касается других тонов, то во всем нашем материале мы наблюдали то же явление, что и при анализе времени звучания медиали - *i* и основной гласной. В зависимости от тона происходят некоторые сдвиги во временных пропорциях конечнослоговых назальных согласных (*-n*, *-ng*) и предшествующих им гласных, и эти сдвиги обусловлены тем, что по времени звучания наиболее вариабельным элементом слога оказывается его основная гласная. Время ее звучания закономерно варьирует от тона к тону. Изменения же (в зависимости от тона) времени звучания

I. There is no marked effect upon the ratio of nasal ending duration preceeding vowel duration as a function of tone. It therefore appears that although vowel duration is a function of tone, the nasal endings are similarly affected by tone so that the ratio is preserved. Цит. соч. стр. 10.

конечнослоговых элементов, как правило, непропорциональны изменению времени звучания гласной. Поэтому и наблюдается некоторая закономерная нестабильность во временных отношениях. Закономерность в большинстве случаев такая же, как и при медиали $-i$: чем продолжительнее тон (основная гласная), тем меньше отношение. Исключение составляет лишь четвертый тон, в котором фиксируется иногда не наибольшее, а наименьшее отношение.

Ср. средние отношения времени звучания конечнослоговых $-n$ и $-ng$ к гласной по тонам.

В первом материале

Во втором материале

			(дикт. I)		(дикт. 2)	
	$-n$	$-ng$	$-n$	$-ng$	$-n$	$-ng$
I тон I, OII		I, 049	I, 430	I, 982	I, 423	I, 306
2 " 0, 879		0, 933	0, 874	I, 198	I, 003	I, 023
3 " 0, 842		0, 829	0, 732	0, 929	0, 87	I, 105
4 " 0, 810		0, 822	0, 637	0, 682	0, 923	I, 367

В связи с приведенными выше данными резонно спросить, почему же в четвертом тоне фиксируется наименьшее отношение? (См. первый материал и второй, данные дикт. I).

Разгадка этого кроется, очевидно, в том, что в четвертом тоне назальные звучат более коротко, чем в других тонах. Время звучания конечных назальных в четвертом тоне таково, что оно оказывается меньшим, чем время звучания гласной в этом тоне, хотя гласная здесь и самая короткая по сравнению с другими тонами.

В принципе, видимо, не исключены реализации и других тонов, а не только четвертого, в которых соблюдается пропорциональность влияния тона на время звучания гласной и конечного назального элемента. Но эта пропорциональность функционально избыточна. Подтверждается это тем фактом, что как медиаль $-i$, так и конечные назальные согласные можно уравнивать по времени звучания во всех тонах, и это не отразится на восприятии ни слога, ни тона. Отношение времени звучания медиали $-i$ или конечных назальных к гласной будет меняться только в зависимости от времени звучания гласной в том или ином тоне.

Описанный выше эксперимент с удалением части звучания конечных назальных элементов показывает, что удаление 70%, а иногда даже 80% звучания $-n$ или $-ng$ не оказывает влияния на восприятие слога. Слог воспринимается правильно и достаточно естественно (не чувствуется срез части назального звучания), если сохранен переходный участок от гласного к назальному (≈ 30 мсек) и 30-40 мсек звучания самого назального. Дальнейшее увеличение времени назального на 30-40 мсек обеспечивает абсолютную естественность реализации слога. Если 30-40 мсек звучания конечного назального представляет собой порог естественной реализации слога, то 70-80 мсек звучания дает ему определенный "запас прочности". Причем, в процентном отношении 70-80 мсек звучания часто составляли в наших примерах не половину, а только 40 или 30% всего звучания назального; 30 же или 40 мсек звучания назального составляли примерно 20% всего звучания, а иногда и меньше.

Восприятие конечных назальных в исследуемых слогах в очень большой степени предопределяется переходным участком звучания от гласного к назальному. В этом убеждают те реализации, в которых оказались полностью удаленными сами - n или - ng и остались только переходные участки - 30-40 мсек. Только эти участки, без дальнейшего какого-либо подкрепления воспринимаются уже как более или менее ясный намек на - n или - ng. Если же после переходного участка оставались неудаенными еще хотя бы 15-20 мсек звучания, то эти реализации воспринимались как слоги с конечным назальным элементом. Правда, их нельзя квалифицировать как естественные реализации, в них чувствуется срез назального звучания.

Из приведенных выше данных о средних значениях отношений видно, что в изолированно произнесенных слогах и связной речи назальные окончания - n и - ng также продолжительны, как и слогообразующие гласные, а нередко они даже превосходят гласные по времени звучания. Причем абсолютные временные значения этих элементов весьма значительны. В реализациях изолированных слогов t̄an, t̄an, t̄an, t̄an, t̄ang, t̄ang, t̄ang время звучания - n и - ng варьировало в пределах: (дикт. I)

t̄an	- 170 - 380 мсек	t̄ang	- 200-440 мсек
t̄an	- 120 - 280 "	t̄ang	- 160-290 "

I. "Усечение" конечных назальных происходило и в процессе исторического развития китайского языка. Известно, что в некоторых диалектах отпали конечные назальные, осталась только назализация гласного.

tān	- 160-280 мсек	tāng	' - 200-340 мсек
tān	- 80-160 " 2)	tāng	- 80-190 "

В связной речи (в стереотипной конструкции, приведенной на стр. 376).

Дикт. 2.

tān	- 110-140 мсек	tāng	- 110-150 мсек
tān	- 110-130 "	tāng	- 110-140 "
tān	- 100-150 "	tāng	- 120-170 "
tān	- 100-130 "	tāng	- 110-160 "

Эксперимент с удалением части звучания назальных окончаний показывает, сколь избыточными оказываются временные реализации этих элементов, на какой "запас прочности" рассчитаны их акустические манифестации. Примечательно то, что большая накладка "лишнего" времени - n и - ng происходит в тех реализациях, в которых слогаобразующая гласная звучит продолжительно. При достаточно большом времени звучания гласной мы не зафиксировали ни одного случая (из 1500 произнесений), когда бы время звучания конечных элементов слога -n и - ng было близко к предельному. При малом же времени звучания гласной такие реализации встречаются.

-
- 2) Эти данные говорят о том, что конечнослоговое - n по времени звучания оказывается не соотношенным с начальнословым - n .

Таблица I

Слог	Количество реализаций	Время звучания
		гласной -а от до
Tān	100	130 - 220
tān	100	170 - 300
tān	100	200 - 340
tān	100	120 - 250
tāng	100	120 - 200
tāng	100	120 - 240
tāng	100	220 - 340
tāng	100	140 - 240

В таблице I указано время, в пределах которого реализовалась гласная -а в каждом из четырех тонов. Данные таблицы отвечают на вопрос, какое время имеется в виду, когда мы говорим о достаточно большом времени звучания гласной. Именно при таком времени звучания гласной, которое указано в табл. I, мы не зафиксировали ни одного случая, когда бы звучание -п и -ng было близко к предельному времени.

Время звучания конечнослоговых -п и -ng, равное 80, 70, 60 и 50 мсек, наблюдалось нами в тех слогах, в которых гласная имела, примерно, такое же время звучания - от 50 до 100 мсек.

Сопоставление реализаций с большим и малым временем звучания гласных и конечнослоговых -п и -ng показывает, что временные пропорции между гласной и -п, -ng в любом

случае сохраняются, примерно, одни и те же. Например, звучания гласной в 180 мсек и $-n$ в 220 мсек дают отношение 1,222, и, примерно, то же отношение (1,2) представляют звучания гласной в 50 мсек и $-n$ в 60 мсек. Различие этих реализаций заключается в том, что одна из них предельная по времени звучания тона в целом (50+60) и близкая к предельной по звучанию конечнослогового $-n$. Другая реализация с большими запасами как в отношении времени звучания тона (180+220), так и конечнослогового $-n$.

Следует, однако, иметь в виду, что в тонированном слоге рассматриваемые проблемы осложнены тем, что надежность акустической манифестации должна быть обеспечена сразу по двум аспектам: восприятию слога и восприятию тона. И здесь не всегда можно наблюдать полную согласованность. Если, например, для восприятия слогов $t\acute{a}n$ и $t\acute{a}ng$ всегда достаточно 30-40 мсек звучания назального $-n$, $-ng$ (слоги при этом будут хорошо восприниматься как состоящие из трех элементов), то для надежного восприятия второго (восходящего) тона в этих слогах 30-40 миллисекундного времени может быть недостаточно.

Это связано с тем, что акустические признаки тона локализируются не только в звучании гласного элемента слога, но и во всей финали. В назальных окончаниях, в частности, продолжат нарачиваться частотные интервалы тонов - восходящие и нисходящие, поэтому сокращение времени звучания назальных окончаний чревато уменьшением интервалов тонов и, следовательно, уменьшением надежности их восприятия. В нашем эксперименте безболезненное для восприятия тона сокращение

времени назальных до порогового значения фиксировалось в слогах с достаточно большим временем звучания слогообразующей гласной. Это и понятно; если за время звучания гласной "успевали" сформироваться достаточно четкие акустические признаки тона, не опасным становилось и значительное сокращение времени звучания назальных окончаний.

ВРЕМЯ ЗВУЧАНИЯ КОНЕЧНОСЛОГОВОГО - i

Время звучания конечнослогового -i выяснялось, в основном, на реализациях слогов *bāi, bái, bǎi, bāi*. Для каждого тона от одного диктора было получено по 50 реализаций. Правильность сегментирования участков звучания проверялась той же операцией, какая применялась и в других случаях (например, при сегментировании слогов *tān* и *tāng*) : во всех реализациях от конца звучания конечного - i удалялись отрезки различной продолжительности. Затем производилось аудирование оставшегося звучания и сравнение интонограмм полных и урезанных реализаций, соотнесение результатов слухового восприятия с количественными данными урезанных реализаций. По времени звучания конечнослоговой элемент - i обычно менее продолжителен, чем основная гласная. Отношение времени звучания конечнослогового - i к слогообразующей гласной в большинстве случаев меньше, чем соответствующее отношение к гласной конечнослоговых - n и - ng. Ср. средние значения отношений:

Дикт. I

Дикт. 2

	- n	- ng	- n	- ng	- i
I тон	1,430	1,982	1,423	1,306	0,922
2 тон	0,874	1,198	1,003	1,023	0,763
3 тон	0,732	0,929	0,870	1,105	0,934
4 тон	0,637	0,682	0,923	1,367	0,751

Конечнослоговые -n, -ng, и -i похожи друг на друга тем, что в отличие от конечных -o и -u у них не так жестко определен предел их временного распространения; они легко растягиваются, особенно в первом и третьем тонах. Мы попытались определить предел временного распространения этих элементов слога, указать тот порог, за которым начнутся подчеркнутые либо неестественные для обычной речи (утрированные) реализации, пение.

Наши данные свидетельствуют о том, что в изолированно произнесенных слогах отношения 2,5 и даже 3 лежат еще в пределах естественных реализаций, причем такая продолжительность конечных -n, -ng, -i обусловлена надежностью их восприятия в слоге или тоном. Так же как -n и -ng конечнослоговое -i хорошо воспринимается в слоге, если сохранен переходный участок от слогообразующей гласной к конечному -i и еще 30-40 мсек звучания самого -i. Некоторая порция времени (50-60 мсек) сверх этого минимума гарантирует надежное восприятие тона. Все же остальное время конечнослогового -i с точки зрения восприятия избыточно. Следует иметь в виду, что переходный участок от слогообразующей гласной к конечному -i довольно продолжителен - 50-80 мсек.

Слоги типа *кӣ* в классификационной таблице

А.А. и Е.Н. Драгуновых соотносятся с безинициальным слогом *wei* (*uei*), в котором элемент *-и* является медиалью.¹ Авторы отмечают, что при наличии инициали реализация финали *-uei* зависит от тона; "в полном виде она произносится в 3-м и 4-м тонах - [*-uei* 3,4]; в 1-м и 2-м тонах вершина слога смещена на [*и*], а компонент [*э*] играет подчиненную роль, так что финаль произносится [*-uei* 1,2]"².

Анализ временных отношений между гласными элементами слогов типа *кӣ* показывает, что такая интерпретация этих слогов вполне справедлива. Следует, пожалуй, сказать только, что даже в слогах первого и второго тонов основным слоогообразующим элементом является не *-и*, а *-ei*. Главным образом, в этом дифтонгальном элементе формируются тональные характеристики слога.

По времени звучания элемент *-ei* намного превосходит *-и*. В нашем материале (50 реализаций слога *кӣ*) среднее отношение времени звучания *-ei* к *-и* оказалось равным 3,3. В рассматриваемой группе реализаций временные отношения варьировали в пределах от 2-х до 5. Такое количественное преобладание *-ei* над *-и* резко контрастирует с теми временными отношениями, которые мы наблюдали в слогах типа *вāi*, хотя и там переходный участок от *-а* к *-i* характеризуется э-образной реализацией.

1. А.А. и Е.Н. Драгуновы, Структура слога в китайском национальном языке. Советское востоковедение № I, 1955, таб.17.

2. Там же, стр.67.

ВРЕМЯ ЗВУЧАНИЯ КОНЕЧНОСЛОГОВЫХ - о и - у

Время звучания конечнослоговых - о и - у выяснялось на реализациях слогов iāo, iǎo, iǎo, iǎo, tiāo, tiǎo, tiǎo, tiǎo, tōu, tōu, tōu, tōu (по 50 реализаций на каждый тон).

В первом и третьем тонах конечнослоговые - о и - у в рассматриваемых реализациях по времени звучания оказались, приблизительно, равными основной гласной. В четвертом тоне - о и - у самые короткие; они почти в два раза короче основной гласной. Во втором тоне время звучания - о и - у приближается ко времени звучания гласной (см. средние значения отношений).

Конечнослоговое - о

Конечнослоговое - у

I тон	1,260	I,18
2 тон	0,856	0,876
3 тон	0,931	I,284
4 тон	0,689	0,676

Если по отношению времени звучания конечнослогового элемента к гласной сравнивать -i, -o и -u, то можно констатировать, что в фактических реализациях - о и - у фиксируются отношения, близкие к тем, которые дает конечнослоговой - i. Но конечнослоговые - о и - у строго ограничены в своей протяженности. Их нельзя растягивать так, как - n, - ng или -i. Энергетическое распределение при произнесении этих элементов таково, что оно исключает большую протяженность зву-

чания. По своему спектру конечнослоговые -о и -и заметно отличаются от -о и -и слогообразующих. Попытка протянуть конечнослоговые -о и -и придает им качество слогообразующих, что влечет за собой расщепление слога.

Эксперимент с удалением части звучания конечных - о и -и показывает, что они хорошо воспринимаются в слоге при сохранении переходного участка звучания (от слогообразующей гласной к -о или -и) и 30-40 мсек звучания самих элементов. Переходный участок звучания здесь так же, как и у слога *bai* (*baí*, *bái*, *baí*) весьма продолжителен - 50-80 мсек.

Rumyantzev M.K.

L.E.Ph., I.O.L. M.S.U

TIME OF SOUNDING OF SYLLABLE ELEMENTS
(on Modern Chinese material)

The article sets forth the results of an experimental study of the time of sounding of the syllable elements: medials - i, -u, final nasals -n, -ng, final elements of the syllables -i, -o, -u. Mean values are given for ^{ratio of the} ~~the~~ sounding time of the elements indicated to the sounding time of the main vowel.

The time proportions of the syllable elements were established: 1) on more or less neutral syllable realisations, the experimental material consisting of separate syllables pronounced in a definite order according to prearranged instructions; 2) on syllable realisations in connected speech of various rates.

Within the bounds of the material studied the author attempts to make clear whether change of the time proportions between the elements of the syllable depends on tone.

Е.К. ФЕДОРОВА

ДЭФ ИВЯ при МГУ

ФОНЕТИЧЕСКАЯ РЕАЛИЗАЦИЯ НЕКОТОРЫХ ГЛАГОЛОВ И
СООТВЕТСТВУЮЩИХ ИМ ПО ЗВУЧАНИЮ ПРЕДЛОГОВ В
СОВРЕМЕННОМ КИТАЙСКОМ ЯЗЫКЕ

Принято считать, что формализация знаменательных слов в современном китайском языке сопровождается редукцией то-
на¹⁾. Такие формальные элементы, как суффикс совершенного
вида *le*, суффикс неопределенно-многократного вида *guo*,
суффиксы имен существительных *z* и *tou* произносятся лег-
ким (редуцированным) тоном.

В современном китайском языке грамматисты выделяют це-
лую группу так называемых бифункциональных слов (комплексов),
выступающих в речи то как глаголы, то как предлоги.²⁾ В эту
группу входят: *yòng, ná, tì, bǎ, gěi, xiàng, wǎng, shàng,*
duì, gēn, dào, cháo, zài.

А.А. Драгунов отмечал семантическую близость знаменатель-
ной и служебной функции некоторых глаголов.³⁾ Отсутствие
семантического разрыва между глагольным и предложным значе-
ниями у вышеперечисленных слов отмечает в диссертации
С.Б. Янкивер.⁴⁾ Бифункциональные комплексы подразделяются ея
на две группы. В первую группу включаются слова, характери-

1) А.А. Драгунов "Исследования по грамматике современного
китайского языка" М-Л. 1952 г., стр. 13, 125.

2) А.А. Драгунов, там же, стр. 124.
С.Б. Янкивер "Предлоги в китайском языке" канд. дисс. М. 1965 г.

3) А.А. Драгунов, там же, стр. 124-127.

4) С.Б. Янкивер, диссертация, стр. 15, 44.

зующиеся единым значением в глагольной и предложной функциях, т.е. бифункциональные слова *dao*, *shang*, *wang* и *chao*. Все остальные бифункциональные комплексы С.Б.Янкивер относит к единицам переходного типа, обладающим полисемией в предложной функции и моносемией в глагольной, причем предположные значения бифункциональных комплексов "группируются не вокруг одного из предложных значений, а вокруг глагольного значения".¹⁾

В синологической литературе высказываются различные мнения о категориальной принадлежности слов, выступающих то как глаголы, то как предлоги; по разному оценивается и степень их формализации. Небезынтересно в связи с этим рассмотреть фонетическую реализацию этих слов, то есть выяснить, происходит ли утрата этимологического тона у этих слов, когда они выступают в качестве предлога?

Для этого необходимо проанализировать основные акустические характеристики данных слов в тех случаях, когда они выступают как глаголы и как предлоги; иными словами сравнить акустические характеристики глаголов и предлогов по: времени звучания, частоте колебаний основного тона и амплитудам слоговой интенсивности.

Время звучания (длительность произнесения слога) выступает как один из существенных признаков тона. Оно представляет собой основу для других его признаков: во времени происходит движение частоты F_0 , формируются определенные частот-

1) С.Б.Янкивер, диссертация, стр.39, 44, 52.

ные интервалы, характеризующие тот или иной тон.¹⁾ Установлено, что редукция тона сопряжена со значительным сокращением времени звучания слога и утратой частотных интервалов.²⁾

Сопоставление временных и частотных характеристик исследуемых объектов может предоставить нам существенную информацию о том, как происходит в них тонообразование. При описании китайских слогов, важным признаком, характеризующим тон, является и интенсивность. Каждый из четырех тонов китайского языка характеризуется определенным распределением интенсивности.³⁾

Сравнение акустических характеристик осуществлялось на материале таких синтаксических структур, в которых интересующие нас слова находились в одинаковом лексическом окружении. Например:

Tā ná zhè zhǐ gāngbǐ	"Он берет эту ручку" и
Tā ná zhè zhǐ gāngbǐ xiě zì	"Он пишет этой ручкой".
Tā néng yòng cídiǎn	"Он может пользоваться словарем" и
Tā néng yòng cídiǎn fānyì	"Он может переводить со словарем".
Wǒ tì jiějie	"Я буду вместо старшей сестры" и
Wǒ tì jiějie zuò fàn	"Я приготовлю еду вместо старшей сестры".

1) М.К.Румянцев, "Тон и интонация в современном китайском языке", Сб. "Спорные вопросы грамматики китайского языка", М., 1963г., стр.36.

2) М.К.Румянцев, "Время звучания слога в тональных языках (на материале китайского языка)", Сб. "Материалы colloquiuma по экспериментальной фонетике и психологии речи", МГУ, 1966г.

3) М.К.Румянцев, "Интонация предикативных и атрибутивных синтагм", Сб. "Спорные вопросы строя китайского языка", М.1965г., стр.21-22.

Т.П.Задоевко, "Краткий очерк системы тонов современного китайского языка", Сб. "Вопросы китайской филологии", М. 1963 г., стр.194-196.

Хотя все бифункциональные слова обладают общим свойством, могут в предложении выступать и в качестве самостоятельного слова, и в качестве служебного, однако, каждому из этих слов присущи и свои особенности. Так одна из групп бифункциональных слов употребляется в качестве предлога, как правило, только в дистантной препозиции к глаголу. В данном разделе мы и рассмотрим акустические характеристики некоторых слов этой группы: *ná, yòng, tí, cháo* в качестве глаголов и соответственно в качестве дистантных препозиционных предлогов.

Акустическая реализация *ná* глагола и *ná* предлога

Глагол *ná* обладает всеми грамматическими особенностями глаголов китайского языка. Его значение - "брать", "взять", "держат в руках", "нести", "схватить".¹⁾ Выступая в качестве предлога, *ná* может перед сказуемым вводить: 1) инструментальное дополнение (как и предлог *yòng*), например: *shāobi xiě zhōngguózi* "писать кисточкой по-китайски"; 2) прямое дополнение (как и предлог *bǎ*), например: *ná ta dāng fùqin* "считать его отцом".²⁾ Для того, чтобы сравнить акустические характеристики *ná* глагола и *ná* предлога на интонографе были записаны 9 предложений, где *ná* выступает в качестве глагола, и 9 предложений, в которых *ná* является предлогом. Эти 18 предложений были записаны в произнесении трех дикторов. В результате мы получили и проанализировали 54 реализации слова *ná*. Ниже приводятся данные анализа интонограмм.

1) Китайско-русский словарь, ред. И.М. Ошанин, М. 1952, стр. 327

2) Там же, стр. 327-328; С.Б. Янкивер, диссертация, стр. 50.

Время звучания. Как глагол, так и предлог на́ произносятся вторым (восходящим) тоном. Второй тон по времени звучания следует непосредственно за самым продолжительным третьим тоном.¹⁾

При исчислении времени звучания на́ глагола и на́ предлога мы измеряли как финаль, так и инициаль слога. В большинстве произнесений инициаль (п) давала в интонаграфической записи синусоиду с периодическими (квазипериодическими) колебаниями.

По времени звучания на́ глагол отличается от на́ предлога. Глагол в большинстве произнесений звучит продолжительнее, чем предлог. Различие во времени звучания глагола и предлога в среднем составляет 31 мсек от общего среднего звучания глагола 230 мсек и предлога 196 мсек ; наибольшее расхождение, зафиксированное в нашем материале, составляет 90 мсек. Небольшая часть реализаций (4 случая из 54) характеризовалась обратным соотношением долгот: предлог имел большее время звучания, чем глагол ; различие во времени звучания предлога и глагола в среднем составило всего лишь 17,5 мсек от общего среднего звучания предлога 240 мсек и глагола 212,5 мсек ; в этом случае наибольшее расхождение по времени звучания равно 60 мсек. Из 27 сравниваемых пар глагол звучит продолжительнее в 19 случаях, в 4 случаях глагол равен по времени звучания предлогу, в 4 случаях время звучания у предлога большее, чем у глагола.

¹⁾ М.К.Румянцев, "Тон и интонация в современном китайском языке", стр.36-37.

Характеристики F_0 . 1) Выше мы отмечали, что второй тон по движению частоты F_0 — восходящий, но акустически в его реализации может фиксироваться и начальное падение частот. 2) Значительно варьирует он и по интервалу восхождения. 3)

Мы исследовали движение частоты основного тона у глагола $на$ и у предлога $на$. Примерами повышения второго тона у глагола $на$ без нисходящей части в начале звучания могут служить следующие предложения: (дик.1) № 1 W_0 $на$ $kuai z$ "Я беру палочки для еды". Глагол $на$ в этом предложении произносится в интервале 150–210 гц (ув.кварта) с последовательностью нарастания частот: 150–160–175–175–190–200–210–210 гц⁴); (дик.2) № 2 Ni $на$ $tā d qian$ "Ты берешь его деньги". Глагол $на$ реализуется в интервале 210–300 гц (ув.кварта) с последовательностью частот: 210–210–220–220–225–240–260–285–290–300 гц. Такой тип реализации второго тона (без нисходящей части в начале звучания) зафиксирован нами и у предлога $на$, например: (дик.2) № 3 W_0 $на$ $kuai z chī fan$ "Я кушаю палочками"; предлог реализовался в интервале 200–310 гц (м.секста) с последовательностью нарастания частот: 200–205–210–210–210–225–240–250–260–290–310 гц. (дик.2) № 4 $Tā$ $zhī$ $на$ $kuai z chī fan$ "Он кушает только палочками"; предлог $на$ реализовался в интервале 190–310 гц (секста) с нарастанием частот: 190–210–210–210–225–240–250–260–290–310 гц.

-
- 1) Символом F_0 обозначается основная частота в речевом сигнале — основной тон (fundamental frequency).
 2) М.К.Румянцев, "Интонационное выражение завершенности и незавершенности высказывания". "Интонация и звуковой состав", МГУ, 1965 г., стр.121.
 3) М.К.Румянцев, "Время звучания слога в тональных языках".
 4) Каждое значение представляет 20 мсек участок звучания.

Интервалы повышения тона в приведенных примерах: ув.кварта (глагол) и м.секста, секста (предлог).

Как у глагола, так и у предлога при данном типе реализации второго тона наблюдались и меньшие интервалы восхождения частоты F_0 ; в частности и у глагольной, и у предложной реализаций зафиксированы интервалы в м.терцию: (дик.1) № 5 $n\acute{i} \ n\acute{a} \ t\bar{a}d \ q\acute{i}an$ "Ты берешь его деньги" ; глагол $n\acute{a}$ реализуется в интервале 175-210 гц с последовательным нарастанием частот: 175-175-175-175-190-190-200-200-210-210-210 гц. № 6 $T\bar{a} \ zh\acute{i} \ n\acute{a} \ ku\grave{a}iz \ ch\bar{i}f\grave{a}n$ "Он кушает только палочками" ; предлог $n\acute{a}$ также реализовался в интервале 175-210 гц (м.терция) с нарастанием частот: 175-175-175-175-175-190-210-210 гц.

Интервалы повышения тона меньшие, чем м.терция наблюдаются, как правило, в другом типе реализаций второго тона, а именно там, где повышению частоты F_0 предшествует некоторое ее падение. В рассматриваемых случаях начальное падение частоты F_0 приходится на звучание инициали n . Примерами таких реализаций могут служить предложения: № 7 $T\bar{a} \ xi\bar{a}n \ n\acute{a} \ b\grave{a}ozh\acute{i}$ "Он сначала берет газеты" ; № 8 $T\bar{a} \ lu\grave{a}n \ n\acute{a} \ h\acute{o}ng\grave{a} \ zu\grave{o} \ zh\grave{e}ge$ "Он делает это, хватая красное ...". В этих предложениях интервал повышения второго тона у глагола $n\acute{a}$ (предложение № 7) у дик.3 равен секунде, а у I и 2 дикторов - м.секунде ; у предлога $n\acute{a}$ (предложение № 8) у 2 диктора интервал восхождения тона также равен м.секунде, а у I диктора - секунде.

Второй тип реализации второго тона (с начальным падением частоты F_0) статистически доминирует. Он дает наиболее вариативную картину по интервалам: от самых больших до самых малых.

У глагола *ná* наибольший частотный интервал в 3 случаях составляет м.сексту (8 полутонов), а наименьший частотный интервал в 3 произнесениях составляет м.секунду (I полутонов). Между этими крайними реализациями по частотному интервалу располагаются реализации с интервалами меньшими, чем м.секста и большими, чем м.секунда. В 2 случаях интервал составляет квинту, в 6 случаях - ув.кварту, в 3 случаях - кварту, в I случае - терцию, в 7 случаях - м.терцию и в 2 случаях - секунду.

У предлога *ná* полярными частотными интервалами являлись: секста (9 полутонов, I раз) и м.секунда (I полутонов, 5 раз). Варьирование по интервалу внутри указанной полосы было следующими: м.секста - I раз, ув.кварту - 4 раза, кварта - I раз, терция - 4 раза, м.терция - 6 раз и секунда - 3 раза.

Интервалы повышения тона у глагола *ná* и предлога *ná* в целом ряде случаев различны. Так, в предложении № 9 (дик. I) *Wó ná kuàiz* "Я беру палочки для еды" интервал повышения у глагола *ná* равен ув.кварту (6 полутонов), а в предложении № 10 *Wó ná kuàiz chīfan* "Я кушаю палочками" интервал повышения у предлога *ná* составляет м.терцию (3 полутона). В предложении № 11 *Nǐ ná tād qián* "Ты берешь его деньги" (дик. 3) у глагола *ná* интервал повышения тона равен ув.кварту (6 полутонов), а в предложении № 12 *Nǐ ná tād qián mǎi zhè běn shū* "Ты покупаешь эту книгу на его деньги" у предлога *ná* интервал повышения - м.терция (3 полутона). В 15 сравниваемых парах из 27 глагол *ná* характеризуется большими интервалами повышения тона, чем

предлог на́. Только в 3 реализациях интервал повышения оказался большим у предлога. В 7 случаях он одинаков.¹⁾

Если оценить акустические реализации на́ глагола и на́ предлога с точки зрения регистра, то надо отметить, что в большинстве произнесений предлог на́ реализуется в более высоком частотном регистре. Из 27 сравниваемых пар в 14 случаях частотный регистр был выше у предлога на́ и только в 4 произнесениях – у глагола на́. В 6 случаях регистр глагола и предлога примерно одинаков, хотя в одном случае при одинаковом регистре в средней части, начало и конец звучит у предлога выше, чем у глагола.

Результаты сопоставления акустических характеристик глагола на́ и предлога на́ говорят о том, что в реализации глагола и предлога имеются некоторые различия (эти различия были ранее зафиксированы нами и у глагола – предлога gei²⁾). Сводятся они к следующему:

1) глагол на́ в большинстве произнесений звучит продолжительнее, чем предлог на́ ;

2) интервалы повышения тона во многих произнесениях у глагола на́ большие, чем у предлога на́ ;

3) амплитуды слоговой интенсивности у глагола на́ в большинстве случаев меньше, чем у предлога на́ .

Для выяснения вопроса о том, существенны ли эти различия в языковом отношении, был произведен эксперимент на вос-

1) В 2 реализациях движение частоты основного тона у дик. I и 2 оказалось понижающимся.

2) Е.К. Федорова, "Фонетическая реализация gei глагола и gei предлога в современном китайском языке". Сб. "Материалы коллоквиума по экспериментальной фонетике и психологии речи", М. 1966 г.

приятие. Он заключался в следующем: глагол и предлог на́ вычленились в магнитофонной записи из контекста и предъявлялись для опознавания аудиторам-китайцам. Это позволило нам соотнести акустические характеристики и их восприятие и, таким образом, выяснить, ассоциируются ли в современном языке акустические характеристики предлога на́ со 2 тоном.

При эксперименте на восприятие глагол на́ и предлог на́ вычленились соответственно из тех предложений, где они имели:

1) достаточно большое время звучания (глагол от 220 мсек до 280 мсек; предлог от 190 мсек до 260 мсек);

2) меньшее время звучания (глагол - 180 мсек, предлог - от 160 мсек до 180 мсек);

3) значительный интервал повышения тона (от кварты до м.септимы);

4) меньшие интервалы повышения тона (секунда-терция).

Для аудитивного восприятия были выбраны реализации предлогов с большим временем звучания и значительным интервалом повышения тона потому, что при малом времени звучания может происходить нивелирование акустических признаков тона¹⁾ безотносительно к тому, выступает данное слово в служебной функции или нет. Достаточное время звучания слога представляет собой первейшее условие для акустической развертки тональных характеристик.

1) М.К.Румянцев. "Время звучания слога в тональных языках".

В эксперименте принимали участие 4 аудитора. Все аудиторы легко узнавали второй тон при реализации глагола *на* со временем звучания от 180 до 280 мсек при интервале повышения тона секунда - м.секста.

Все аудиторы сразу определяли второй тон у предлога *на* со временем звучания от 160 до 260 мсек, при интервале повышения тона секунда - м.септима.

В предложении № 13 (дик.3) *Tā xiān nà bàozhǐ* "Он сначала берет газету" второй тон глагола *на* реализовался как тон нисходяще-восходящий. Причем, нисходящая часть была образована последовательностью частот: 360-340-340-275-250-225-200 гц, восходящая часть образована последовательностью частот: 200-200-225 гц. Интервал начального падения тона - м.септима, а интервал конечного повышения - секунда. Эту реализацию три аудитора восприняли как 4 тон^{I)}, один аудитор как второй тон. Аналогичная "ошибка" при восприятии второго тона у предлога *на* произошла и в следующем случае: второй тон предлога *на* у диктора 3 в предложении № 14 *Tā xiān nà bàozhǐ dāng cǎiliào* "Он прежде всего считает материалом газету" также реализовался как нисходяще-восходящий тон. Нисходящая часть была образована последовательностью частот: 425-400-375-350-325-300 гц, восходящая - 300-340-350 гц. Интервал начального падения тона - ув. кварта, интервал конечного повышения тона-секунда. Эту реализацию один из аудиторов воспринял как первую половину 3 тона (3 низкий тон), три других аудитора приняли ее за третий тон.

I) О случаях подобного восприятия см. И.К. Румянцев "Интонационное выражение завершенности и незавершенности высказывания", стр. 121-122; "Время звучания слога в тональных языках", стр. 210.

Вместе с тем, в некоторых реализациях, где второй тон глагола и предлога *на́* также реализовался как нисходяще-восходящий тон с преобладающей нисходящей частью, аудиторы все же узнавали второй тон. Например, в предложении № 15 (дик.2) *Tā néng ná cídiǎn* "Он может взять словарь" у глагола *на́* время звучания 240 мсек, нисходящая часть образована последовательностью частот: 360-360-325-290-280-280-270 гц, восходящая часть - (270)-280-280-280-280-290 гц. Интервал начального падения тона-кварта, а интервал конечного повышения тона-м.секунда. Три аудитора восприняли это *на́* как произнесенное вторым тоном, один - как I тоном. Возможно, что такое восприятие данной реализации глагола связано с распределением интенсивности¹⁾, - линия интенсивности заметно смещена в сторону восходящей части (см.график1).



В предложении № 16 (дик.1) *Tā luàn ná hóngdào zuò zhège* "Он делает это, хватая красное" у предлога *на́* время звучания 260 мсек, нисходящая часть образована последовательностью

1) М.К.Румянцев, "Время звучания слога в тональных языках", стр.210-211.

частот: 260-250-250-225-210-200-175 гц, восходящая часть - 175-175-190-190-200-200 гц. Интервал начального падения тона-квинта, интервал конечного повышения тона - м.терция. Все четыре аудитора восприняли это как второй тон. Конечная фаза в реализации предлога здесь также более интенсивна, чем начальная.

Амплитуды слоговой интенсивности. Каждый тон характеризуется определенным распределением интенсивности. Так, у второго тона интенсивность возрастает к концу произнесения.

Сравнивая амплитуды слоговой интенсивности у глагола *па́* и у предлога *па́*, мы можем отметить, что как у глагола, так и у предлога интенсивность к концу произнесения значительно повышается. Большие амплитуды интенсивности фиксируются у предлога *па́*.

Сопоставляя данные 27 сравниваемых пар, мы можем заключить, что в 16 случаях амплитуды слоговой интенсивности больше у предлога; в 5 случаях - у глагола и в 6 случаях они одинаковы. В каких же случаях амплитуды слоговой интенсивности у предлога *па́* больше, чем у глагола *па́*? Если соотнести амплитуды интенсивности со временем звучания *па́* глагола и *па́* предлога, то можно отметить следующее: большие амплитуды у предлога *па́* фиксируются, как правило, при меньшем времени звучания.

Таким образом, данные интонограмм по основным акустическим характеристикам *па́* глагола и *па́* предлога дают нам возможность отметить следующее:

1. Глагол *па́* звучит продолжительнее предлога *па́*.
2. Амплитуды же слоговой интенсивности у глагола *па́* меньшие, чем у предлога *па́*.

3. Этимологический тон глагола и предлога *на́* (второй тон) по движению частоты основного тона реализуется по обычным правилам, как восходящий тон.

4. Глагол *на́* характеризуется большими интервалами повышения тона нежели предлог *на́*.

5. В большинстве произнесений предлог *на́* реализуется в более высоком частотном регистре.

Акустическая реализация *yòng* глагола и *yòng* предлога.

Глагол *yòng* означает: "пользоваться", "употреблять", "применять", "использовать", "расходовать", "потреблять".¹⁾

А.А.Драгунов, говоря о глаголах, "которые могут употребляться в двух функциях – знаменательной и служебной, в роли глагола-предлога", отмечает, что глагол-предлог *yòng* лишен индивидуального тона.²⁾

Для эксперимента на интонографе были записаны 8 предложений, где *yòng* является глаголом, и 8 предложений, в которых *yòng* используется как предлог. В записи участвовало 2 диктора. В общей сложности мы располагали 32 реализациями морфемы *yòng*.

Ниже мы приводим данные анализа интонограмм.

Время звучания. И глагол, и предлог *yòng* произносятся 4 (нисходящим) тоном. По времени звучания – это самый короткий тон. Анализируя время звучания глагола *yòng* и предлога *yòng*, мы можем констатировать, что продолжитель-

1) Китайско-русский словарь, редак. И.М.Ошанин, М.1952 г., стр.361.

2) А.А.Драгунов, "Исследования по грамматике современного китайского языка", М. 1952 г., стр.124-125.

ность звучания глагола *yòng* в большинстве произнесений бо́льшая, чем у предлога. Из 16 сравниваемых пар глагол *yòng* звучит продолжительнее предлога в 13 случаях; предлог звучит дольше глагола только в одном случае (дик.1). В двух случаях время звучания глагола и предлога *yòng* одинаково. Различие по времени звучания глагола *yòng* и предлога *yòng* составляет в среднем, примерно, 30 мсек при общем среднем звучании глагола 210 мсек и предлога 190 мсек. Наибольшее расхождение во времени звучания составляет 60 мсек. Наиболее значительное расхождение во времени звучания глагола и предлога *yòng* наблюдается в следующих сравниваемых парах: (дик.2) № 1 *Tā zhèng yòng hǎo qiānbǐ* "Он как раз пользуется хорошим карандашом". Глагол *yòng* звучит 220 мсек, № 2 *Tā zhèng yòng hǎo qiānbǐ huìhuà* "Он как раз рисует хорошим карандашом", продолжительность звучания предлога *yòng* 170 мсек. (дик.2), № 3 *Wǒ luàn yòng hóngd* "Я по ошибке пользуюсь красными", глагол *yòng* звучит 220 мсек, а в № 4 *Wǒ luàn yòng hóngd zuò zhège* "Я по ошибке делаю это из красных", предлог *yòng* звучит 160 мсек.

Частота колебаний основного тона. Проведя электроакустическую запись предложений, где *yòng* встречается то как глагол, то как предлог, мы получили возможность по интонограммам проследить направление движения основного тона. Как у глагола, так и у предлога оно характеризуется нисходящим направлением движения частот, причем, в большинстве произнесений тон падает весьма значительно. В предложении № 5 (дик.1) *Wǒ luàn yòng hóngd* "Я по ошибке пользуюсь красными"

глагол *yòng* реализуется в диапазоне 260-140 гц

(септима) с последовательностью частот: 260-260-260-250-225-210-200-175-160-160-140 гц; а в предложении № 6 *Wǒ luàn yòng hóngd zuò zhège* "Я по ошибке делаю это из красных" предлог *yòng* реализуется в диапазоне 260-190 гц (кварта) с последовательностью частот: 260-250-240-225-200-200-190-190-190 гц. Интонограмма предложения № 7 (дик.2) *Tā néng yòng cídiǎn* "Он может пользоваться словарем" фиксирует произнесение глагола *yòng* в диапазоне 300-160 гц (септима) с последовательностью частот: 300-300-300-275-250-225-200-190-175-160-160 гц, а в предложении № 8 *Tā néng yòng cídiǎn fānyì* "Он может переводить со словарем" предлог *yòng* реализуется в диапазоне 325-210 гц (м.секста) при последовательном падении частот: 325-325-310-275-275-260-240-225-210-210-200-210 гц.¹⁾ Приведенные выше примеры свидетельствуют о том, что предлог *yòng* не лишается своей тональной характеристики. В большинстве реализаций интервалы падения тона у предлога весьма значительны. (Дик.2) № 9 *Wǒ xiān yòng zhè zhī gāngbǐ xiě zì* "Я сначала пишу этой ручкой" и № 10 *Tā zhǔn yòng tād qián mǎi zhè běn shū* "Он разрешает купить эту книгу на его деньги", интервалы понижения четвертого тона у предлога *yòng* составляют сексту.

- 1) В одном из произнесений наблюдается менее резкое падение тона, как у глагола, так и у предлога *yòng*: (дик.1) *Tā běn yòng kuàiz páochǎn* и *Tā běn yòng kuàiz chīfàn* "Он в основном пользуется палочками" и "Он в основном кушает палочками". В обоих произнесениях и глагол *yòng* и предлог *yòng* реализуются в интервале 240-210 гц (секунда). Но это незначительное падение четвертого тона в слове *yòng* можно объяснить тем, что за словом *yòng*, произносимым четвертым тоном, следует слово *kuàiz*, также произносимое четвертым тоном, что и влияет на диапазон предыдущего *yòng*. (См. Т.П. Задоевко, "Краткий очерк системы тонов современного китайского языка", стр. 200; М.К. Румянцев, "Интонация предикативных и атрибутивных синтагм", стр. 44).

Следует отметить, однако, что интервалы падения четвертого тона у глагола *yòng* и у предлога в большинстве случаев различны. Из 16 сравниваемых пар в 8 парах глагол *yòng* характеризуется значительно большими интервалами падения, чем предлог *yòng*.¹⁾ Так, в предложении № 12 (дик.2) *wǒ luàn yòng hóngd* "Я по ошибке использую красные" четвертый тон глагола *yòng* реализуется с интервалом падения, равным септима, т.е. II полутонам, а в предложении № 13 *wǒ luàn yòng hóngd zuò zhège* "Я по ошибке делаю это из красного" интервал между начальной и конечной точками четвертого тона предлога *yòng* составил кварту, т.е. 5 полутонов. В предложении № 14 *Tā běn yòng kuàiz* "Он в основном пользуется палочками" глагол *yòng* имеет интервал падения равный м.сексте (8 полутонов), а в предложении № 15 *Tā běn yòng kuàiz chīfan* "Он в основном кушает палочками" предлог *yòng* реализуется с интервалом падения, составляющим терцию (4 полутона).

У глагола *yòng* наибольшим частотным интервалом является септима (II полутонов, в 2 случаях), а наименьшим - секунда (2 полутона, в I случае). В других реализациях у глагола *yòng* фиксируются следующие интервалы: секста - 2 раза, м.секста - 6 раз, ув.кварта - 4 раза, кварта - I раз. У предлога *yòng* наибольший частотный интервал - секста (9 полутонов) в 2 случаях, а наименьший - секунда (2 полутона) в одном случае. Остальные частотные интервалы предлога *yòng* следующие: м.секста - 5 раз, квинта - I раз, ув.кварта - 2 раза, кварта - 3 раза, терция - 2 раза.

1) В 4 случаях интервалы падения тона больше у предлога, но различие в интервале между предлогом и глаголом составляет всего лишь один и максимум три полутона.

Наш материал не дает нам основания сделать какие-либо определенные выводы о различии глагола и предлога по частотному регистру, так как из 16 сравниваемых пар, в 5 случаях глагол *yong* произносился в более высоком частотном регистре, а в 6 случаях – предлог *yong*. В 5 случаях они произносились на одинаковой высоте.

Для аудиторского анализа мы вычленили слог *yong* из 23 предложений. Для восприятия аудиторам были предложены глагол *yong* и предлог *yong* со следующими акустическими характеристиками:

- 1) с достаточно большим временем звучания (глагол от 190 мсек до 250 мсек; предлог от 180 мсек до 230 мсек);
- 2) с меньшим временем звучания (глагол – 170–180 мсек, предлог – 160–170 мсек);
- 3) со значительным интервалом падения тона (от кварты до септими);
- 4) с меньшим интервалом падения тона (от секунды до терции).

В эксперименте принимали участие пять аудиторов. Все аудиторы сразу узнавали четвертый тон при реализации глагола *yong* со временем звучания от 180 мсек до 250 мсек при интервале падения тона кварта-септима.

Без затруднений воспринимался 4 тон и у предлога *yong* со временем звучания от 160 мсек до 230 мсек при интервале падения тона терция – секста.

В предложении № 16 (дик. I) *Tā zhūn yong tād qián*
 "Он как раз использует его деньги" четвертый тон глагола *yong* реализовался как восходяще-нисходящий. Восходящая

часть образована последовательностью частот: 210-250-275-290 гц, нисходящая часть - 290-250-240-210-210-210 гц. Интервалы начального повышения и конечного падения тона - одинаковы (ув.кварта), время же звучания различно. Начальное повышение тона занимает 80 мсек, конечное падение - 130 мсек. Четыре аудитора данную реализацию квалифицировали как 4 тон, один аудитор воспринял как I тон.

Расхождения в восприятии тона аудиторами наблюдались и в следующих двух случаях:

А) (дик. I) № I7 Tā bēn yǒng kuaiz "Он в основном пользуется палочками для еды", глагол yǒng, вычлененный из этого предложения, имел: время звучания - 170 мсек, тон реализовался как восходяще-нисходящий. Восходящая часть была образована последовательностью частот: 210-240-240 гц, нисходящая часть - 240-225-210-210-210 гц. Интервалы начального повышения и конечного падения тона одинаковы (секунда); время звучания восходящей части - 60 мсек, нисходящей - 110 мсек. Три аудитора восприняли данную реализацию, как тон более похожий на 4, один аудитор - как 4 или I тон, а один аудитор - как 3 тон.

Б) (дик. I) предложение № I8 Tā bēn yǒng kuaiz chīfan "Он в основном кушает палочками", вычленялся предлог yǒng, акустические характеристики которого следующие: время звучания - 170 мсек; тон реализовался как восходяще-нисходящий. Восходящая часть - 210-225-240 гц (секунда), нисходящая - 240-225-210-210-210 гц (секунда); время звучания восходящей части - 60 мсек, нисходящей - 110 мсек. Три аудитора определили данный тон как больше всего похожий на

4 тон, один аудитор определил его как нечто среднее между 4 и I тонами и один аудитор - как I тон, но с неестественно низким по звучанию регистром.

Амплитуды слоговой интенсивности. Для 4 тона характерно убывание интенсивности к концу слога.¹⁾ Амплитуды интенсивности уòng действительно убывают к концу слога. У глагола уòng амплитуды слоговой интенсивности меньше, чем у предлога уòng . В 10 сравниваемых парах (из 16) амплитуды слоговой интенсивности были меньшими у глагола уòng , и только в 3 случаях они были больше, чем у предлога. В 3 случаях они были одинаковы.

Если сопоставить время звучания глагола и предлога уòng с амплитудами слоговой интенсивности, то из 10 пар, в которых предлог имеет большую амплитуду, в 8 случаях это падает на те предложения, в которых время звучания предлога уòng меньше, чем у глагола уòng .

Выводы:

I) Акустические реализации предлога уòng по основным своим характеристикам в общем не выходят за пределы обычного четвертого тона, хотя в ряде случаев и имеются некоторые особенности:

- а) глагол звучит продолжительнее предлога уòng ;
- б) амплитуды слоговой интенсивности у глагола уòng меньше, чем у предлога уòng ;
- в) глагол уòng отличается большими интервалами падения тона, чем предлог уòng .

¹⁾ М.К.Румянцев, "Интонация предикативных и атрибутивных синтагм".

2) Существенного различия в частотном регистре у глагола *yǒng* и предлога *yǒng* на данном материале нами не отмечено.

Акустическая реализация *tí* глагола и *tí* предлога.

К числу бифункциональных комплексов относится и морфема *tí*. В современном китайском языке *tí* возможно в функции самостоятельного глагола-сказуемого; употребляется *tí* и в качестве предлога.

На интонографе были записаны 20 предложений в произнесении двух дикторов. В 10 из них *tí* употреблен как глагол, а в других 10 является предлогом. Таким образом, в общей сложности было получено 40 реализаций *tí* в 20 сравниваемых парах.

Ниже приводятся данные анализа интонограмм.

Время звучания. Как глагол, так и предлог *tí* произносятся 4 (нисходящим) тоном.

Время звучания *tí* глагола и *tí* предлога - различно. Продолжительность звучания больше у глагола *tí*. В нашем эксперименте в 13 сравниваемых парах из 20 глагол *tí* звучит продолжительнее предлога *tí*. Различие по долготе в среднем составляет 36 мсек от общего среднего времени звучания глагола 233 мсек и предлога 200 мсек; наибольшее расхождение, констатируемое нами, составляет 90 мсек. Только в 4 случаях оказалось обратное соотношение долгот: предлог *tí* звучал дольше глагола *tí*. Различие в этих случаях было во всех 4 примерах одинаково - 20 мсек, при общем среднем времени звучания предлога 260 мсек, глагола - 250 мсек. В трех сопоставляемых парах продолжительность звучания глагола *tí*

и предлога *tí* была равной. Наибольшее расхождение по времени звучания глагола и предлога отмечено в следующих сравняемых парах: (дик.1) № 1 *Wǒ néng tí Chén tóngzhì*

"Я могу заменить товарища Чэня", глагол *tí* звучит 250 мсек; № 2 *Wǒ néng tí Chén tóngzhì zǒu* "Я могу пойти вместо товарища Чэня", предлог *tí* звучит 160 мсек. (Дик.2) № 3 *Tā kěn tí nǐ* "Он хочет заменить тебя", глагол *tí* звучит 280 мсек. № 4 *Tā kěn tí nǐ jiānghuà* "Он хочет прочитать лекцию вместо тебя", продолжительность звучания предлога *tí* 220 мсек.

Характеристики *Fo* . Частотные характеристики глагола *tí* и предлога *tí* представляют собой примеры типичные для реализации четвертого тона. Как у глагола, так и у предлога в большинстве произнесений фиксируются значительные интервалы в падении частот. Например: (дик.2) № 5 *Wǒ xiān tí tā* "Я сначала заменю его", глагол *tí* произносится в диапазоне 350-150 гц (м.децима) со следующей последовательностью частот: 350-275-260-240-210-190-150 гц; у первого диктора этот же глагол *tí* реализуется в диапазоне 290-175 гц (секста) с последовательностью частот: 290-250-200-175-190 гц. В предложении № 6 *Zhāng tóngzhì méi lái, tā tí Zhāng tóngzhì gōngzuò* "Товарищ Чжан не пришел, он будет работать вместо него", предлог *tí* (дик.2) реализуется в диапазоне 325-140 гц (м.децима) с падением частот: 325-275-250-240-225-210-200-175-175-160-140 гц; у первого диктора - в диапазоне 300-200 гц (квинта) с последовательностью частот: 300-290-260-240-200-200 гц.

У глагола *tí* наибольший частотный интервал - малая

децима (в I случае) и один раз - октава, а наименьший - малая терция (I раз). Между этими двумя крайними значениями идут реализации с интервалами меньше октавы, но больше м.терции. В 2 случаях интервал равен септима, в I случае - м.септима, в 3 случаях - сексте, в 4 случаях - м.сексте, в I случае - квинте, в 3 случаях - ум.квинте, в I случае - кварте и в 2 случаях - терции.

У предлога *tí* крайними частотными интервалами были: интервал, превышающий октаву (м.децима) - I раз, и м.терция - I раз. Внутри крайних точек интервалы варьировал следующим образом: м.септима - 3 раза, секста - I раз, м.секста - 2 раза, ув.кварта - 2 раза, кварта - 2 раза, терция - 4 раза.

Интервал падения тона у глагола *tí* и предлога *tí* во многих случаях различен. Например, в предложении № 7 (дик.I) *wǒ xiān tí ta* "Я сначала замени его", интервал падения тона у глагола *tí* составляет сексту (9 полутонов), а в предложении № 8 *wǒ xiān tí ta xiě zì* "Я сначала напишу вместо него", интервал падения тона у предлога *tí* равен терции (4 полутона). В предложении № 9 (дик.I) *nǐ zhè cǐ tí ta* "Ты на этот раз замени его" интервал падения тона у глагола *tí* равен квинте (7 полутонов), а в предложении № 10 *nǐ zhè cǐ tí ta cānjiā* "На этот раз ты будешь участвовать вместо него" падение 4 тона у предлога *tí* равно м.терции (3 полутона).

В 9 случаях из 20 глагол *tí* характеризуется большими частотными интервалами падения тона по сравнению с предлогом *tí*. Причем, различие в интервале доходит до 5 полу-

тонов. В 5 случаях предлог *tí* обладает большим интервалом, чем глагол *tí*, но различие между интервалом падения тона у глагола *tí* и предлога *tí* составляет в 4 случаях всего лишь один-два полутона. В 5 случаях интервал падения равный.

Из 19 пар в 8 случаях частотный регистр предлога *tí* выше, чем у глагола *tí*. В 3 случаях частотный регистр выше у глагола *tí*. В 8 случаях у глагола и предлога *tí* частотные регистры почти одинаковы.

Для аудитивного восприятия вычленялся слог *tí* из 26 предложений. Для опознавания аудиторам были предложены глагол *tí* и предлог *tí* со следующими акустическими характеристиками:

- 1) с достаточно большим временем звучания (глагол от 220 мсек до 280 мсек, предлог - от 200 мсек до 300 мсек);
- 2) с меньшим временем звучания (глагол 200-210 мсек, предлог 150-170 мсек);
- 3) со значительным интервалом падения тона (от ув.кварта до м.децимы);
- 4) с меньшим интервалом падения тона (от м.терции до терции).

В эксперименте принимали участие пять аудиторов.

Все аудиторы легко определяли 4 тон при реализации глагола *tí* со временем звучания от 200 мсек до 280 мсек при интервале падения тона - ув.кварта - м.децима.

Аудиторы узнавали 4 тон и у предлога *tí* со временем звучания от 160 мсек до 300 мсек при интервале падения тона терция - м.децима.

Некоторые расхождения при определении тона были в сле-

дующих случаях: (дик.2) № II Ní zhè cì tì ta "На этот раз ты замени его". Вычленялся глагол tì . Его время звучания 220 мсек, интервал падения - м.секста. Четыре аудитора восприняли его как 4 тон, но с оговорками: два аудитора колебались в выборе между 4 и 2 тонами, два аудитора отмечали сходство данной реализации с 4 и I тоном. Один аудитор определил его как 3 тон. Неуверенность в определении 4 тона в рассматриваемом случае можно, по-видимому, объяснить тем, что движение частоты основного тона здесь оказалось двунаправленным, т.е. наблюдается начальный участок падения тона и затем некоторая константная часть (ровный отрезок): 250-225-200-175-160-160-160 гц.

Тон предлога tì , вычлененного из предложения № I2 (дик.I) Ní zhè cì tì ta cānjiā "На этот раз ты будешь участвовать вместо него", также воспринимался двумя аудиторами как 4 тон похожий на I , двумя - как I тон и только один аудитор определил его как 4 тон. В этом произнесении предлог tì имел время звучания 150 мсек, интервал падения тона - м.терция, со следующей последовательностью частот: 250-250-210 гц.

В предложении № I3 (дик.I) Ní tì wó "Ты замени меня" вычленялся глагол tì . Время звучания - 230 мсек, интервал падения тона - м.терция с последовательностью частот: 300-290-275-260-250 гц. Эта реализация глагола tì одним из аудиторов была воспринята как 4 тон, другим - как 4 тон похожий на I, и тремя аудиторами - как обычный первый тон.

При малом времени звучания и сравнительно небольшом

интервале падения или при интервале падения, имеющем некоторую константную часть в конце реализации, 4 тон воспринимается иногда как I тон. Это связано, очевидно, с тем, что и I тоне на конце предложения тоже наблюдается некоторое падение частот, связанное с понижающейся интонацией в конце предложения.

Амплитуды слоговой интенсивности. Сопоставляя амплитуды слоговой интенсивности у глагола *tí* и у предлога *tí*, мы можем отметить, что амплитуды у предлога *tí* больше, чем у глагола *tí*. Из 20 сравниваемых пар в 12 случаях у предлога *tí* отмечаются большие амплитуды слоговой интенсивности, в 2 случаях у глагола *tí*; в 6 случаях амплитуды глагола и предлога равны.

Сопоставление длительности произнесения глагола и предлога *tí* с амплитудами их слоговой интенсивности показывает, что большая интенсивность предлога, в большинстве случаев, соотносится с меньшим времени его звучания.

Выводы:

1. В большинстве случаев глагол *tí* звучит продолжительнее предлога *tí*.
2. Как глагол, так и предлог *tí* произносятся четвертым тоном с ярко выраженным падением частоты тона.
3. Глагол *tí* отличается большими интервалами падения тона, чем предлог *tí*.
4. Амплитуды слоговой интенсивности у глагола меньше, чем у предлога *tí*.
5. В большинстве произнесений предлог *tí* реализуется в более высоком частотном регистре.

Акустическая реализация cháо глагола и cháо предлога.

Для сравнения акустических характеристик cháо глагола и cháо предлога на интонографе было записано 8 предложений, из которых в 4 предложениях cháо является глаголом и в 4 - предлогом. Эти 8 предложений были записаны в произнесении трех дикторов. Таким образом мы располагали 24 реализациями слова cháо. Ниже приводятся данные анализа интонограмм.

Время звучания. Этимологический тон cháо - второй. Время звучания глагола cháо большее, чем у предлога. Из 12 сравниваемых пар глагол cháо звучит продолжительнее предлога в 8 случаях. Предлог cháо звучит дольше глагола только в одном случае. В 3 случаях глагол cháо по времени звучания равен предлогу cháо.

Различие глагола и предлога cháо по времени звучания составляет в среднем 26 мсек, при общем среднем звучании глагола - 242 мсек и предлога - 237 мсек. Наибольшее расхождение, зафиксированное нами, составляет 50 мсек.

Из 8 случаев, где глагол cháо звучит продолжительнее предлога, наибольшее расхождение наблюдается в следующих сравниваемых парах:

(дик.1) № 1 Dà mén cháо běi "Ворота смотрят на север", глагол cháо звучит 270 мсек, № 2 Dà mén cháо běi kāizhe "Ворота открыты в сторону севера", предлог cháо звучит 220 мсек;

(дик.3) № 3 Dà mén cháо nán "Ворота смотрят на юг". Продолжительность времени звучания глагола cháо - 230 мсек, № 4 Dà mén cháо nán kāizhe "Ворота открыты в сторону

юга", продолжительность звучания предлога *cháo* - 190 мсек.

Характеристики F_0 . В разделе, касающемся акустических характеристик глагола и предлога *ná*, мы говорили о двух возможных вариантах движения частоты во втором тоне.¹⁾ Мы отмечали, что повышение частот во 2 тоне может и не сопровождаться их начальным падением. Такие же случаи мы можем привести и у глагола *cháo*, например: (дик.3) № 5 *Pàokōu cháo zhèbiān* "Жерла пушки направлены в эту сторону". Глагол *cháo* произносится в интервале 175-275 гц (м.секста) с последовательностью частот: 175-175-190-210-240-250-275 гц. (дик.3) № 6 *Dà mén cháo nán* "Ворота смотрят на юг". Глагол *cháo* в этом предложении реализуется в интервале 160-290 гц (м.септима) с последовательно нарастающими частотами: 160-190-210-250-260-290 гц.

Реализация 2 тона без нисходящей части в начале звучания отмечена нами и у предлога *cháo*, например, в предложении № 7 *Dà mén cháo běi kāizhē* "Ворота открыты в сторону севера" у диктора первого предлог *cháo* произносится в интервале 200-240 гц (м.терция) с последовательностью частот: 200-200-225-240-210 гц, у диктора третьего предлог *cháo* реализуется в интервале 160-225 гц (ув.кварта) с последовательностью частот: 160-160-190-200-225 гц. Интервалы повышения тона в приведенных примерах: м.септима, м.секста (глагол) и м.терция, ув.кварта (предлог).

Наименьшие интервалы восхождения частоты F_0 при данном типе реализации 2 тона глагола и предлога *cháo* зафик-

1) См.раздел о глаголе и предлоге *ná*, стр. 458-460

сированы в следующих реализациях: (дик.3) № 8 Dà mén
 chāo běi "Ворота смотрят на север". Глагол chāo реализо-
 вался в интервале 190-240 гц (терция) с последовательностью
 частот: 190-210-225-240 гц; (дик.1) № 9 Dà mén chāo nán
 kāizhe "Ворота открыты в сторону юга". Предлог chāo -
 240-250 гц (м.секунда); последовательность частот: 240-240-
 250-240 гц.

Интервалы повышения тона у глагола chāo меньше, чем
 терция наблюдаются в другом типе реализаций 2 тона -
 там, где повышению частоты F₀ предшествует некоторое ее
 падение. Примерами таких реализаций могут быть предложения:
 № 10 Dà mén chāo běi "Ворота смотрят на север", в
 котором у дикторов 1 и 2 интервал повышения тона у глагола
 chāo равен секунде; (дик.1) № 11 Dà mén chāo nán
 "Ворота смотрят на юг" и № 12 Tā chāo dōng "Он обра-
 щается на восток". В этих предложениях интервал повышения
 2 тона у глагола chāo также составляет секунду.

У глагола chāo наибольший частотный интервал м.септи-
 ма (10 полутонов, 1 случай), а наименьший - секунда (2 по-
 лутона, 4 случая). Между этими полярными точками по частот-
 ному интервалу располагаются реализации с интервалами мень-
 шими, чем м.септима и большими, чем секунда. В одном случае
 интервал равен м.сексте, в одном случае-кварте, в 4 случаях -
 терции и в 1 случае - м.терции.

У предлога chāo самым большим интервалом является
 м.секста (8 полутонов) - 1 раз, самым малым - м.секунда
 (1 полутон) - 3 раза. Между этими крайними реализациями от-
 мечены следующие частотные интервалы: ув.кварта - 1 раз,

м.терция - 2 раза, секунда - 5 раз.

Интервалы повышения тона у глагола *cháo* и предлога *cháo* в подавляющем большинстве случаев различны. Так, в предложении № 13 (дик.3) *Pàokǒu cháo zhēbiān* "Жерла пушки направлены в эту сторону" интервал повышения у глагола *cháo* равен м.сексте (8 полутонов), а в предложении № 14 *Pàokǒu cháo zhēbiān gōngjī* "Пушки стреляют в эту сторону" интервал повышения у предлога *cháo* составляет секунду (2 полутона). В предложении № 15 (дик.3) *Tā cháo dōng* "Он обращается на восток" интервал повышения у глагола *cháo* равен ув.кварте (6 полутонов), а в предложении № 16 *Tā cháo dōng zǒu* "Он идет на восток" интервал повышения у предлога *cháo* равен секунде (2 полутона).

В 9 сравниваемых парах из 12 глагол *cháo* характеризуется большими интервалами повышения тона, чем предлог *cháo*. Предлог же *cháo* только в 2 случаях оказался с большим интервалом повышения нежели глагол *cháo*. В 1 случае интервал повышения тона был равным у глагола и у предлога.

Если сравнить реализации глагола *cháo* и предлога *cháo* по регистру, то мы можем сказать, что в большинстве случаев в более высоком частотном регистре реализуется предлог *cháo*. Из 12 сравниваемых пар в 7 у предлога *cháo* регистр более высокий, в одном случае частотный регистр был выше у глагола *cháo* на протяжении всей реализации, в 1 случае при одинаковой средней части начало и конец произнесения были выше у глагола. В 3 случаях регистр глагола *cháo* и предлога *cháo* практически почти одинаков.

Для аудиторского анализа мы вычленили слог *chaó* из 16 предложений. Для восприятия аудиторам были предложены глагол *chaó* и предлог *chaó* со следующими акустическими характеристиками:

- 1) с достаточно большим временем звучания (глагол от 220 мсек до 270 мсек ; предлог от 200 мсек до 400 мсек) ;
- 2) с меньшим временем звучания (глагол - 200 мсек, предлог - 180-190 мсек) ;
- 3) со значительным интервалом повышения тона (от кварты до м.септими) ;
- 4) с меньшим интервалом повышения тона (м.секунда - терция).

В эксперименте принимали участие четыре аудитора. Все аудиторы легко узнавали 2 тон при реализации глагола *chaó* со временем звучания от 200 мсек до 270 мсек при интервале повышения тона секунда - м.септима.

Также воспринимался 2 тон и у предлога *chaó* со временем звучания от 190 мсек до 230 мсек при интервале повышения тона м.секунда - м.секста.

В предложении № 17 *Đã mén cháó nán* "Ворота смотрят на юг" второй тон у глагола *cháó* реализовался как тон нисходяще-восходящий. Нисходящая часть была несколько значительней, нежели восходящая. Нисходящая часть была образована последовательностью частот: 200-175-170 гц, а восходящая часть: 170-190 гц. Интервал начального падения тона равен м.терции, а интервал конечного повышения равен секунде. Эта реализация двумя аудиторами была воспринята как 1 тон, двумя аудиторами как 2 тон. Как и у морфемы *na* такое вос-

приятие данной реализации глагола *cháo* связано, по-видимому, с распределением интенсивности: интенсивность значительно увеличивается на том участке звучания, где тон повышается.

Амплитуды слоговой интенсивности. Второй тон, обычно, характеризуется возрастанием интенсивности к концу произнесения. Большие амплитуды интенсивности отмечаются у предложения *cháo*.

Сравнивая показатели интенсивности 24 реализаций глагола и предложения *cháo*, мы можем сказать, что в 7 случаях (из 12) амплитуды слоговой интенсивности больше у предложения *cháo*, в 3 случаях — у глагола *cháo* и в 2 случаях амплитуды глагола и предложения *cháo* одинаковы.

Как и в предыдущих морфемах большие амплитуды слоговой интенсивности предложения *cháo* соотносятся с меньшим временем его звучания.

Выводы:

1. Глагол *cháo* звучит в большинстве случаев продолжительнее предложения *cháo*.
2. Амплитуды слоговой интенсивности у глагола *cháo* в большинстве произнесений меньше, чем у предложения *cháo*.
3. Тон, присущий слогу *cháo*, реализуется как у глагола, так и у предложения *cháo* как обычный 2 тон.
4. Глагол *cháo* характеризуется большими интервалами повышения, чем предлог *cháo*.
5. Предлог *cháo* чаще реализуется в более высоком частотном регистре.

Выше было отмечено, что сравнение акустических характеристик глаголов и предлогов осуществлялось в таких синтак-

ических структурах, в которых интересующие нас морфемы находились в одинаковом синтаксическом окружении. Но синтаксические структуры и их актуализация (внутреннее членение) - различны. В одном случае мы имеем синтаксическую структуру, состоящую из трех членов П-С-Д, в другом - структуру из 5 членов; к подлежащему, сказуемому и дополнению добавляется предлог со своим объектом (П - предлог + объект - С - Д).

В структуре П-С-Д глагол-сказуемое является одним из главных членов предложения. В высказывании, представляющем собой обычную констатацию факта, сказуемое вместе с дополнением является предикатом сообщения. Предикат же, как правило, интонационно выделен. Этим в значительной степени могут быть объяснены более яркие акустические характеристики глагола-сказуемого и дополнения.

В синтаксической структуре П - предлог + объект - С-Д сочетание предлог + объект, будучи второстепенным членом предложения (обстоятельным определением), при обычной констатации факта не является столь весомым в смысловом отношении, как П и С в структуре П-С-Д.

Следовательно, уже сама синтаксическая структура (ее актуализация) предопределяет акустически более слабую реализацию группы предлог + объект. В связи с этим важное значение приобретают два момента:

1) насколько существенно давление синтаксической структуры (ее актуализации) на акустическую реализацию морфем, выступающих в предложении в разных значениях, и

2) какова динамика акустических характеристик в тех членах синтаксических структур, которые окружают исследуе-

мые морфемы. Для этого необходимо рассмотреть не только акустические характеристики самих глаголов и соответствующих им предлогов, но и акустические характеристики слов, окружающих эти глаголы и предлоги.

Наиболее существенны для нас характеристики, связанные с временем звучания и амплитудами слоговой интенсивности подлежащих сравниваемых предложений и объектов глагола и предлога.

Коротко эти характеристики сводятся к тому, что подлежащие и объекты акустически оказываются более сильными (как по времени звучания, так и по интенсивности) в тех синтаксических структурах, где исследуемые морфемы выступают как глаголы.

Отмеченные закономерности, как мы могли убедиться, относятся в равной степени ко всем четырем исследуемым морфемам. Основной смысл их сводится к тому, что в настоящее время еще не представляется возможным делать вывод об утрате тона теми морфемами, которые выступают как предлоги, хотя о некоторых сдвигах в их акустическом облике можно уже, по-видимому, вести речь. Не исключено, однако, что отмеченные нами сдвиги в акустической реализации предлогов еще не представляют собой их внутренней сущности, а объясняются, в той или иной степени, свойствами той интонационной структуры, частью которой они являются.

E.C.Fedorova

L.E.Ph., I.O.L. M.S.U.

PHONETIC REALIZATION OF SOME VERBS AND PREPOSITIONS
CORRESPONDING TO THEM IN SOUND IN MODERN CHINESE

Grammarians distinguish a group of so-called bifunctional words (complexes) in Modern Chinese that are used in the language sometimes as verbs sometimes as prepositions.

It is assumed that the formalization of significant words in the Chinese language is followed by reduction of tone. In this connection it is not without interest to study the phonetic realization of bifunctional words that is to establish whether these words lose the etymological tone, while being used as prepositions.

In this particular work we are concerned with the acoustic realization of the four bifunctional words: ná, yò ng, tí, cháo.

Our study of the problem showed that in most of the pronouncings:

- 1) the verbs sound longer than the corresponding prepositions,
- 2) the amplitude of syllabic intensity of the verbs is less than that of the prepositions,
- 3) the verbs are characterized by longer intervals in the rise and fall of the tone that the prepositions are,
- 4) the prepositions are realized in a higher frequency range.

But despite the mentioned differences in the acoustic realization of the verbs and prepositions it is not possible yet to speak of the prepositions as having lost their tone structure, one can observe only some changes in it.

It is even possible to say that the mentioned changes in the acoustic realization of the prepositions are not due to the rules of their own, but can be explained to a great extent by the characteristics of the intonation pattern, of which they are part.

О Г Л А В Л Е Н И Е

В.А.Артемов	- Некоторые вопросы спектрального анализа звуков речи	3-12
С.П.Баронин	- Спектральный анализ и проблема описания речевых сигналов	13-30
Г.С.Слуцкер	- К вопросу о фонетическом кодировании речевого сигнала	31-54
В.Н.Соболев	- Анализ спектров речи при помощи ЦВМ ..	55-75
Б.М.Лобанов	- Многофункциональный анализатор спектра речевых сигналов типа "Спектр-1"	76-82
Б.М.Лобанов	- Синтез схемы компенсации амплитудных прогешностей для анализаторов спектра с типовыми частотными характеристиками	83-95
Л.П.Блохина Р.К.Потапова	- Речевой сигнал и фонемный образ	96-101
М.Г.Каспарова В.Д.Лаптев	- О речевых вариантах как о форме существования языкового инварианта ...	102-112
Р.Г.Мегрелишвили	- Спектральный анализ щелевых согласных французского языка в сочетании с гласными (э, ё, а) в позиции СГ	113-154
Т.Е.Плакунова	- Некоторые данные спектрального анализа согласных литовского языка (результаты статистической обработки)	155-168
Я.Урбелене	- Некоторые данные спектрального анализа долгих гласных литовского языка в ударном, предударном и заударном слогах ..	169-181

- И.Б.Думкова - Тон и ударение в китайских двусложных
и трехсложных словах 182-206
- А.М.Карпетьянц - Опыт формального подхода к изучению
интонации путунхуа 207-274
- М.И.Каплун - Интонационная дифференциация предика-
тивной и атрибутивной (определяющей)
связи в современном русском языке ... 275-314
- А.П.Павленко - Опыт экспериментально-фонетического
исследования словесного ударения в
индонезийском языке 315-346
- В.И.Петрянкина - Долготная характеристика гласных языка
суахили 347-359
- В.И.Петрянкина - Время звучания согласных языка
суахили 360-367
- М.К.Румянцев - Время звучания элементов слога (на
материале китайского языка) 368-400
- Е.К.Фёдорова - Фонетическая реализация некоторых
глаголов и соответствующих им по
звучанию предлогов в современном
китайском языке 401-436

ПОДП. К ПЕЧАТИ 16/1-89 Г. Л-83013. Ф. 60х90/16
БУМ. ОФС. № 2. ФИЗ.П.Л. 27,5. УЧ.-ИЗД Л. 16,02.
ЗАКАЗ 1858. ТИРАЖ 500 ЭКЗ. ЦЕНА 93 КОП.

ОТПЕЧАТАНО НА РОТАПРИНТАХ В ТИП. ИЗД. МГУ
МОСКВА, ЛЕНГОРЫ.