

В. М. Лемская, С. В. Глушков, Я. А. Глухий

ТЕОРЕТИЧЕСКАЯ ФОНЕТИКА АНГЛИЙСКОГО ЯЗЫКА

Учебно-методическое пособие



V. M. Lemskaya, S. V. Glushkov, Ya. A. Glukhiy

THEORETICAL PHONETICS OF ENGLISH

Study & methods guide

МИНИСТЕРСТВО ПРОСВЕЩЕНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Томский государственный педагогический университет»
(ТГПУ)

В. М. Лемская, С. В. Глушков, Я. А. Глухий

ТЕОРЕТИЧЕСКАЯ ФОНЕТИКА АНГЛИЙСКОГО ЯЗЫКА

Учебно-методическое пособие

Электронное издание локального распространения

Томск
Издательство ТГПУ
2026

© Томский государственный педагогический университет, 2026
ISBN 978-5-907791-81-7

MINISTRY OF EDUCATION OF THE RUSSIAN FEDERATION
Tomsk State Pedagogical University

V. M. Lemskaya, S. V. Glushkov, Ya. A. Glukhiy

THEORETICAL PHONETICS OF ENGLISH

Study & methods guide

Electronic edition of local distribution

Tomsk
TSPU Publishing House
2026

© Tomsk State Pedagogical University, 2026
ISBN 978-5-907791-81-7

УДК 811.111.34(075.8)
ББК 81.432.1-1я73
Л44

Рекомендовано к изданию
редакционно-издательским советом
Томского государственного
педагогического университета

Рецензент:

доктор филологических наук, доцент, декан факультета
русской и чувашской филологии и журналистики
ФГБОУ ВО «Чувашский государственный университет им. И. Н. Ульянова»
А. М. Иванова

Лемская, В. М.

Л44 Теоретическая фонетика английского языка : учебно-методическое пособие [Электронный ресурс] / В. М. Лемская, С. В. Глушков, Я. А. Глухий. – Электрон. текстовые дан. (3,24 Мб). – 1 электрон. опт. диск (CD-ROM). – Загл. с титул. экрана. – Томск : Издательство Томского государственного педагогического университета, 2026. – 151 с.

ISBN 978-5-907791-81-7

Издание представляет собой обобщение сведений об основных принципах теоретической фонетики английского языка, ее методах и современных инструментах исследований.

Рассчитано на обучающихся высших учебных заведений по направлению 45.03.02 Лингвистика, 45.05.01 Перевод и переводоведение и 44.03.05 Педагогическое образование (с двумя профилями подготовки), а также на широкий круг лиц, интересующихся теоретическими лингвистическими вопросами.

УДК 811.111.34(075.8)
ББК 81.432.1-1я73

Системные требования:

ПК не ниже класса Pentium II; RAM 512 Mb; Windows XP/7–10 (32-разрядная или 64-разрядная версии); разрешение экрана 1 024 × 768 (768 × 1 024); CD-ROM-дисковод, мышь; Adobe Acrobat Reader DC (либо другое, открывающее PDF-файлы).

ISBN 978-5-907791-81-7

© Лемская В. М., Глушков С. В., Глухий Я. А., 2026
© Томский государственный
педагогический университет, 2026

Contents (Содержание)

Contractions (Сокращения)	8
Preface (Предисловие)	9
1. Phonetics as Science (Фонетика как наука)	11
1.1. The Subject and Tasks of Phonetics (Предмет и задачи фонетики).....	11
1.2. The Branches of Phonetics (Разделы фонетики).....	11
1.3. The Connection of Phonetics with Other Branches of Linguistics (Связь фонетики с другими разделами лингвистики).....	14
1.4. Interconnection of Phonetics with other Branches of Science (Взаимоотношения фонетики с другими науками).....	15
1.5. Methods of Phonetic Research (Методы фонетического исследования)	18
Revision: Section 1 (Повторение: Раздел 1).....	27
2. The Phonetic Basis of English: Articulatory & Perceptual Aspects (Фонетическая база английского языка: артикуляционный и перцептивный аспекты)	28
2.1. The Concept of the Articulatory Basis in a Language (Понятие артикуляционной базы языка).....	28
2.2. The Articulation Basis of English (Артикуляционная база английского языка).....	29
2.3. Comparison of Articulatory Bases / Articulatory Differences between Russian & English (Сопоставление артикуляционных баз / артикуляционные различия русского и английского языков).....	29
2.4. Literary Pronunciation (Литературное произношение).....	31
2.5. Phonetic Interference in Bilingualism (Фонетическая интерференция при билингвизме).....	33
Revision: Section 2 (Повторение: Раздел 2).....	34
3. The Phonetic Structure of English (Фонетический строй английского языка)	35
3.1. A Brief History of the Phonetic System of English (Краткая история развития фонетического строя английского языка).....	35
3.2. The Theory of the Phoneme (Теория фонемы)	36
3.3. The Phoneme as a Language Unit, Sound as a Speech Unit: the Concept of the Phoneme, Differentiation of the “phoneme” & “sound” concepts (Фонема как единица языка, звук как единица речи: понятие фонемы, дифференциация понятий «фонема» и «звук»)....	36
3.4. Phonemes, their definition: the concept of phonemes & allophones, aspects & functions of phonemes (Фонемы, их дефиниция: понятие фонемы и аллофонов, аспекты и функции фонемы).....	37
3.5. The Composition of Speech Flow: Phonetic Word, Syntagm, Phrase, Phono-Paragraph, Text (Состав потока речи: фонетическое слово, синтагма, фраза, фоноабзац, текст).....	39
3.6. Interpretation of the Phoneme in Works by Prominent Linguists (Трактовка фонемы в работах выдающихся лингвистов).....	39

3.7. The Phoneme Structure of English (Фонемный состав английского языка)	40
3.8. Principles of Classification of (English) Phonemes (incl. Articulatory Classification): Vocalism, Consonantism (Принципы классификации (английских) фонем (включая артикуляционную классификацию): вокализм, консонантизм)	43
Revision: Section 3 (Повторение: Раздел 3)	60
4. Phonetics vs. Phonology (Фонетика и фонология: отграничение)	61
4.1. Phonological Features (Фонологические признаки)	63
4.2. Phonological schools: Moscow, Leningrad and Prague Linguistic Schools (Фонологические школы: Московская, Ленинградская и Пражская лингвистические школы)	64
4.3. Methods of Phonological Analysis: Development of the Main Principles (Методы фонологического анализа: разработка основных принципов)	68
4.4. The Notion of the Phonological Opposition; the System of Phonological Oppositions (Понятие фонологической оппозиции; система фонологических оппозиций)	70
4.5. Types of Transcription (Виды транскрипции)	71
4.6. International Phonetic Alphabet, IPA (Международная фонетический алфавит, МФА)	71
Revision: Section 4 (Повторение: Раздел 4)	72
5. Modifications of Phonemes in Speech, Types of Modifications (Модификация фонем в речи, типы модификаций)	73
5.1. Allophonic Variations & their Causes (Аллофонные вариации и их причины)	73
5.2. Positional & Combinatorial Changes of Phonemes (Позиционно-комбинаторные изменения фонем)	74
Revision: Section 5 (Повторение: Раздел 5)	84
6. The Syllable / Syllabic Structure in English (Структура слога / слоговая структура в английском языке)	86
6.1. Syllable as a Phonological-Phonetic Unit / as a Phonetic & Phonological Unity (Слог как фонолого-фонетическая единица / как фонетическое и фонологическое единство)	86
6.2. Types of syllables (Типы слогов)	86
6.3. The Structure of the English Syllable (Структура английского слога)	87
6.4. Syllable-Forming Sounds in English (Слогообразующие звуки в английском языке)	87
6.5. The Rules of Syllable Division in English (Правила выделения слога / слогоделения в английском языке)	88
6.6. Syllable Formation & Syllable Division (Теории слогообразования и слогоделения)	88
6.7. The Relationship Between Phonetic & Orthographic Syllables (Соотношение фонетических и орфографических слогов)	90
Revision: Section 6 (Повторение: Раздел 6)	90
7. The Accent Structure of English Words (Акцентная структура английских слов)	91
7.1. The Notion of Stress (Понятие ударения)	91
7.2. Word Stress (Словесное ударение)	91
7.3. The Linguistic Nature of Word Stress (Лингвистическая природа словесного ударения)	93

7.4. Degrees / Types of Word Stress: Dynamic, Musical, Quantitative & Qualitative (Степени / типы словесного ударения: динамическое, музыкальное, количественное и качественное)	94
7.5. Factors that Determine the Place / Position & Degree / Strength of Word Stress in an English Word (Факторы, определяющие место / положение и степень / силу словесного ударения в английском слове)	95
7.6. The Functions of Word Stress in Language (Функции словесного ударения в языке).....	96
Revision: Section 7 (Повторение: Раздел 7).....	97
8. Intonation & Prosody (Интонация и просодия)	98
8.1. Approaches to Defining Intonation & Its Components (Подходы к определению интонации и ее компонентов)	98
8.2. The Essence & Bases of Intonation Structure: the Components of the English Intonation, their Functions & the Prosodic Characteristics Of Speech (Сущность и основы интонационной структуры: компоненты английской интонации, их функции и просодические характеристики речи).....	99
8.3. Prosody, Prosodics & Prosodemics (Просодия, просодика, просодемика).....	105
8.4. Functions of Intonation & Prosody / Prosodemics (Функции интонации и просодики / просодемики)	105
Revision: Section 8 (Повторение: Раздел 8).....	106
9. Phonostylistics (Фоностилистика).....	107
9.1. The Notion of Pronunciation Type & Pronunciation Style (Понятие типа произнесения и стиля произношения)	107
9.2. The Classification of Pronunciation & Speech Styles (Классификации произносительных и речевых стилей).....	107
9.3. Speech Culture / Standard of Speech (Культура речи).....	109
9.4. Speech Etiquette (Речевой этикет).....	109
9.5. Phonetic Styles (Фонетические стили).....	110
Revision: Section 9 (Повторение: Раздел 9).....	111
10. Variation / Variants in English Pronunciation (Вариативность / варианты английского произношения)	112
10.1. Pronunciation norm: orthoepy & the orthoepic norm; literary pronunciation, the concept of; the emergence of pronunciation norms (Произносительная норма: орфоэпия и орфоэпическая норма; литературное произношение, возникновение нормы произношения)	112
10.2. Regional & Social Variantology (Региональная и социальная вариантология).....	113
10.3. National Pronunciation Variants of English: British, American, Australian (Национальные произносительные варианты английского языка: британский, американский, австралийский).....	114
10.4. Regional Pronunciation Types of British English: Southern and Northern English Pronunciation, Standard Scottish Pronunciation, Pronunciation in Wales & (Northern) Ireland (Региональные произносительные типы британского варианта английского языка: южноанглийское и североанглийское произношение, общепринятый шотландский тип произношения, произношение в Уэльсе и (Северной) Ирландии).....	116

10.5. American English Pronunciation (Американское произношение)	117
10.6. Australian English Pronunciation (Австралийское произношение)	122
Revision: Section 10 (Повторение: Раздел 10).....	123
11. Present Day Changes in Modern English Pronunciation (Изменения в произношении, происходящие в современном английском языке)	124
11.1. Causes for Changes in Modern English (Причины изменений в современном английском языке)	125
11.2. Modern Types of RP (Received Pronunciation) (Современные виды RP, нормативного произношения)	127
11.3. The Main Changes in the Systems of the English Vowels & Consonants (Основные изменения в системе английских гласных и согласных фонем)	131
Revision: Section 11 (Повторение: Раздел 11).....	131
12. Revision: Final Test (Повторение: итоговый тест).....	133
Keys (Ключи)	137
Revision: Section 1 Answers (Повторение: ответы, Раздел 1).....	137
Revision: Section 2 Answers (Повторение: ответы, Раздел 2).....	138
Revision: Section 3 Answers (Повторение: ответы, Раздел 3).....	138
Revision: Section 4 Answers (Повторение: ответы, Раздел 4).....	139
Revision: Section 5 Answers (Повторение: ответы, Раздел 5).....	140
Revision: Section 6 Answers (Повторение: ответы, Раздел 6).....	141
Revision: Section 7 Answers (Повторение: ответы, Раздел 7).....	141
Revision: Section 8 Answers (Повторение: ответы, Раздел 8).....	142
Revision: Section 9 Answers (Повторение: ответы, Раздел 9).....	143
Revision: Section 10 Answers (Повторение: ответы, Раздел 10).....	143
Revision: Section 11 Answers (Повторение: ответы, Раздел 11).....	144
Final Test Answers (Ответы итогового теста)	146
References (Литература)	147

Contractions

(Сокращения)

AmE – American English

AusE – Australian English

BrE – British English

ModAmE – Modern American English

ModAusE – Modern Australian English

ModBrE – Modern British English

Rus – Russian

Preface

(Предисловие)

Phonetics is a unique and specific branch of linguistics, especially, in the sense that there is, probably, no such other branch, or segment, or discipline within the limits of the science about language and speech that could encompass with such an obviousness and with such a contradictoriness both the material and the ideal circles of phenomena that we usually unite under the term – language. Phonetics (and phonology, be the latter separate or a part of the former) is the field where linguistics suggests an immediate application to other spheres of science and life, and where it embraces openly knowledge from other spheres of science (like mathematics, statistics, physics, computer technology, telecommunication) and even from art (such as music or poetry).

Though the modern linguistics turned a little away from the phonetic considerations (which is, after all, true only to a certain extent, and which is partially due to comparatively recent appearance of many other offsprings of linguistics, which demand more care from their mother-science to survive), and though the phonetic studies attract less attention nowadays than, for instance, grammar (due to its vastness), and cognitive and corpus linguistics (due to their recency and glamour) the issues that were traditionally studied by phonetics (and phonology) still retain their charm, and will always do. It can be added that the charms of phonetics are positive not because all the problematic issues were solved, but rather because they were not.

A reader of the hereto volume may notice that the essays collected here differ much in their size, and some may seem to be too expanded and eloquent, whereas others – too concise and short. In any case it is our only one suggestion that a moderate student should find everything necessary in the collection of essays to pass the exam in the course of the *Theory of English Phonetics* lectured at the faculties of foreign languages. Besides, and yet, we sincerely hope (and this is the solely real aim of the hereto essays without which the whole collection has hardly any value at all) that the essays will find a more intellectual and devoted reader for whom the essays will serve as a threshold to stimulate his / her striving for search and quest in the field of phonetic studies. In these essays we tried to enlighten the most interesting, the most crucial, the most paradoxical and, certainly, the very basic and key points belonging to the domain of phonetics (and phonology). Whether we gained success in performing the task or not – there are our students and readers to judge, to whom we, using this as one more opportunity, would like to express our greatest love, respect, trust, and gratitude, and many thanks. It is for the sake of our students and readers, and for the sake of their interest and curiosity in phonetics that we present the hereto essays.

We need to add also that it is our second hope that the long dozen of essays collected in this volume will seem for its readers to be not at all long, and that it will rather provide for its readers a short and sweet, and fast access to the most important subjects of phonetics.

The both theoretical and practical approaches to the questions of phonetics have boosted rapidly in the recent decades-due to an easier access to large corpora and collections of authentic data via various video and audio hosting services, and to an excellent development in technology and methods available for ‘research at hand’. Theoretically, many issues have been (and still are) under revision both by the Russian and foreign phoneticians. Some put much effort in raising public awareness on the linguistic, including phonetic, questions, using the modern channels of media communication, like podcast interviews, their own websites and blogs and video channels. We would like to point out the British phonetician Dr. Geoffrey Lindsey whose fruitful work has led to our deep reconsideration of many topics while preparing the re-issue of this study manual. We’ve chosen to give some brief account of Dr. Lindsey’s and some other scholars’ ideas in special extra sections throughout the text of this book with hope that the readers will find opportunities to get acquainted with his original materials available in a free access mode. We are also thankful to the ideas on articulatory settings while teaching English phonetics presented in lectures and publications by Piers Messum. Another round of thanks goes to the Russian dialectologist phonetician Igor Isaev (the Institute of Linguistics, Russian Academy of Sciences) for his numerous interviews filled with professional approaches presented in a simple way and for his personal consultation on a number of issues.

Finally, we would like to say that we are open for any kind of discussion or questions, or any other communication on the hereto matters, and we want to express our strong belief that among various branches of linguistics the most interesting is, of course, phonetics.

*The authors – Yaroslav A. Glukhiy,
Sergei V. Glushkov,
Valeriya M. Lemskaya*

P.S. Here we must also express our deep grief and missing of a dear colleague and mentour, Yaroslav Andreevich Glukhiy (1941-2014), who was an amazing phonetician himself, field explorer discovering the mysteries of the Enets phonetics (a minority language of the Russian North), and a true devoted teacher famous among and loved by his students setting an example of perfect English pronunciation and virtuosity whom we both have learned much from.

Valeriya M. Lemskaya, Sergei V. Glushkov

1. Phonetics as Science **(Фонетика как наука)**

Phonetics is a branch of linguistics, like the other branches, such as lexicology and grammar. Lexicology studies the vocabulary of language, the origin and development of words, their meaning and word-building. Grammar (divided into morphology and syntax) studies the regularities in modification of words and in combination of words into sentences.

1.1. The Subject and Tasks of Phonetics **(Предмет и задачи фонетики)**

The term *phonetics* comes from the Greek word φωνή (phōnē) meaning *sound, voice*. It is rather difficult to give an exact and precise definition of the term. The oldest and most concise definition is that phonetics is a science of speech sounds. The definition is rather loose partly because it has very little consideration of the meaningful and functional aspects of speech sounds, partly because speech sounds are also studied in acoustics and physiology. The definition of phonetics as “the science of speech sounds considered as elements of language” (Webster’s Dictionary, cit. by Vassilyev, 1970 : 7) is seemingly better but does not take into account that phonetics also studies stress, syllable, intonation.

Another definition states that phonetics is the science that “studies the sound system of the language, that is segmental phonemes, word stress, syllabic structure and intonation” (Соколова, 2010 : 6). This obviously includes segmental phonemes within the sphere of phonetic studies, which is in conformity with the Russian (Soviet) school of phonetics, but, and it will be clear later, is in complete disagreement with the Western school of phonetics. To say that phonetics studies speech sounds means that it explores only sound sequences produced by a human vocal apparatus which are carriers of organized information of language and which are meaningful.

1.2. The Branches of Phonetics **(Разделы фонетики)**

Phonetics is divided into several branches in accordance with its aspects and tasks based on various criteria.

1.2.1. Auditory, Acoustic & Articulatory Phonetics **(Аудитивная, акустическая и артикуляционная фонетика)**

The common approach is that phonetics has three branches: articulatory phonetics, acoustic phonetics and auditory phonetics. This division is grounded, roughly, on the speaking-

hearing criterion. Articulatory phonetics is the branch of phonetics “which is concerned with the study, description and classification of speech sounds as regards their production by the human speech apparatus” (Vassilyev, 1970 : 10). It is the most productive, developed and the oldest branch of phonetics. “From the articulatory point of view every speech sound is a complex of definite, finely coordinated and differentiated movements and positions of the speech organs” (op. cit.). Acoustic phonetics is the branch of phonetics that studies the physics of the air vibrating between the mouth of the speaker and the ear of the hearer and producing sound waves of different character and acoustic effect, i.e. different speech sounds. “From the acoustic point of view, a speech sound, like any other sound in nature, is a physical phenomenon, a kind of moving matter and energy” (Vassilyev, 1970 : 12). Auditory phonetics is the branch of phonetics which main concern is the investigation of the hearing process. At present time it mainly deals with the brain activity rather than with the physiological process of signal delivery and transfer through the nervous system of the hearer. Its interest lies more in the sphere of experimental psychology not linguistics. Till lately some linguists have not distinguished auditory phonetics as a separate branch of phonetics at all. But, still an outstanding Polish-Russian linguist, Baudouin de Courtenay (full name – Jan Niecisław Baudouin de Courtenay, 1845-1929) in St. Petersburg known as Ivan Aleksandrovich Baudouin de Courtenay, 1845-1929) has invented the term anthropophonics, which united the articulatory and the auditory aspects of speech sounds. In the Soviet (Russian) linguistics the term has changed to physiological phonetics due to close connection between the articulatory and the auditory aspects of speech sounds. When our speech organs work, we at the same time hear what we speak, thus receiving feedback from our hearing system. Step by step, through accumulation of instrumental data, auditory phonetics finds its place in the sphere of phonetic science. It is necessary to comment that Baudouin de Courtenay was the first one who introduced the term phonology, as the science of purely linguistic (functional) aspect of speech sounds, opposed to anthropophonics (the proper phonetics in modern terms), see also an example of auditory word perception modelling in (Shillcock, 1992).

1.2.2. General & Special Phonetics (*Общая и частная фонетика*)

Phonetics can be divided into general and special sections due to its scope of its study – universal principles across all languages or the specific systems of a single language (Соколова, 2010 : 7):

1) general phonetics is a section studying all the sound-producing possibilities of human speech apparatus (organs) in various languages. It studies the laws that govern the changes which speech sounds undergo in the flow of speech and finds out the types of such changes in various languages;

2) special phonetics is the branch of phonetics concerned with analysing and describing the phonetic structure of a particular, individual language; it examines specific phonemes, their allophones, and stress, rhythm, and intonation patterns unique to a language and provides data for general phonetics to build theories, and applies broader theories and methods to its analysis.

1.2.3. Descriptive, Historical & Comparative Phonetics

(Описательная, историческая и сравнительная фонетика)

Based on the temporal perspective or dimension of the analysis, phonetics can be:

1) historical phonetics – is a section that traces and establishes the successive changes in the phonetic system of a given language or a language family at different stages of its historical development. It helps to understand how this modern phonetic system came to be and what changes may take place in the future;

2) descriptive phonetics is a branch of special phonetics that analyses and describes the phonetic system of a language at a single point in time, focusing on its current state and structure; it is aimed at systematically documenting the present-day phonemes, stress patterns, rhythm, and intonation of language (Соколова, 2010 : 7).

Descriptive phonetics can also be complemented or even opposed to comparative phonetics depending on the application of the primary methods or research. Thus, comparative phonetics studies the correlation between the phonetic systems of two or more languages, especially kindred ones. It finds out the correspondences between their speech sounds. E.g., it is the comparative phonetics that tells us that to every initial prevocalic [z] of Modern German there corresponds the consonant [s] in Modern English: *senden* – *send*, *silber* – *silver*, *Saltz* – *salt*, *singe* – *sing*, *Seite* – *side*.

1.2.4. Theoretical, Applied & Experimental Phonetics

(Теоретическая, прикладная и экспериментальная фонетика)

With regard to the primary objective and end use of the phonetic knowledge, phonetics can be divided into (Евстифеева, 2022 : 16):

1) theoretical phonetics – deals with theoretical problems of a particular language. It gives students the latest theories and views on many phonetic problems. It describes some moot points in a particular language;

2) applied phonetics which is concerned with the practical utilization of phonetic knowledge and methods to solve real-world problems across various professional, educational, and technological fields.

Applied phonetics may include experimental phonetics which uses tools, methods, and empirical data to solve practical problems. Experimental phonetics – is a section that studies various phonetic phenomena in the laboratory conditions by means of all sorts of devices or within situations by observation of reactions to certain stimuli with or without use of special equipment. It has gained far-reaching results in the recent decades with development of various technologies.

1.3. The Connection of Phonetics with Other Branches of Linguistics (Связь фонетики с другими разделами лингвистики)

Phonetics also studies the relation between written and spoken language and is connected with other linguistic disciplines such as grammar, lexicology, stylistics, etc.

1.3.1. Grammar (Грамматика)

Phonetics is connected with grammar through the system of rules of reading, which make it possible to pronounce correctly the past tense forms of regular verbs, the singular and plural forms of nouns, and etc. Please consider the examples:

1) the differences in pronunciation of morpheme -ed after voiced and voiceless consonants in past forms of verbs:

[d] after voiced consonant	[t] after voiceless consonants
beg – begged	stop – stopped
clog – clogged	knock – knocked

2) the differences in pronunciation of root consonants observed in singular and plural forms of nouns:

leaf – leaves
house – houses
bath – baths

3) the vowel interchanges helping to distinguish the singular and plural forms of nouns and the tense forms of irregular verbs:

nouns	verbs
basis – bases [si:z]	write – wrote – written
crisis – crises [si:z]	ride – rode – ridden
thesis – theses [si:z]	bite – bit – bitten

1.3.2. Lexicology (Лексикология)

Phonetics is also connected with lexicology, since the presence of stress in the right place helps to distinguish nouns/adjectives from verbs. Homographs may also be differentiated only by pronunciation, because they are identical in spelling. Please consider the examples:

nouns	verbs	adjectives	verbs	homographs	
an 'accent	to ac'cent	'separate	to sepa'rate	row [rau]	row [rəu]
a 'desert	to de'sert	'predicate	to predi'cate	lead [li:d]	lead [led]
an 'object	to ob'ject	'graduate	to gradu'ate	wind [wind]	wind [waɪnd]

1.3.3. Stylistics (Стилистика)

Stylistics is another linguistic discipline phonetics is connected with through intonation, speech-melody, rhythm and word stress. Regular repetition of stressed and unstressed words gives a rhythmic coloring to the English sentence. Logically accented words in a sentence help to give special prominence to a sentence, and speed of delivery helps to differentiate formal and informal styles which express different emotions and attitudes of the speaker. Also intonation which possesses definite phonetic features helps to distinguish various types of utterances, e.g. (Леонтьева, 1988 : 7-9):

He 'went to his \office.	(a statement of fact)
He 'went to his /office.	(a question)
↓He went to his √office.	(an implication: surprise, correction)

1.4. Interconnection of Phonetics with other Branches of Science

(Взаимоотношения фонетики с другими науками)

Speaking-hearing (and human speech in general) as a phenomenon is a result of a complicated event series many of which are still and probably never will be known.

1.4.1. Acoustics (Акустика)

The production of speech involves the vocal organs disturbing and modulating a stream of air to generate sound waves. This phase is referred to as the physical or acoustic stage – acoustics being a subfield of physics. Consequently, speech sounds can be analysed from the perspectives of both acoustics and physiology.

1.4.2. Anatomy (Анатомия)

The connection between **phonetics and anatomy** is fundamental and direct (Соколова, 2010 : 10). Phonetics is the scientific study of *how* speech sounds are produced, transmitted, and perceived. This production is entirely dependent on the anatomical structures and physiological functions of the **vocal apparatus**. Phonetics analyses how the biological system of organs whose primary roles are breathing and swallowing is co-opted to generate sound. The phonetic process maps directly onto specific anatomical components like respiration, phonation, etc. In other words, phonetics is the functional study of speech anatomy. The anatomical structure sets the physical possibilities and constraints for the phonetic system of any human language.

1.4.3. Physiology (Физиология)

The coding activity initiates particular signals that come from the brain of the speaker to his / her articulating organs which move to produce particular speech sounds – being physiological in its nature the stage is called articulatory. From thereon the whole process goes backwards – fourth – the sound waves reach the ear of the hearer where they initiate signals to be transmitted to the brain of the hearer – being also physiological in its nature the stage is called auditory.

1.4.4. History (История)

Phonetics is linked to history primarily through its subfield of historical phonetics, which studies the evolution of a language's sound system over time (Евстифеева, 2022 : 11). By tracing sequential changes in pronunciation across different periods, it explains how modern phonetic systems developed from their earlier forms. Furthermore, by identifying established patterns of sound change, this diachronic branch of phonetics allows linguists to make informed predictions about potential future developments, thereby directly connecting the analysis of speech sounds to the broader history of a language.

1.4.5. Psychology (Психология)

Phonetics is connected with psychology via its focus on the mental processes of encoding and decoding speech. The connection is established because the cognitive information (a concept, an idea, a cognit, etc.) is coded into the linguistic one (it can be said also that the linguistic code comprises an integral part of cognitive information), i.e. concepts, ideas, etc. are coded into linguistic signs (patterns of speech sounds / phonemes, for instance) – this stage of the speaking-hearing process, in the broad term, is called psychological. Furthermore, the transmitted signals initiate decoding of the linguistic information (particular linguistic signs) into the original concepts, ideas – this stage, in the broad term, is also psychological. It is also noted that the interest of auditory phonetics lies both in the spheres of experimental **psychology** and linguistics. Therefore, phonetics is inherently linked to psychology in its study of how the mind produces and comprehends the sound structure of language (Беляева, 2019 : 18).

1.4.6. Sociology (Социология)

The connection between phonetics and sociology can be characterised as follows. Phonetic features (how sounds are produced) are not just biological or linguistic abstractions but are socially indexical. They function as powerful markers of social stratification, perception and change, group identity and affiliation, etc. Phonetics provides empirical data on *how* people speak, while sociology explains *why* they speak so in a given context linking sounds to social structure, identity, and power.

1.4.7. Logic (Логика)

The connection between phonetics and logic is established primarily through the concept of logical stress and can be characterised as follows (Евстифеева, 2022 : 11). Phonetics, specifically intonation and stress, serves as the auditory tool for expressing logical relations within an utterance. A speaker phonetically marks the logical focus of the message, thereby clarifying intent, contrast, or new versus given information. Phonetics provides the expressive means for realising logical distinctions that are fundamental to coherent communication.

1.4.8. Other Branches of Science (Другие науки)

Phonetics is connected with a range of other sciences, which can be summarized into three main groups (Соколова, 2010 : 13):

✓ **Mathematics, Statistics, and Technological Sciences:** phonetics openly integrates knowledge from mathematics, statistics, computer technology, physics, and telecommunications (mathematics and physics provide the theoretical framework, statistics enables rigorous data analysis, computer technology offers the instrumental toolkit, and telecommunications represents a key area of applied innovation. This transforms phonetics from a descriptive discipline to a precise, quantitative, and technologically engaged science).

✓ **Neuroscience and Cognitive Science:** it is linked to modern neuroscience and cognitive science, which study the neural connections in the brain underlying cognitive and linguistic information.

✓ **Semiotics:** the field is also connected to semiotics, as the stages of speech communication are viewed as transformations of semiotic signs (a process of semiosis).

Summarizing our understanding of the speaking-hearing as a phenomenon, we can list several stages of the process. A few points need to be mentioned here: the psychological stages are the least studied and usually provide support for various and far too often opposite views; the five-stage approach is only a model and as every model may be quite different from reality; our knowledge of the processes going at every stage may differ sharply from reality and sometimes suggests dubious interpretations; the so-called “coding” and “decoding” terms are inherited from the times when language and intellect were regarded as different (even opposite) objects (whereas the former is obviously a part of the latter); modern neuroscience and cognitive science regard the cognitive information, i.e. intellect / mind, and therefore the linguistic information, i.e. language / linguistic code, to be the function of connections activated between assemblies and networks of neurons in the human brain (Fuster, 2003) – yet no definite answer as to how the language and mind correlate with the neurons and brain seem to be even discerned at the scientific horizon; all the five stages can be regarded as representing different transformations of semiotic signs and thus can viewed in terms of semiosis (Fig. 1):

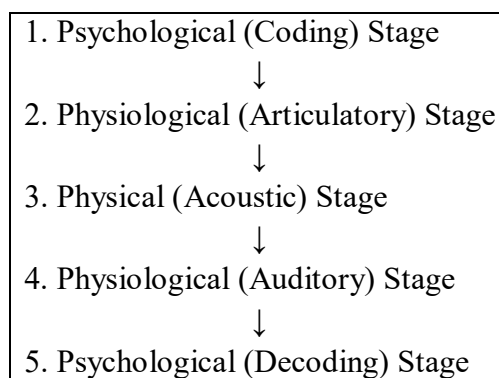


Fig. 1. Speaking-Hearing Process

1.5. Methods of Phonetic Research

(Методы фонетического исследования)

It cannot be said that each branch of phonetics has its own absolutely unique methods. Partly this happens due to intricate and intimate relation between phonetics and phonology. For instance, the experimental method and the method of direct observation are applied in articulatory, and in acoustic and in auditory phonetics as well.

We will distinguish between the following **methods of phonetics** (*phonetics* is understood in terms of the Soviet / Russian school): 1) *direct observation method*; 2) *experimental (instrumental) method*; 3) *instrumental methods*; 4) *method(s) of phonological analysis* (see Section 4.3).

The direct observation method is the method of observing the facts of a language in their natural surrounding. In phonetics it comprises three important modes: *observation by ear, by sight and by muscular sensation*. The method can be effectively employed only if a researcher has been specially trained to differentiate between “the minutest movements of their own and other people’s speech organs, and to distinguish the slightest variations in sound quality” (Дикушина, 1965 : 16).

The experimental method is the method of obtaining data and facts of a language through constructing and modelling special artificial conditions. It is more economic than the direct observation method and makes a researcher able to find out not only what is allowed but also what is prohibited in a certain language. Consider the following example from Russian: the task is to prove the statement that in Russian the sequence of /st/ and /n/ divided by morphological boundary is pronounced as /sn/ (let us schematically represent this as: /st/+ ⊥ +/n/= /sn/). If we apply the direct observation method we need to analyse a great volume of speech material looking for the examples demonstrating or refuting the statement. And still we cannot be sure as some exclusion could well avoid our analysis by accidentally missing to appear in the speech material. The experiment allows us to identify the group of words subject to analysis (because we know what shall be on the left side of our “equation”) and listening to the Russian speaker pronouncing them we can quickly define whether the statement is correct or incorrect. Experiments are especially good in identification of exclusions and what is not allowed in a

language. Consider the pronunciation of the Russian words that prove the above statement: *честный, местный, возрастной* – here no phoneme /t/ is pronounced which supports the original statement and – moreover – the statement that the pronunciation of /t/ under the circumstances is prohibited in Russian.

1.5.1. Instrumental & Experimental Methods of Phonetic Research

(Инструментально-экспериментальные методы фонетического исследования)

Instrumental methods are sometimes called **experimental**, which is not quite correct because, on the one hand, an experiment can be implemented **without any tool** or apparatus and, on the other hand, the usage of a technical device does not obligatory imply an experiment. If one makes a record of speech in natural conditions with the help of a tape or CD recorder – he/she applies the **instrumental method without an experiment**. Yet, very often the usage of a technical device, tool, or instrument is combined with an experiment run for a certain purpose, under certain conditions and to obtain certain goals. There are various and different instruments applied in phonetic studies. Among the old ones there deserve to be mentioned:

- the **laryngoscope**,
- the **artificial palate**, and
- the **kymograph**.

“**The laryngoscope** is a small circular mirror that is introduced into the pharynx as far back as possible” (Дикушина, 1965 : 20), and that allows to see the upper part of the pharynx and the vocal cords. The artificial palate allows seeing “the exact tongue-palate contacts in articulating sounds” and is to be made for each informant specifically “according to the exact shape of the individual hard palate” (Дикушина, 1965 : 19). A fine white powder is sprinkled onto the artificial palate, which is, then, to be fitted into the person’s mouth and the sound is to be articulated. In the place of the contact between the tongue and the palate the powder (chalk or talc for instance) will be removed, and the artificial palate, being extracted, can be then photographed or diagrammed. The drawing or the photo of the artificial palate is called **palatogram** and it shows the articulatory features of the sound pronounced, such as the place and point of articulation and strength and muscular tension.

The kymograph is an apparatus that transmits “the variations in the air pressure of the air stream emitted from the mouth” (Дикушина, 1965 : 21) through the thin rubber tube to a stylus, which leaves tracings on the blackened paper. The kymograms reflect acoustic characteristics of speech sounds, and allow to identify and to distinguish between vowels and consonants, between different vowels, between voiced and voiceless consonants, between different consonants. The basic schematic design of the kymograph can be seen in Fig. 2.

There are other, comparatively new, **methods** of phonetic research with the help of special apparatus. They are *cinematography* and *photography*, *X-ray photography (still and moving)*, *electromyography*, *tomography*, *oscillography*, and *spectrography*.

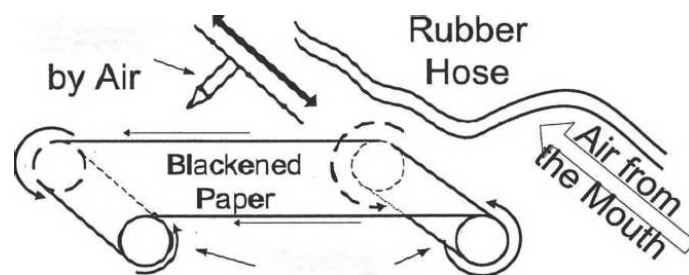


Fig. 2. The Basic (Principle, Schematic) Design of the Kymograph

Oscillography & spectrography. The latter two enable to obtain acoustic characteristics of sounds like frequency, amplitude, spectrum; and auditory characteristics like pitch, loudness, timbre; and prosodic characteristics like length and intonation. Oscillograms (Fig. 3) and spectrograms (Fig. 4) bear indirect information about the articulatory characteristics of speech sounds as well. Therefore, oscillography and spectrography are the most widely applied techniques in speech analysis. The oscillograph is an apparatus that transforms sound waves of different frequencies into electric signals of different frequencies that can be viewed on the screen, recorded, photographed or diagrammed. Oscillograms reflect all the acoustic characteristics of a speech sound.

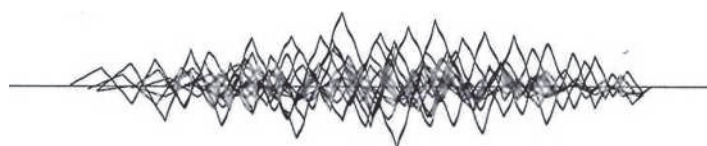


Fig. 3. The Oscillogram (a schematic drawing)

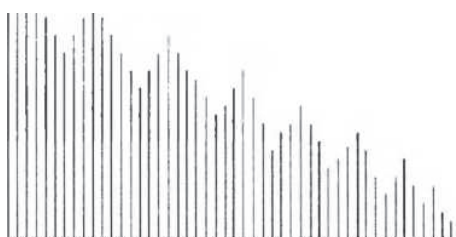


Fig. 4. The Sound Spectrum (a schematic drawing)

Note: formants, including the zero-formant, can be seen in Fig. 4 as high peaks continuously stepping down. Both Figures are only principle schematic drawings. The spectrum of a real sound, for example, a vowel, pronounced by a particular person has a unique “formant picture” – the distance between the formants, their height and number differ considerably from vowel to vowel, and significantly enough from speaker to speaker, the latter fact makes it possible to identify the speaker by the formant picture of his / her voice. The table to follow (Table 1) represents the correlation between some acoustic (physical) and auditory (physiological) characteristics of speech sounds.

Table 1

Correlation Between Acoustic and Auditory Characteristics of Speech Sounds

№	Acoustic Characteristics	Auditory Characteristics
1	Amplitude / intensity (амплитуда, сила)	Loudness (громкость)
2	Frequency (periodical oscillations / musical tone / (or just) tone; non-periodical oscillations / noise) (частота)	Pitch (высота / тон / высота тона / высота голоса)
3	Spectrum (спектр)	Timbre (тембр)

Sometimes the acoustic characteristics are called objective and the auditory characteristics are called subjective which mean that the former are physical properties independent of the auditory apparatus of the hearer and the latter are, to certain extent, individual depending on the perception of hearing apparatus and nervous system of a particular person (what is “loud” for one person can be “muffled” or “not so loud” for another person).

As frequencies of different cavities of a human speech apparatus overlap each other in oscillograms it is difficult to read them to discern waves of different frequencies. To overcome the difficulty in reading the spectrum of a speech sound the method of spectrography is used. The spectrograph is an apparatus that divides the spectrum of a speech sound (comprised of sound waves of different frequencies) into its components (i.e. into particular frequencies) by adjusting a special filter to a special frequency.

As spectrography and oscillography are based on strict physical phenomena and processes they can be definitely and precisely described by mathematics and imitated by an automatic processor. This is why modern oscillography and spectrography of speech sounds are effectively processed by computers, which makes the two instrumental techniques very convenient and handy. The most well-known software products, specially utilized for the purposes of instrumental analysis of speech, are *Praat* (by Paul Boersma and D. Weenink from Institute of Phonetic Sciences, University of Amsterdam), *Speech Analyser* (by SIL International), *Speech Filing System* (from the University College London), *Signalize* (by LinguistList Plus Inc.), *Sona* (from the University of Bonn), *SpeechStudio* (by Laryngograph Ltd.) and *Computerized Speech Lab* (by Kay Elemetrics).

Ultrasound. This methodology employs ultrasound-based articulatory research using the **Articulate Assistant Micro** portable system (Лемская, 2025), cf. Fig. 5.

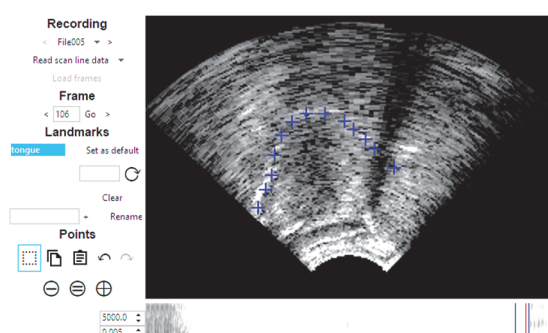


Fig. 5. An ultrasound screening of the tongue within the AAA software

This equipment comprises a scanner with a transducer, a headset microphone, and **Articulate Assistant Advanced (AAA)** software, cf. (Washington, 2020). Data collection involves presenting pre-determined stimuli on-screen to participants who articulate each stimulus three times at a comfortable pace (Зияданова, 2025). The transducer fixed under the chin using a stabilising helmet captures mid-sagittal tongue images in real time, whilst a synchronised microphone records the audio (Лемская, 2025). The subsequent analysis utilises a multi-software pipeline. Acoustic boundaries for target sounds (e.g., the /k/ stop) can be first annotated in **Praat** using spectrograms. These time-aligned intervals can be processed in **UltraTrace**, where researchers manually trace the tongue contour, or spline (Washington, 2020), Fig. 6, on the clearest ultrasound frame for each annotated segment. The traced coordinate data can be exported for statistical processing using custom **Python** scripts which generate averaged tongue shape profiles with standard deviation plots, allowing for a comparative analysis of articulation across different vocalic contexts and word positions (Токмашев, 2023).

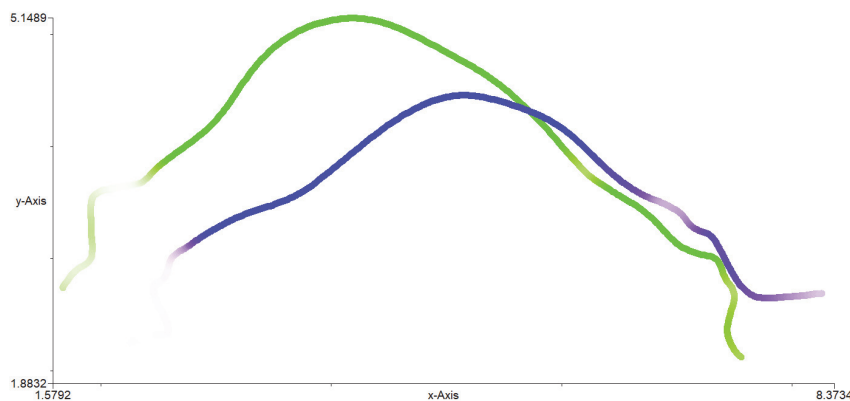


Fig. 6. Comparison of the average of the /a/ phoneme by two speakers analyzed by the AAA software

Some **methods of phonological analysis** (the linguistic / functional method) must be mentioned here, too (for more detail see Section 4.3). This includes several interconnected methods aimed at establishing the inventory of phonemes: 1) the *distributional method* (analysis establishing the distribution of speech sounds in the words of a language in all positions and combinations); 2) the *statistical method* (analysis establishing the frequency, probability, and predictability of occurrence of phonemes and their allophones); 3) the *semantic method* (*method of minimal pairs, commutation test*; analysis establishing the ability of sound segments to differentiate meaning by substituting sounds in a constant phonetic environment).

!

The source-filter model of the human voice

The human voice can be explained using the source-filter model, established by G. Fant. In this model, the source is the vocal folds in the larynx, which vibrate to create a basic sound. The filter is the vocal tract, which shapes that sound into distinct speech, cf. (Lindsey, 2011).

The source determines the pitch and intonation, while the filter creates the resonances (formants) that produce different vowels and much of our unique voice quality. Although both parts grow as we age, they do so at different rates.

During adolescence, boys' voices 'break' as testosterone causes a sudden lengthening of the vocal folds, lowering their pitch range. The software Praat has a 'Change gender' feature which can modify the pitch and the formants independently. Interestingly, a voice can sound higher even when the fundamental pitch contour remains unchanged.

1.5.2. Basics of Articulatory, Acoustic & Auditory Aspects of Speech Sounds

(Основы артикуляционного, акустического и аудитивных аспектов звуков речи)

Philosophically speech sounds and phonemes relate to each other as form and content. They relate as substance / matter and idea also. Speech sounds are material form of phonemes. Phonemes are ideal content of speech sounds. As both belong to reality (or better say we think they belong to reality) one can study speech sounds and phonemes within one branch of linguistics, or within different branches, which can be seen in different and various approaches onto the relation between phonetics and phonology. For the purpose of this essay we will consider speech sounds and phonemes within two different branches of linguistics – proper phonetics and phonology correspondingly.

A speech sound is the impact of acoustic wave, produced by human organs of articulation, onto the human hearing apparatus. Thus speech sounds have three aspects – the articulatory, the acoustic and the auditory, cf. (with the definition in (Vassilyev, 1970 : 71)).

In the articulatory aspect “every speech sound is a complex of definite, finely coordinated and differentiated movements and positions of the various speech organs” (Vassilyev, 1970). Acoustically a speech sound is a wave movement of matter and energy in the air medium. From the auditory point of view a speech sound is a series of very precise and involuntary movements and interactions of hearing organs.

All the three aspects are inseparable from each other in the sense that they are intrinsic to the nature of a speech sound and none of the latter can exist without each of the three. Yet, the analysis of speech sounds can be done for every aspect separately, exploiting separate and special terminology, depending on what point we view the speech sounds (being objects of reality) from and what side of speech sounds we take into account.

Groups of organs that take part in articulation of speech sounds in this or that way are as follows:

- 1) the lungs, the bronchial tubes, the trachea or the windpipe,
- 2) the larynx with vocal cords,
- 3) the pharynx,
- 4) the mouth cavity (including the epiglottis, the soft palate with the uvula, the hard palate, the tongue (having the tip, the blade, the middle part, and the back) the alveolar ridge, the teeth, the lips),
- and 5) the nasal cavity. The original and primary functions of the organs are biological ones (see Fig. 7 for organs of articulation).

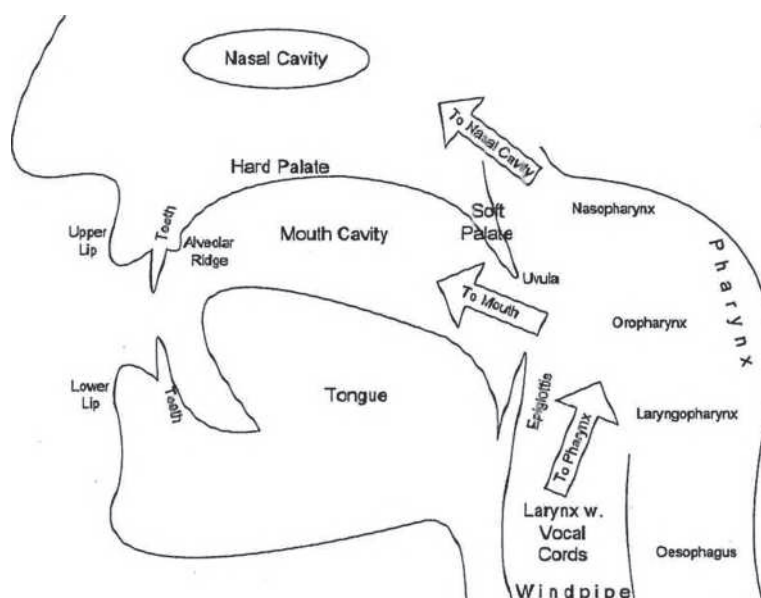


Fig. 7. Human Organs of Articulation

Some of the terms here need to be explained. The **pharynx** or the **pharyngeal cavity** is the cavity that connects the mouth cavity, the nasal cavity and the top of the larynx and oesophagus. The pharynx includes the laryngopharynx, the oropharynx, the nasopharynx. The **epiglottis** (Rus. *надгортанный хрящ*) is a thin flap of tissue at the back of the tongue that covers the windpipe during swallowing to prevent food or drink from entering the lungs. The **soft palate** with the uvula (Rus. *небная занавеска с маленьким язычком или увулой*) is a flap of tissue that hangs from the top of the mouth in its back part. In its raised position the soft palate covers the nasal cavity directing the air stream into the mouth cavity. The **larynx** or the **voice-box** is a movable frame of cartilage held in place by membranes and muscles (Vassilyev, 1970 : 75). It can move over the space of about a centimetre, thus changing its own volume and causing the vocal cords to change their position. **Vocal cords** are a pair of folds or muscles placed in the middle of the larynx and connected to it with their edges reaching from front to back. In their back parts each fold is attached to one of a pair of adjustable cartilages that can be moved by muscles to bring the inner edges of the vocal cords completely together or spread them apart (Fig. 8).

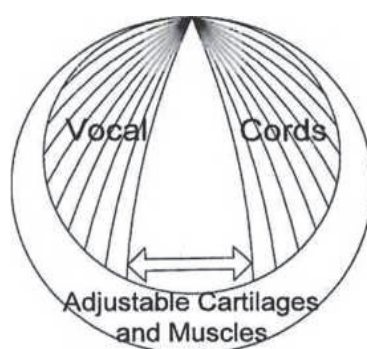


Fig. 8. Vocal Cords (top view)

The inner edge of the vocal cords is about 23 mm in adult men and about 18 mm in women (Vassilyev, 1970 : 75), cf. with L. R. Zinder's data: 15 mm and 12 mm correspondingly (Зиндер, 2007 : 91). It is still not quite clear how the vocal cords work. According to one of the theory the vocal cords work like a siren, which is that they vibrate in their whole length and change the pressure of the air stream alternately, and the alternate rarefactions and condensations of the air are the sound waves that reach the ear. According to another theory (which had more adherents), vocal cords work like strings, which means that the passing air makes them vibrate not in their full length only but in their parts also, which again alternates the pressure of the air stream. The vibrations of the parts of the vocal cords are proportional to the vibrations of the whole length of the cords. In any case a man's and a woman's voices differ because the man's vocal cords are longer and thicker.

The tongue is the most flexible organ of articulation. There can be distinguished the following parts of the tongue, depending on what fixed organs lie against what part of it: the front of the tongue – lies opposite the upper teeth, the middle or the central part – lies opposite the hard palate, the back part – opposite the soft palate. In English the front part of the tongue is usually divided into the tip and the blade.

The pharynx, the mouth and the nasal cavities are combined within the term supra-glottal cavities, which determine the formant characteristics of speech sounds. But here the acoustic aspect of speech sounds starts.

According to both theories on the work of the vocal cords stated above, the vibration of the whole length of the vocal cords creates the sound wave which frequency is called the frequency of the main tone or the main frequency, or the fundamental frequency of sounds of speech. It is denoted as f_0 . All other frequencies of the spectrum of a sound are denoted as f_1 , f_2 , f_3 and so on. If $f_0=x$, then $f_1=2x$, $f_2=3x$, $f_3=4x$ and so on because the vibrations of the parts of the vocal cords are proportional to the vibrations of the whole length of the cords (see above). Frequency is the number of double oscillations against the fulcrum (the point of rest) at a unit of time (1 sec). Frequency is measured in hertz (Hz, kHz). Formants are the frequencies of the maximums of the spectrum of a speech sound, which are the maximums of the concentration of energy in the spectrum (Златоустова, 1997 : 58-59). Though frequencies and formants of speech sounds are different notions, formants, being the maximum peaks of frequencies in the spectrum, follow the scale of frequencies, and correspond to them. Formants are denoted as F_0 (zero formant), F_1 , F_2 , F_3 and so on.

Amplitude (or intensity) of sound wave is the sound energy passing through a unit of square (1 sq. cm) at a unit of time (1 sec). Amplitude or intensity is measured in decibels (dB). The **acoustic spectrum** is “the complex range of frequencies of varying intensity which constitute the quality of a sound” (Vassilyev, 1970 : 13).

The two approaches to the work of the vocal cords have a far going influence onto the acoustic theory of speech sounds. If we adhere to the first theory we have to assume that the vocal cords create the fundamental frequency only (the maximum of which is the zero formant). All other frequencies are created in the supra-glottal cavities and are not dependent on the fundamental frequency. The author of the theory was L. Hermann, a German scientist, who in the end of the 19th century suggested an idea that in the speech sound spectrum all the

frequencies higher than the main frequency appear due to the air stream passing through the cavities located above the vocal cords. This happens because a tube (and human supra-glottal cavities can be approached as a number of tubes) creates the wave of a certain frequency if air, or water, or any other substance passes through it, or if the tube moves in the substance.

The frequency of the wave depends on the size and length of the tube, on the volume of air held inside it, on the entrance and orifice of the tube, on the form of the inside walls and the material which the inside walls of the tube are made of. According to the theory, the supra-glottal cavities modulate the air stream creating their own self-frequencies or tones, which overlap the fundamental frequency or tone (see Fig. 3 for oscillogram and Fig. 4 for spectrogram). An important conclusion of the theory is that all the formants above the zero formant are not harmonically proportional to it.

The second theory is based on the approach that the vocal cords work like a string, i.e. that they vibrate not only in their full length, but in their parts also. The vibrations of parts of the vocal cords are harmonically proportional to the vibration of the whole length of the cords, which has been already mentioned. Still in 1780 the Russian academician H. Kratzenstein said that the spectrum of vowels is formed by the amplifying of certain partial (Rus. *парциальный*) tones or frequencies, created by the vibrations of the parts of vocal cords, in the supra-glottal cavities. This theory was further developed by H. Helmholtz, another outstanding German scientist. He suggested the following mechanism for the formation of the spectrum of a speech sound: the vibrations of the whole length of the vocal cords and their parts create a great number of harmonics, or periodical or proportional frequencies, or periodical tones, or just tones (whichever one calls them) and the supra-glottal cavities work like resonant chambers amplifying some of the frequencies (i.e. increasing their amplitude) and damping others. A tone which corresponds to the self-frequency of any of the supra-glottal cavities (which can be regarded as a number of tubes) is amplified, whereas a tone which does not correspond to it is subdued. The supra-glottal cavities respond solely as filters letting certain frequencies pass by and fading the others. As the walls of human supra-glottal cavities are soft and wet the resonant filters have small selectivity, which is that they amplify not only the waves precisely corresponding to the self-frequency of the cavities but also the waves which frequencies fall within certain range (like the range of frequencies between 350 and 450 Hz). The maximums of the frequencies comprise the formants of the spectrum, the number of which in a vowel, for instance, can be up to seven, though the most important are the zero formant and the consequent three or four. The conclusion of the theory is that the frequencies higher than the fundamental one are proportionally dependent on the main tone.

For a long time the two theories were regarded as intransigent and uncompromising to each other. The drawback of the first theory is that the formants (and frequencies) of a speech sound are periodical in reality (or, better say, tend to be periodical) to the zero formant. That was long considered to be a contradiction to the theory. The weak point of the second theory is that it is not clear how and what supra-glottal cavities bear responsibility for a particular formant. The latter theory has also a weak explanation for non-partial, non-periodical tones / frequencies in the spectrum of speech sounds.

In the middle of the 20th century G. Fant published his book in the acoustic theory of speech which is one of the main books to start the study of the acoustic aspect of speech sounds.

In 1964 the book was translated and published in Russian. Nowadays no research in acoustics can be implemented without the reference to this investigation. It might be of an interest to know that the work has been ordered by Swedish Air Force and performed in RTI (Royal Technological Institute, Sweden) with the support of the Rockefeller's Fund and the MIT (Massachusetts Institute of Technology, USA). Much of the material studied in this work is from the Russian language. G. Fant pointed out that every formant is formed by all supra-glottal cavities though to different extent (Фант, 1964 : 33) and that every cavity influences all resonant frequencies to this or that extent (Фант, 1964 : 202). It is partially to the investigation made by G. Fant that nowadays the two theories are considered to be not so uncompromising as neither of them contradict to mathematic or to phonetic data.

The auditory aspect of speech sounds is the least studied. Within the auditory aspect we will describe the human auditory apparatus. A human's **ear** has three parts: the outer ear, the middle ear and the inner ear. The **outer ear** consists of the auricle and the aural passage, the latter, being a hollow vessel, serves as a resonant chamber to amplify sound waves. The aural passage is about 2.5 cm long and it ends in the ear-drum that divides the outer ear from the middle ear. The **middle ear** is a cavity in the temple bone of about 1 cu cm, filled with air, and connected to pharynx through Eustachian tube. The ear-drum is a thin pellicle (10 by 9 mm size, 0.1 mm thick) consisting of three layers. In the middle ear and behind the ear-drum there are located three aural bones – the malleus / mallet, the incus/anvil and the stirrup-bone. The vibrations of the ear-drum are transferred by the three bones to the **inner ear**, covered by the stirrup-bone and filled with lymphatic liquid. The inner ear has the cochlea, responsible for balance and hearing systems of a human being. Along the whole length of the cochlea (in its hearing channel) there are two membranes with a variety of different cells, including, the so-called, hair cells (their number is 25 thousand) that receive the vibrations of the lymphatic liquid and transfer them to the neurons of the hearing nerve (which numbers 30 thousand neurons).

It is not quite clear yet how the hearing apparatus works. It is definite that the hearing range is to some extent individual and has certain limits. A human being hears sounds only within a certain range of frequencies and amplitudes. Also it needs to be mentioned that acoustic characteristics (mostly frequency and amplitude) do not have linear correlation with the auditory sensations though within certain boundaries sounds with bigger amplitude are perceived as more loud and sounds of higher frequencies – as sounds of higher pitch.

Revision: Section 1 (Повторение: Раздел 1)

1. What is phonetics as a branch of linguistics?
2. What are the three branches of phonetics defined by their aspect of studying speech sounds?
3. What is the difference between general phonetics and special phonetics?
4. What is studied by descriptive, historical, and comparative phonetics?
5. What is the distinction between theoretical, applied, and experimental phonetics?
6. How is phonetics connected with grammar, lexicology and stylistics?
7. Why is it difficult to give an exact definition of phonetics?
8. What are the main methods of phonetic research?
9. What sciences outside linguistics is phonetics connected with?
10. What stages are distinguished in the speaking–hearing process?

2. The Phonetic Basis of English: Articulatory & Perceptual Aspects (Фонетическая база английского языка: артикуляционный и перцептивный аспекты)

2.1. The Concept of the Articulatory Basis in a Language (Понятие артикуляционной базы языка)

In spite of the fact that there are many languages in the world, many nationalities have identical articulatory organs and as a result in many languages we have identical sounds. Compare English and Russian [p], [b], [t], [d], [s], [z], [k], [g]. But! Being typologically identical, these sounds are different from the articulatory point of view. These articulatory differences account for the fact that each language has its own tendencies which are characteristic of this language only. These articulatory differences and peculiarities are part and parcel of all native speakers and represent the essence of the *articulation basis*. What is it? How can we define it?

The articulation basis is usually defined in general phonetics as the unity of habitual movements and positions of the articulatory organs for the given language, cf. with the definition by O. S. Akhmanova: «...совокупность произносительных тенденций, характеризующая речедвижения говорящих на данном языке», lit. ‘a total of articulatory tendencies and speech-motor of the speakers of a given language’ (Ахманова, 2009 : 55). The articulation basis comprises the study of differences in: the tension, the tongue shape, the pressure of the articulators, the lip and cheek and jaw posture and movement, which run through the whole articulatory process (O’Connor, 1978 : 289). It may be considered as a language tradition and should be understood as communication in the given language from one generation to another.

!

Articulation: Speech Organs

Using MRI scanning, the complex internal mechanics of speech can be visualised (Lindsey 2011). Speech is essentially modified exhalation, with quick breaths taken between phrases. The larynx (voice box) houses the vocal folds, which vibrate to create sound.

The mouth acts as a gateway, opening more for vowels and closing or obstructing for consonants – for example, with the lips for a /b/ or the lower lip against the teeth for a /v/. Meanwhile, the soft palate at the roof of the mouth lifts to block the nasal cavity during most speech, only lowering briefly to produce nasal sounds.

The tongue is the most active component, with both its visible tip and its crucial rear section moving intricately. The back of the tongue can arch or flatten to shape sounds like /j/, /k/, /g/, and various vowels.

Producing several words per second requires highly coordinated, learned muscular control. This explains why altering one’s accent or mastering a foreign language’s pronunciation demands significant commitment and practice.

2.2. The Articulation Basis of English (Артикуляционная база английского языка)

!

The Articulation Basis of the English Language explained

Pronunciation is a crucial gateway skill for English learners, affecting performance in exams and job interviews. However, traditional teaching often relies solely on the ineffective “listen and repeat” method. A more successful alternative is the *articulatory approach*, which treats pronunciation as a physical motor skill, akin to a sport, cf. (Messum, 2010, 2012, 2017, 2019).

The articulatory setting of English is, foremost, characteristic by its unique default tongue position. The tongue is spread at the back, lightly touching the rear molars. The tip of the tongue is retracted compared to many other languages (e.g., French, Japanese, or Russian). The tip of the tongue constantly taps or raps against the alveolar ridge to produce sounds like /t/, /d/, /n/, and /l/. This setting differs fundamentally from other languages and must be adopted to speak English naturally.

In short, the method of articulatory approach for teaching English involves the following factors: 1) *developing sensitivity*: students first explore their own mouths with their tongues to understand their articulators (like the alveolar ridge) and build physical awareness; 2) *mastering stress*: learners are taught that the defining feature of English is its stress system. They practice creating stressed syllables using abdominal muscles to push air out forcefully, making syllables louder, longer, and more prominent; 3) *understanding aspiration*: the approach teaches how English uses relaxed lips and a strong abdominal push of air to create aspirated sounds (like the ‘p’ in ‘pin’); 4) *adopting the English articulatory setting*: each language has a default mouth position. For English, this involves the tongue being retracted and spread at the back, with the *tip* (not the blade) constantly tapping the alveolar ridge to make sounds like /t/, /d/, /n/, and /l/. Learners must visualise the tongue position and compare it to other languages.

Ultimately, learning pronunciation this way is seen as adding a new motor skill, which does not interfere with existing language habits, very much like training the body for a new sport.

2.3. Comparison of Articulatory Bases / Articulatory Differences between Russian & English (Сопоставление артикуляционных баз / артикуляционные различия русского и английского языков)

There are peculiarities in each language if we compare English and Russian, e.g.: the mid part of the tongue is used very often in Russian and as a result we have palatalization; the tongue-tip is tenser in English and more used than in Russian, where the blade is dominant (as a result we have dorsal articulation of forelingual consonants in Russian); compare lip-rounding in English and Russian – *lot* and *лодка*; sounds present in one language are absent or differ from that in the other language – [ŋ] is present in English and there is no such sound in Russian, English glottal [h] differs severely from Russian backlingual [x].

The Articulation Basis of each language depends on the phonemic system in the language or to be more precise rests on the phonological principles, i.e. what distinguishes one sound from another in the given language and the same sound in a foreign language. What phonological principles are dominating in English and in Russian? Let us try to answer the question first on the example of vowels and then on consonants.

In the **subsystem of vowels**:

1) only two parameters coincide in English and Russian concerning classification of vowels – the tongue position (classification according to the vertical and horizontal positions of the tongue), and the lip position. The rest (length of vowels, degree of tenseness (tense/lax), and stability of articulation (monophthongs / diphthongs / diphthongoids)) are characteristic only of English;

2) in the production of Russian vowels [o], [y] the lips are considerably rounded and protruded, but English lips for [ɒ], [u], [ɔ:], [u:] are more tense. Differential principle in Russian – rounded / unrounded; in English slightly rounded / unrounded;

3) in English the bulk of the tongue is more often in the back part of the mouth cavity. In Russian – vice versa;

4) horizontally in English – 5 classes of vowels; in Russian – 2 classes;

5) there are many checked vowels in English (no diminution to the end of the vowel) – all short vowels, diphthongs, and long vowels are checked before voiceless consonants, e.g.: *bet, not, speak, type, coat*.

6) free (unchecked vowels) occur when long vowels and diphthongs are in open syllables or when followed by voiced consonants, e.g.: *sea, window, card, bee*, and etc.

7) as there are long and short vowels in English, they ARE divided in tense/lax vowels. No such division can be found in Russian (ЛеоНТЬева, 1988 : 43; Cook, 2017 : 7).

In the **subsystem of consonants**:

1) in English the neutral position of the tongue is flat and lowered. Easy to pronounce [ŋ], [h];

2) consonants in English have six zones of articulation, whereas in Russian only four (see Table 2 below);

3) correlation based on the principal voiced / voiceless does exist in both languages, but in word-final position is neutralized in Russian;

4) apical articulation dominates in English, dorsal in Russian;

5) more flat and low position of the tongue limits the system of soft consonants in English and provides the basis for “dark articulation”. Softened are only [ʃ], [ʒ], [tʃ], [dʒ] which are pronounced with the secondary focus, i.e. mid part directed towards the hard palate or alveolar teeth ridge). In Russian [л'] mid part comes higher to the hard palate; in palatalized Russian [к'], [л'], [х'] mid part tends to the hard palate (not soft), while in English [k], [g], and [h] the back part tends to the soft palate (ЛеоНТЬева, 1988 : 30-32);

6) no aspiration in Russian;

7) stress in English shows the tendencies to be recessive, and rhythmic (see next essay).

The articulation basis is important in foreign language teaching: better results are achieved when the learner gets the basis of articulation right rather than trying for the foreign sound sequences from the basis of his own language.

Table 2

Articulation Matrix of Consonants in English and Russian

	English	Russian
Labial	+	+
Interdental	+	–
Alveolar (Dental)	+	+
Midlingual	+	+
Backlingual	+	+
Pharyngeal (Glottal)	+	–

2.4. Literary Pronunciation

(Литературное произношение)

Literary pronunciation is a standardized form of pronunciation accepted as the norm for public speaking, broadcasting, and educated speech. For the discussion on the literary (norm) pronunciation in the regional variants of English see Section 10.1.

Literary pronunciation represents the **standardised, normative form of a language**, which is explicitly contrasted with dialects, accents, and regional variants (Беляева, 2019 : 8-12). Its key characteristics and functions can be summarised as follows:

1) **the Normative Standard**: literary pronunciation is presented as the central, accepted norm; It is the default against which other forms are measured;

2) **Opposition to “Evolutionary” Variation**: dialects can be viewed as natural “evolutionary illnesses” that lead to language diversification; literary pronunciation, by contrast, is the **stable, codified form** that resists this fragmentation; it serves as a unifying force across a large territory where such “illnesses” or variations naturally arise;

3) **a Marker of Education & Urbanity**: literary pronunciation is implicitly linked to education and non-regional identity; one can differentiate between a foreign “accent” and a regional dialect / subdialect («говор»); to speak with the literary standard is to sound educated and “urban” («городской»), as opposed to someone whose speech clearly identifies a regional background;

4) **Functional Clarity and Authority**: while dialects have their own rich history and internal logic, the literary standard provides a **common framework for unambiguous communication**; it is the form taught, used in official contexts, and considered “correct” by authoritative sources (like a language coach or teacher who can explain why a dialectal innovation is non-standard).

In essence, the **literary pronunciation is not merely one variant among many, but is the consciously cultivated, prescriptive standard that maintains linguistic unity, clarity, and social authority amidst the natural diversity and constant evolution of living speech.**

!

Pronunciation: Outdated Dictionaries

English spelling is often challenging for native speakers to recall accurately. For second-language learners, however, the difficulty is reversed: they typically encounter words in written form and must work out how to pronounce them – a task made harder by the irregular nature of English spelling. Traditionally, learners have turned to dictionaries that include pronunciation guides, though such references can quickly become outdated.

Pronunciation evolves in two key ways: through systematic changes in a language's vowel and consonant systems, and through shifts in how specific words are spoken, cf. (Lindsey, 2019 : 47-51). For instance, the vowel system in standard southern British English has changed considerably since the mid-1900s, yet most dictionaries continue to use older phonetic representations.

As a result, many dictionary-recommended pronunciations no longer reflect contemporary usage. What appears as the primary pronunciation may now be uncommon, especially among younger speakers, while newer, widely used variants might be listed as alternatives – or omitted entirely. In some cases, dictionaries even caution against pronunciations that have become standard.

The online dictionary CUBE was developed from an existing database that underwent significant updates. Despite these revisions, outdated pronunciations are occasionally still found, often identified by users who can provide feedback directly through the platform. Compare contemporary pronunciation with that published in most dictionaries:

Word	ModBrE	Dictionary
ate (Lindsey, 2019 : 112)	/et/	/et/
interest (Lindsey, 2019 : 128)	/ˈɪntrest/	/ˈɪntrəst/ ¹
homage	/həʊˈmɑːʒ/, /oʊˈmɑːʒ/	BrE /ˈɑːmɪdʒ/, AmE /ˈhɑːmɪdʒ/
garage (Lindsey, 2019 : 123)	/ˈɡærɪdʒ/	AmE /ɡəˈrɑːʒ/ ²
koala	ModAusE /kəˈwɑːlə/	/koʊˈɑːlə/
harass (Lindsey 2019 : 125)	/həˈræs/	/ˈhærəs/
processes	ModAmE /ˈprəsəˌsɪz/	BrE /ˈprəʊˌsɛsɪz/, AmE /ˈprɑːˌsɛsɪs/
aeon (Lindsey 2019 : 111)	/ˈeɪən/	/ˈiən/
cure	/kjʊː/	/kjʊə/
mature	/məˈtʃʊː/	/məˈtʃʊə/ ³
air	/ɛː/	/eə/

¹ Cf. (Lindsey, 2019) – many words that have changed are those where traditional RP used a weak vowel schwa or /ɪ/, while younger speakers today use a full vowel; likely due to spelling influence or similar.

² Americans tend to approximate the source language more closely; relatively recently, American English has developed ways of marking loanwords as foreign.

³ In D. Jones publication (Jones, 1917), [ɔː] is given as an alternative: /məˈtʃɔː/.

divisive	/də'vaɪsɪv/, /də'vɪsɪv/	ModAmE /də'vaɪsɪv/, /də'vaɪzɪv/, /də'vɪsɪv/
debris (Lindsey, 2019 : 116)	AusE (also BrE) /də'bri/	AusE /'dɛbri/, /'dɛbri/
version (Lindsey, 2019 : 143)	/'vɜːʒən/	/'vɜːʃən/
Asia (Lindsey, 2019 : 112)	/'eɪʒə/	/'eɪʃə/
to contrast	ModAmE /'kəntræst/	AmE /kən'trast/
to protest	ModAmE /'proutest/	AmE /prə'test/
mischief	AmE /'mɪstʃɪf/, BrE / 'mɪs,ʃiːf/	/'mɪstʃɪf/
mischievous	/mɪs'tʃɪvɪəs/, AmE /'mɪstʃɪvəs/	/'mɪstʃɪvəs/, AmE /'mɪstʃɪvɪəs/
drawing (Lindsey, 2019 : 119)	/'drɔːrɪŋ/; /wɪð'drɔːrɪŋ/, /wɪð'drɔːrəl/	/'drɔːrɪŋ /
gotten (Lindsey, 2019 : 124)	“It's gotten /'gɒtən/ better”	traditional “It's got better”
trip	/ʈrɪp/	/trɪp/
drip	/ɖrɪp/	/drɪp/

!

Pronounce the Word *Pronunciation*

The word “pronunciation” presents three common difficulties, cf. (Lindsey, 2011). First, its spelling differs from the verb “pronounce” – it is written with a “u”, not “ou”. Second, the stress falls on the second syllable, making the first syllable weak (with a schwa sound: /prə/), unlike the overemphasized “PRO-” often heard from learners. Third, it has five distinct syllables, with the “-cia-” part pronounced as two separate syllables: “ci-a” (like “-si-ya-”), not merged into one. Correctly said, it sounds like /prəˌnʌn.sɪ'etʃən/.

2.5. Phonetic Interference in Bilingualism

(Фонетическая интерференция при билингвизме)

Language interference in bilingual people refers to the influence of a speaker’s native or dominant language (their “base” language) on their use and acquisition of a second language (Ким-Малони, 2014 : 129). This occurs when the linguistic rules, structures, or patterns of the first language are incorrectly applied to the second language, leading to errors and difficulties.

This interference creates barriers, particularly in academic and formal contexts, even when conversational ability is adequate. Effective language teaching strategies to overcome this involve direct comparison between the two languages, highlighting both similarities and differences (e.g., in grammatical categories like gender, case, verb aspects, and agreement), and providing targeted exercises to reinforce correct usage. Teaching materials which ignore the learner’s specific base language and the resultant interference are less effective than those designed to address it directly.

In the broader study of bilingualism, **phonetic interference** (often called phonological or accent interference) is a very common type (Соколова, 2010 : 160). It refers specifically to:

– the influence of the native language’s sound system on the pronunciation of a second language;

– this includes the substitution of sounds, the application of native stress and intonation patterns, and difficulties with sounds that do not exist in the speaker’s first language.

Thus, phonetic interference is a key component of language interference overall. It is the subconscious application of the native language’s *sound system* that leads to a foreign accent and can sometimes cause misunderstandings in the second language. The text’s core argument—that effective teaching must address interference from the specific base language—applies equally to phonetic issues as it does to the grammatical ones it highlights.

!

Language interference in pronunciation

Language interference can emerge in diverse contexts, including popular culture, cf. (Lindsey, 2011). This phenomenon is illustrated by analysing non-native English singers, such as K-Pop artists, whose performances often reveal the influence of their first language.

The Korean language contains several unique phonological traits not present in English, like an aspirated /s/. On the other hand, English has sounds absent in Korean, such as /v/, which Korean speakers typically render as “bwi”. Another articulatory distinction concerns the /w/ sound: in English, it is produced with lip-rounding, unlike in Korean.

Furthermore, Korean does not phonemically contrast certain English sounds. Similar to Japanese, it does not distinguish between /r/ and /l/. Another difficult contrast involves the vowel pairs in words like “live” / “leave” and “chip” / “cheep”. These are commonly labelled “short i” and “long i”, yet length is an insufficient marker for differentiation, particularly in song. The instances listed above are frequently observable in performances by Korean artists singing in English and can be attributed to native language interference.

Revision: Section 2 (Повторение: Раздел 2)

1. What is meant by the articulation basis of a language?
2. What articulatory parameters are included in the articulation basis?
3. How is the articulation basis connected with language tradition?
4. Why can typologically identical sounds differ in different languages?
5. What is the characteristic articulatory setting of the English language?
6. What is meant by the articulatory approach to teaching English pronunciation?
7. What are the main articulatory differences between English and Russian vowels?
8. What are the principal articulatory differences between English and Russian consonants?
9. What is meant by literary pronunciation?
10. What is phonetic interference in bilingualism?

3. The Phonetic Structure of English (Фонетический строй английского языка)

3.1. A Brief History of the Phonetic System of English (Краткая история развития фонетического строя английского языка)

!

The English Great Vowel Shift

The Great Vowel Shift was a major series of sound changes beginning ca. 1400 that fundamentally altered the pronunciation of long vowels in English. Its effects are still evident today, explaining why English vowel letters are pronounced so differently from other European languages (e.g., 'A' as /eɪ/ instead of /a:/). The shift involved phenomena: **raising** (vowels moving higher in the mouth) and **diphthongization** (pure vowels becoming gliding sounds). It operated as a **chain shift** (the movement of one vowel prompted changes in others).

The **Historical Progression** of the Great Vowel Shift can be viewed within the following stages: **Middle English (c. 1400)**: long vowels were pure, e.g., *pri:s* (price), *fle:s* (fleece), *mu:θ* (mouth). **Early Stages (1400-1500)**: vowels like PRICE and MOUTH began diphthongizing, starting the chain reaction. **Modern English**: the process continued for centuries, culminating in what can be termed the “**Great Vowel Shift Plus**”, which eventually turned most of the original long vowels into the **closing diphthongs** characteristic of modern accents.

There are some modern outcomes and accent variation which must be enlisted here, cf. (Lindsey, 2019 : 17-21). The result is a modern **natural class** of closing diphthongs (e.g., FLEECE /ɪj/, FACE /eɪ/, GOOSE /uɔw/, GOAT /əw/). Dialects vary in how far this shift progressed. **Southern British accents** (e.g., Received Pronunciation, Standard Southern British) fully diphthongized these vowels. **Northern England and Scotland** often retained monophthongs for FACE and GOAT. **Cockney and related accents** (influencing Australian English) underwent a further “Diphthong Shift”, creating wide glides (e.g., PRICE /æj/).

With the facts listed above, it has to be mentioned that the “blunder” theory of Modern English retaining phonemically paired short and long vowels (e.g., KIT vs. FLEECE) is not justified. The Great Vowel Shift destroyed such pairs, replacing them with qualitatively different vowels and diphthongs. This has led to persistent confusion in how English vowel systems are described and taught. The Great Vowel Shift was not a single, confined event but an ongoing phonological revolution that shaped the English sound system. Understanding its continuous nature is key to analysing both historical changes and contemporary accent patterns, moving beyond outdated descriptive frameworks.

3.2. The Theory of the Phoneme (Теория фонемы)

In the business of learning English you have noticed that the orthographic form of a word may not always correspond with the sequence of sounds which we hear when the word is spoken aloud, e.g.: the initial consonant in *no*, *gnome*, *know* is [n] in each case despite orthographic variety. Compare Russian *снег* [к], *легкий* [х], *его* [h]. On the other hand, *boot* & *foot* are pronounced with different vowels despite orthographic similarity, cf. [bu:t] & [fut]. Thus, we may say, that speakers of all languages produce a far wider range of different phonetic sounds than phonemic descriptions would suggest. These sounds are united in a small number of sound types, which are capable of distinguishing the meaning and the form of words. Such sound types are called phonemes.

There are many conceptions of the phoneme theory current among Russian and foreign scientists and none of the definitions of the phoneme are satisfactory from all points of view. We have already made a reference to the view by V. A. Vassilyev, which we are going to follow within the scope of this book. However, other points of view must be also mentioned, for further detail see Section 3.5 and 4.3.

The difficulty of giving a definition of the phoneme lies in the fact that the phoneme has several aspects and functions that must be combined in one definition. We think it is worthwhile to quote again the best phoneme definition, given by V. A. Vassilyev and favoured by many other phoneticians: “the segmental **phoneme** is the smallest (i.e. further indivisible into smaller consecutive segments) language unit (sound type) that exists in the speech of all the members of a given language community as such speech sounds which are capable of distinguishing one word from another word of the same language or one grammatical form from another grammatical form of the same word” (Vassilyev, 1970 : 136).

3.3. The Phoneme as a Language Unit, Sound as a Speech Unit: the Concept of the Phoneme, Differentiation of the “phoneme” & “sound” concepts (Фонема как единица языка, звук как единица речи: понятие фонемы, дифференциация понятий «фонема» и «звук»)

As it was mentioned before, the inherent complexity in defining “phonetics” stems from the field’s foundational focus on two distinct yet interrelated elements: phonemes and speech sounds. These elements are viewed through different theoretical lenses. One perspective, associated with Soviet-Russian scholarship, considers them two aspects of a single reality, where the phoneme represents the abstract, ideal side and the speech sound its material manifestation. An alternative Western perspective treats them as separate objects, yet still characterises the phoneme as abstract and ideal, in contrast to the concrete, material nature of the speech sound. Consequently, regardless of the theoretical stance, the phoneme is understood as an abstract, functional unit of the language system, and the speech sound (or allophone) as its specific, physical realisation in actual speech.

3.4. Phonemes, their definition: the concept of phonemes & allophones, aspects & functions of phonemes (Фонемы, их дефиниция: понятие фонемы и аллофонов, аспекты и функции фонемы)

The various speech sounds that we actually pronounce while speaking or reading are called phonemic variants (or allophones). The phoneme manifests itself and does exist in the speech of all the members of the given language community only in the forms of allophones. Allophones are such speech sounds that have one or more articulatory and acoustic features in common (and thus produce similar auditory effect) and at the same time differ from each other in some degree. Allophones of one and the same phoneme, when pronounced one instead of the other are incapable of differentiating words. Allophones of the same phoneme never occur in the same phonetic surrounding, thus being unable to differentiate words and grammatical forms (they are in complementary distribution).

Allophones are divided into two groups: 1) principal allophones; 2) subsidiary allophones.

Principal allophones are usually associated with the sound pronounced in isolation which: 1) is free from the influence of neighbouring sounds; 2) has the greatest number of articulatory features among all the allophones of the phoneme. Consider the case with English /t/ for example: /t/ is 1) forelingual; 2) occlusive; 3) voiceless (fortis). But [t] in English *tea* is palatalized; in *not this* – dental; in *not quite* – not plosive; in *twice* – labialized; in *stay* – not aspirated. These features (those that are not numbered) are irrelevant – they may be and may be not, though: forelingual, occlusive, voiceless – are distinctive features and are always present.

Subsidiary allophones are divided into: 1) combinatory; and 2) positional.

Combinatory allophones appear as a result of adaptation (assimilation or accommodation) when one phoneme influences another, e.g.: 1) English alveolar [t], [d], [n], [l] become dental before interdental consonants (regressive assimilation); 2) English (plosive) stops lose their plosion before (plosive) stops, like in *black cat*, *start talking* (accommodation).

Positional allophones – are those which are used traditionally, according to the norms of the language, e.g.: in English /l/ before front vowels and before /j/ has clear (soft) allophone, whereas before back and central vowels and in word final position it has dark variant; in Russian sonorous sounds are devoiced word finally – *рубль*, *вихрь*, *мой*, and different vowel allophones are used depending on the word stress, like in *мал* – *малыш* – *малыши* (different vowel allophones).

Why is it necessary to speak about subsidiary allophones? They are of great practical importance to language teaching. Allophonic mistakes give foreign accent. We develop special phonetic exercises to avoid mistakes in principal and subsidiary variants.

There can be discerned three aspects of the phoneme: 1) Material; 2) Abstractional (generalized); 3) Functional. The phoneme is a unity of all these aspects. They can not be separated.

Let us consider the three aspects in more details below:

1) The phoneme has material aspect because it exists in the forms of allophones which are definite speech sounds.

2) The phoneme has abstractional (generalized) aspect as native speakers abstract themselves from non-distinctive features of allophones and at the same time generalize quite subconsciously those features which are functionally important. How does this process go?

The process of unconscious phonemic abstraction and generalization exists as follows: native speakers quite subconsciously abstract themselves from differences between similar speech sounds, i.e. between allophones of the same phoneme. In other words, native speakers do not notice the differences between sounds in different phonetic contexts, because they do not differentiate the words, e.g. a phonetically untrained Englishman does not notice the difference between variants of /t/ in *eight*, *eighth* and *try* because the dental variant of /t/ in *eighth* or post-alveolar variant of /t/ in *try* does not distinguish words in English. But thinking that he pronounces one and the same phoneme /t/, this Englishman in reality pronounces a different variant of it in each word. This is an example of phonetic abstraction.

The process of abstraction goes alongside with the process of generalization. Native speakers abstract themselves from non-distinctive features of allophones, but at the same time they generalize quite subconsciously those features, which are functionally important, i.e. those features, which can not be changed without changes in meaning. E.g. – what common articulatory features can be found in the allophones of /t/. For certain they are: occlusive, forelingual, voiceless. These three features are generalized in the mind of the speaker and they form a bundle of distinctive features. Let us prove it, e.g.: 1) if we replace occlusive [t] in the word *eight* by an affricative we get another word – *eighth*, 2) if we change voiceless (fortis) articulation of [t] by a voiced (lenis) one we get the word *aid* [eid]; 3) if we change forelingual articulation by backlingual – we get *ache* [eik] (Vassilyev, 1970 : 141-143).

3) The functional aspect of the phoneme is reflected in its definition – the phoneme functions because it constitutes words, it functions because we understand these words (though they are not pronounced correctly). It functions, because it makes one word or grammatical form distinct from the other, e.g.: *ask* – *asks* – *asked*; *storm* – *stormy*; *sleep* – *sleepy*.

Let us consider functions of the phoneme by (Vassilyev, 1970 : 41-47). In speech the phoneme performs three functions:

1) Constitutive – because sounds constitute words, phrases and sentences which are meaningful, e.g.: Russian – *a* – *ад* – *ада* – *Адам* – *Адама* – *Адамам* and etc.;

2) Distinctive – because sounds help to distinguish one word from another, e.g.: *мом* – *дом* – *ром* – *ком* – *сом*;

3) Recognitive – because sounds (or allophones) help to recognize words. This function consists in making the words and their grammatical forms easily recognizable as the result of the use of the right allophones in the right places. It is the function of allophones but not that of phonemes as such. But since allophones are concrete instances of phonemes, we may say that the recognitive function of the right allophones in the right places is performed by the phonemes.

Y. S. Maslov says that recognitive function – is a hardly noticeable screw in a working machine as a whole, but it does not mean that the screw has no function in the wholeness (i.e. in the language). If we use the wrong allophones in our speech – the comprehension is hampered and the (hearer's) attention is drawn to the strange sounding of words and not to their meaning («это слабо заметный винтик в работающей машине в целом, но это не значит, что этот

винтик не имеет функции в целом (т.е. в языке). Если мы используем в речи неправильные аллофоны, понимание речи затруднено и концентрируемся и обращаем внимание на странное звучание слов, а не на их смысл»).

The most important of all these functions is distinctive function, because thanks to this function a phoneme is opposed to another phoneme or no sound in the same position thus differentiating words, e.g. *рама – рам – рамы* and etc. All these functions work in a spoken language simultaneously together. We separate them only for phonetic analysis.

3.5. The Composition of Speech Flow: Phonetic Word,

Syntagm, Phrase, Phono-Paragraph, Text

(Состав потока речи: фонетическое слово, синтагма, фраза, фоноабзац, текст)

The oral form of language–speech–manifests as a relatively continuous flow of linguistic units, each possessing an acoustic form and bearing various meanings. However, this continuous movement of the speech stream is perceived by humans as segmented. The primary phonetic units for dividing the sound stream are the phrase, the speech measure (or syntagm), the phonetic word, the syllable, the sound, among others (Беляева, 2019 : 93).

During the speech act, a text is generated–defined as any “verbal production” (oral or written) created by a speaker (Современный..., 2012 : 13). A text is a complex object within which three zones can be distinguished: 1) the semantic (meaning) zone; 2) the grammatical zone; and 3) the phonetic zone. The semantic and phonetic zones of a text are connected to extra-linguistic spheres of speech activity: the semantic zone relates to the conceptual sphere (where the speaker’s communicative intention is formed), while the phonetic zone relates to the motor-auditory sphere (where the acoustic signal itself is produced and its auditory description is constructed).

A **phono-paragraph** is an integral fragment of text, distinguished not only on semantic grounds but also on phonetic criteria: a pause longer than those within the phono-paragraph; a sharp rise in pitch on the initial phrase compared to the final phrase of the preceding phono-paragraph; a slowing of tempo on the final measures of the phono-paragraph compared to a noticeable acceleration on the initial measures of the subsequent one (Васильева, 2017 : 13-14).

A **phrase** is the largest phonetic (speech) unit, characterised by relative semantic and intonational completeness and bounded by two sufficiently lengthy pauses.

A **syntagm** (or speech measure) is a phonetic unit smaller than a phrase. It is pronounced as an unbroken series of sounds, carries an intonation of completeness, and is bounded by two brief pauses.

A **phonetic word** is a unit smaller than a speech measure, defined by bearing a single stress. A single phonetic word may correspond to several orthographic words.

3.6. Interpretation of the Phoneme in Works by Prominent Linguists

(Трактовка фонемы в работах выдающихся лингвистов)

Here we give more, though far from complete list (or rather an outline; cf. also Section 3.1 & 4.3), to represent the vivid mosaic of colours on the phoneme, cf.:

L. R. Zinder: «фонема – это кратчайшая, т.е. неделимая во времени и линейно единица языка», lit. ‘the phoneme is the shortest, i.e. indivisible in time and linearly, unit of a language’;

The Prague Phonological School: the phoneme is an abstraction and has no material character;

R. I. Avanesov: «класс звуков, которые функционально являются идентичными», lit. ‘a class of sounds that are functionally identical’;

F. de Saussure: an element of the phonological system capable of being distinguished from all other elements of this system;

N. S. Trubetzkoy: a unity of phonologically distinctive features;

L. Bloomfield: a bundle of distinctive features;

D. Jones: is a family of uttered sounds in a given language, which are related in character and are used in such a way that no one member ever occurs in a word in the same phonetic context as any other member, e.g.: in English words *keep* & *call* the initial consonants are distinct speech sounds, but belong to the same phoneme /k/. The first [k] occurs before [i:], the second doesn’t occur in this position.

The first scientist who formulated the theory of the phoneme was *J. Baudouin de Courtenay* (1845-1929) who worked in St. Petersburg University. But he gave a mentalistic view on the phoneme defining it as “mental image of a sound” («психический эквивалент звука, нечто, связанное с нашими представлениями»).

L. V. Scherba – a pupil of Baudouin de Courtenay, a great Russian philologist – defined the phoneme – as the shortest element of general acoustic manifestations which are capable to be associated in a language with meaningful representations.

But later in his book on the phonetics of French he gave a materialistic definition of the phoneme: “In actual speech we pronounce a much greater variety of sounds than we think. These sounds are united in a comparatively small number of sound types which we will define as phonemes. The various sounds that we actually pronounce which are individual representing the universal are called phonemic variants” (Щерба, 1963 : 19).

Thus a phoneme can distinguish the meanings of words & their forms: *том – там, тома – томам*. A phoneme has a certain amount of independence. It can: 1) express a meaning by itself, e.g.: commands of silence and different emotions: [ʃ], [au]; 2) elements of semantic perceptions are often associated with elements of sound perception, e.g.: in *пил, был, взял* [л] – is associated with past tense; in *корова, вода, лодка* [а] – with the Subject in a sentence; in *корову, воду, лодку* [у] – with the Object in a sentence.

3.7. The Phoneme Structure of English **(Фонемный состав английского языка)**

The phoneme theory and the three aspects of speech sounds considered in the previous essays make it possible to pass onto the classification of the English sounds as representatives of the English phonemes, i.e. those speech sounds that are capable of

distinguishing one word from another, or one grammatical form from another one (see definition by Vassilyev, 1970 : 136).

The most general and the old distinction is between vowels and consonants, which, however, does not make the explanation of the difference between them an easy one. The thing is that none of the criteria seems to be good enough to address the difference with due expressiveness and unambiguity, be these criteria of articulatory, acoustic, auditory or functional (phonological) character.

Empirically (and traditionally) it is obvious that sounds within each of such groups as: 1) [p], [t], [k], [b], [d], [g]; [tʃ], [dʒ]; [f], [v], [ʃ], [z], [ʒ], [θ], [s], [z], [h]; 2) [m], [n], [ŋ]; [l]; [w], [r], [j]; and 3) [i], [i:], [e], [ɜ:], [æ], [a:], [ʌ], [ə], [ɒ], [ɔ:], [u], [u:] have more in common between themselves than sounds from without each of the group. Yet, the phenomenon still seeks for the exact scientific explanation and theoretical substantiation as the formula “have more in common”, used above, is of little scientific accuracy and value. Much of the controversy deals with the proper classification of sounds in group 2, as such sounds as [r], [j], and [w], being articulatory, acoustically and, hence, auditory very different from the rest of the sounds of the second group and from the sounds of the first group, are sometimes referred to the group of semi-vowels or vowel glides. So the question arises: how many groups or classes of sounds can be discerned in English – two, three or four (not to mention the English diphthongs, which may comprise quite a separate class or group of sounds)?

In these essays, differentiating the English speech sounds into main groups we will try to consider all the aspects of speech sounds. As the speech sounds, considered below, cannot be anything else but the representatives of the English phonemes we will apply, where it is necessary, the functional (phonological) criterion to refer this or that sound form to this or that group / class of sounds.

Articulatory all English speech sounds from groups 1 and 2 are characterized by having the “so-called close articulation” (Соколова, 2010 : 60). This means that the sounds are formed by means of some kind of obstruction or closure created by an organ or organs of articulation. The obstruction or closure can be complete, partial or intermittent, which leads to further differentiation between occlusive (complete obstruction) and constrictive (partial or intermittent obstruction) consonants.

(Note: In case of [h] the closure is made by the voice-box which, being a frame of cartilages is very flexible and agile. In case of such sounds as [r], [j], and [w] it is very disputable if any obstruction, even the partial one, is made, which significantly distinguishes these sounds from others in group 2 and from those in group 1. The articulation of the English speech sounds belonging to group 3 (including the diphthongs) is characterized by the complete absence of obstruction, which is essential for complete distinguishing of this group from group 1 and group 2. The sounds of this group are unanimously called vowels. If we recognize that there is no obstruction in articulation of [r], [j], and [w] in English (and [j] in Russian) then not only these sounds are to be regarded separately from all the above groups, but the very absence or presence of obstruction cannot be regarded as a constituent feature to differentiate between vowels and consonants. Hence the distinction between vowels and consonants (especially according to this principle) may seem to be very vague and the existence of these two groups of the English sounds as such may seem to be dubious).

Another general articulatory feature that may be used to differentiate between vowels and consonants was suggested by Baudouin de Courtenay and later supported by L. R. Zinder. This feature is the presence or absence of a certain focus of a speech sound. The focus of a speech sound is the tension and strength (напряжение) of a certain part of speech apparatus at the moment of production of the speech sound. As the formation of consonants is characterized by the tension in a certain part of speech apparatus (localized articulation) and the formation of vowels is characterized by the tension of the whole of speech apparatus (nonlocalized articulation), L. R. Zinder suggested to call consonants those speech sounds that have a certain focus of articulation, and vowels – those that do not have one (Зиндер, 2007 : 114).

The drawbacks of the whole idea, both theoretical and practical, are as follows:

1) the focus of a speech sound is a too vague phenomenon to be objectively measured;
2) it can hardly be that speech sounds either have a certain focus or completely lack it, it is more correct to speak of a number of more certain, less certain, more less certain, less more certain and so on foci;

3) vowels have a certain focus also, though it is not localized as compactly as in case of consonants, it is the focus of the vowels that makes it possible to articulatory distinguish between back, central and front vowels;

4) diphthongs in many languages are traditionally classified as vowels having two foci of formation and it will be very difficult to find any other objective feature to differentiate between monophthongs and diphthongs unless regarding them as vowels having one focus or two foci correspondingly;

5) finally for English [r], [j], and [w] case, from which this whole discussion started, to say that these sounds have a certain focus of articulation which differs them from such English sounds as [i:], [u:] and, probably, [ɜ:] (all latter ones are extremely tense in English) is hardly better than to say that it is the presence or absence of obstruction that differs them.

With certain approximation it may be said that the English consonants (groups 1 and 2) and vowels (group 3) can be effectively differentiated articulatory on the basis of presence or absence of active organ or organs in their formation. If this is true, articulatory the English consonants shall be regarded as such speech sounds which are formed by prepositioning and further movement of some active organ, or its part, or organs of human speech apparatus, whereas the English vowels – as such speech sounds which are formed by prepositioning and further retention of the whole of speech apparatus in certain configuration.

Traditionally there are two major classes of sounds distinguished by phoneticians in any language. They are termed consonants and vowels. Consonants are known to have voice and noise combined. In case of consonants various obstructions are made. These obstructions, as has already been mentioned, may be 1) complete; 2) partial; 3) intermittent. Vowels consist of voice only and no obstruction is made when we articulate a vowel.

3.8. Principles of Classification of (English) Phonemes (incl. Articulatory Classification): Vocalism, Consonantism (Принципы классификации (английских) фонем (включая артикуляционную классификацию): вокализм, консонантизм)

3.8.1. Consonants (Согласные)

The classification of the English consonants may be said to encompass certain main principles.

As to the **first principle of the classification of English consonants** there are few ways of seeing the situation. V. A. Vassilyev considers that primary importance should be given to the **type of obstruction and manner of noise production**. On this ground he distinguishes two large classes of consonants:

- a) occlusive, in the production of which a complete obstruction is formed;
- b) constrictive, in the production of which an incomplete obstruction is formed.

Each of the classes is subdivided into noise consonants and sonorants. Noise consonants are divided into plosive consonants (or stops) and affricates. The system (based on the above-mentioned articulatory characteristics – type of obstruction and manner of noise production) may be presented in the following scheme (Table 3).

Table 3

Type of Obstruction and Manner of Noise Production to Classify the English Consonants (after V. A. Vassilyev)

Consonants	Occlusive	Noise	Plosives (Stops)	/p, t, k, b, d, g/
			Affricates	/tʃ, dʒ/
		Sonorant	/m, n, ŋ/	
	Constrictive	Noise		/f, v, ʃ, ʒ, θ, ð, s, z, h/
		Sonorant	Medial	/w, r, j/
			Lateral	/l/

Another point of view is shared by a group of Russian phoneticians including M. A. Sokolova, K. P. Gintovt, I. S. Tikhonova, R. M. Tikhonova (Соколова, 2010 : 61). They suggest that the first and basic principal of classification should be the **degree of noise** (Table 4). According to this principle, English consonants are divided into:

- 1) noise consonants;
- 2) sonorants.

Noise consonants are subdivided into: occlusive, constrictive, and occlusive- constrictive. Sonorants are subdivided into: occlusive, and constrictive (medial and lateral).

Table 4

Degree of Noise to Classify the English Consonants (after M. A. Sokolova et al.)

Consonants	Noise	Occlusive		/p, t, k, b, d, g/
		Constrictive		/f, v, ʃ, ʒ, θ, ð, s, z, h/
		Occlusive-constrictive		/tʃ, dʒ/
	Sonorant	Occlusive		/m, n, ŋ/
		Constrictive	Medial	/w, r, j/
			Lateral	/l/

Both above-mentioned classifications are valid, consistent and complete. But for teaching purposes the latter given by M. A. Sokolova et al. is more concise and symmetrical.

The **second principle of classification** is **according to the active organ of speech, or to the place of articulation**. According to this principle, the English consonants are classed into: 1) labial; 2) lingual; 3) glottal. The class of labial consonants is subdivided into: a) bilabial; b) labio-dental. The class of lingual consonants is subdivided into: a) forelingual; b) mediolingual; c) backlingual. There is only one glottal (or pharyngeal consonant) representing a class – [h]. For the second principle of classification see Table 5 below.

Table 5

Active Organ of Speech or Place of Articulation to Classify the English Consonants

Consonants	Labial	Bilabial	/p, b, m, w/
		Labio-dental	/f, v/
	Lingual	Forelingual	/t, d, tʃ, dʒ, ʃ, ʒ, θ, ð, s, z, n, l, r/
		Mediolingual	/j/
		Backlingual	/k, g, ŋ/
	Glottal	/h/	

The **third principle of classification** is **according to the work of vocal cords**. Here we must note that it has long been believed that from the articulatory point of view the distinction between such pairs of consonants as [p]-[b], [t]-[d], [k]-[g], [s]-[z], [f]-[v], [tʃ]-[dʒ], [ʃ]-[ʒ], [θ]-[ð] is based on the absence or presence of vibrations of the vocal cords, or on the absence or presence of voice or tone component. However, many experimental data on physiological and acoustic aspects of these sounds showed that this is not the only difference between them. It is obvious now that there is also energy difference. All voiced consonants are weak (lenis) and all voiceless consonants are strong (fortis). It is worth noticing that now there is a considerable controversy about what phonetic feature is involved in the abovementioned oppositions. In the intervocalic position, for example, *latter* – *ladder*, the voicing difference is important, since it is the distinctive feature of the consonants. In word-initial and word-final positions the pronunciation of traditionally voiced consonants may well be voiceless. In these positions it is

the energy difference that serves as a differentiating feature. In initial position aspiration would be a more important feature for stops, for example, *tick – Dick, cap – gap, pit – bit*. In a word-final position it is the length of the preceding vowel that would constitute the chief difference (e.g. the vowel of *bead* is longer than that of *beet*). J. D. O'Connor says: "The vowel in "site" is considerably shorter, and in "side" considerably longer, and this is regularly so before fortis/lenis consonants in final position" (O'Connor, 1978 : 133).

In RP allophones of /p/, /t/, /k/ are invariably voiceless, those of /b/, /d/, /g/ may be and are often voiced but they do not generally have full voicing (i.e. from closure right through to release) unless they are surrounded by other voiced sounds. In *table, under, anger* [b, d, g] are fully voiced, but in *abscess, bedsore* (пролежень), *bagpipe* (волынка), where they are in contact with a voiceless sound, or when they occur immediately before or after a pause, as in isolated word *dog*, voicing does not go right through. Indeed it may not be present at all (O'Connor, 1978 : 130-131).

In view of what has just been said the oppositions /p/-/b/, /t/-/d/, /k/-/g/, /s/-/z/, /f/-/v/, /tʃ/-/dʒ/, /ʃ/-/ʒ/, /θ/-/ð/ are primarily based on energy difference, that is on fortis-lenis articulation, which are their phonologically relevant features. It is for this reason that such characteristics as voiceless-voiced have given place to "fortis- lenis" terms in English (Соколова, 2010 : 60-65).

The fourth principle of classification is according to the position of the soft palate. According to this principle, consonants can be oral and nasal. There are only three occlusive sonorants [m], [n], [ɱ], which require the lowered position of the soft palate. If we compare the system of Russian consonants with English ones, we must underline one very peculiar feature of Russian consonant. In the Russian language the middle part of the tongue is very active in articulating consonants; so when it is raised high towards the hard palate the air passage is narrowed or constricted and this gives the consonant soft colouring. Palatalization is phonemic in the Russian language and, consequently all Russian consonants are divided into two classes: a) palatalized; b) non-palatalized.

In the English consonantal system there is the problem of affricates, that is their phonological status and their number. English phoneticians say that in English there are 6 affricates – D. Jones: [tʃ], [dʒ], [ts], [dz], [tr], [dr]. A. C. Gimson adds two more affricates [tθ], [dθ], as in *eighth, width* (Gimson, 1964 : 166, cit. by Vassilyev, 1970 : 190). But if we approach this phenomenon (8 affricates) from the morphological and phonological points of view, we must categorize [tʃ], [dʒ] as monophonemic units and [ts], [dz], [tr], [dr], [tθ], [dθ] – (as in *width, eighth, out set, beds* and so on) as biphonemic complexes since there is a phoneme boundary between the two consonants.

Of course, if we approach the phonetic system of the English language from "teaching pronunciation" point of view it is natural that we should know some additional information about the articulation of consonants (if the consonant is apical or dorsal; if it is dental or alveolar; if the narrowing is flat or round, etc.). Note in this respect that, for example, such sounds as [f], [v] are characterized by flat narrowing, whereas such sounds as [s], [z] – by round narrowing. These characteristics are of no importance from phonological point of view but they provide necessary information for comparison between the English consonants and those of Russian and so are considerably important for teaching purposes. (Flat narrowing = when the articulator (the

tip) is held close to the point of articulation [f], [v], Round narrowing = when a groove (желебок) is formed in the mid part of the tongue which is opposite the teeth-ridge. The groove is short + the sides of the tongue are slightly raised).

Summarizing what has been described in this essay we could state that together with the majority of specialists in English phonetics we may consider relevant the following articulatory features: 1) type of obstruction; 2) place of obstruction and the active organ of speech; 3) force of articulation.

!

The English Consonant /h/

Speakers of languages such as French and Italian, which lack an /h/ sound, may omit it in English (e.g., /ed/ for “head”). Other languages, like Spanish and Russian, possess a similar but not identical sound, which speakers may substitute. To produce the English /h/ correctly, one must combine breath with a whisper, as in the word “Hi!”. Hypercorrection can also occur, where /h/ is incorrectly added to words, cf. (Lindsey, 2011). Furthermore, some languages use alternative sounds for /h/, such as a back-of-the-throat fricative (heard in Liverpool English for final /k/) or a smoother velar fricative (as in Scottish “loch”), which Russian speakers might use instead.

!

The English Consonant /r/

English accents in Britain are broadly divided into **northwestern** (rhotic) varieties, where /r/ is pronounced in all positions, and **southeastern** (non-rhotic) varieties, where /r/ is only pronounced before a vowel sound.

In non-rhotic speech, words like *sore*, *Spar* and *layer* end in a vowel, making them homophones with *saw*, *spa* and *Leia*. However, when such a word is followed by a vowel, a **linking /r/** is often inserted to separate them (e.g., “a sore /r/ in the mouth”).

Historically, dropping final /r/ was criticised as a lower-class feature in the 18th century, but it became accepted once adopted into Received Pronunciation in the 19th century. The related **intrusive /r/** – which appears after a vowel even where no <r> is written (e.g., “idea /r/ of”) – has attracted more criticism, despite being the same phonetic process. It has, however, been used routinely by public speakers for over a century, cf. (Lindsey, 2019 : 71-72).

!

Vowel Length and Linking Sounds

In the materials on the English language, there are common inaccuracies in how major dictionaries represent English pronunciation, particularly regarding vowel length and linking sounds in different accents. Some points on the topic must be mentioned here.

Myth of “Linking Y” and “Linking W”: the idea that speakers of British and Australian English “insert” a /j/ or /w/ sound between words like *two/three* and a following vowel (e.g., *two o’clock*) is false. These sounds are not inserted; they are the natural glide at the end of the diphthongal vowels /i:/ and /u:/ (what Henry Sweet called “consonantal diphthongs”, cf. Lindsey, 2019 : 23-26).

Key Rhythmic Difference: in British and Australian English, a phrase like *four o'clock* is rhythmically longer than *two o'clock* or *three o'clock*. This is because *four* contains a genuinely long vowel followed by a **separately inserted** linking /r/ to avoid a vowel clash (hiatus). In General American, all three phrases have similar rhythm because *four* does not contain a distinctively long vowel; the /r/ is an integral part of the vowel sound.

Vowel Length Distinction: a fundamental difference between the accents is that British and Australian English maintain a phonemic distinction between short and long vowels (e.g., the vowel in *gather* vs. *father*), whereas General American does not. Dictionaries often misrepresent this.

Dictionary errors: leading learner's dictionaries incorrectly transcribe General American *four* as a long vowel plus /r/, and an Australian dictionary wrongly claims the accent lacks vowel length distinctions. Acoustic measurements confirm American vowels in words like *story* are significantly shorter than their British/Australian counterparts.

The nature of intrusive /r/: the inserted /r/ in phrases like *law(r) and order* is the same phonological process as linking /r/ in *four(r) o'clock*. It is used to avoid hiatus (two adjacent vowels) and is phonetically motivated, not spelling-based. The spelling of the preceding word (e.g., *law* vs. *four*) is irrelevant.

Problem with phonetic symbols: the common use of the symbols /i:/, u:/, ɔ:/ for British English is misleading. It groups together /i:/ and /u:/ (which are actually diphthongs ending in glides) with genuinely long monophthongs like /ɔ:/, obscuring the true phonological patterns.

Thus, more accurate phonetic descriptions must be made that reflect the actual rhythmic patterns, vowel lengths, and phonological processes of each accent, rather than perpetuating outdated or oversimplified dictionary conventions.

!

The English Consonant /s/

The consonant sound most associated with the letter 'S' (/s/, /z/, /ɪz/) is arguably the most important consonant in English due to its high functional load. It carries crucial grammatical meanings, including indicating plurality (e.g., *cats*), possession (e.g., *cat's*), third-person present tense (e.g., *she goes*), and contracted forms of 'is' or 'has' (e.g., *it's*). These endings are frequent and semantically vital.

A key challenge is that these critical '-s' suffixes occur at the very end of words, where non-native (L2) speakers, operating on a "last-in, first-out" basis, often under-articulate or drop them, unlike native (L1) speakers who consistently retain them.

Acoustically, the English 'S' is notably **louder** (a sibilant) than the relatively quiet 'TH' sounds (/θ/ and /ð/). This loudness difference is crucial for distinction, especially in awkward clusters like *months* or *is the*. Here, native speakers often simplify by weakening or dropping the 'TH' to preserve the clear 'S', whereas L2 speakers frequently do the opposite, prioritising the unfamiliar 'TH' and losing the meaningful 'S'.

The articulation of English 'S' is complex. True sibilants require a high-velocity air jet directed at an obstacle like the teeth. However, many speakers, perhaps increasingly, produce it by directing air between the tongue tip and teeth to strike the **lower lip**, creating the characteristic sharp sound.

The text contrasts English with other languages. For instance, Spanish ‘s’ is acoustically between English /s/ and /ʃ/, while Japanese ‘s’ is between English /s/ and /θ/. This leads learners to underestimate the distinct, stronger quality of English sibilants. Furthermore, the vulnerability of English ‘TH’ is shown by dialectal phenomena like TH-fronting (e.g., pronouncing *think* as *fink*), though L2 speakers often incorrectly substitute it with a sibilant (e.g., *sink*), cf. (Lindsey, 2019 : 75-76).

In essence, the English ‘S’ is paramount due to its grammatical weight, acoustic prominence, and the tendency for its meaningful final position to be misprioritised by learners.

!

The English Consonant /t/

The common notion of widespread “T-dropping” in English is often perceived as elision (i.e. ‘dropping’), but it is actually the replacement of the standard [t] sound with other articulations, primarily **flapping** in American English and **glottaling** in British English.

Flapping (or tapping): predominant in North American English (AmE), this involves replacing /t/ (and /d/) with a voiced, quick flap of the tongue tip [ɾ] between vowels when the following vowel is weak (e.g., *city*, *thirty*). It can also occur across word boundaries (e.g., *not only*).

Glottaling: famous in British accents, this involves replacing /t/ with a **glottal stop** [ʔ]. It occurs in similar phonetic environments as American flapping (e.g., *city*, *not only*) but is also common before consonants (e.g., *football*, *atmosphere*) and before a syllabic /n/ (e.g., *Britain*, *certain*). Its social acceptance in Britain varies, being more stigmatised in some positions than others.

The core argument focuses on a **linguistic “conspiracy”** – where separate, unrelated changes lead to the same result – driving the spread of glottal stops **before vowels** in younger General American speech, a traditionally non-glottaling environment.

Two conspiring strategies are identified:

1) **the unpacking of syllabic N:** younger speakers insert a vowel (e.g., [ɪ]) after /t/ in words like *certain* [ˈsɜːt.n] → [ˈsɜːʔ.ɪn], cf. (Lindsey, 2019 : 65-66). This places an existing glottal /t/ before a newly pronounced vowel;

2) **the increase of hard attack:** the more frequent use of a glottal stop to begin vowel-initial words (e.g., [ʔ]angle) blocks the traditional flapping of a preceding /t/ (e.g., in *right angle*), cf. (Lindsey, 2019 : 91-93). The glottal stop of the hard attack can then be reinterpreted as the realisation of the final /t/ itself.

In summary, while “T-dropping” is a misnomer, the articulations of /t/ are highly variable. A significant recent trend is the intrusion of British-style pre-vocalic glottal stops into AmE via two distinct phonological pathways, constituting a linguistic conspiracy.

There are homographs in English that have a variant of pronunciation that includes /t/ and a variant that doesn’t. The word *cursed* can be pronounced in two ways: as a single syllable or as two syllables (“curs-ed”). The two-syllable version is often preferred for reasons of euphony, as it creates a clearer rhythm of alternating stressed and unstressed syllables, which can help prevent grammatical confusion.

Furthermore, if the one-syllable version is used in “cursed child”, the final ‘t’ sound runs directly into the ‘ch’ of “child”. In natural speech, native speakers frequently omit this ‘t’ in such a context. For these reasons, the two-syllable pronunciation, “curs-ed child”, is generally considered slightly more successful.

English features a distinct 'k' sound known as an **ejective**. In the International Phonetic Alphabet, ejectives are marked with an apostrophe (e.g., [k']). It is important to note that, despite sometimes being mistaken for them, these are not click sounds. While the ejective /k'/ is the most common, English also has less frequent ejective versions of /t/ and /p/, cf. (Lindsey, 2019 : 56-57).

The **production** of ejective sounds relies on a **glottal stop**. This closure of the vocal folds traps a pocket of air in the vocal tract above the larynx. Whereas ordinary /k/, /t/, and /p/ are made with outgoing lung air, ejectives are made by compressing this trapped air using the throat muscles, often by raising the larynx like a piston. When the oral closure (for /k/, /t/, or /p/) is released, the compressed air escapes with a noticeable 'pop'. The ejective /k'/ is typically the easiest to produce, as the air pocket involved is the smallest and requires the least effort to compress.

Ejectives also occur in Global Languages, they are not unique to English; they function as regular consonants in many languages, standing in contrast to their non-ejective counterparts. For example, **Hausa**, a major language of West Africa with tens of millions of speakers, has several ejectives. Its ejective /k'/ is often represented in writing by a hooked letter. **Quechua**, the most widely spoken indigenous language family in the Americas, also uses ejective sounds.

A key reason for the **increasing use** of ejectives in contemporary English is the growing prevalence of **hard attack** – the use of a glottal stop [ʔ] to sharply begin a word starting with a vowel (e.g., "[ʔ]I [ʔ]always [ʔ]arrive [ʔ]early"). Speakers will often employ an ejective consonant immediately *before* such a vowel, as the articulatory setup for the glottal stop seamlessly facilitates the production of the preceding ejective sound.

3.8.2. Vowels (Гласные)

The various qualities of English vowels are determined by the oral resonator – its size, volume and shape. The resonator is modified by the movable speech organs – the tongue and the lips. The quality of a vowel depends on whether the speech organs are tense or lax, whether the force of articulation is stable or it weakens at the end of articulation, cf. also (Harris, 2000). So vowel quality could be thought of as a bundle of definite interconnected and interdependent articulatory characteristics. Usually vowels are classified according to the following: 1) the tongue position; 2) the lip position; 3) the degree of tenseness; 4) the length of the vowel; 5) the stability of articulation. Let us consider each of the principles in more details.

First – according to the tongue position. In vowel production the tongue moves in a horizontal direction (forward ↔ backward) and in vertical direction (high ↑ low). *According to the horizontal movement*, Russian phoneticians distinguish five classes of English vowels (Table 6):

- 1) front [i:], [e], [æ];
- 2) front-retracted [ɪ];
- 3) central [ɜ:] (see [ə:] in Table 6 below), [ʌ], [ə];
- 4) back-advanced [u], [a:];
- 5) back [u:] [ɔ:], [ɒ].

Table 6

Horizontal and Vertical Arrangement of the English Vowels (Y. A. Glukhiy)

Front	front-retracted	central	back-advanced	back
i:				u:
	ɪ		ʊ	
e		ə:		
			ə	
		ʌ		ɔ:
æ				ɑ:

Cf. the vowel chart given in the International Phonetic Alphabet, IPA (Fig. 9):

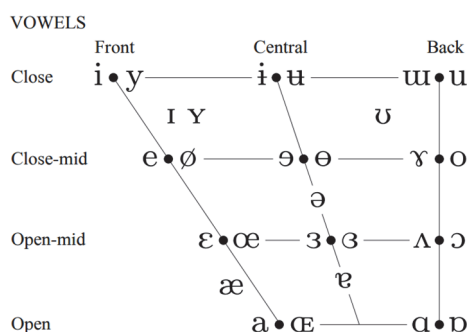


Fig. 9. The vowel chart given in the International Phonetic Alphabet, IPA (open sources)

Cf. the English vowel chart in terms of IPA (Fig. 10).

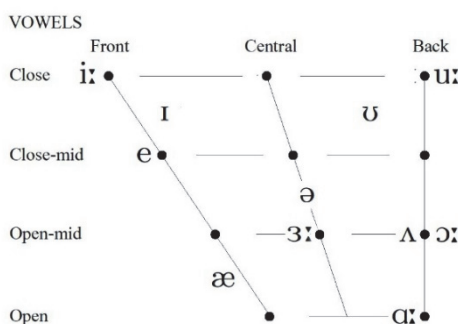


Fig. 10. The English vowel chart in terms of IPA (open sources)

British phoneticians do not single out the classes of front-retracted and back-advanced vowels. So both [i:] and [ɪ] are classed as front, and both [u:] and [ʊ] as back (Соколова, 2010 : 82). They consider them as variants of one phoneme, in spite of the fact that quality of these two vowels is different due to the raised part of the tongue.

The other articulatory characteristic of vowels as to the *tongue position* is its *vertical movement*. According to this principle, British phoneticians distinguish three classes of vowels:

high (or close), mid (or half-open), and low (or open) vowels. Russian phoneticians made the classification more detailed distinguishing two subclasses in each class, i.e. broad and narrow variations of the three vertical positions of the tongue. Thus we may distinguish the six groups of vowels (Table 7).

Table 7

The groups of the English vowels

high (close)	a) narrow	[i:], [u:]
	b) broad	[ɪ], [ʊ], [ɪə], [ʊə] – <i>fierce, tour</i>
mid (half-open)	a) narrow	[e], [e:] (see [ɜ:] in Table 6), [eɪ], [əʊ] – <i>paid, load</i>
	b) broad	[ə], [ʌ]
low (open)	a) narrow	[ɔ:], [ɜə] (also denoted as [eə]), [ɔɪ] – <i>source, loin</i>
	b) broad	[æ], [ɒ], [a:], [aɪ], [aʊ] – <i>tide, gown</i>

Why do we distinguish narrow and broad variations? For example, the vowels [ɜ:] (see [ə:] in Table 6) and [ə] are both central and mid-open, but in the production of [ɜ:] the central part of the tongue is raised a little higher than in the production of [ə]. For this reason the vowel [ɜ:] is defined as central mid narrow (with no possibility to vary or it will become “a close vowel” (i.e. referring to another class)) and [ə] is defined as mid broad.

Second – according to the lip position (lip rounding). In English lip rounding is not relevant phonologically since no two words can be differentiated on its basis. According to the lip position (rounding), English vowels are divided into:

- 1) rounded [ɒ], [ɔ:], [ʊ], [u:] + nuclei of [əʊ], [ɔɪ];
- 2) unrounded [i:], [ɪ], [e], [æ], [a:], [ʌ], [ɜ:] (see [ə:] in Table 6), [ə] + all diphthongs except [əʊ], [ɔɪ].

Third – according to the degree of tenseness. Tenseness, according to L. R. Zinder, is not to be mixed up with intensity (Зиндер, 2007 : 182). According to the degree of tenseness (muscular tension), English vowels are classified as tense and lax. So far there exist different opinions on this question. D. Jones says that only close vowels [i:], [u:] may be determined as tense. G. P. Torsuev defines all English long vowels as “tense” as well as [æ], and all short vowels are considered by him as “lax”. It is believed that greater tenseness of long vowels is closely connected with their length. When we pronounce a long vowel, the organs of speech are kept in a certain position for a longer period of time (as compared with short vowels), i.e. tenseness may be considered as relevant indispensable (concomitant) feature («обязательно сопутствующая») of English long vowels (Соколова, 2010 : 86). Special instrumental analysis showed that historically long vowels are tense while historically short vowels are lax. But this characteristic must be included into classificatory description of English vowels because it might be helpful in teaching English to Russian students since there are no tense vowels in Russian.

Fourth – according to length of the vowel. According to the length, vowels are subdivided into: historically long and historically short. But vowel length depends on the

following factors: 1) its own length; 2) the position of a vowel in a word, e.g.: *bee* – *bead* – *beat*, 3) the number of syllables in a word, e.g.: *'verse* – *uni'versity*, *The man laughed* – *The manager laughed*, 4) the character of syllabic structure (V, CV-, CCV, -VC, CVC, and etc.), compare *dew* [dju:] and *duly* ['dju:li]; 5) the position in a phrase; 6) the type and style of pronunciation and other factors; 7) one and the same vowel is longer when pronounced with the Fall-Rise tone, e.g.: \fit[, \far[.

Length is a non-phonemic feature in English but it may serve to differentiate the meaning of a word. This can be proved by minimal pairs, e.g.: *beat* – *bit*, *port* – *pot*, *deed* – *did*, *pool* – *pull*. It is a moot point in English vowel system. But it is necessary to point out here that the five long vowels in English are different from the seven short vowels. Compare: long – [i:], [ɜ:], [a:], [ɔ:], [u:], and short – [i], [e], [æ], [ʌ], [ɒ], [u], [ə]. They are different not only in length but also in quality. If we compare some similar pairs of long and short vowels, e.g.: [i] and [i:], [u] and [u:], [æ] and [a:] we can see distinct difference in quality (resulting from difference in tongue shape and position and lip position) as well as in length. For this reason, all long vowels have symbols which are different from those of short vowels (it is sometimes not so evident in these essays due to technical reasons); we can see that the long and short vowel symbols would be still different from each other even if we omitted the length mark ([:]): so it is important to remember that the length mark is used not because it is essential, but because it helps learners to remember the length difference (Roach, 1990 : 18).

Fifth – according to the stability of articulation, which specifies the actual position of the articulating organ in the process of articulation of a vowel. There are two possible varieties: a) the tongue position is stable; b) it changes, that is the tongue moves from one position to another.

In the first case when the tongue is stable the articulated vowel (a monophthong) is relatively pure, in the second case a vowel (a diphthong) consists of two clearly perceptible elements (the tongue moves from one position to another). There may be a third variety (a diphthongoid), an intermediate case, when the change in the tongue position is fairly weak. So, according to this principle, the English vowels may be subdivided into: a) monophthongs; b) diphthongs; c) diphthongoids. According to this principle, there may be distinguished twelve (or ten) monophthongs: five (or three, if exclude [i:] and [u:]) long ones – [i:], [ɜ:], [a:], [ɔ:], [u:] and seven short ones – [i], [e], [æ], [ʌ], [ɒ], [u], [ə]; eight diphthongs: [iə], [ɜə], [uə], [ei], [ai], [ɔi], [əu], [au]; whereas [i:], [u:] may be considered to be diphthongoids in English (D. Jones, S. F. Leontyeva). In allophonic transcription they can be represented as [i^{i(j)}], [u^{u(w)}]. A diphthongoid is a vowel that ends in a different element (there is another point of view on the subject, namely, that there are consonantal diphthongs with a vowel and a closing glide present, not change of articulation within a single vowel phoneme, see the discussion below). The tendency for diphthongization is becoming stronger in modern English.

Of course there are five triphthongs: [eɪə] – *player*, [aɪə] – *fire*, [ɔɪə] – *loyal*, [əʊə] – *lower*, [aʊə] – *power*. But in actual speech the middle sounds in these triphthongs (i.e. [i] or [u]) are dropped or can hardly be heard and the resulting sound is difficult to distinguish from some of the diphthongs and long vowels (Roach, 1990 : 22).

Now, let us try to define the phonemic status of the English diphthongs. Diphthongs are complex entities and we have to say whether they are monophonemic or biphonemic units. It is impossible to find a simple and logic criterion which might serve a basis for a decision. Russian scholars consider them to be monophonemic on the basis of articulatory, morphological and syllabic indivisibility, as well as on the basis of duration and commutability. English diphthongs are complex monophonemic entities, because there is no morpheme or syllable boundary that separate the nucleus and the glide, e.g.: *saying* ['sei-iŋ], *crying* ['krai-iŋ], *poorer* ['puə-rə], *slower* ['sləu-ə]. The fact that neither morpheme, nor syllable boundary, that separate the nucleus and the glide, can pass within a diphthong proves the articulatory indivisibility of the diphthongs. The present study of the duration of diphthongs shows that the length of diphthongs is the same as that that characterizes the English long monophthongs in the same phonetic context, cf.: [sait] and [si:t], [kəut] and [kə:t] (Соколова, 2010 : 81). May be because of this fact British phonetician A. C. Gimson considers English diphthongs as long phonemes with different glides. And finally the commutation test proves the monophonemic status of diphthongs because any diphthong could be commutated with practically any vowel, e.g.: [bait] and [bit], [bait] and [bat], [bait] and [bɔ:d], and etc. Besides, monophonemic character of English diphthongs is proved by native speakers, who perceive these sounds (diphthongs) as monophonemic entities (Соколова, 2010 : 79-82). In Russian there are no diphthongs as phonemic entities. Such combinations of sounds as [ой], [ай], [яа], [йо], [ау], [яа] – *пой, чай, яд, йод, каустик, вуаль* are biphonemic clusters, consisting of a vowel and consonant [й] or two vowels. Both elements in Russian clusters are energetic and distinct. But in English diphthongs – the first element, the nucleus, is strong and distinct only, the second, the glide, is weak and indistinct.

In American English (AmE) the three diphthongs ending with a glide to [ə] in British English (BrE) are no longer used. They are [iə], [ɜə], [uə], e.g.: *dear* – in AmE sounds [dir], *dare* – in AmE sounds [deir], *tour* – in AmE sounds [tur], G. Brown says that diphthongs [ɜə] and [uə] are disappearing in BBC English (Received Pronunciation, or RP) in the speech of younger speakers (Brown, 2017 : 36-37). While describing English vowels we must say that the articulation bases of English and Russian vowels are different:

1) the lips: in the production of Russian vowels [o], [y] the lips are considerably rounded and protruded than in the production of English vowels [ɒ], [u], [ɔ:], [u:]. The lips of the Russian people are more lax and their comers are lowered. Englishmen have the so called “flat-type” position of the lips, their lips are more tense and the comers of the lips are raised;

2) the bulk of the tongue: in the production of the English vowels the bulk of the tongue is more often in the back part of the mouth cavity. In the production of Russian vowels the bulk occupies mostly the front part of the mouth cavity (compare consonants alveolar vs. dorsal). Besides, in the articulation of English vowels the bulk occupies more positions horizontally (we have here five classes of vowels, see above);

3) the length of the vowels in Russian is an irrelevant feature (no long vowels which can be opposed phonemically to short vowels in Russian);

4) concerning the stability of articulation we may say that there are monophthongs and diphthongs (diphthongoids) in the Russian vowel system, but these diphthongs are not phonemic

unities and there is a morpheme boundary within the diphthong-like sound complexes, e.g.: чай – ча-йу, мой – мо-йу, and etc. (Леонтьева, 1988 : 43-44).

To sum up, phonological analysis of articulatory features of English vowels allows to consider functionally relevant the following characteristics: a) stability of articulation (monophthongs, diphthongs, diphthongoids); b) tongue position (five classes horizontally and three classes vertically).

!

Contrasting Vowels /ɛ/ and /æ/

The Singaporean accent closely aligns with standard British and American English in a key phonetic respect: it maintains a clear distinction between the TRAP vowel (as in ‘Hadley’) and the DRESS vowel (as in ‘Hedley’). For native speakers of these varieties, this contrast is semantically important, differentiating many pairs of words: to specify, the DRESS vowel is /ɛ/ (as in *Hedley*), and the TRAP vowel is /æ/ (as in *Hadley*), cf. (Lindsey, 2011).

This distinction, fundamental to British and American English, is often not perceived or produced by millions of speakers of other languages, including many skilled linguists and interpreters from regions such as Germany, Italy, and Russia, who may be unaware these represent two separate vowel sounds.

!

Vowels: North England & New Zealand

The accent of Northern England and the broad accent of New Zealand, while sounding very different, share a remarkably similar underlying system of short vowels. This system has remained consistent for around five centuries, built upon five vowels positioned at the edges of the phonetic vowel space, exemplified by the words KIT, DRESS, TRAP, LOT, and PUT, plus the neutral schwa sound (as in *comma*), cf. (Lindsey, 2011).

In Northern England, the vowel in PUT is also used in words like *strut* and *run*, preserving a pre-Shakespearean pattern before these words developed a separate vowel in Southern British English. Over time, outside Northern England, these vowels shifted. They tend to move in a coordinated manner, like a queue maintaining distance to avoid merging and causing confusion. When an entire set shifts just the right amount, the overall structure of the system can remain similar, but each vowel now corresponds to different words.

This is precisely what happened in New Zealand English relative to Northern England. The front vowels (TRAP, DRESS, KIT) have each shifted forward, while LOT remains largely unchanged and STRUT has a complex history. The result is that both accents use essentially the same set of short vowel sounds, but with the exception of LOT and schwa, those sounds are applied to different sets of words. Therefore, the same phonetic transcription represents different pronunciations: one set of words in Northern England and a different set in New Zealand, explaining acoustic difference despite a shared structural foundation.

transcription	North of England	New Zealand
ɔ̃ɹɪmz kɹɑʃt ən bægz fə bʊks	Jim’s crashed, and begs for bucks.	Gems crushed in bags for books.

!

Vowels: with Online Translators

There is a challenge for non-native learners in distinguishing the English KIT (/ɪ/) and FLEECE (/i:/) vowels, a contrast with high functional load absent in many languages. The traditional phonetic symbols, particularly the outdated “short / long I” (/ɪ/ vs. /i:/ or even /ɪ/ vs. /i:/) notation, must be criticized for misleadingly suggesting the primary difference is length, cf. (Lindsey, 2011). In reality, the distinction is predominantly one of vowel quality.

There may be a teaching strategy proposed to bypass abstract symbols and instead locate equivalent sounds within the learner’s native language. For instance, a French speaker can map English FLEECE to the French sequence *-ille* (as in *fille*) and English KIT to the French close-mid vowel *é*. Using tools like online translators to vocalise these forms provides an accessible model, allowing learners to produce a clear distinction by leveraging familiar sounds from their own phonetic inventory. The method emphasises using the ear over the written form to overcome the obstacle of English spelling inconsistencies.

!

Vowels: Australian Lowering of F3

There has been articulatory and acoustic evolution of the GOAT vowel (/əʊ/~/oʊ/) in contemporary Australian English, with distinctive realisations and a recent innovative variant. The diphthong in the GOAT lexical set (e.g., *go*, *know*, *show*) is a well-documented site of phonetic variation across English dialects. This report provides a structured analysis of its development in Australian English, charting its progression from the General Australian (GenAus) standard to a novel, socially salient variant. The analysis addresses transcriptional inaccuracies, describes articulatory shifts, and explores the sociophonetic factors driving change.

Transcriptional convention and the GenAus standard: the common dictionary practice of transcribing the GOAT offglide with /ʊ/ is phonologically inaccurate, as it misrepresents the sound and contradicts the restriction against word-final checked vowels in English. The phonetically appropriate transcription for the glide is /w/ (e.g., *going* as /gəʊwɪŋ/), a convention followed by linguistically informed resources. In GenAus, when not followed by a vowel, the diphthong’s offglide is characteristically fronted, articulated similarly to the French /ɥ/ (as in *huit*), resulting in a realisation approximating [əɥ].

Emergence of a rhotacised variant: a significant innovation observed among younger Australian speakers is the rhotacisation of the GOAT offglide. This variant involves a glide towards a bunched/molar /r/ articulation, transcribed as [ɹ], characterised by a raised tongue body rather than a retroflex tip. Acoustically, this correlates with a pronounced lowering of the third formant (F3), the primary acoustic cue for rhoticity.

There are some **distributional and contextual constraints** regarding the innovation in question, namely, the **critical environment**, where the fronted [ɥ] and rhotacised [ɹ] variants occur in pre-pausal or pre-consonantal positions being **blocked** before a directly following vowel. Then, there is **interaction with hard attack:** before a vowel, only the standard [w] offglide occurs. However, if a glottal stop (hard attack) intervenes, the environment is blocked, allowing the fronted or rhotacised variants to surface (e.g., *go* [ʔ] *away*). The /w/ in sequences like *going* is not the **so-called ‘linking**

/w/, but an inherent component of the diphthong, not an epenthetic “linking” sound inserted between vowels, cf. (Lindsey, 2019 : 63-64).

Some **systemic phonetic comparisons** must be made here. First, /ɪ/ vs. **Intrusive /r/**: this development is categorically distinct from intrusive /r/ (e.g., *law[r]* and *order*). Intrusive /r/ occurs **only** pre-vocally and arises from phonemic merger, whereas the GOAT rhotacisation occurs in **non-pre-vocalic** contexts and has no historical /r/ source. Then, while fronting of the GOAT vowel is attested in other dialects (e.g., Southern England, New Zealand), the systematic, socially indexed rhotacisation of the offglide appears to be a distinctive contemporary Australian feature, cf. (Lindsey, 2019 : 27-28).

The trajectory from a fronted glide [əʏ] to a rhotacised glide [əɪ̯] represents a natural phonetic progression via increased F3 lowering. This exemplifies a community-driven phonetic enhancement, making a distinctive local feature more pronounced. Its association with younger speakers and visibility in digital media suggests it functions as a sociolinguistic marker within Australian English.

Thus, the Australian GOAT vowel demonstrates dynamic phonetic change. The established GenAus fronted variant has provided the articulatory foundation for a novel rhotacised offglide, governed by specific phonological constraints and acoustically defined by F3 lowering. This innovation highlights the ongoing evolution of Australian English and underscores the importance of accurate phonetic transcription and acoustic analysis in understanding dialectal variation. Further empirical study is warranted to quantify its spread and social indexing.

! Vowels: Contrasting /ə/ and /ʌ/

There is a key vowel contrast in English: between the **STRUT** vowel (/ʌ/, as in *cup*) and the weak **commA** vowel (schwa /ə/, as in *the*). This distinction – a hallmark of Received Pronunciation (RP) and accents of Southeast England – is a **global minority**. It is consistently maintained only in: the Southeast of England (the source of RP), Ireland, and Australia (due to its historical roots in Southeast England).

The majority of English accents worldwide do not maintain this contrast (Lindsey, 2011). Instead, they typically follow one of these patterns:

- 1) **STRUT-commA merger**: the STRUT vowel merges with schwa, using the same vowel in both stressed and unstressed positions (e.g., *above* pronounced /ə'bəʊ/); this is found in General American English, Scottish English, Welsh English, the West Country of England, and many Caribbean, African, and Asian varieties (e.g., Nigeria, India, the Philippines);
- 2) **STRUT-FOOT merger**: the STRUT vowel merges with the FOOT vowel (/ʊ/, as in *book*), as in traditional accents of Northern England (e.g., *above* pronounced /ə'bʊʊ/);
- 3) **stressable Schwa**: in many varieties (e.g., Welsh, New Zealand, South African English), the schwa vowel (/ə/) can be used in fully stressed syllables, eliminating the need for a separate, strong STRUT phoneme.

Therefore, the **FOOT-STRUT split** – the historical development that created a separate STRUT vowel – is not a standard feature of global English. The more common phenomenon is **STRUT-centring**, where STRUT words join the schwa set /ə/. Even within Southern England, the STRUT-commA system may be less stable than often assumed.

!

Vowels: (Un)stressed Schwa /ə/

There is a widespread but inaccurate practice in pronunciation dictionaries of transcribing AmE with two separate vowel symbols – /ʌ/ (STRUT) and /ə/ (schwa) – when in fact it possesses only **one** phoneme for both. In Received Pronunciation (RP), a genuine contrast exists between a strong STRUT vowel (/ʌ/) and a weak, unstressed schwa (/ə/), cf. (Lindsey, 2011). However, General American English merges these; the same central vowel sound (schwa) is used for both stressed positions (as in *cup* or *love*) and unstressed ones (as in *comma*).

The prominent British-led dictionaries, such as the Longman Pronunciation Dictionary, perpetuate confusion by applying the RP two-symbol distinction to AmE, often noting only in small print that they are allophones of a single phoneme. In contrast, several authoritative resources – including Merriam-Webster, Lexico, and modern phonetic textbooks – correctly use only the schwa symbol (/ə/) for American transcriptions.

The core fallacy identified is the **dogma that “schwa cannot be stressed”**. This rule is artificial and leads American speakers, who naturally stress the schwa sound daily, to make transcriptional errors. The International Phonetic Alphabet (IPA) itself places no such restriction on the schwa symbol.

Thus, the insistence on two symbols for American English is a prescriptive inaccuracy. The reality, shared by many global English accents (including Northern English, Welsh, and Scottish), is that schwa is a fully stressable vowel, and its transcription should reflect the phonological system of the accent being described, not an imported RP-based rule.

!

Vowel System (Vowel Space) of English

There is a common perception of vowels as discrete, fixed sounds, arguing that phonetically they exist within a **continuous, multi-dimensional space**, analogous to colours within a chromaticity diagram, cf. (Lindsey, 2011). Just as one can blend seamlessly from one colour to another, vowels can glide into one another (e.g., from /i/ to /a/, passing through /ε/). This vowel space is roughly triangular, bounded by three maximally distinct “corner” vowels corresponding to extreme mouth shapes: spread (/i/), open (/a/), and rounded (/u/). The central, “colourless” vowel is schwa (/ə/).

Historically, the **articulatory model** of vowel description dominated, pioneered by figures like Alexander Melville Bell and later systematised by Daniel Jones with his **Cardinal Vowels**. This model attempted to chart vowels on a two-dimensional quadrilateral based on the estimated highest point of the tongue (high/low, front/back). Jones established primary and secondary sets, differing in lip-rounding. However, this traditional **International Phonetic Alphabet (IPA) vowel chart** is fundamentally flawed: it ignores the critical role of **lip position** and misleadingly suggests that vowels described by the same tongue position (like primary /i/ and secondary /y/) are identical, when acoustically they are not, cf. (Lindsey, 2011).

The modern **acoustic-auditory model** gains its importance, which describes vowels not by guessed tongue shapes but by their measurable **formant frequencies**. The first two formants (F1 and F2) are most critical for vowel differentiation. F1 corresponds roughly to vowel height (lowest

for /i, u/, highest for /a/), and F2 corresponds roughly to frontness/backness (highest for /i/, lowest for /u/). This acoustic space accurately maps how vowels are perceived and explains why the primary Cardinal Vowels are more common globally: when the lip positions of the primary set are reversed to create the secondary set, the vowels become more centralised and less distinct in this acoustic space, making them more marked.

The parallels with **colour terminology** are drawn to illustrate linguistic relativity: just as languages divide colour space differently (e.g., Russian two basic terms for “blue”, while some languages have one covering blue / green), so do languages divide the vowel space. Spanish may have three front vowels (/i, e, a/), while Italian distinguishes four (/i, e, ε, a/). One cannot assume symbols like <o> represent the same vowel quality across languages.

Moreover, higher formants like **F3** are crucial for certain distinctions. This is exemplified by the rhotic vowel /ɹ/, which has similar F1 and F2 to schwa but is characterised by a drastically **lowered F3**. This demonstrates the utility of a three-dimensional vowel model (F1, F2, F3) for a complete picture, though the two-formant space suffices for most vowels.

In conclusion, the traditional articulatory description is a “pre-Galilean” approximation. Modern technology – using tools like **Praat** for formant measurement and synthesis – allows for objective, scientific analysis and generation of vowels. The overarching message is that phonetics should move beyond subjective guesswork about tongue placement and embrace the precise, measurable acoustic reality of the vowel space.

!

Vowels: Canadian Raising

The reason the vowels in words like “eye” and “ice” sound different, despite often sharing the same dictionary transcription, can be attributed to two key phonetic phenomena, cf. (Lindsey, 2011). First, for most English speakers, there is a significant difference in **vowel length**. The vowel in “eye” is typically quite prolonged, whereas in “ice” the vowel is noticeably shorter. This shortening is caused by the following voiceless (or fortis) consonant /s/, which truncates the vowel’s duration.

Second, for many speakers – particularly in Canada and parts of the United States – there is also a distinct difference in **vowel quality**, a feature known as Canadian Raising (though it extends beyond Canada). For these speakers: in “eye”, the vowel begins with a very open quality, similar to the vowel in “arm”: [aɪ]; in “ice”, the vowel begins with a raised, more centralised quality, closer to the hesitation sound “uh”: [əɪ]. This change in quality is also triggered by a following voiceless consonant, such as the /s/ in “ice”, the /f/ in “knife”, or the /k/ in “like”. Thus, the distinction arises from a combination of durational shortening and, for many, a systematic shift in the vowel’s starting point.

!

Vowel: Essentially Wrong Phonetic Symbols

The traditional set of phonetic symbols used to transcribe British English vowels must in fact be criticized because the conventional chart is fundamentally flawed in its categorisation and symbolism. A more accurate analysis rooted in the observable phonological behaviour of the vowels gives reason for this revision.

The core argument is to re-categorise the English vowels, i.e. not by abstract length or tongue height, but by their phonological properties, cf. (Lindsey, 2019 : 31-33). The proposed correct grouping is:

1) seven closing diphthongs: /ɪj/ (as in *see*), /ɛj/ (*day*), /aj/ (*price*), /oj/ (*choice*), /əw/ (*goat*), /aw/ (*mouth*), and /ɰw/ (*goose*). These end in glides (/j/ or /w/);

2) six long monophthongs (or “linking R vowels”): These are the true & steady vowels in words like *palm*, *nurse*, *square*, *start*, *north*, and *force*;

3) six short lax (checked) vowels: Such as /ɪ/ (*kit*), /ɛ/ (*dress*), /æ/ (*trap*), /ɒ/ (*lot*), /ʊ/ (*foot*), and /ʌ/ (*strut*). These cannot occur at the end of a word and must be followed by a consonant;

4) schwa (/ə/): In a category of its own; it is weak and short but, unlike the other short vowels, can occur word-finally and can trigger Linking R.

Several phonological phenomena demonstrate **evidence** that the seven vowels listed above form a natural class of **closing diphthongs**, not monophthongs or centring diphthongs.

Linking (or “separating”) R: only the six long monophthongs can trigger a linking /r/ when followed by a vowel (e.g., *star/r/ing*). The seven closing diphthongs never do (e.g., *paying*, not *pay-ring*), because they already end in a glide which separates them from a following vowel.

Phonotactics deals with the question of avoiding **hiatus**: only these seven can occur directly before another vowel within a word or morpheme (e.g., *chaos*, *poem*). This is because they end in a glide, avoiding the vowel-vowel hiatus that English disfavors. Loanwords like *karaoke* are then adapted to fit.

Pre-L breaking: these seven vowels can have a schwa inserted before a final /l/ (e.g., *file* pronounced as [fajəl]). This is analysed as a vowel breaking apart the final glide of the diphthong and the /l/.

Pre-fortis clipping: these diphthongs undergo the most drastic shortening before voiceless (fortis) consonants (e.g., the vowel in *safe* is much shorter than in *save*). This is a crucial teaching point to avoid confusion between pairs like *peace* and *peas*.

Historical and theoretical context suggest support for this interpretation. This “glide analysis” goes back centuries, citing works by John Wallis (1653) and Thomas Batchelor (1809). It was firmly established by Henry Sweet, who called them “consonantal diphthongs”, cf. (Lindsey, 2019 : 23-26). However, Daniel Jones later discarded this analysis in favour of symbols like /i:/ and /u:/, based on a theoretical “chroneme” (length) distinction and influenced by an early-20th-century upper-class fashion for monophthongal pronunciations, now considered outdated.

There is, however, also **critique** of the “high / non-high” explanation. The common alternative explanation that uses tongue height to group the vowels (e.g., labelling /i:/ as “high”) must also be refuted. This fails to explain: why “high” vowels in *fill* and *full* do not undergo pre-L breaking, while those in *file* and *foul* do (the glide analysis explains it: the latter have a glide to break apart); and why linking R does or does not occur. The monophthongisation of the vowel in *near* to [nɪ:] shows that a consistently high tongue position does not block linking R.

The data listed above has **implications for teaching**. The suggested “glide insertion” is a misnomer: Teachers often instruct students to “insert” a /j/ or /w/ between words like *he___earns*. This is incorrect; the glide is already inherent in the pronunciation of the first word (*he* = /hɪj/). Adding an extra glide creates a different phrase (*he yearns*).

Update on resources used for reference must also be made. Modern resources that use or acknowledge the glide-based transcription are: The monograph *English After RP* by G. Lindsey (Lindsey, 2019); the free online **CUBE** dictionary, which offers both accurate (glide) and conservative transcriptions; the *Improve Your Accent* website for comparing transcription systems; *YouGlish* for hearing real-world pronunciations.

Thus, the traditional vowel chart and its symbols perpetuate an inaccurate, early-20th-century view of pronunciation. Adopting a system that recognises the seven closing diphthongs for what they are – vowel+glide sequences – provides a phonologically coherent framework that elegantly explains key features of English connected speech, such as Linking R, clipping, and hiatus resolution, and offers clearer guidance for learners.

Revision: Section 3 (Повторение: Раздел 3)

1. What is meant by the phonetic system (phonetic structure) of a language?
2. What is the general definition of the phoneme?
3. What are allophones?
4. Into which two main groups are allophones divided?
5. What are the three aspects of the phoneme?
6. What are the three functions that the phoneme performs in speech?
7. What is the main difficulty in giving a definition of the phoneme?
8. What is the difference between a speech sound and a phoneme?
9. What was J. Baudouin de Courtenay's view on the phoneme?
10. How did L. V. Scherba's definition of the phoneme evolve?
11. What is the most general distinction in the phonemic composition of a language?
12. What role did the Great Vowel Shift play in the development of the English phonetic system?
13. How are English vowels classified according to the degree of tenseness (muscular tension), length and the stability of articulation?
14. Why does English spelling often not correspond to pronunciation?
15. What units are distinguished in the phonetic structure of the speech continuum (i.e. sound stream)?

4. Phonetics vs. Phonology (Фонетика и фонология: отграничение)

The actual complicity of the definition of the term “phonetics” lies in the fact that from the very origin of this branch of linguistics it dealt with two different objects: phonemes and speech sounds, related to each other as content and form, cf. (Harris, 2000). The field of science that studies phonemes is called phonology (the term phoneme is said to be coined in 1873 by the French linguist A. Dufriche-Desgenettes). There are too many various views as to the relation and the difference between the phonetics and the phonology. As an example, consider the following sentence from the book by M. A. Sokolova et al.: “Phonetics has two main divisions; on the one hand, phonology, the study of the sound patterns of languages, of how a spoken language functions as a ‘code’, and on the other, the study of substance, that carries the code” (Соколова, 2010 : 7). Here phonology is included into phonetics. This view is traditional for Soviet (Russian) linguists, cf. (for instance with the view expressed in: Якобсон, 1962). Europeans prefer to completely distinguish between phonetics and phonology. N. S. Trubetzkoy (another spelling is Trubetskoy, Moscow 1890 – Vienna 1938), the author of *Grundzüge der Phonologie (Principles of Phonology)* – his book was first published in German, see (Трубецкой, 2012)), without reference to whose book no phonological analysis can be done, supposed that a complete division between phonetics and phonology is principle and practical. Americans rather divide between proper phonetics (which deals with physical characteristics of speech sounds) and phonology. It should be mentioned here that in many cases where the publications of American linguists have the word phonology in their title they often deal with the same spheres and aspects of speech sounds that are considered in papers and books of Soviet (Russian) linguists which have the term phonetics on the cover.

The relation between phonetics and phonology is very intricate and intimate. If for theoretical purposes we regard the phonemes and speech sounds as two sides of one and the same object of reality (a view propagated by the Soviet-Russian scholars), we shall have to consider phonemes representing its abstract, ideal side, whereas speech sounds – representing its substantial, material side. If, theoretically, we regard phonemes and speech sounds as completely different objects of reality (the view propagated by Western scholars), we still have to consider the former being abstract and ideal, whereas the latter – being substantial and material. So, as you see the extremes meet if not completely, yet partially, which happens often in science, and the differences between Russian and Western schools show their partial relativity to certain extent and in certain context.

The relation between phonetics and phonology can be understood by saying that phonetics is “the science which studies the characteristics of human soundmaking, especially those sounds used in speech, and provides methods for their description, classification and transcription”, whereas phonology is “a branch of linguistics which studies the sound systems of languages” (Crystal, 2024 : 349-351).

The phonology (also known as the functional or linguistic phonetics) deals with phonological properties of phonemes, syllables, accent, intonation and interprets them as socially significant elements. It studies the ways how speech sounds function in the language, how many or how few of all the sounds of language are utilized in that language, and what part they play in manifesting the meaningful distinctions of the language. It also studies variations of pitch since pitch variations may be quite different in different languages. In English, pitch changes are not part of the shape of a word: that is to say, we can pronounce such word as “No” with a variety of pitch patterns, level, rising, falling or combinations of these, so as to add overtones of doubt, certainty, apathy, interrogation and the like, but the word remains the same old basic negative. But in Chinese there are four different words all of which are pronounced rather like English *ma*, and they are distinguished by their patterns of pitch, *ma* with high level pitch means “mother”; with a rise from medium to high pitch the meaning is “hemp”; a rise from low to medium gives “horse”; and a fall from high to low gives “scold” (O’Conner, 1978 : 18).

When we talk about how phonemes function in a language, and the relationship among the different phonemes – (when, in other words,) we study the abstract (the ideal) side of the sounds of a language – we are studying a related but different subject that we call phonology. Only by studying both the phonetics and the phonology of English is it possible to acquire a full understanding of the use of sounds in English speech (Roach, 1990 : 35). To study the sounds of a language from the phonological point of view means to study the way they function, to find out which sounds a language uses as part of its pronunciation system, how sounds are grouped into functionally similar units, termed phonemes.

It is common knowledge that different languages have a different number of phonemes and different allophones representing them. So each language has its own system of phonemes. The social value of articulatory and acoustic qualities of sounds for the language as a means of communication is different in different languages. In one language two physically different sounds (units) are identified as “the same” sounds, because they have similar functions in the language system. In another language community they may be classified as different because they perform different linguistic functions.

Compare dark & clear “l” sounds in English – they do not affect the meaning, whereas in Russian similar sounds affect the meaning – “лук / люк”.

There are many other differences which are not important on the phonological level of analysis. Cf., for example: *pie*, *spy*, and *lamp-post* – the [p] sounds are different here because of the phonetic context in which they occur: in *pie* – it is aspirated, in *spy* – it is non-aspirated, in *lamp-post* – the first [p] is replaced by a glottal stop and not a plosive sound. Yet, phonologically these sounds are the same (Соколова, 2010 : 53; and earlier editions).

So the aim of phonology is firstly to determine which differences of sounds are phonemic and which are non-phonemic, and secondly, to find the inventory of phonemes of a language. In doing this, we must arrange sounds into functionally similar groups. We do not know yet what sounds are contrastive in this language and what sounds are merely allophones of one and the same phoneme. These are achieved through application of methods of phonetic (and phonological) studies.

4.1. Phonological Features (Фонологические признаки)

As it was mentioned above, the aim of phonology is firstly to determine which differences of sounds are phonemic and which are non-phonemic (and further, what the phonetic or phonological features are that differentiate phonemes), and secondly, to find the inventory of phonemes of a language.

Phonetic features are likewise divided into segmental and suprasegmental features. **Segmental features** (such as voicing, place of articulation, etc.) characterise sounds, sound types and phonemes, whereas **suprasegmental features** (which include stress and intonation) extend over groups of segments and organise segmental units into larger complexes—suprasegmental units. An important property of all features is that they cannot exist independently, without a carrier; phonetic units themselves serve as this carrier. Moreover, features are not linear entities—they do not follow one another in sequence, but form integrated complexes. In other words, each phonetic unit in a given language possesses its own specific set of phonetic features.

For example, the phonetic means of Russian with a distinctive function include sounds, stress (both word and phrasal) and intonation, which often operate jointly or in combination (Современный..., 2012: 14–15, 43).

4.1.1. Aspiration (Аспирация)

!

Aspiration: English Consonants /p/, /t/, /k/

The phonetic feature of **aspiration** – a burst of breath accompanying certain consonants. It is present in English, German, and Scandinavian languages but absent in many others, including Spanish, French, Japanese, Russian, and Dutch.

In English, aspiration occurs specifically with the voiceless plosives /p/, /t/, /k/ when they appear at the beginning of a word or a stressed syllable. Acoustically, this aspiration is similar to the sound /h/ or whispering, which is why it is denoted in the International Phonetic Alphabet with a superscript ^h.

The English /t/ is alveolar (not dental) and involves **affrication** (a faint /s/-like quality), cf. (Lindsey, 2019 : 55-56). The degree of aspiration varies across accents. Some older Received Pronunciation (RP) speakers may use less aspiration. Certain Northern English and strong South African accents often have weak or absent aspiration. Many speakers of Indian English, despite Hindi having aspirated sounds, may not aspirate /p, t, k/ in English.

Consequently, learners from non-aspirating languages must actively learn to produce aspiration in English, while English-speaking actors may need to suppress it to authentically mimic accents where it is not a feature.

Aspiration with Approximants

The sounds /p, t, k/ are aspirated when they occur before a vowel. This means the onset of the vowel is accompanied by a breathy release. A word such as ‘cat’, for example, has aspiration both at its start and its finish.

The three sounds /p, t, k/ are not invariably followed by a vowel. They may instead be followed by one of the ‘approximant’ sounds – /w, r, l, j/ (spelt as ‘w’, ‘r’, ‘l’ and ‘y’, or represented in the International Phonetic Alphabet by /j/), cf. (Lindsey, 2019 : 59-62).

First, /w/. If air is passed through this sound, the result is rather like the noise made when blowing out a candle. Passing air through the /j/ (‘y’) sound produces something akin to the consonant at the end of the German word ‘ich’, or the sound found at the start of English words like ‘hue’ and ‘human’. Passing air through an /l/ gives a sound similar to the Welsh ‘ll’, as heard at the beginning of place names starting ‘Llan’. Passing air through an English /r/ yields a sound resembling the Russian consonant in the middle of ‘Rashan’ or ‘Natasha’.

sound	cluster	examples
/w/	/t/ with /w/	‘tweet’, ‘twin’ and ‘twenty’.
	/k/ with /w/	‘quick’, ‘quite’ and ‘question’.
	/p/ and /w/ (in borrowings)	‘Pueblo’ or ‘Puerto Rico’.
/j/	/p/ with /j/ (most often appears before the vowel /u/)	‘pupil’, ‘puma’ and ‘piano’.
		‘cue’, ‘cube’ and ‘cute’.
	/k/ and /j/	‘tune’, ‘tube’ and ‘Tuesday’; AmE omit the /j/: “toon”, ‘toob’; BrE merge /t/ plus /j/ into /tʃ/: ‘chune’, ‘chube’ and ‘Chewsday’
	/t/ with /j/	‘please’, ‘play’ and ‘plan’.
/l/	/p/ and /l/	‘click’, ‘claim’ and ‘clue’.
	/k/ and /l/	-
	/t/ plus /l/	‘price’, ‘print’ and ‘press’.
/r/	/p/ plus /r/	‘try’, ‘true’ and ‘train’.
	/t/ and /r/	‘crime’, ‘crop’ and ‘crowd’.
	/k/ plus /r/	

4.2. Phonological schools: Moscow, Leningrad and Prague Linguistic Schools (Фонологические школы: Московская, Ленинградская и Пражская лингвистические школы)

Now we come to, probably, the most interesting and question generating part of the theory of the phoneme – the interrelationships among the phonemes of a language. When the phonemes of a language are established, their phonologically relevant features are determined, there arise

other phonological problems and namely: 1) can different phonemes have common allophones? 2) can allophones of a phoneme lose any of their phonologically relevant features? 3) if so, how to define the phonemic status of the grammatically correlated sounds? There are three points of view on these problems expressed by the representatives of the Moscow Phonological School, the Leningrad Phonological School, and the Prague Phonological School (Linguistic Circle). Consider their main postulates.

4.2.1. Methods of the Moscow Phonological School

(Методы Московской фонологической школы)

The representatives of the Moscow Phonological School – R. I. Avanesov, V. N. Sidorov, P. S. Kuznetsov, A. A. Reformatsky, M. V. Panov, N. F. Yakovlev – claim that there may be two positions for a phoneme: strong position (in which the phoneme retains its all distinctive (or phonologically relevant features, and weak position (in which a phoneme may lose one of its distinctive features, e.g.: Russian *пот* – *под*; English *active* – *activity*). In Russian voiced consonants lose their voiced character (a phonologically relevant feature) in final position and are pronounced as voiceless, e.g.: *дуб* – *лед* – *глаз* – *луз* and etc. This fact leads to the loss of the distinction between /б/ – /п/, /д/ – /т/, /з/ – /с/, /г/ – /к/ in Russian, or we may say that in the final position, the phonological oppositions based on phonologically relevant features (voiced//voiceless) are neutralized in Russian. The Moscow Phonological School scientists term this phenomenon – “neutralization of phonological oppositions”. But these scientists say that a phoneme is morphologically bound with all the derivatives of one and the same word, and, if placed in a strong position, the phoneme retains its status, e.g.: *кос* – *косы* – *коса*; *коз* – *ко'зы* – *коза* – *'козы*. Thus the sound [с] in *коз*- must be assigned to the phoneme /з/ and the sound [с] in *кос*- must be assigned to the phoneme /с/ (Vassilyev, 1970 : 168-171).

4.2.2. Methods of the Leningrad Phonological School

(Методы Ленинградской фонологической школы)

The representatives of the Leningrad Phonological School – L. V. Scherba, M. I. Matusevitch, L. R. Zinder, O. I. Dickushina, D. Jones, K. Pike – reject the notion of “neutralization of phonological oppositions”. They consider that an allophone of a phoneme cannot lose any of its distinctive features. They advocate the autonomy of the phoneme and its independence of the morpheme. And if an allophone loses any of its distinctive features, it becomes an allophone of the phoneme, the distinctive features of which it acquires, i.e. speech sounds in a phonologically neutral or weak position belong to that phoneme with whose principal variant they completely coincide. Thus the [с] sound in *коз* should be assigned to the phoneme /с/, because it fully coincides with the principal variant of the phoneme /с/. The advocates of the Leningrad Phonological School (and their western associates) suggest that even one case of phonemic independence of alternating sounds is enough to consider the sounds as representatives of different phonemes (Зиндер, 2007 : 64), and that “Once a phoneme – always a phoneme” (Лайонз, 2010 : 129).

4.2.3. Methods of the Prague Phonological School

(Методы Пражской лингвистической школы)

As it was mentioned, the phoneme (according to the Prague Phonological School (Linguistic Circle) – H. Martinet, N. S. Trubetzkoy, R. Jakobson) is an abstraction and has no material character. In the ideas of the Prague Linguistic Circle, a concept was formulated that in the present stage, which is in front of you, there are always traces of the past, the future and some kind of more or less stable present. In accordance with this school's view, there exist types of phonemes higher than the unit phoneme. They call this higher unit the archiphoneme. They say that the archiphoneme is a combination of distinctive features common to two phonemes (/c/ and /з/ in our case). According to this viewpoint, the [c] sound in *коз* and the [c] sound in *кос* is the archiphoneme (denoted as C – capital Russian letter – introduced by N. S. Trubetzkoy), which is neither voiced or voiceless. It is irrelevant to the phonological opposition voiced / voiceless («эта фонема безразлична к признаку глухой / звонкий»).

Later the Moscow Phonological School also introduced a similar higher phonological unit and they called it hyperphoneme *с/з*, differing from Trubetzkoy's archiphoneme only in the way of symbolizing it (NB: actually there is a more principle way of differentiating between the archiphoneme and the hyperphoneme, which lies in the fact that the former represents the case when certain position (weak position) "neutralizes" the difference between phonemes, yet the strong position for the phonemes can be easily found, whereas the latter represents the case when there is no weak or strong position as such), cf.: Trubetzkoy [c] – Moscow Phonological School [с / з]. Cf. in Russian *во'да* and *'воды* – according to Moscow Phonological School vowels of the first syllables are variants of one phoneme, whereas according to Leningrad Phonological School – they represent different phonemes (consider also Russian *на воду*). As in Russian *собака* there is no strong position and it is impossible to identify what phoneme is represented by the first vowel – it is the hyperphoneme used here – *a/o* according to Moscow Phonological School. The final consonant in Russian *ход* phonemically can be identified as phoneme /д/ through its strong position in *хода*; whereas for the *вом* and *вом бы* there is no strong position (and this is the difference between the hyperphoneme and the archiphoneme) – yet we know that in *вом* the final consonant is not /к/, /с/, or /ш/, we even know that word final phoneme in *вом* belongs to т/д group – though we cannot identify which of the two phonemes is used here – i.e. we have here the bewitched archiphoneme т/д. It is important that the archiphoneme т/д having no individual distinctions between т or д retains the group distinction – it is definitely т/д group and not п/б, к/д, or с/з groups.

4.2.4. N. S. Trubetzkoy's "Basics of Phonology"

(«Основы фонологии» Н. С. Трубецкого)

Prince Nikolai Sergeyevich Trubetzkoy (1890–1938), a leading figure of the Prague Linguistic School, is rightfully considered the founder of modern phonology. Using data from over

200 languages and building upon theoretical concepts of the phoneme, he developed a rigorous framework for studying and describing phonemic systems.

Trubetzkoy's theory is termed 'functional' because he posited that a phoneme comprises only phonologically significant features and serves to distinguish word meanings, thus fulfilling a sense-discriminating function (Беляева, 2019 : 53-54). He defined the relationships between phonemes as 'oppositions' and categorised their various types in his seminal work, *Principles of Phonology*, published posthumously in German in 1939. This oppositional method became a powerful tool for linguistic analysis, extending beyond the study of speech sounds. Trubetzkoy created a typology of oppositions—the structural relationships between phonemes. Key distinctions include:

- **one-dimensional vs. multidimensional:** oppositions based on a single distinctive feature (e.g., /T/ and /D/, differing only in voicing) versus several features (e.g., /E/ and /U/, differing in tongue position, height, and lip rounding);

- **proportional vs. isolated:** oppositions found in multiple phoneme pairs (e.g., the hard/soft pairs /Z/–/Z'/, /S/–/S'/) versus unique, non-repeating pairs (e.g., /L/–/L'/);

- **privative, equipollent, and gradual:** these differ in the logical relationship between members;

- **privative oppositions** are based on the presence or absence of a single feature (e.g., voicing in /D/–/T/);

- **equipollent oppositions** involve phonemes that are logically equivalent and independent (e.g., /F/, /P/, /T/);

- **gradual (non-binary) oppositions** show a stepwise progression of a feature, such as the Russian vowels /I/, /Y/, /U/, which differ along the front-central-back axis.

Let us summarize the main differences between the views of the Moscow and the Leningrad Phonological Schools onto the phoneme theory in the following table 8:

Table 8

The Conceptual Differences between Moscow Phonological School and Leningrad Phonological School in their Views onto the Phoneme Theory

Aspect	Moscow Phonological School	Leningrad Phonological School
1) <i>definition</i>	the phoneme – the shortest possible unit of the phonetic structure of a language is	
2) <i>phoneme vs. morpheme</i>	for morphological structure of a language, the shortest possible unit – the morpheme (that of the system of language associated with the meaning: Rus. <i>нѣс</i> – <i>нѣс</i> , <i>бил</i> – <i>нул</i>)	the autonomy of the phoneme and its independence from the morpheme; denying morphology from phonological analysis (departure from Baudouin de Courtenay)
3) <i>neutralization</i>	the theory of strong and weak positions for phonemes (the notion of “neutralization of phonological oppositions”)	denying “neutralization” of the phoneme, which can not lose its distinctive features (if yes – it becomes the phoneme, whose distinctive features it acquired)

Aspect	Moscow Phonological School	Leningrad Phonological School
4) <i>strong / weak positions</i>	strong / weak positions not for phonemes, but for phonological principles (for some phonemes may be strong, for others weak), e.g.: Rus. word final position is weak for <i>код – кот</i> , but strong for <i>плоть – плот</i>	no support of this view
5) <i>investigation of positions</i>	commitment to investigate all possible positions for sounds (linear, combinatory, contextual); positional and combinatory changes of allophones of one and the same phoneme	paid little attention; only positional and combinatory changes of phonemes; the changes of phonemes (Rus. <i>мена фонем</i>) lead to differentiation of meaning: e.g. Rus. <i>дуб, дубки – дуба</i>
6) <i>hyperphoneme</i>	the notion introduced, e.g.: Rus. <i>собака, вода</i> = [a/o]	this notion is not recognized

4.3. Methods of Phonological Analysis: Development of the Main Principles (Методы фонологического анализа: разработка основных принципов)

Here we discuss the methods of phonological analysis, whereas the methods of analysis in phonetics were discussed in Section 1.5 above.

Thus we may say that any language exists in its sounds and to master some foreign language we must know not only the orthography, but we must be able to recognize these sounds in connected speech and be able to reproduce these sounds correctly in order to contribute to the process of communication. To acquire this ability we must know the articulation of the sounds (articulatory phonetics), their acoustic properties (acoustic phonetics), their auditory features (auditory phonetics), and their phonological properties – their position; distribution, variation and how they are contrasted in speech – functional phonetics or phonology.

Note: It has been pointed out in the beginning of this essay that speaking about methods of phonetics we understand the latter one in terms of Russian / Soviet school. If we set phonetics and phonology completely apart from each other (as most Western researchers do) we need to exclude the method(s) of phonological analysis from the sphere of phonetics.

The method(s) of phonological analysis (sometimes called *proper linguistic method* or *functional method of phonetics*) includes *the distributional method, the statistical method, the method of minimal pairs* (semantic method, commutation test, method of substitution). All the methods are interconnected (that is why it is possible to speak about the phonological method or the functional method as it is) and are aimed at establishing the inventory of the phonemes of the given language.

The phoneme has several functions and can be viewed from different aspects. That is why there are many definitions of the phoneme in linguistics depending on what aspects or functions of the phoneme are emphasized by a researcher or schools of researchers. It needs to be mentioned here that a phonemic (phonological) system of any language is far from being stable;

instead the units of phonological systems tend to change their status and relation to other units constantly. Certainly, it is no easier to identify in general what a phoneme is when every special substance of such a generalization results from one change and aims at another (for evolution and changes of phonological systems see for instance: Плоткин, 1982).

The definition of the phoneme by V. A. Vassilyev is as follows: “the segmental phoneme is the smallest (i.e. further indivisible into smaller consecutive segments) language unit (sound type) that exists in the speech of all the members of a given language community as such speech sounds which are capable of distinguishing one word from another word of the same language or one grammatical form of a word from another grammatical form of the same word” (Vassilyev, 1970 : 136).

The distributional method is the analysis “establishing the distribution of speech sounds, i.e. all the positions or combinations in which each speech sound of a given language occurs (or does not occur) in the words of that language” (Vassilyev, 1970 : 24). **The statistical method** is very close to the distributional method and can be defined as the analysis “establishing the frequency, probability, and predictability of occurrence of phonemes and their allophones in different positions in words” (Vassilyev, 1970 : 25).

Though the methods resemble each other, which can be seen from the definitions, the results of their applications, so to say, are different, which should be mentioned here:

1) the initial data and the material of the statistical analysis are mostly phonemes and their allophones, already identified as such, whereas the material of the distributional analysis is the non-segmented speech flow which is to be segmented into speech sounds;

2) the statistical method shows in what positions phonemes occur, whereas the distributional method also shows in what positions the sounds in question can never occur;

3) the statistical analysis provides numerical data of occurrence of phonemes and their allophones in certain positions and predicts the occurrence of a certain phoneme in a particular linear succession of phonemes, the distributional analysis investigates the interrelation of speech sounds between each other and substantiates whether they are allophones of the same phoneme or of different phonemes.

The distributional method states that:

1) allophones of different phonemes occur in the same phonetic context, e.g.: *мом – дом, pit – bit*, i.e. that they are in contrastive distribution; or

2) allophones of the same phoneme never occur in the same phonetic context, e.g.: *call – keep*. They are here in complementary distribution.

Yet there are cases when two sounds are in complementary distribution but are not referred to the same phoneme. Consider English sounds [h] & [ŋ] for example: [h] – occurs only initially or before a vowel while [ŋ] occurs only medially or finally and never occurs initially. Here the method of distribution is modified by addition of the criterion of phonetic similarity / dissimilarity. Articulatory features are taken into account in this case also (Соколова, 2010 : 55).

The semantic method is the analysis establishing the ability of sound segments to differentiate the meaning. It consists in finding minimal pairs (also called quasi-homonyms) of words and their grammatical forms in a language. For this the sound segments distinguished as a result of the distributional analysis are substituted one by one in a certain phonetic environment to see if different substitutions lead to different semantic meanings. This procedure

is also called commutation test, e.g.: pin – bin, win, sin, wins. Here allophones [p, b, w, s] represent different phonemes, because the meaning is different in all words, though the phonetic context is the same. The semantic method of identifying the phonemes of a language attaches great significance to meaning. This method is based on a phonemic rule that phonemes can distinguish words and morphemes when opposed to one another. If two or more linguistic units of a level higher than the phonemic one (i.e. words or morphemes) differ solely in one of their segments occurred in the same position whereas the rest of phonetic environment is the same – the linguistic units are called a minimal phonological pair and the segments in question are regarded as different phonemes or, better say, representatives of different phonemes. Consider one more example of the commutation test in English words: *pen, den, ten, then*, and etc.

4.4. The Notion of the Phonological Opposition; the System of Phonological Oppositions

(Понятие фонологической оппозиции; система фонологических оппозиций)

In phonology (using the method(s) of phonological analysis) we must also establish the system of oppositions. There are three kinds of oppositions. If members of the opposition differ in one feature the opposition is said to be single, e.g.: *pen – ben*. Common features: occlusive – occlusive, labial – labial. Differentiating feature: fortis – lenis. A double opposition takes place if two features are marked (i.e. different), e.g.: *pen – den*. Common features: occlusive – occlusive. Differentiating features: labial – lingual, fortis – lenis. If 3 distinctive features are marked the opposition is triple, e.g.: *pen – then*. Differentiating features: occlusive – constrictive, labial – dental, fortis – lenis. Consider the above mentioned oppositions represented in Table 9 below. For more details see (Соколова, 2010 : 57-58).

Table 9

Three Kinds of Oppositions in Phonology

№	Number of features	opposition	common feature(s)	differentiating feature(s)	examples
1	1	single	occlusive labial	fortis	pen
				lenis	ben
2	2	double	occlusive	fortis	pen
				labial	
				lenis	den
				lingual	
3	3	triple	noise	fortis	pen
				labial	
				occlusive	
				lenis	then
				lingual (dental)	
				constrictive	

4.5. Types of Transcription (Виды транскрипции)

The abstractional and material aspects of the phoneme (see Section 3.3.) have given rise to two basic types of transcriptions: 1) phonemic – connected with abstractional aspect; 2) phonetic (allophonic) – connected with material aspect.

A phonemic (linguistically broad transcription) is based on the principal of one symbol per phoneme. Each of its symbols denotes a phoneme as a whole, e.g.: /t/ – this symbol in slanting lines stands for all allophones of the phoneme /t/ in English – alveolar, dental, post-alveolar, half-voiced, unaspirated and etc.

An allophonic (phonetic) or linguistically narrow transcription is based on the principal of one symbol per allophone. This transcription provides special signs for each variant, e.g.: [ɫ], [l'], [ɫ̥], [t], [t'], [ä], [å], [ö]. The texts in phonetic transcription are hard to read; too many symbols. Phonemic transcription is used more extensively than phonetic, because we not underline minute details, e.g.: [ɫ] – dark, and [l'] – soft (clear). It is easier and quicker to remember, because most of its symbols coincide in shape with the letters of the alphabet; English [p], [t], [k] are aspirated before stressed vowels if not preceded by [s], so there is no need to give such phonetic symbols as *ph*, *th*, *kh*. Phonetic transcription is used in field work to denote minute details of speech sounds (Vassilyev, 1970 : 145).

4.6. International Phonetic Alphabet, IPA

(Международная фонетический алфавит, МФА)

The recording of spoken speech can be achieved using the International Phonetic Alphabet (IPA). This Latin-based alphabet was developed on by the International Phonetic Association, which has existed since 1886, with its membership including 11 countries, such as the USA, in its first year. The most recent refinements to the alphabet were made in 1989. The current system comprises 115 individual symbols, though phonetic dictionaries in Europe and the USA most frequently use only a portion of these—typically 40 to 45 symbols. Diacritical marks are also employed to denote specific qualities such as palatalisation or length (Беляева, 2019 : 30). Building upon this foundation, for instance, L. V. Shcherba created a transcriptional system for the Russian language (Современный..., 2012 : 25). In fact, any human language may be transcribed by means of IPA.

!

Phonetics: IPA Upside Down

When people first come across the International Phonetic Alphabet (IPA), they frequently question how its different symbols came to be, cf. (Lindsey, 2011). A common puzzle is why so many of the letters appear upside down. Most of these inverted characters were not created by simply flipping a letter vertically, as one might do on a computer today.

To understand the reason behind this, we must return to the late 19th century and the origins of the IPA. Its founders based the system on the familiar Roman alphabet, but their goal was to create

a universal alphabet capable of representing every sound found in the world's languages – far more than the 26 letters of the standard Latin script used for English. They needed new symbols that would harmonise with the existing letters, but at that time, casting new metal type was a costly process. Phoneticians therefore devised a clever and far more economical solution.

The eminent Henry Sweet, who laid the groundwork for the IPA, outlined this approach in 1877: “There is, however, one simple method of forming new letters without casting new types, which is very often convenient. This is by turning the letters, thus”. This explains why so many of the IPA's upside-down symbols are, in fact, rotated versions of basic Roman letters.

Revision: Section 4 (Повторение: Раздел 4)

1. What is the principal difference between phonetics and phonology?
2. What are the main principles of the Prague phonological school?
3. What characterises the approaches of the Moscow & St.-Petersburg (Leningrad) phonological schools?
4. What is meant by the method of phonological analysis?
5. What does the distributional method of phonological analysis establish?
6. What does the statistical method of phonological analysis establish?
7. What is the purpose of the semantic method (method of minimal pairs, commutation test)?
8. What are phonological (distinctive) features?
9. What is a phonological opposition?
10. What types of transcription are distinguished, and what is the role of IPA?

5. Modifications of Phonemes in Speech, Types of Modifications (Модификация фонем в речи, типы модификаций)

5.1. Allophonic Variations & their Causes (Аллофонные вариации и их причины)

Such phenomena as modifications of consonants and vowels in connected speech must also be considered.

Language in everyday use is performed in connected sequences of larger units than separate sounds, i.e. in words, phrases and longer utterances. And we must know what happens to sounds not only in words pronounced in isolation but also in a block of connected speech. These changes are mostly quite regular and predictable. Your speech is too complicated because of the numerous modifications of sounds in speech. These modifications are observed both within words and at word boundaries, because speech sounds influence each other in the flow of speech.

!

‘An’ or ‘A article’?

The phonological behaviour of the indefinite article (‘a’/‘an’) and definite article (‘the’) in English is in fact about how speakers manage **hiatus** – the undesirable adjacency of two vowel sounds. The core principles and variations are summarised below.

Standard rule and hiatus: the indefinite article has two **allomorphs** (conditioned variants): ‘a’ (/ə/) before consonants and ‘an’ (/ən/) before vowels. The ‘n’ in ‘an’ serves as a **hiatus resolution** strategy, preventing two vowels from clashing (e.g., “an end” avoids the hiatus in “a end”). A parallel, though less commonly noted, variation exists for the definite article: /ðə/ before consonants and /ði:/ before vowels.

There are **hiatus resolution** strategies: when speakers use the schwa forms (/ə/, /ðə/) before a vowel, hiatus is typically avoided not by the ‘n’, but by **Hard Attack**. This involves inserting a **glottal stop** (/ʔ/) between the vowels (e.g., “ðə ʔ end”, “tə ʔ end”). This strategy is prevalent in languages like German and is increasingly common in English, cf. (Lindsey, 2019 : 91-93).

The dialectal and social **variation** of this phenomenon is as follows. First, **non-standard use of ‘A’**: many English varieties, historically and today, use “a” before vowels (e.g., “a end”). This is documented in historical English dialects (per Joseph Wright, 1905); contemporary **London English**, particularly among some speakers who frequently use glottal stops (e.g., /əʔend/); **South African English**, influenced by Afrikaans; **African American Vernacular English (AAVE)**; **General American English** (a noticeable increase in using “a” before a vowel is observed, especially on online platforms. This often occurs before a pause or hesitation, where the hiatus is negated by the break itself. The strong form /eɪ/ (lengthened during hesitation) or forms with a glottal stop (/əʔ/, /æʔ/)

are used); **British English note:** when the strong form /eɪ/ is used before a vowel with a glottal stop (e.g., /eɪʔ/), it can phonetically merge with the word “eight” when the final /t/ is glottalled.

Observations on certain trends must be made. There is increasing hard attack: the use of glottal stops to resolve hiatus (e.g., /ðəʔend/) appears to be growing, notably among younger American women in presentation-style speech (e.g., on online platforms); the form ‘an’ is **statistically** less frequent than ‘a’ because English words begin with consonants more often than vowels; speakers may have flexibility and employ multiple strategies simultaneously, and the use of “strong forms” of typically weak words (like /eɪ/ for “a”) is strengthening in more formal or self-conscious contexts.

In essence, while the standard rule of ‘a’ / ‘an’ is well-known, actual usage reveals a complex landscape of phonological strategies (like the intrusive ‘n’ or glottal stop) to manage vowel clash, with significant variation across dialects, speech styles, and social groups.

5.2. Positional & Combinatorial Changes of Phonemes

(Позиционно-комбинаторные изменения фонем)

As a result of the intercourse between consonants and vowels and within each class (i.e. in Consonants & in Vowels) there appear such processes of connected speech as assimilation, accommodation, elision, vowel reduction, and sound alternation.

There are sound changes in words, their derivatives and grammatical forms. They are called sound alternations, e.g.: *sing – sang – sung*, *advice – advise*, *breath – breathe*, *child – children*, *important – importance*, *speak – speech*, *spell – spelling*, *петь – пеку*.

The alternation of sounds can be a **reflection of language history**. Let’s illustrate it by the example of some vowels in Russian (Беляева, 2019 : 40-41). Specifically, whether a vowel alternates with a zero sound (i.e., disappears) in certain grammatical forms depends on its **etymological origin**.

Alternation is a **historical marker** which acts as a diagnostic tool to identify a vowel’s historical source. A vowel that alternates with zero (e.g., **o** in *сон / сна* or **e** in *день / дня*) reveals its origin from an old, ultra-short vowel (specifically, the letters **ерь** or **ерь**).

Conversely, **non-alternating vowels**, i.e. vowels that do not alternate with zero (e.g., **o** in *дом / дома* or **e** in *лес / леса*) have a different historical origin (not from **ерь/ерь**, but from other sources, such as the letter **ять** for the vowel **e**).

Thus, vowel alternation with zero is not random but is a **grammatical fossil** preserved in modern Russian that points directly to the historical development of the language’s vowel system.

5.2.1. Assimilation (Ассимиляция)

In consonants in most cases we distinguish assimilation. Assimilation (ассимиляция, уподобление) and accommodation (приспособление) are both phonetic adaptive mechanisms. Assimilation is the adaptive modification of a consonant by a neighbouring consonant in speech

flow. In case of assimilation the assimilated consonant becomes similar to the assimilating one, or both modify so that they become resembling each other (in case of double assimilation), e.g. the alveolar [t] followed by the interdental [θ] or [ð] becomes dental. In case of accommodation the modification of one or both adjacent sounds does not lead to resemblance or similarity between them.

Assimilation may be:

1) complete – when the assimilated phoneme coincides with the assimilating phoneme, e.g.: *horse-shoe* [hɔːʃ fuː];

2) partial – when the assimilated phoneme retains its main phonemic features and becomes partly similar to the assimilating phoneme, e.g.: [t], [d], [n], [l], [s], [z] before [θ] or [ð] become dental;

3) intermediate – when the assimilated phoneme changes into a different phoneme, but does not coincide with the assimilating phoneme, e.g.: *Congress* – /n/ – becomes /ŋ/ (but not /g/).

Assimilation may be:

1) progressive – when the assimilated phoneme is influenced by the preceding phoneme, e.g.: *place* – [l] is partly devoiced; [z] in *is* – [iz] – in abbreviated form becomes [s] if preceded by a fortis (voiceless) – *your aunt's coming* – the [t] fortis makes [s] in *is coming*;

2) regressive – when the preceding phoneme is influenced by the following one, e.g.: *news* + *paper* gives ['njuːs, peɪp ə], though *news* (новости) is [njuːz]; in Russian – *сдала экзамен*;

3) double (reciprocal) when the adjacent phonemes influence each other, e.g.: *twenty* – here /t/ is labialized (labialized allophone [tʷ] is used), whereas /w/ is devoiced (voiceless allophone is used).

!

Assimilation: Consonants /s/+/p^h/ > /sp/

Aspiration is a critical feature of English pronunciation, as its absence can lead to misunderstandings; unaspirated /p, t, k/ are often perceived by native speakers as /b, d, g/. However, there are specific contexts where aspiration is **blocked** for native speakers, cf. (Lindsey, 2011).

The primary blocking sound is /s/. When /p, t, k/ follow an /s/ **within the same word or morpheme**, they are not aspirated. Consequently, if the /s/ is removed, the remaining sound is heard as its voiced counterpart: 'speech' without /s/ sounds like 'beach'; 'store' without /s/ sounds like 'door'; 'screen' without /s/ sounds like 'green'; 'splendour' without /s/ sounds like 'blender'.

This occurs because English /b, d, g/ are not fully voiced in most positions, making them acoustically similar to unaspirated /p, t, k/. This lack of contrast after /s/ can create homophones in rapid speech, such as 'it's blended' reduced to "'s blended", which sounds identical to 'splendid'.

This prohibitive effect extends to other **voiceless fricatives** (e.g., /ʃ/, as in 'shtick'), though examples are less common. German exhibits the same phenomenon.

Critically, this blocking only applies **within a single word or morpheme**. If the /s/ belongs to a separate word or prefix, aspiration remains: e.g., 'Miss Piggy', 'plus tax', 'miscalculate'.

Historical word fusion can obscure this rule. Over time, prefixed words can fuse, allowing the /s/ to de-aspirate a following plosive: ‘discuss’ (from *dis-* + *cuss*) now has a /k/ that sounds like /g/. This leads to variability: ‘dystopia’ may be pronounced with /t/ or /d/, and ‘displease’ shows conflicting treatments in dictionaries.

Dialectal differences exist, for instance, Many British speakers fuse numbers like ‘fifteen’ and ‘sixteen’, pronouncing them as ‘fifteen’ and ‘sixteen’. Many American speakers often keep the ‘teen’ morpheme distinct, preserving an aspirated /t/.

For learners, particularly from languages like Italian or Slavic where /s/ voices before voiced consonants, it is important to avoid pronouncing clusters like /sp/ as /zb/. Transcription systems may offer the option to display these clusters as /sb, sd, sg/ for clarity, alongside indicating aspiration where it correctly occurs.

Further complexities regarding aspiration before unstressed vowels and in word-final position are beyond the scope of this summary.

!

Assimilation: Consonants /s/+/t^h/>/ʃt/

There is a phonological process in English whereby the /s/ sound in certain consonant clusters assimilates to a /ʃ/ (“sh”) sound, particularly before /r/. This explains why some speakers pronounce words like *strong* and *street* as “shtrong” and “shreet”.

The phonetic basis: English /r/ and post-alveolar sounds

The key to this change lies in the nature of the English /r/ sound. Unlike the trilled or tapped ‘r’ of many languages, English /r/ is a **post-alveolar** sound, produced with the tongue body retracted, cf. (Lindsey, 2011). This is the same area of the mouth used to articulate /ʃ/ (‘sh’), /tʃ/ (‘ch’), and /dʒ/ (‘j’). These sounds also commonly involve lip rounding.

There are established **assimilation patterns** in English. English already exhibits a preference for post-alveolar consistency within clusters. First, **absence of */sr/**: no native English words begin with /sr/ (e.g., *srimp*, *sred*). When encountering such spellings, speakers instinctively pronounce them as /ʃr/ (e.g., *Sri Lanka* as “Shri Lanka”).

A widespread phenomenon, especially among younger speakers, involves the assimilation of /tr/ and /dr/.

- **Train Changing:** /tr/ → /tʃr/ (e.g., *train* pronounced as “chrain”, not “chain”).
- **Drum Majoring:** /dr/ → /dʒr/ (e.g., *drum* pronounced as “jrum”).
- **The emergence of “Street Shopping”:** the logical extension of this trend is the assimilation of the /str/ cluster, termed *Street Shopping*. This process changes /str/ to /ʃtr/ (e.g., *street* → “shreet”, *strong* → “shtrong”).

It is seen as creating phonetic “elegance”, making the entire cluster post-alveolar, just like /ʃr/, /tʃr/, and /dʒr/, thereby eliminating the awkward tongue movement from an alveolar /s/ to a post-alveolar /r/. *Street Shopping* is less established than Train Changing and Drum Majoring but is well-attested in some accents (speakers like Michelle Obama are cited) and appears to be increasing. Research indicates it is typically used by speakers who already employ the other two assimilation processes.

The shift from /str/ to /ʃtr/ is not an isolated oddity but part of a broader, ongoing pattern of assimilation in English. Driven by a phonetic preference for consistency in the place of articulation, it represents a tidying up of word-initial consonant clusters, particularly amongst younger generations, leading to a system comprising the consistently post-alveolar clusters /ʃr/, /tʃr/, /dʒr/, and /ʃtr/.

!

Assimilation: Consonants /t^h/+ /j/ > /tʃj/

There are historical origins and modern variations in pronouncing sequences involving an alveolar consonant followed by a /j/ sound (a “yod”), particularly before the /u:/ vowel. This results in the variation between pronunciations like “tube” (with /j/) and “toob” (without /j/), or “tube” and “chube”, cf. (Lindsey, 2019 : 59-62).

The /j/ sound in words like “music”, “few” and “cute” **historically originates** from the Norman French influence on Middle English after 1066. French words containing a front rounded vowel /y/ were adapted by English speakers, who split it into a diphthong /iu/. Over time, the /i/ element evolved into the consonant glide /j/, which is now hidden within spellings using the letter ‘u’ (e.g., *cute, music, few*).

Pronouncing an alveolar consonant (/t, d, n, s, l, z/) directly followed by /j/ is phonetically awkward, leading to two main **simplification strategies**:

1) **Yod-Dropping**: the complete omission of the /j/ sound. This is widespread in **American English** after alveolar consonants (e.g., *tube* /tu:b/, *new* /nu:/, *dune* /du:n/), though its application is inconsistent (e.g., *cube* keeps the /j/, *puke* keeps it, *Luke* drops it). It is taken further in some **Eastern English accents**, where yod-dropping can occur after *all* consonants.

▪ **Yod-Coalescence (or “Tuna Chewing”)**: the merging of the alveolar consonant and the /j/ to form a new, intermediate post-alveolar sound. This is very common **after a stress** in many accents: *statue* (/t/ + /j/ → /tʃ/), *graduate* (/d/ + /j/ → /dʒ/), *issue* (/s/ + /j/ → /ʃ/), *visual* (/z/ + /j/ → /ʒ/). In **stressed positions** (e.g., *tune, dune, suit*), accents vary significantly: some **keep the yod** (e.g., /tju:n/); some **drop the yod** (e.g., /tu:n/ - common in American English); some use **coalescence** (e.g., /tʃu:n/ - “chune”, common in some British accents).

There is extensive **variation and modern complications** both between accents and from word to word. Furthermore, the history of a word influences its pronunciation. **Older borrowings** were often “nativised” and given a /j/ (e.g., *Cuba* /ˈkju:bə/). **Recent borrowings** or those taken from accents that drop yods typically resist the /j/ sound (e.g., *Putin* not *Pyutin*, *dude* not *dyude*, *jacuzzi*, *Sudoku*).

So, the pronunciation of words like *tube*, *tune*, and *dune* is governed by a complex interplay of historical sound change, phonetic simplification processes (dropping or merging the /j/), and sociolinguistic factors relating to a word’s origin and the speaker’s accent. This leads to the three-way variation between forms like *tyube*, *toob*, and *chube*.

The variation in pronouncing words like *strong* as “shtrong” stems from a pervasive but often overlooked process of **assimilation**, where sounds change to become more like neighbouring sounds. This is distinct from and more widespread than the isolated “SH” substitution it initially suggests.

The **primary change** involves the **retraction** of alveolar sounds (/t, d/) when followed by /r/. In accents like Southern British English: /t/ and /d/ are pulled back from their standard alveolar position towards the **post-alveolar** position of the following /r/; this retraction makes /tr/ sound similar to /tʃr/ (“ch-r”) and /dr/ sound similar to /dʒr/ (“j-r”), processes termed **Train Changing** and **Drum Majoring** respectively (e.g., *train* → “chrain”, *drum* → “jrum”).

There is **dialectal variation** connected with /r/. This shift is driven by the specific **post-alveolar** quality of English /r/. In **Scottish** accents that use a tapped or trilled /r/ (which is not post-alveolar), this retraction does not occur. Paradoxically, some Scots accents show a different assimilation where /r/ affects a *following* sound (e.g., in *force*). The pattern is not a simple American innovation but a phonetic process linked to the nature of the /r/ sound in an accent.

Word boundaries have their importance in the question of assimilation. Assimilation is typically blocked by a morpheme or word boundary. The influencing /r/ must be within the same basic word unit. Compare *bedroom* (high frequency, boundary faded) where “drum majoring” applies (/d/ → /dʒ/, “bedgeroom”) with *boardroom* (lower frequency, boundary intact) where it does not. Similarly, an /s/ will only de-aspirate a following /t/ or /k/ within the same morpheme (e.g., *speech* sounds like “beach”) but not across a clear boundary (e.g., *miscalculate* retains an aspirated /k/). Boundaries fade at different rates across accents: many Americans retain a boundary in *sixteen* (with aspirated /t/), while many Brits do not.

The pronunciation “shtrong” (or “**Street Shopping**”), a current trend, is a further assimilation where /str/ becomes /ʃtr/. Crucially, evidence suggests that for many younger speakers, the change is not simply /s/ → /ʃ/ “at a distance”. Instead, the intervening /t/ has first undergone Train Changing to /tʃ/. Therefore, a common modern pronunciation may be better transcribed as /sʰtʃrɒŋ/ (“s-chrong”), featuring both the changed stop and a retracted /s/, rather than a direct /ʃtr/.

The **practical implication** and academic oversight of the information above is as follows. Teaching learners to use /tʃr/ and /dʒr/ in words like *train* and *drum* can be a straightforward way to help them achieve a more native-like pronunciation. Despite the scale and significance of this ongoing phonetic change, it remains largely unaddressed in mainstream textbooks and dictionaries.

5.2.2. Accommodation (Аккомодация)

Cases of accommodation have been already touched upon when we discussed the phoneme theory. In English it is the accommodation that triggers such phonetic phenomenon as a loss of plosion, when in a cluster of two plosive consonants the first one discards its plosive character and sometimes gives place to a completely different sound – a glottal stop, e.g.: *black cat*, *start talking*, *lamp-post*, *bad boy*, and etc. Also accommodation is observed in English in such cases

as loss of aspiration of voiceless stops before a vowel after /s/, e.g.: *speak, spy, sty* and etc. Also cases of progressive relative shortening of phonologically long vowels before nasal consonants, before voiced stops, and before voiceless stops as compared to their length in word final position may be treated probably as examples of accommodation, e.g.: *see – seen – seed – seat* (phonologically long vowel /i:/ is shorter in every consecutive word).

NB: historically the usage of short vowels with certain consonants and/or with their combinations in certain positions results in English in lengthening or diphthongization of the vowel, which may also be treated as accommodation, though historical, cf.:

Old English short vowel		Middle English l ong vowel		New English diphthong or long vowel
<i>nǣma</i>	>	<i>nāme</i>	>	<i>name</i>
<i>mēte</i>	>	<i>mēte</i>	>	<i>meat</i>

For the similar reason the English *child* has a diphthong and *children* has a short vowel, for more examples see (Hock, 2021 : 139) et sequential. Probably, a complex accommodative-assimilative phenomenon can be observed in the word too [tu:], where /t/ is labialized under the influence of /u:/ and /u:/ is a little advanced under the influence of /t/.

!

Glottal Stop = “Hard Attack”

The glottal stop has been increasing in English pronunciation, particularly in a form of a so-called “hard attack” before vowel-initial words, contrasting with its traditional role and its use in languages like German, cf. (Lindsey, 2019 : 91-93).

Anatomical Basis: the glottis and glottal stop. The hard attack occurs within the following anatomy. The **larynx (voice box)** houses the **vocal folds (cords)**, which open for breathing and vibrate for speech. The space between them is the **glottis**. A **glottal stop (/ʔ/)** occurs when the folds are firmly closed, blocking the airstream. It is a full consonant in some languages.

The Glottal Stop in Traditional English. In English, the glottal stop has historically been a marginal sound: as reinforcement for voiceless stops (/p, t, k/) in some accents, or as a replacement for /t/ in non-prevocalic positions (e.g., “Greaʔ Britain”), often mislabelled as “T-dropping”.

“Hard attack”: glottal stop before vowels. **Hard attack** refers to starting a vowel sound with a preceding glottal stop (e.g., “the /ʔend/”) (Lindsey, 2011). In German (**Glottisschlag** or **Knacklaut**), which is in fact a language directly related to English, this is the standard, neutral way to begin a word-initial vowel, contributing to the language’s distinctive rhythm. Traditionally in English, it was used **only for special emphasis**. On the other hand, descriptions of “smooth” (old) Received Pronunciation of the mid-20th century show minimal use of non-emphatic hard attack, cf. (Lindsey, 2011).

The increase of hard attack in contemporary English. Hard attack has become significantly more common in modern English, often **without emphatic intent**. A key driver is its use to **avoid hiatus** (the direct clashing of vowels). Speakers increasingly use the weak forms /ðə/ and /tə/ even before vowels (instead of /ði:/ and /tu:/), inserting a glottal stop to separate the vowels: e.g., /ðə ʔend/, /tə ʔend/. Some British speakers may even use both a linking /r/ **and** a glottal stop to prevent hiatus.

The use of hard attack also varies in accordance with the **style** of speech. Heavy use of non-emphatic hard attack is particularly characteristic of **presentational speech** (e.g., lectures, explanations to an audience), where it may serve to clarify word boundaries for extra clarity. Its use typically decreases in casual, one-to-one conversation. It is also less common before weak vowels (e.g., in /ə/, of). While notably prevalent among younger American women, the trend is widespread among younger speakers in both the US and Britain. It can also occur before certain consonants, like /l/.

In short, the glottal stop, from its traditional roles in English, is now increasingly employed as a German-style hard attack. This shift is primarily to resolve vowel hiatus in connected speech and is especially prominent in formal, presentational speaking styles, marking a significant change in the rhythm and phonetic detail of contemporary English.

5.2.3. Elision (Элизия)

Elision is a complete loss of sounds, both vowels and consonants in rapid speech. The nature of elision may be stated quite simply: under certain circumstances sounds disappear, because in any language (English is not an exception) there is a universal law of economy of pronouncing efforts. The speaker and the listener are two participants of communication. The speaker's aim is to inform the listener of something. The listener's aim is to comprehend the idea. Elision (or simplification) goes quite unnoticed by the listener as it does not affect the meaning. The listener is mainly interested in the meaning the speaker wants to convey and not in the precise phonetic organization of a block of connected speech.

Producing elisions is something which foreign learners do not need to learn to do, but it is important for them to know that. When native speakers of English talk to each other quite a number of phonemes that foreigner might expect to hear are not actually pronounced. Below are some examples of elision:

1) loss of weak vowels after [p], [t], [k]: *potato* [p^hteitəʊ], *today* [t^hdeɪ], *canary* [k^hneəri], *perhaps* [p^hhæps] (first syllable may disappear completely);

2) weak vowel before [n], [l], [r] becomes a syllabic consonant ($V_{\text{weak}} + [n]/[l]/[r] = C_{\text{syll}}$): *tonight* [təɪt], *police* [pɒlɪ:s], *correct* [kɹɛkt];

3) avoidance of complex consonant clusters. In clusters of three plosives or two plosives plus a fricative, the middle plosive may disappear: *acts* [æks], *looked back* [lʊk bæk], *scripts* [skrips];

4) loss of final [v] in *of* before consonants: *lot of them* [lɒt ə ðəm], *waste of money* [weɪst ə mʌni];

5) other cases of elision of vowels, difficult to explain: *interest* ['ɪnrəst], *different* ['dɪfrənt], *political* ['plɪtɪk], *similar* ['sɪmlə], *fanatics* [f'nætɪks].

Most often elided consonants in English seem to be [t], [d], [k], [h], consider the following examples: *Voters will elec(t) city officials. Deduc(t) seven dollars from the hotel. He attemp(t)s to save money. Las(t) year. Mos(t) recent. Wes(t) German. Did (h)e get the promotion! Tell (h)er I'm busy. Put the desks [des:s] in the room. She asks [æ:s:s] too many questions*, see more in (Roach, 1990 : 108-109; Grant, 1993 : 49, 107).

5.2.4. Reduction (Редукция)

In English as well as in Russian vowels in unstressed syllables are usually reduced. Reduction is a historical process of weakening, shortening or disappearance of vowels in unstressed positions. Reduction is closely connected not only with word stress but also with rhythm and sentence stress. Stressed words are pronounced with greater energy or breath. The neutral sound [ə] appears most frequently in unstressed positions. This frequency of [ə] is the result of the rhythmic pattern. English rhythm prefers a pattern in which stressed syllables alternate with unstressed ones (e.g.: even in single words – *analyse* ['ænləlaɪz] and *analysis* [ə'nælɪsɪs]). Here a full vowel becomes [ə], and [ə] becomes a full vowel.

Since unstressed syllables between stressed ones are pronounced in the same period of time, no matter whether there are many or few, the length of the vowels in unstressed syllables depends on their number. So, when there are many unstressed syllables, they are pronounced quicker and their vowels are shorter than when there are few of them, (stress-timed rhythm in English), e.g.: a book / a nice book / a very nice English book/a very nice English coloured book, and etc.

In Russian there is a well-developed system of unstressed vowels. All Russian vowels are subjected to reduction, the vowels [o], [и] more often than others, e.g.: *молоко, малыши, конституция, воды, на воду*, and etc.

There are three types of reduction of vowels:

- 1) quantitative reduction – when the quantity of a vowel changes, e.g.: *I 'looked 'for it* – [fə'rit], and *I 'looked for the 'book* – [fə ðə 'buk];
- 2) qualitative reduction – when the quality of a vowel changes, cf.: *I 'must fly now* – [mʌst], and *I must 'fly now* – [məʃ];
- 3) zero reduction – consists in the omission of a vowel or a consonant, e.g.: *I 'can go* – [kn], *He has 'done it* – [əz], *I must 'go* – [məʃ].

A peculiar feature of reduction must be mentioned here. While, RP extensively employed /ɪ/ and /ʊ/ in unstressed syllables preceding consonants, in SSB, these reduced vowels are progressively substituted with the further step of qualitative reduction, schwa /ə/, cf. (Lindsey, 2019 : 39-40).

!

Weak Forms: CAN vs. CAN'T

Weak forms – the reduced pronunciation of common function words – have a crucial role in English phonetics, rhythm, and meaning, cf. (Lindsey, 2011).

The **core concept** of weak forms lies in the fact that weak forms are reduced pronunciations (often using the schwa /ə/) that make common words sound like affixes to neighbouring, stronger words. They are fundamental to the distinctive rhythm of native English speech and are used even in careful, formal contexts. Critically, they serve an important linguistic function by differentiating between words that are otherwise homographs (same spelling) or near-homographs.

Homographs are **differentiated** via weak forms. Key examples where a **strong form** signals one meaning and a **weak form** signals another include:

That	Demonstrative (strong: /ðæt/)	conjunction (weak: /ðət/, often disappearing entirely).
Have	Main verb (strong: /hæv/)	auxiliary verb (weak: /(h)əv/, often contracting further e.g., <i>I've</i>).
There	Demonstrative adverb (strong: /ðeə/)	existential pronoun (weak: /ðə/).
To / For	Prepositions (typically weak: /tə/, /fə/)	numerals/quantifiers <i>two</i> / <i>too</i> and <i>four</i> (always strong: /tu:/, /fɔ:/).

These pairs can appear adjacent to each other in speech (e.g., *to two, for four*), with the pronunciation difference maintaining clarity.

Typically **weak words** can be pronounced **strongly** in specific contexts: in **citation forms** (e.g., “the word *for*”), at the **end of a clause or before a pause** (e.g., “What did you do that **for**?”), when the word is **emphatically stressed**.

The importance of contrast: words without weak forms

Some words, like *off* and *on*, **never** have weak forms. This is primarily to maintain a clear auditory contrast with their weak-form homophones *of* (/əv/) and *and* (/ən/), preventing misunderstandings (e.g., *cut it **off*** vs. *cut it **of***).

There is a fundamental positive/negative contrast in English. The most vital communicative function of weak forms is distinguishing positive and negative statements. **Negative words** (*not*, *never*, *can't*, *isn't*, etc.) are **strong and stressed**. **Corresponding positive words** (auxiliary and modal verbs) are **typically weak or contracted** (e.g., /kən/, /wəz/, *I'm*, *he's*). The contrast is often marked more by the **strength of the negative** versus the **weakness of the positive** than by the barely audible negative /t/.

This **crucial contrast** of *can* vs. *can't* is a major source of misunderstanding for learners. The positive ‘**can**’ is almost always weak (/kən/ or /kɪn/) and can reduce to just /kn/ or even /k/. The negative ‘**can't**’ is strong (/kɑːnt/ in BrE, /kænt/ in AmE). The final /t/ is often a glottal stop or barely released, making the **full, stressed vowel** the primary cue for the negative.

Therefore, a conclusion can be made that natives rely on the pattern: **weak vowel = positive, strong vowel = negative**. The strong form /kæn/ in American English is most commonly heard in the negative, not the positive.

!

Weak Forms: Contractions

Contractions in English display various types and stylistic implications in speech and writing. Contractions are extremely common in native English conversation, accounting for over 5% of words in spoken corpora. The most frequent is ‘**it's**’, followed by ‘**that's**’, ‘**there's**’, and pronoun+auxiliary combinations like ‘**you're**’. A contraction typically involves a reduced function word (often an auxiliary verb) losing a syllable and attaching to a preceding word, indicated in writing by an apostrophe (e.g., *cannot* → *can't*).

Two main **types** of contractions can be distinguished, namely: **negative contractions**: formed by attaching a reduced ‘not’ (-n’t) to an auxiliary verb (e.g., *isn’t*, *can’t*, *won’t*). These are stressed and retain their own syllabic value; and **auxiliary contractions**: formed by attaching a reduced auxiliary verb (like ‘is’, ‘are’, ‘will’) to a preceding word, usually a pronoun or ‘there’ (e.g., *it’s*, *you’re*, *there’s*). These are usually unstressed.

Auxiliary contractions have a longer **history**, being prevalent even in Shakespeare’s work (e.g., ‘*all’s well*’), though some older forms like ‘*tis*’ have faded. Modern speakers often have a choice between two negative constructions (e.g., ‘*it isn’t*’ vs. ‘*it’s not*’), with usage governed by frequency and context: ‘*isn’t*’ and ‘*aren’t*’ are heavily used in questions, while ‘*haven’t*’ and ‘*won’t*’ are generally preferred over ‘*I’ve not*’ or ‘*I’ll not*’, cf. (Lindsey, 2011).

Contractions are typically **avoided** in two key scenarios when the auxiliary verb stands alone without its following predicate (e.g., “Yes, I **would**”, not “Yes, I’d”) or for deliberate emphasis, where the full, uncontracted form is used.

There are **stylistic considerations** concerning formality vs. “coldness”. The common belief that contractions are inherently informal is outdated. While they are less frequent in highly scripted, formal oratory (e.g., Churchill’s speeches), they are standard in contemporary public speaking, used even by figures like the monarch. Their primary avoidance is in deliberately impersonal or “cold” prose, such as in certain academic journals. Therefore, the stylistic question is not whether a context is casual *enough* for contractions, but whether it is formal and impersonal *enough* to prohibit them. This explains their relative scarcity in non-native English and their deliberate omission in the speech of fictional robots or aliens to sound artificial.

!

Weak Forms: by natives vs. non-natives

A few dozen of the most frequently used words in English – which together account for approximately twenty percent of our speech – are ordinarily not pronounced as their spelling would imply, but in a distinct, reduced manner. These Weak Forms play a significant role in why many learners do not sound native-like. Employing them is standard, not informal or slang, cf. (Lindsey, 2011).

English contains two distinct words spelt ‘that’. One is a demonstrative, pointing to something present, as in “I’d rather have that”, and it is always pronounced in its strong form. The other is a grammatical connector, which is generally weak: /ðət/.

These often-weakened words may sometimes be articulated in their strong form, for instance when emphasised, when they occur at the very end of an utterance, or frequently when they appear at the very beginning. However, within the flow of natural speech, they are typically weak. It is a phonetic reality that English maintains a remarkably pronounced distinction between its strongest and weakest words, a feature which substantially contributes to the language’s characteristic rhythm.

! English features online presenters and AI get wrong: Common Phonetic Features Overlooked in Online Presentation Styles

The distinctive speaking style found in many online platform presenters often arises from a combination of deliberate style and inadvertent error. While some patterns are conventional, others resemble mistakes also made by synthetic AI speech, compromising naturalness.

Two fundamental phonetic features of English that are frequently mishandled are **Weak Forms** and **Deaccenting**, cf. (Lindsey, 2011).

1) Weak forms and “Weak Form Strengthening”: weak forms are the reduced pronunciations of common function words (e.g., /ə/, *the*, *to*, *and*), making them sound like prefixes to stronger words. They are crucial for native rhythm and for distinguishing meaning (e.g., the strong vs. weak versions of *that*).

The common error for many presenters would be **Weak Form Strengthening** – using the strong form of words (e.g., “A article”) even where a weak form (e.g., “an article”) is standard in conversational speech. This can occur due to a perceived need for punchy, emphatic, and overly clear delivery, or frequent hesitation and pauses, which naturally trigger strong forms at utterance beginnings, creating a staccato rhythm.

2) Decaccenting is a rhythmic rule where “given” or repeated information loses its stress, shifting the accent to newer information (e.g., “My name is **Bond**. James **Bond**” – where the second “Bond” is deaccented).

The common error for presenters is to often incorrectly accent every word, including given information. This makes speech sound monotonous, unnatural, and can blur intended contrasts. Contributing factors include: reading from teleprompters without considering the spoken context; filming in disjointed segments, making it easy to forget which information has already been verbally established; a general principle that accenting all words equates to maximum clarity.

These features are notably challenging for both highly proficient second-language speakers and **AI text-to-speech systems**. Synthetic voices, such as those used by Google Translate or in some automated channels, consistently fail to correctly implement deaccenting and the appropriate stress in compound nouns, resulting in a robotic and repetitive cadence. Some AI voices even mimic human-like imperfections (e.g., intrusive ‘R’, audible breaths) while still lacking these core prosodic skills.

While the informal, innovative style of online platforms has normalised certain features like jump cuts, the widespread neglect of Weak Forms and Decaccenting by presenters – mirroring the limitations of current AI – often results in speech that is less natural, less rhythmically sophisticated, and more fatiguing to listen to than native conversational English.

Revision: Section 5 (Повторение: Раздел 5)

1. What is meant by modification of phonemes in speech?
2. What are allophonic variations?
3. What are the main causes of allophonic variation?
4. What is meant by positional and combinatory changes of phonemes?

5. What are the four specific types of positionally-combinatory changes?
6. What is assimilation?
7. What is accommodation?
8. What is elision?
9. What is reduction?
10. Why are phoneme modifications important for the study of speech?

6. The Syllable / Syllabic Structure in English

(Структура слога / слоговая структура в английском языке)

6.1. Syllable as a Phonological-Phonetic Unit / as a Phonetic

& Phonological Unity (Слог как фонолого-фонетическая единица / как фонетическое и фонологическое единство)

A syllable is a segment of a phonetic word, comprising sounds with varying degrees of sonority; it is a combination of a syllabic sound with one or more non-syllabic sounds (Беляева, 2019: 27).

The syllable is the minimal suprasegmental unit. From a physiological perspective of production, a syllable consists of one sound or several sounds produced by a single expiratory impulse (Современный..., 2012: 43).

6.2. Types of syllables (Типы слогов)

The pronunciation of a vowel letter in a stressed syllable depends on the syllable type, which is determined by the letters following the stressed vowel. Stressed syllables are conventionally categorised into four types, cf. (Аракин, 2017).

1) This is an **open syllable**, meaning it ends with a vowel. The stressed vowel here is pronounced as its name in the alphabet, typically representing a long vowel or a diphthong. There are two variations:

- an **absolutely open syllable**, where its written form matches its phonetic realisation (its sound), e.g., *me* [mi:], *no* [nəʊ];

- a **conditionally open syllable**, which is only graphically open but phonetically closed; words like *note* [nəʊt] and *Pete* [pi:t] historically derive from disyllabic words where the first syllable was open; in modern English, these are monosyllabic words, as the final letter /e/ is silent; it serves only as a conventional marker that the syllable is open and its vowel should be pronounced as in an absolutely open syllable.

2) This is a **closed syllable**, where one or more consonants (any except the letter /r/) follow the stressed vowel. The vowel letter in a closed syllable represents a short vowel sound. Consonant letters preceding the vowel usually do not affect its pronunciation. For instance, the letter /e/ in a stressed closed syllable represents the short vowel [e] regardless of how many consonants precede it. For example: *en* [en], *pen* [pen], *spend* [spend].

3) This syllable type is defined by the stressed vowel being followed by the letter /r/. The /r/ itself is not pronounced; it only indicates that the preceding vowel has a long pronunciation.

4) This type graphically resembles the first type (the conditionally open syllable). The key difference is that in a Type IV syllable, the letter standing between the stressed vowel and the silent /e/ is specifically the letter /r/.

6.3. The Structure of the English Syllable (Структура английского слога)

A syllable consists of a mandatory syllabic element, usually a vowel, or more rarely a syllabic consonant (Современный..., 2012: 43). This element is termed the **nucleus** or **peak** (e.g., the sound [a] in the word *срѣх*). The part of the syllable preceding the nucleus is the **onset** (e.g., [str] in *срѣх*); the part following the nucleus is the **coda** (e.g., [x]). The nucleus and coda together form the **rime** (e.g., [ax]). Syllables are classified as **covered** or **uncovered** based on the presence or absence of an onset, and as **closed** or **open** based on the presence or absence of a coda.

Some facts must be reminded of here, cf. (Section 3.7.2). The duration of a vowel sound in English is influenced by several key factors: 1) its inherent, or intrinsic, length; 2) its position within a word (e.g., *bee* – *bead* – *beat*); 3) the total number of syllables in the word, cf. (*'verse* and *uni'iversity*, or *The man laughed* and *The manager laughed*); and 4) the surrounding syllabic structure (such as V, CV, CCV, VC, CVC), as seen in the contrast between *dew* [dju:] and *duly* ['dju:li].

English contains many **checked vowels**, where the sound is curtailed and cannot glide towards a following element. This category includes all short vowels, all diphthongs, and long vowels when they occur before a voiceless consonant, for example: *bet*, *not*, *speak*, *type*, *coat*.

Conversely, **free** (or **unchecked**) **vowels** are those which are not phonetically shortened. This occurs when long vowels and diphthongs appear either in an open syllable or are followed by a voiced consonant, as in *sea*, *window*, *card*, and *bee*.

6.4. Syllable-Forming Sounds in English (Слогообразующие звуки в английском языке)

Different languages employ different phonemes as syllabic nuclei—in Russian, only vowels serve this function, whereas in English, both vowels and certain consonants in specific combinations can be syllabic (Евстифеева, 2022 : 67). In English, when one of the two consonants separating a stressed syllable from an unstressed one is a sonorant consonant (such as /l/, /r/, /m/, /n/), that consonant forms the onset of the second syllable. Consequently, the preceding consonant attaches to that second syllable as well, leaving the first, stressed syllable open. For example, in the word *ta|ble* ['teɪ.bəl], the consonant **b** joins the second syllable, which begins with the sonorant /l/, resulting in an open first syllable.

A weak vowel preceding the sounds [n], [l], [r] can be absorbed to form a syllabic consonant, too (V~weak~+[n]/[l]/[r]=C~syll~), which is demonstrated in the following examples: *tonight* [tənaɪt], *police* [pəli:s], *correct* [kɹɛkt].

6.5. The Rules of Syllable Division in English

(Правила выделения слога / слогоделения в английском языке)

In a disyllabic word, the vowel in the stressed syllable follows the same reading rules as a vowel in a monosyllabic word. To determine the syllable type, however, one must first establish the written boundary between the stressed and unstressed syllables. When determining this boundary in disyllabic words, the following principles apply, cf. (Аракин, 2017).

If a single consonant letter (excluding the letter **r**) follows the stressed vowel, it belongs to the subsequent unstressed syllable. For example: **stu** | **dent** ['stju:dənt]. Consequently, the stressed syllable is open, and the vowel within it is read according to **Type I**, long vowel sound (*exception: when the consonant **r** follows the stressed vowel in a disyllabic word, the vowel is read according to Type IV*. For example: **during** ['djʊərɪŋ], **Mary** ['mɛ:ri]).

It should be noted that there are exceptions in English where a stressed vowel in an open syllable is pronounced with a short sound, as in **city**, **pity**, **copy**, **very**, etc.

If the stressed vowel is separated from the following vowel by two or more consonant letters, including a double **r**, then the first of these consonants closes the stressed syllable. In this case, the vowel is read according to **Type II** (short vowel sound). For example: **din** | **ner** ['dɪnə], **mar** | **ry** ['mæri].

This rule does not apply if one of the two consonants separating the syllables is a sonorant consonant (like /l/, /r/, /m/, /n/). For example: **ta** | **ble** ['teɪbl]. Here, the consonant **b** attaches to the second syllable, which begins with the sonorant /l/, leaving the first stressed syllable open.

6.6. Syllable Formation & Syllable Division

(Теории слогообразования и слогоделения)

The following can be said regarding **syllable formation and syllable division**, cf. (Современный..., 2012; Беляева, 2019).

Core principle: the sonority theory

The predominant framework for understanding the syllable (especially in the Russian linguistics) is the **Sonority Theory**, notably developed by R. I. Avanesov and M. V. Panov, which can be summarized in the following points.

A syllable is conceptualised as a wave or peak of sonority. **Sonority** is defined as the relative loudness of a sound, inversely related to the degree of constriction in the vocal tract. Thus, the sonority scale, from least to most sonorous, is typically as follows: **plosive (stop) consonants** → **fricatives** → **nasals / liquids (sonorants)** → **vowels**.

The **vowel**, as the most sonorous segment, invariably acts as the **syllabic nucleus (peak)**. The number of vowels equals the number of syllables. Sonorant consonants (**р, л, м, н**) can sometimes form a syllabic nucleus in certain contexts, due to their relatively high sonority.

Several universal **principles of syllable formation** that govern the preferred syllable structure are discriminated, though they are not absolute rules (Соколова, 2010 : 53; Евстифеева, 2022 : 66-68).

- **Principle of open syllables:** there is a preference for **open syllables** (ending in a vowel) over closed ones in non-final positions.

- **Principle of rising sonority:** sounds within a syllable ideally arrange themselves from lower to higher sonority towards the nucleus (e.g., *май-ка*). This principle is frequently **violated** in some languages, like modern Russian (e.g., *второй, с лон-дон-цем*), complicating syllable division.

- **Principle of sonority dispersion:** the rise in sonority at the syllable onset should be as sharp as possible (e.g., *tra* is preferred, as /t/ and /r/ are far apart on the sonority scale), while the fall in sonority from the nucleus to the coda should be gradual.

Two main **theories of syllable division** are distinguished. First, the **sonority theory** which provides specific rules for dividing words into syllables in Russian (Соколова, 2010 : 53-54; Евстифеева, 2022 : 68-70):

- **Open syllable rule:** All non-final syllables are **open** (end in a vowel), **except** when an intervocallic consonant cluster follows the pattern “**sonorant + obstruent**” (e.g., *на-ма*). Here, the sonorant closes the preceding syllable.

- **Rising sonority rule:** A non-initial syllable must begin with the least sonorous segment possible, adhering to the principle of rising sonority from its start.

Second, the **implosion/explosion theory** of syllable division, based on an articulatory criterion, defines a syllable as a sound sequence produced by a single **expiratory impulse**. Each sound has an implosive (closing) and explosive (opening) phase. The **syllable boundary** is placed where an implosive phase is followed by an explosive phase (i.e., where closure gives way to opening).

Concerning the **types of syllable division**, it is essential to differentiate between three separate processes, which often yield different results (Беляева, 2019 : 28-29):

- **Phonetic syllable division:** governed by the sonority principles above (e.g., *за-пу-беж-ный*).

- **Orthographic division (hyphenation):** follows rules for word splitting at the end of a written line, often respecting morpheme boundaries (e.g., *за-пу-беж-ный*).

- **Morphemic division:** the segmentation of a word into its constituent meaningful parts (morphemes) (e.g., *за-пу-беж-н-ый*).

In essence, the Russian is primarily analysed as a **sonority peak**. Its formation tends to follow **universal preferences** for open structures and rising sonority, though these are often breached. The **Sonority Theory** provides a language-specific set of rules for syllable division, prioritising open syllables and phonotactic constraints. A critical analytical point is the clear separation of **phonological syllabification** from **orthographic hyphenation** and **morphological analysis**.

6.7. The Relationship Between Phonetic & Orthographic Syllables (Соотношение фонетических и орфографических слогов)

A **phonetic syllable** is a unit of **speech sound organisation**. It is defined by **actual pronunciation** and the **sonority wave** of spoken language, as discussed in previously (Евстифеева, 2022 : 71-72). The **spoken syllable** is governed by **phonetic principles** and directly observable in acoustic analysis. An **orthographic syllable**, on the other hand, is a unit of **writing and spelling conventions**, cf. (Сафронова, 2023). It is defined by the **rules of orthography** in a given language and is used primarily for **hyphenation** (word division at the end of a line). Its boundaries are determined not by pronunciation alone but by the interplay of the main spelling principles: the **morpho-phonemic** and the **phonetic** one. The former is dominant and reflects the underlying **phonemic** composition of a **morpheme** (the smallest meaningful unit), not its variable phonetic realization, and the latter dictates that spelling should sometimes directly match the **actual pronunciation**, which can be in contradiction with the morpho-phonemic principle.

Revision: Section 6 (Повторение: Раздел 6)

1. What is a syllable and what is its role in speech?
2. What approaches to syllable definition are distinguished?
3. What is the articulatory (chest pulse) theory of the syllable?
4. What is the sonority theory of the syllable?
5. What are the key phonetic features of a sentence that rely on syllable organization?
6. What types of syllables are distinguished according to their structure?
7. What is syllable division and what are its principles?
8. What sounds are syllabic in English?
9. How is a syllabic consonant formed in English? Provide the formula and an example.
10. What does the articulatory indivisibility of a diphthong imply for syllable or morpheme boundaries?

7. The Accent Structure of English Words (Акцентная структура английских слов)

7.1. The Notion of Stress (Понятие ударения)

To stress a syllable is to strike it, highlight it, or give it emphasis. Specifically, it's the prominence given to one or more syllables in a word through greater force or duration. But this applies mainly to languages like Russian, which have dynamic or force-based stress. In Russian, the stressed vowel in a word—take “вода” (water), for example—is both the longest and the loudest. However, there are other types of languages where stress, as we understand it, doesn't really exist. Instead, they use pitch accent—variation in musical tone (Соколова, 2010 : 58). Mandarin Chinese, for instance, isn't concerned with stress in the same way; what matters is the tone movement: a falling “pà”, a rising “pá”, or a high level “pā”. For them, the crucial element is the tonal component. Force-based emphasis takes a back seat. So, yes—striking, highlighting, making stronger.

7.2. Word Stress (Словесное ударение)

Stress is a greater prominence given to one or more words in a sentence as compared with the other words of the same sentence. In English nouns, adverbs, verbs, semantically important prepositions are stressed. An important role for intonation plays a logical stress, when the communicative centre of the sentence is shifted to any word in the same sentence, e.g.: *'Can you 'break an 'apple in /two?* (stress each word in this sentence with logical stress – *you, break, two.*).

!

Stress: Compound Nouns (Exemplified by Superhero Names)

The pronunciation of superhero names varies according to stress patterns, cf. (Lindsey, 2019 : 79-81). Some place emphasis on the first element, such as **WONDER** Woman, while others stress the second, like Invisible **WOMAN**. The classic superhero name is comprised of two words, a structure which accounts for approximately 70 percent of the top 100 comic book characters, cf. (Lindsey, 2011).

To determine the correct stress, the key question is: **What grammatical role does the first word play?** Is it a **noun** (a thing), or is it an **adjective** (a descriptive quality)? If the first element is a **noun**, primary stress falls on the **first word**: e.g., **BATH**room, **DOG** food, **CELL** phone, **GEOGRAPHY** professor. If the first element is an **adjective**, primary stress falls on the **second word**: e.g., small **ROOM**, great **FOOD**, mobile **PHONE**, nutty **PROFESSOR**.

Applying this rule: *ghost* (noun) → **GHOST** Rider; *savage* (adjective) → Savage **DRAGON**.

The first group are **compound nouns**, while the second are **noun phrases**. The distinction involves more than stress; compounds tend to be more metaphorical, whereas noun phrases are more literal. Many superheroes use colours (e.g., Green **LANTERN**, Scarlet **WITCH**), but compounds are very common (e.g., **SPIDER**-Man, **CAT**woman).

Titles and full names follow different rules. With formal titles or honorifics, the **name** is stressed: Doctor **LINDSEY**, Doctor **STRANGE**. Full names also take stress on the second element: Nick **FURY**, Luke **CAGE**.

Puns or plays on words can create exceptions. *Harley QUINN* is stressed as a full name, though it deliberately echoes the compound **HARLE**quin, leading some to use the first-word stress.

Some borderline cases exist where a noun functions more like an adjective, leading to second-word stress (e.g., spring **BREAK**, child **PRODIGY**, winter **SOLDIER**). Dialectal variation occurs: in British English, *boy SCOUT* parallels *child PRODIGY*, whereas American English treats it as the compound **BOY** scout (like **BAT**man).

Occasionally, usage is variable even for a single speaker, though standard compounds often prevail (e.g., **ORANGE** juice, **IRON** Man, not *orange JUICE* or *iron MAN*). Even adjectives can become part of a compound when their meaning becomes non-literal (e.g., a *green HOUSE* vs. a **GREEN**house). Pronouns and verbs can also form compounds, following the first-word stress pattern: **SHE**-hulk, **WATCH**men, **DARE**devil.

! Stress: Non-English Names

The pronunciation of ‘Kamala’ – a Sanskrit name meaning “lotus” – illustrates the challenge English faces with unfamiliar borrowings. As the name is not yet common in the English-speaking world, speakers typically adopt one of two strategies, cf. (Lindsey, 2011):

1) Assimilation (nativisation): adapting the word to fit English phonetic patterns. This yields two common anglicised versions: **Option A:** /ˈkæmələ/ (stressed on the first syllable, rhyming with *camel*), following models such as *Pamela* or *Amazon*; **Option B:** /kəˈmə:lə/ (stressed on the second syllable, rhyming with *koala*), following patterns like *impala* or *masala*, cf. (British English *gala*).

2) Non-assimilation: attempting to approximate the source language. This has become the dominant approach in **American English**, guided by a general rule that a stressed ‘a’ in foreign words takes the /ɑ:/ vowel, as in *father* or *palm*. This results in the pronunciation /ˈkɑ:mələ/, which is now the majority usage in the United States and is employed by the Vice President herself, cf. (*pasta*, *taco*, *Han Solo*).

A key phonetic difference between the dialects is **vowel length**. British English maintains a phonemic distinction between long and short vowels (e.g., the long /ɑ:/ in *calm* versus the short /ɒ/ in *comma*). American English has largely lost this distinction, merging the vowels in pairs like *father* and *bother*. Consequently, the American /ɑ:/ is a more readily available and functionally different sound for such borrowings.

For most British English speakers, accurately reproducing the original short vowel of ‘Kamala’ – which falls phonetically between the sounds in *calm*, *camel*, and *comma* – is difficult. Using the British long /ɑ:/ feels unnatural for a short vowel, while the short /ɒ/ (as in *comma*) is both too short and has an inappropriately rounded quality. It is generally unrealistic to expect speakers to

consistently use a sound outside their native phonetic inventory for a single name, just as it would be to expect Americans to adopt the British short 'o' in "James Bond". Therefore, British speakers are more likely to use one of the assimilated forms, whereas the non-assimilated American approximation has become standard in the United States.

!

Stress: "Words" in Film Music

The rhythmic patterns of language can be discerned within music, particularly in theme tunes, due to established conventions governing how lyrics and melody are aligned. The primary principle governing this alignment is **stress**: stressed syllables within words are typically matched to accented beats in the music, cf. (Lindsey, 2011). Musical metre is generally regular, built upon a repeating pulse, while the stress patterns of English words are documented in dictionaries for polysyllabic terms. With single-syllable words, content words are usually stressed, whereas most function words are not.

Consequently, a specific musical phrase will naturally suit the rhythm of certain words but not others. This interplay extends to **theme music**. Even in instrumental pieces, a melody is often composed to fit the rhythm of a title. This has been a frequent practice in television and film themes over the years. Consider the classic sitcom *Bewitched*. Its theme employs a two-note motif where the second note aligns with the first beat. Historically, during the 1960s, it was common for film producers to desire a promotional song sharing the film's title. Composers would therefore often ensure the main tune could accommodate the title's lyrics.

However, directly singing the title words within the theme – a practice increasingly viewed since the 1970s as naïve, corny, or deliberately comical – can also rely on the humour of a perceived coincidental match. The most striking examples involve longer titles. While two or three-note motifs are musically commonplace, a tune that perfectly matches a lengthy title seems to surpass mere coincidence, as setting many syllables to music is less straightforward. Notably, one composer successfully embedded a nine-syllable title into one of his most famous film themes.

7.3. The Linguistic Nature of Word Stress

(Лингвистическая природа словесного ударения)

Stress has a linguistic role in any language. The importance of stress in languages can be well illustrated by some peculiarities of the Russian language. A defining feature of Russian phonetics is the behaviour of vowels in unstressed syllables, a process known as vowel reduction, cf. (Беляева, 2019 : 65). In the standard literary language, a **qualitative reduction** occurs where distinct unstressed vowels (like **а, о, э**) in the first pre-stressed syllable merge into a single sound, typically an **и**-like vowel after palatalised consonants, a system called "ikanie" (Rus. *икание*), e.g., /питух/ for *нетух*. This merger creates orthographic challenges, as words like *леса* 'forests' and *луса* 'fox' can become homophonous.

Crucially, the literary standard represents just one simple system within a vast dialectal spectrum. Russian dialects exhibit significantly more complex and varied patterns of unstressed vowel pronunciation, including:

- **“yekanie” (еканье):** an older Moscow norm where vowels reduce to an e-like sound.
- **full vowel distinction:** found in some Northern Russian dialects, where all vowels remain distinct in unstressed positions, eliminating the homophony of the literary standard.
- **moderate “yakanie” (яканье):** in some Central Russian dialects, unstressed **a** is pronounced only before non-palatalised consonants; it becomes an **и**-like sound between palatalised consonants.
- **dissimilative yakanie / akanie:** a highly complex system prevalent in Southern dialects (e.g., Lipetsk, Voronezh). Here, the quality of the unstressed vowel depends not only on neighbouring consonants but also on the identity of the stressed vowel in the word, following a principle of dissimilation (making sounds different).

This multi-variant system, with around 20 identifiable types of pre-stress vocalism, is unique to Russian dialects. Furthermore, despite expected linguistic change over time, dialectological expeditions show these intricate vowel reduction patterns have remained remarkably stable and well-preserved for over a century.

7.4. Degrees / Types of Word Stress: Dynamic, Musical, Quantitative & Qualitative (Степени / типы словесного ударения: динамическое, музыкальное, количественное и качественное)

The following types of stress or accentual systems in languages are described (Соколова, 2010 : 59; Евстифеева, 2022 : 74-75).

Dynamic (or force) stress: this is the type found in languages like Russian. The stressed syllable is made prominent through increased loudness (intensity) and greater duration. It serves an organising function, acting as an anchor (“hook”) for the phonetic word. It can also have a lexical (meaning-distinguishing) function, differentiating words like *мУка* (torment) and *мука* (flour).

Pitch accent (tonal prominence or ‘musical stress’): this type is found in languages like Mandarin Chinese. Here, the primary feature is variation in musical tone (e.g., rising, falling, level pitch) and the force or duration is secondary. Thus, the pitch contour on a syllable can change the meaning of a word (e.g., *pā*, *pá*, *pà*).

With consideration of **positional distinction**, there is a further classification relevant to dynamic stress languages. The stress can be **fixed**, when it consistently falls on a particular syllable (e.g., always the first, always the last), or it can be **free (or mobile)**, where stress can fall on different syllables in different words, as in Russian, which is what gives it its meaning-distinguishing role.

7.5. Factors that Determine the Place / Position & Degree / Strength of Word Stress in an English Word (Факторы, определяющие место / положение и степень / силу словесного ударения в английском слове)

Stress in disyllabic and polysyllabic words

In two-syllable (disyllabic) English words, the stress typically falls on the **first syllable**, provided it is not a prefix that has lost its semantic meaning cf. (Аракин, 2017). For example: *dinner* ['dɪnə], *office* ['ɒfɪs]. In words containing a prefix, the stress usually falls on the **second syllable**. For example: *return* [rɪ'tɜ:n], *begin* [bɪ'ɡɪn].

In three-syllable and polysyllabic words, the stressed syllable is most often the **third syllable from the end**. The vowel in this syllable is pronounced according to **Type II** (with a short sound), regardless of the number of consonants separating it from the following unstressed syllable—even if that consonant is the letter /r/. For example: *factory* ['fæktəri], *family* ['fæmɪli], *experiment* [ɪk'sperɪmənt].

However, when the stressed third-from-end syllable contains the sound [ju:], it is read as **Type I** (an open syllable). For example: *funeral* ['fju:nərəl].

When determining the stressed syllable, grammatical endings (i.e. inflection) and productive suffixes are **not taken into account**. For example: *demonstrate* ['demənstreɪt], *demonstrated* ['demənstreɪtɪd], *demonstrating* ['demənstreɪtɪŋ].

When a noun is formed by adding the suffix **-ion** to a verb, the primary stress shifts to the syllable **immediately preceding the suffix**. The original verbal stress is retained as a **secondary stress**. For example: *demonstration* [ˌdemən'streɪʃn]. A secondary stress is indicated in transcription by a low stroke [ˌ]. Some words carry two primary stresses. This applies to the numerals *thirteen* to *nineteen*—e.g., *thirteen* ['θɜ:ti:n]—as well as to words with prefixes that impart a new meaning, such as *disappear* ['dɪsə'piə] and *rewrite* ['ri:'raɪt]. The so-called **tertiary stress** is also distinguished in English. Thus, English distinguishes b

!

Stress-Shift: Stress, Pitch, Accent, and Rhythm

The concepts of word stress, pitch-accent, and rhythmic timing in English are interconnected, cf. (Lindsey, 2019 : 79-83). **Word Stress** is the pattern of prominent syllables within a word, as indicated in dictionaries (e.g., *Piccadilly* has primary stress on 'dil-' and secondary on 'Pic-'). **Pitch-Accent** is a noticeable pitch movement attached to a stressed syllable within a spoken phrase. The most important is the **nuclear (tonic) pitch-accent**, which by default falls on the last content word in a phrase. **Rhythmic Beats** stand for the timing of prominent syllables in the flow of speech, which tends towards regular spacing.

The phenomenon of “Stress Shift” (iambic reversal): when words are combined, their dictionary stress patterns often adapt to avoid a “stress clash” (two strongly stressed syllables occurring side-by-side). This is frequently perceived as the primary stress moving to an earlier syllable, e.g., *thirTEEN* becomes *THIRteen MEN*; *PiccaDILly* becomes *PICcadilly Circus*. This

“shifted” pattern can become fixed for common phrases (e.g., *COvert OPs*, *Piccadilly Circus*) or even for single words derived from common usage (e.g., *HEATHrow Airport* leading many to say *HEATHrow* even alone).

The role of intonation and rhythm: the perceived shift is less about changing lexical stress and more about the application of **intonation** (pitch) and **rhythm** (timing) in phrases. In this context, the following phenomena must be defined. **Nuclear pitch-accent:** in a phrase, only the final content word typically receives the strong, falling nuclear pitch-accent heard in dictionary recordings. Earlier stressed syllables may receive a simpler **onset pitch-accent** or no pitch prominence at all. **Rhythmic beat skipping:** English rhythm commonly skips a rhythmic beat on a middle stressed syllable in a sequence of three. For example, in “Little Jack Horner”, the stress on “Jack” often loses its beat, squeezing “Little Jack” into one rhythmic unit. **Combined effect:** the tendency for phrases to have just two pitch-accents (onset + nucleus) and to skip middle rhythmic beats reinforces the auditory impression of stress shifting in connected speech.

Variability and Dictionary Inconsistencies must also be reported. **Speaker/Accent Differences** are somehow interrelated. Processes like stress shift are not universal, for instance, while many say *HEATHrow*, some British speakers retain *HEATH ROW* when the word is isolated. There are **dictionary limitations:** dictionaries are often inconsistent or outdated in marking stress, particularly for words with variable patterns (e.g., *thirteen*, *overt*). They typically only show the pronunciation when the word carries a nuclear pitch-accent in isolation, which is not representative of its sound in fluent speech. **Pretonic syllables**, i.e. syllables immediately before the primary stress exhibit variability, with some speakers allowing a pitch-accent on them (e.g., *EN-tire COUNtry*) and others not (e.g., *en-TIRE COUNtry*).

In essence, the fluent rhythm of English is created by a dynamic interplay where dictionary stresses may or may not be realised with pitch prominence or rhythmic beats, leading to the flexible and context-dependent patterns of spoken language.

7.6. The Functions of Word Stress in Language (Функции словесного ударения в языке)

In different languages, stress serves different roles. In Russian, for example, it has a meaning-distinguishing (distinctive) function, as well as what we call an organising function (Беляева, 2019 : 23-26). It helps you understand whether you mean “flour” (мука) or “torment” (мука), or whether you’re saying things are “bad” (плохо) or that you want to “bake some bread” (испекУ хлеба). Historically, it also gives phonetic structure to a word. In Russian, the stressed syllable acts like a hook on a coat-stand—the whole phonetic word hangs from it. In standard Russian, the stressed syllable is the strongest; the syllable immediately before it is about two-thirds as loud or long, and the others are about one-third. This dynamic structure—known as Potebnya’s formula, after the remarkable linguist Afanasy Potebnya—assigns an index of 3 to the stressed syllable for loudness/duration, 2 to the first pre-stressed syllable, and 1 to the others. Stress helps us organise speech (constitutive function), to understand where a word begins and ends phonetically in the flow of conversation (recognitive function). When I look at

you speaking, I don't see individual words dropping out; I have to segment this continuous sound stream somehow. And you'll notice, if you listen to a language you don't know at all, it's very hard to tell where one word ends and another begins. You might perceive phrases bound together by intonation, but the boundaries within them are blurred.

So, dividing speech into words is a tricky business, and stress can play very specific—and emotionally significant—roles in different languages. Stress holds a special place in language.

Revision: Section 7 (Повторение: Раздел 7)

1. What is word stress?
2. What types of word stress are distinguished according to position?
3. What functions does word stress perform?
4. How does word stress distinguish meaning?
5. According to the analysis of compound nouns versus noun phrases, what is the general rule for determining the primary stress based on the grammatical class of the first element?
6. Under what circumstances can the original primary stress be retained as the secondary stress?
7. What is the main difference between the dynamic and the musical stress in languages?
8. When determining the stressed syllable, what is not taken into account?
9. What type of vowel reduction occurs in the pre-stressed syllables in the literary Russian language?
10. What is common behaviour of the English stress in disyllabic and polysyllabic words?

8. Intonation & Prosody (Интонация и просодия)

8.1. Approaches to Defining Intonation & Its Components (Подходы к определению интонации и ее компонентов)

The most important speech unit, complete and independent enough to function as a unit of communication is the sentence. It can perform its function not only because it consists of words that are made up of definite sounds, have definite meanings and follow each other according to the rules of the language, but also because it has definite phonetic features without which the sentence does not exist. These features are closely connected with the meaning of the sentence as a whole and carry important information that the words of the sentence do not express. These features are: the pitch (or speech melody), the rhythm, the stress, the loudness, the tempo and the timbre. Each feature performs a definite task and all of them work simultaneously in a sentence.

Intonation is present in every sentence and sometimes it is more important how we say than what we say. S. Allen (an English phonetician) says that “a reasonably correct speech flow is more important for intelligibility than correct sounds”. Prof. D. Jones says that intonation may be defined as the variations which take place in the pitch of the voice in connected speech. O'Connor says that the “use of pitch to distinguish whole utterances without changing the shape of the component words in an utterance” may be defined as intonation, cf.: *You 'know Mr. \Jones* and *You 'know Mr. \Jones?; \No. 'Twenty \five* and *\No. 'Twenty \five*. Here the component words are the same but the total meaning because of intonation is different.

!

English Intonation: Patterns

Intonation in speech is produced by a combination of varying vocal pitch, stress, and rhythm, which together convey meaning (see also (Lindsey, 2019 : 97-98)). The principal intonation patterns are: falling, rising and falling-rising.

Statements in English most naturally end with a **falling intonation**. This differs from some other languages, which may use a rising pattern in similar contexts. **Open (WH-) questions** begin with a question word (e.g., *what, where, why, how*) and seek information. They also typically end with a **falling intonation**. **Closed (Yes/No) questions** are grammatically distinct, these questions invite a ‘yes’ or ‘no’ response and are characterised by a **rising intonation** at the end. **Tag questions** are formed by adding a short question tag after a comma (e.g., “It’s cold, isn’t it?”). The intonation here depends on context: a **falling intonation** on the tag indicates the speaker is fairly certain and is not genuinely asking for an answer; a **rising intonation** signals a genuine question where the speaker is seeking confirmation.

Deaccenting refers to the pronunciation of a word, particularly at the end of a phrase, without any notable stress or pitch prominence, cf. (Lindsey, 2011). In compound nouns, it is typically the second element that loses its accent. Thus, one says **AIR**port, not air**PORT**, and **FACE**book, not Face**BOOK**. Producing the final word in a phrase without accentuation can be particularly challenging.

A crucial function of deaccenting in English is to signal **given information** – that is, an idea already present in the discourse, such as a recently mentioned word. A classic example is James Bond's introduction: he first says "Bond" (accented), and then in "James Bond", the repeated surname is deaccented because it is now given information.

This pattern does not come naturally to speakers of many other languages. Another famous illustration is the line "To be or not to be", where the repetition is consistently deaccented: "to be or **NOT** to be".

Exceptions exist, particularly with expressions ending in numbers, where the main accent usually falls on the numeral (e.g., **MP3**, hi-**FIVE**). However, if the final number is itself a repetition and therefore given information, it is deaccented, shifting the accent to the preceding element.

This principle also applies to **football scores**. Normally, each team's score is accented (e.g., "Arsenal **1**, Liverpool **2**"). However, in the case of a draw, the repeated score is given information and is deaccented (e.g., "Arsenal **1**, **LIVERPOOL** 1").

8.2. The Essence & Bases of Intonation Structure: the Components of the English Intonation, their Functions & the Prosodic Characteristics of Speech (Сущность и основы интонационной структуры: компоненты английской интонации, их функции и просодические характеристики речи)

Intonation, a fundamental aspect of spoken language, refers to the systematic use of vocal pitch, rhythm, and other prosodic features to convey structure, meaning, and attitude beyond the words themselves. It is a crucial element for effective communication, serving to organise speech into meaningful units, distinguish sentence types (e.g., statements from questions), and express emotions or speaker intent, cf. (Lindsey, 1981). The primary components of English intonation, which work in concert to create its characteristic patterns, include **speech melody (pitch)**, **rhythm**, **phrase stress**, **loudness**, **tempo**, **timbre**, **pausation**, and **duration** (Соколова, 2010 : 69; Евстифеева, 2022 : 88). Each component contributes specific functional and perceptual qualities to the overall prosodic contour of an utterance, guiding listener interpretation and ensuring linguistic coherence.

8.2.1. Speech Melody (Речевая мелодика)

The pitch (or speech melody) of the voice does not stay on the same level while we pronounce the sentence. It fluctuates, falls and rises on vowels and voiced consonants. So when

we speak normally, the pitch of our voice is constantly changing. These falls and rises are not chaotic, but form definite patterns typical of English. And for communicative purposes and linguistic significance these falls and rises must be under speaker's control (compare a horse-rider galloping = sudden rises & falls = are not linguistically important). There are two important parameters that can modify pitch contour ['kɒntə]: the pitch level and the pitch range.

The pitch level of a sense-group is determined by the pitch of its first- stressed syllable. Variations in voice pitch or melody occur within the interval between its upper and lower limits.

Any sense-group may be pronounced on 3 different pitch levels, cf.: *'Where did you \go?*

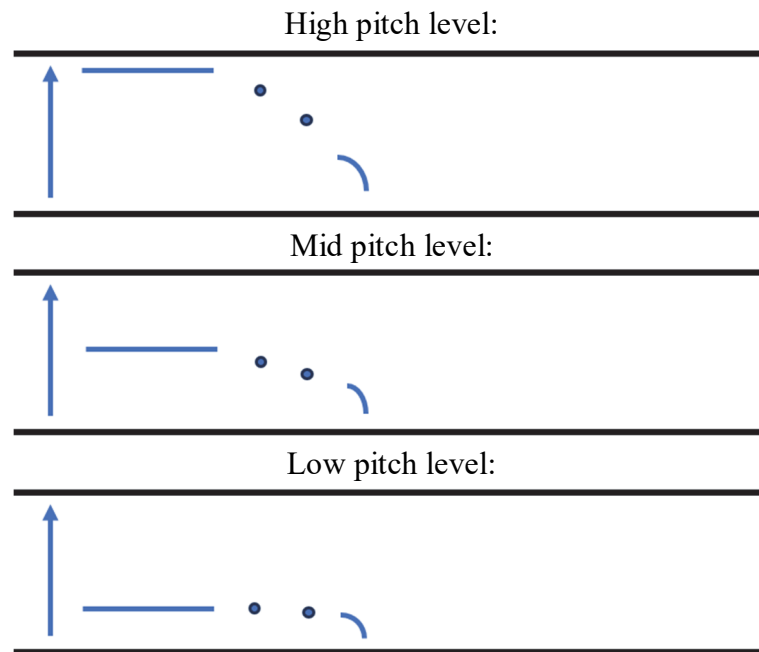


Fig. 11. A depiction for different pitch levels of a sense-group

The higher is the pitch of the first stressed syllable, the higher is the general pitch of the whole sense-group,

The pitch range of a sense-group is the interval between the highest-pitched and the lowest-pitched syllables, e.g.: *At 'this /game § he was a 'perfect \master.*



Fig. 12. A depiction of two sense-groups

Here the range is between the word *this* and the beginning of the syllable *game* in the 1st sense-group and the interval between the 1st stressed syllable in the word *perfect* and the last stressed syllable in the word *master*.

Pitch ranges may be 1) normal; 2) wide; 3) narrow:

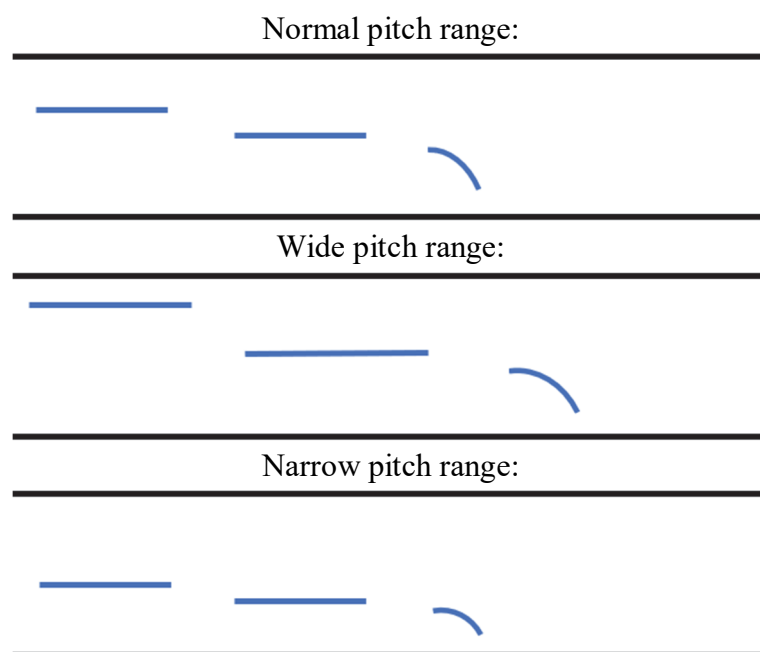


Fig. 13. A depiction of a sense-group with different pitch ranges

The pitch level (the first stressed syllable) and the pitch range (the interval between the highest-pitched and the lowest-pitched syllables) are interdependent – the higher the pitch level, the wider the range (see more in: (Vassilyev, 1962: 135-138)).

The pitch is the most important feature of intonation. Melody delimits sentences into sense-groups. It is present both in spoken and written languages; there are special punctuation characters to mark it within or at the end of a sentence (“; ;:”; “. ! ?”).

8.2.2. Rhythm (Ритм)

Rhythm is a very important component of intonation and may be defined as regular alternation of stressed and unstressed syllables. The ability to move smoothly from one stress to the next and to fit in the unstressed syllables between them forms the basis of a good English accent. English has a stress-timed rhythm. The stress-timed rhythm theory states that the times from each stressed syllable to the next tend to be the same. All unstressed syllables between the stressed ones will be compressed as far as possible to allow the next stresses syllable to come on a regular beat, e.g.: *The 'dogs eat the \bones, The 'dogs'd eaten the \bones, The 'dogs that've eaten the \bones* and *The 'dogs will have eaten the \bones*, see more in (Cook, 2017 : 36-37).

Russian also has stress-timed rhythm, but in Russian all words in a sense- group are stresses (e.g.: *'Наши сту'денты по'ехали на каникулы*). In English there is a different rhythmic pattern (structure): all form words (particles, prepositions, auxiliary verbs, articles) are fitted in between the stressed ones. The incorrect use of rhythm betrays the non-English origin of the speaker even in cases of correct pronunciation.

!

Rhythm & Tempo

Based on the analysis of sperm whale vocalisations, the following key ideas are essential for understanding the role and complexity of **tempo in human speech**, cf. (Lindsey, 2011).

Discrete tempo categories: tempo is not a continuum but can be organised into distinct, categorical speeds (like the whales' five discrete tempos). In human speech, this translates to perceptually distinct rates (e.g., slow, conversational, rapid) that carry functional or stylistic meaning.

Tempo as a combinatorial feature: tempo operates independently and combines with other elements like **rhythm** (stress patterns) to multiply communicative possibilities. Just as whales combine 18 rhythms with 5 tempos to create dozens of codas, humans combine specific speech rates with rhythmic and intonational patterns to create nuanced utterances.

Rubato: expressive tempo modulation: speech tempo is dynamically elastic. The concept of **rubato** – the smooth, gradual, and often coordinated speeding up and slowing down – is a direct analogue to expressive tempo variation in human speech, used for emphasis, phrasing, and conveying attitude or emotion.

Tempo and structural boundaries: changes in tempo often signal structural boundaries in an utterance, similar to “**ornamentation**” or extra clicks marking the edges of whale coda sequences. In human speech, slowing down (deceleration) or pausing frequently marks the end of phrases, clauses, or sentences, functioning like auditory punctuation.

Tempo as a building block of meaning: it is not merely a side effect but a fundamental **building block** of the communicative system. Alongside rhythm & pitch, systematic variation in tempo adds to the combinatorial system that creates meaningful units (words / phrases) and larger structures (sentences / discourse).

Analogy to music: the comparison to music is insightful. Tempo in speech, as in music, is a primary carrier of expressive and structural meaning within a combinatorial system, even if it lacks the precise referential meaning (duality of patterning) found in language. It helps organize the flow of information and convey non-propositional content.

In summary, tempo in human speech is a **structured, categorical, and dynamic feature** that combines with rhythm, functions as expressive rubato, marks linguistic boundaries, and serves as a core component in the combinatorial system that generates meaningful communication.

!

Rhythm: Pink Panther

English distinguishes between stressed **content words** (meaningful, abundant) and unstressed **function words** (grammatical, few). In speech, function words often become **weak forms**: reduced to a schwa (/ə/) or a barely-there prefix that must attach to a following stressed word (e.g., *today* /tə'deɪ/, *to do* /tə'du:/), cf. (Lindsey, 2011). This process, called **cliticisation**, is natural, not lazy.

A common marker of non-native speech is pronouncing these as full words (e.g., saying /ɒv kɔ:s/ rather than the native /əv'kɔ:s/). The rhythm of weak forms mirrors the **Pink Panther theme tune**, where unstressed notes lead into stressed beats. This pattern – a weak syllable followed by a strong one – creates **iambic pentameter**, the classic metre of English poetry. Thus, using weak forms aligns with the language's inherent rhythm.

8.2.3. Phrasal stress (Фразовое ударение)

Phrasal stress is a component of intonation, specifically using **intensity** (loudness), to highlight the most semantically important word within an utterance or phrase. Its primary function is the **phonetic organisation** of a syntagm, i.e. a meaningful word group, cf. (Современный..., 2012 : 53; Беляева, 2019 : 24). It is the prosodic prominence given to a key word within a phrase. Its default position is normally final (syntagmatic stress), but it can move to any word to signal logical emphasis, at which point it fulfils a dual role as logical stress. Phonetic patterns (stress placement) are used to express logical and informational structure within a text.

8.2.4. Loudness (Громкость)

Loudness is also an important component of intonation. Each syllable bears a definite amount of loudness. Pitch movement are inseparably connected with loudness. When people are excited they use the upper part of their pitch range much more often than they do normally. This is usually accompanied by an increase in loudness and tempo. When depressed they use mostly the lower part of their pitch range and talk slowly and quietly.

8.2.5. Tempo (Темп)

Tempo is the rate of speech and pausation. Tempo suits the semantic importance of the sentence and is connected with the rhythm. The rate may be slow, normal, and fast. Important parts of the utterance sound slower, unimportant parts are pronounced at a greater speed than normal, e.g.: *'When I 'went /there – and it was back in 1942. I.* (is unimportant and is marked by reduced loudness and increased speed).

8.2.6. Timbre (Тембр)

Timbre depends on the physiological properties and physiological peculiarities of the speaker. It is little studied.

!

Timbre: Vocal Fry

The distinctive crackle of vocal fry, or creaky voice, arises from an unusually slow and irregular vibration of the vocal folds, producing a pulsing, rattling sound, cf. (Lindsey, 2019 : 95-96). Far from being a modern anomaly, this voice quality has a rich history. It was famously employed in mid-20th century British Received Pronunciation by actors like George Sanders, who used its purring depth to convey villainous authority. In languages such as Burmese, creak serves a fundamental linguistic role as a phonemic tone, essential for distinguishing words, cf. (Lindsey, 2011).

Physiologically, it is a natural and harmless mode of speech. Acoustically, it can be surprisingly efficient, particularly in women, often generating a bright, resonant sound rich in higher

frequencies – a stark contrast to the breathier or duller creak sometimes associated with male speakers.

In contemporary speech, especially online, vocal fry is most frequently encountered within specific, repetitive intonation patterns. To understand this, consider how English intonation works: speech is divided into phrases, each with a central pitch movement, or nucleus, that acts like spoken punctuation. A common media pattern involves a set-up phrase (ending with a comma-like rise), a dramatic pause, and a pay-off phrase ending with a definitive falling ‘full stop’. A key feature of this style is a deliberate **downshift** in pitch, which projects weight and control. This downshift can target different parts of the phrase: a **pre-nuclear downshift** lends gravitas before the final stress; a **nuclear downshift**, famously used by Margaret Thatcher, starts the main fall from a lower pitch for formidable emphasis; a **post-nuclear downshift** is where vocal fry typically occurs, as the voice drops so low after the main fall that it slips into the ‘cellar’ of creaky vibration.

Thus, the pervasive vocal fry in media is often the creaky culmination of this formulaic, falling intonation, cf. (Lindsey, 2011). The repetitiveness of the pattern itself, rather than the fry alone, can trigger listener irritation or even misophonia.

The social perception of vocal fry is deeply entwined with the symbolism of pitch. Low pitch is associated with size, calm, and authority – hence its historical use to project control. When used frequently by young women today, that same low, creaky quality is often stigmatised as signalling undue confidence, boredom, or indifference. This criticism may reveal a bias against young women appropriating a vocal domain traditionally linked to male authority. For many speakers, however, it is simply a neutral feature of their voice.

This phenomenon exists within a broader landscape of repetitive intonation trends, including widespread **Uptalk** (rising tones on statements). Notably, the **fall-rise tone** – once primarily a ‘comma’ in British English – is now increasingly used as a final ‘full stop’, often dipping into vocal fry at its lowest point. This creates a modern soundscape where both rising and falling patterns can culminate in that distinctive creak, securing vocal fry’s place in the rhythm of contemporary media speech.

8.2.7. Pausation (Паузация)

Pausation is a universal means of segmenting the flow of speech into prosodic units – phrases and syntagms, cf. (Васильева, 2017 : 28). It is the placement of pauses and their gradation in duration, which is an important component of intonation. Speech is not a non-stop continuous stream. It contains stops of varying length (i.e. pauses), which facilitate the speaking process and aid the correct transmission of meaning. Pauses serve to facilitate the perception, understanding, and processing of speech; or draw attention to specific parts of an utterance, cf. (Беляева, 2019 : 25). So, pausation is the controlled use of silent and filled breaks in speech. It is a critical phonetic mechanism for organising spoken language, facilitating its production and decoding, and is influenced by linguistic norms, communicative intent, and individual speaker psychology.

8.2.8. Duration (Длительность)

Duration is the temporal or temporal component of intonation, referring to the length of time a speech sound, syllable, or segment is sustained during production, cf. (Беляева, 2019). Duration can be measured acoustically in milliseconds (ms), though many speakers perceive sounds as **long** or **short**. Alongside speech **melody (pitch)** and **intensity (loudness)**, duration is a core prosodic feature of a language, it is also a critical acoustic correlate of linguistic stress. In many languages the stressed syllable is signalled by a **significant increase in the duration of its vowel**. Vowel duration is also considered the main parameter in realising rhythmic stress in a language.

8.3. Prosody, Prosodics & Prosodemics (Просодия, просодика, просогемика)

The term *prosody*, synonymous with *suprasegmental phonetics*, is generally used in relation to all suprasegmental units. The term *prosodic* is most frequently applied to the syllable and the phonetic word. The prosodic structuring of a word is created through stress (Современный..., 2012: 43, 51). The distinction between **prosody** (Rus. *просодия*), **prosodics** (Rus. *просодика*), and **prosodemics** (Rus. *просогемика*) is as follows. **Prosody** is the building material of spoken speech, **prosodics** is the method of its use, and **prosodemics** is its social function (Потапова, 2014: 6).

Prosodic, or intonational, features are conventionally subdivided into tonal, quantitative-dynamic, articulatory, and phonational parameters.

Tonal prosodic parameters (melody) are associated with changes in fundamental frequency (F0). The parameters of the melodic contour (the curve of F0 change) include: the direction of pitch movement (rising, falling, level, and their combinations); the interval of a specific tonal movement (the magnitude of F0 change); the range—the total magnitude of F0 change across an entire syntagm; the level, or register (mid, low, high), at which the tonal changes occur; and the speed of these changes—the steepness of the pitch movement.

Quantitative-dynamic parameters include: pause—a universal means of segmenting the speech flow into prosodic units (phrases and syntagms); duration and tempo; and intensity.

Articulatory parameters include: mouth aperture; prosodic nasalisation; backward/forward shift in lingual articulation; and general contraction or expansion of the resonator.

Phonational parameters encompass various voice types.

8.4. Functions of Intonation & Prosody / Prosodemics (Функции интонации и просодики / просогемики)

Now we have enough basic information to consider the functions of intonations. The intonation has the following functions: the communication function, the sentence-forming (or

constitutive) function, the sentence delimiting function, the distinctive function, and the attitudinal function. Let us make a short review of each of the above functions.

The communication function is based on the fact that one of the aims of communication is the exchange of information between people. The meaning of the sentence is conveyed and understood from variations of intonation. By putting a stress on a particular word, the speaker shows firstly that he is treating that word as a carrier of new information and secondly that the information of the other words (not emphasized) is not new but can be understood from the context (Соколова, 2010 : 148).

The sentence-forming (constitutive) function of intonation helps to join words into sense-groups and sentences with the help of differentiations in pitch and stress. It divides the speech flow into significant utterances making it more intelligible and understandable.

The sentence-delimiting function exists as such to divide longer sentences into shorter units called sense-groups which are closely connected in meaning, depending what the speaker wants to express (finality, hesitation, continuation, and etc.), cf.: *There was no love lost § between them* and *There was no love § lost between them*; *The 'boys § who are ill § 'can't 'come* – all the boys and *The 'boys who are ill § 'can't 'come* – some of the boys; *Я развлекал ego § смухами § моего брата*.

The distinctive function helps to distinguish communicatively different types of sentences, e.g.: *'wait 'here* – order; *'wait /here* – polite request; *'wait ^here* – protest.

The attitudinal functions helps to show the speaker's emotional attitude to the situation, e.g.: *It's 'OK* – statement; *It's /OK* – encouragement; *It's ^OK* – implication.

Revision: Section 8 (Повторение: Раздел 8)

1. What is the basic definition of intonation?
2. How can intonation change the meaning of an utterance without altering the words themselves?
3. What are the key parameters that modify a pitch contour (melody), and how are they defined and interrelated?
4. In English speech, what can be said about the rises and falls of pitch (intonation)?
5. What is the defining characteristic of English rhythm, and what is the phonetic consequence for unstressed syllables between stresses?
6. How are loudness and tempo typically correlated with a speaker's emotional state and the semantic importance of speech segments?
7. What is the primary phonetic characteristic and most common linguistic position of vocal fry (creaky voice) in English?
8. Why is pausation important in the language?
9. How can duration as a component of intonation be measured in a language?
10. List and briefly define the functions of intonation.

9. Phonostylistics (Фоностилистика)

9.1. The Notion of Pronunciation Type & Pronunciation Style (Понятие типа произнесения и стиля произношения)

The **type of articulation / delivery** (literally, “type of pronunciation”) is best understood as accent—the specific, often regional, phonetic and phonological traits in speech (e.g., using [ʃ] in *что*). This refers primarily to the phonetic characteristics of a person’s speech, cf. (Соколова, 2010 : 43).

Foreign accent occurs when one speaks a language you don’t command perfectly, e.g., speaking English with a Russian accent. **Acquired / local accent**: one can easily “catch” an accent by immersion in an authoritative or interesting linguistic environment. For example, students may pick up features of a regional dialect within a fieldwork expedition.

The **style of pronunciation** encompasses a wider range of **sociolinguistic variation**, which includes the continuum from **standard literary pronunciation** to **non-standard vernacular**, the adoption of different registers, from formal to highly informal colloquial speech, and the conscious or unconscious use of **dialectal features** (accent + local grammar / words) depending on context and identity, cf. (Соколова, 2010 : 43).

In essence, **pronunciation is not monolithic** but **varies** along axes of **geography** (accent / dialect), **proficiency** (foreign accent), **formality** (vernacular vs. standard), and **social identity** (adopting or parodying specific features). These axes will be viewed in more detail here below.

BBC English (Received Pronunciation, or RP) was generally accepted throughout the UK as a suitable accent for such occasions as the reading of news bulletins and the making of announcements of nationwide importance. But fashions have changed and the movement in recent years has been towards a more relaxed and more personal style on the air. Most public speakers and BBC news readers adopt an informal “chatty style” to project a friendly image to their audience.

9.2. The Classification of Pronunciation & Speech Styles (Классификации произносительных и речевых стилей)

In contemporary linguistic studies, literary language is defined as a primarily written form of speech characterised by both general literary and specific bookish or written linguistic means. This includes the codification of a set of rules governing language usage, as well as the ordering and systematisation of these rules in textbooks, dictionaries, reference works, monographs, articles, and similar publications. Functional speech styles are identified, encompassing a

neutral style and a bookish style—the latter comprising official-business, scientific, and journalistic substyles (Беляева, 2019 : 10).

Colloquial speech refers to the spoken form used by native speakers of a literary language in relaxed, spontaneous, and unprepared communication. It is marked by expressive language features and extralinguistic traits such as informality, lack of preparation, reduced or contracted forms, and the support of gestures and facial expressions.

Practically every type or style of speech employs its own inventory of units, which are regulated specifically within that variety.

The distinction between styles within a literary language, evident in lexicon and grammar, also extends to pronunciation. Two main pronunciations are recognised: a conversational style and a public (or bookish) style. The conversational style is the neutral, stylistically relaxed speech dominant in everyday communication. The lack of emphasis on impeccable pronunciation in this style leads to the emergence of variant forms.

The bookish style is manifested in various forms of public address, such as broadcasting, lectures, and reports. This style demands flawless linguistic presentation, strict adherence to historically established norms, and the avoidance of pronunciation variants.

Where differences are primarily phonetic, two further sub-styles are distinguished: the full style and the conversational (or incomplete) style. The full style is characterised by clear, distinct sound articulation, typically achieved through a slower speaking rate. The conversational style is marked by a faster tempo and, naturally, less careful articulation of sounds (Современный..., 2012 : 89).

!

Relationship between English accent & social class

There is a complex relationship between accent, social class, and perceived trustworthiness in Britain, focusing on the evolution of so-called “standard” pronunciation.

Historically, **Received Pronunciation (RP)** was the accent of the elite – defined as “public school” (private school) pronunciation (Lindsey, 2019 : 1-6). It was promoted as a tool for social mobility and became the authoritative voice of media and broadcasting. However, from the 1960s onward, its association with poshness and an out-of-touch upper class made it a liability. Within Britain, RP is now often viewed as laughable or deliberately avoided.

What much of the world now hears as a neutral “British” accent – from news presenters, actors, and dictionaries – is more accurately termed **Standard Southern British (SSB)**. Unlike RP, SSB is not defined by elite status but is a modern, geographically based standard. It is predominantly spoken by the **middle and upper-middle classes of Southern England**, who remain dominant in media, business, and politics.

The core argument is that **accent acts as a powerful social shortcut**. Listeners make rapid judgements about trust, authority, and background based on speech. Southern, middle-class accents (like SSB) are subconsciously associated with education, success, and officialdom, leading audiences to perceive them as more trustworthy. This is not because the speakers are inherently more reliable, but because the accent triggers ingrained social biases.

Thus, Britain has moved from a rigid **linguistic monoculture (RP)** to a landscape where **diversity is celebrated**, yet a southern, middle-class accent (SSB) retains a default authority in public life. This perpetuates a system where accent and perceived class continue to profoundly influence who is trusted and heard.

9.3. Speech Culture / Standard of Speech (*Культура речи*)

Speech culture – a discipline concerned with the strict adherence to the standards of the literary language (Маркова, 2025). The qualitative parameters of literary speech respond differently to the demands of various functional styles. The parameter of speech culture known as logicity operates at the syntactic level: the parts of a phrase must exist in a rational interrelationship. Lexical richness is measured by the quantity and variety of linguistic signs employed within communication. This particular characteristic, manifested at a systemic rather than an individual level, distinguishes the artistic, journalistic, and liturgical styles; by contrast, business and scientific styles can appear linguistically monotonous. Another parameter, purity, is expressed through the use of literary lexicon and the avoidance of mixing linguistic means from different styles. Further parameters of speech and communicative culture include grammatical correctness, vocabulary range, intonation, adherence to speech conventions, good manners, facial expression, and so forth.

Furthermore, speech culture is a branch of linguistics which investigates problems related to the refinement of language as an instrument of culture in the broadest sense. The subject matter of speech culture, as an academic discipline, encompasses questions regarding the standardisation of the literary language, the principles and rules of verbal communication, the ethical norms of interaction, and issues pertaining to the contemporary state of a society's speech culture (Воронцова, 2021: 16). The ethics and varieties of oratory are diverse and require from the speaker adherence to principles of moral integrity, cultivated speech, and behavioural decorum.

9.4. Speech Etiquette (*Речевой этикет*)

Speech etiquette, encompassing the nationally specific and socially conditioned conventions governing behaviour during interaction with a member of a different linguistic community; the knowledge of rules for language use within a particular society (verbal etiquette) (Тарева, 2017).

Etiquette is defined by leading scholars as the totality of behaviour rules accepted within each society. However, alongside behavioural features dictated by practicality, there exists speech etiquette, which correlates with the concepts of 'speech' and 'verbal activity'. Speech etiquette serves as the point of intersection which enables the identification of universal aspects in cross-cultural interaction, while simultaneously representing the result of comprehending the specifics of verbal behaviour by members of different linguo-cultural communities. Definitions of speech etiquette may include the following: these are stereotypical verbal units, employed in

etiquette-based situations, which may vary across different stylistic registers; it is a system of established communicative formulas, prescribed by society as rules of verbal conduct for various purposes in a chosen tone, in accordance with social roles and setting. The components of speech etiquette are: 1) Politeness formulas, consisting of expressions for greeting, farewell, gratitude, and requests; 2) Forms of address; 3) Adherence to standards of speech – this entails the competent use of language, avoiding superfluous words, vulgar and coarse expressions; 4) Observance of linguistic and value-based taboos – this involves refraining from using prohibited expressions, words, and topics of conversation; 5) The use, to a greater or lesser extent, of non-verbal means of communication (tone of voice, facial expressions, gestures) (Калинина, 2022). The English language is an international language of communication throughout the world; consequently, it is taught in schools across different countries, and students occasionally employ words from English speech etiquette in communication, such as “Goodbye!”.

9.5. Phonetic Styles (Фонетические стили)

A functional style of any given language is a system of methods for employing, selecting, and combining the means of verbal communication, dictated by the social aims and functions within a people’s communicative practice.

Functional styles are intrinsically linked to phonetic styles. There exists an interconnection between the stylistic means across different levels of the language system: phonetic, lexical, syntactic, and morphological.

A phonetic style, or phonetic style – also referred to as a “type of pronunciation”, “style of pronunciation”, “pronunciatory style”, etc. – is a set of segmental and suprasegmental features applied in a communicative context or extra-linguistic situation to fulfil specific functions. In contemporary linguistic literature, various classifications of phonetic styles are proposed. In modern English, the following phonetic styles are distinguished: informational, scientific, publicistic, artistic, and conversational (Великая, 2009 : 21).

The **informational phonetic style**, or ‘stylistically neutral’, is characterised by a predominance of intonational structures for conveying information. This is a formal style used in newsreader speech on radio and television (also in motivations to act, orders, interrogations, testimonies, etc.), and in its unofficial variant, in oral announcements within informal settings.

The **scientific phonetic style**, employed in discussions, lectures, and seminars (also reports on scientific topics, scientific and technical discussions, all forms of scientific sounding texts), utilises intonational structures to convey intellectual information. As the goal is to persuade the audience, establish contact, and capture attention, the intonation may be individual and emotional.

The **publicistic phonetic style** is marked by volitional or desiderative intonations. Its aim is not only to inform but also to persuade, influencing the thoughts and feelings of the audience. It can be used in public speaking, presentations, interviews, liturgical sermons.

The **artistic style** is distinguished by its emotionality, expressiveness, and use of expressive prosodic means (loudness, tempo, pitch range). It is applied in stage speech, dramatic

reading, radio and television broadcasts, and in literary studies. Its primary objective is to affect the audience's thoughts and emotions.

The **conversational phonetic style** is used in daily life and informal situations (i.e. in all forms of casual communication). Its intonational structures are diverse and dependent on the utterance's purpose, ranging from intellectual to emotional.

Phonetic styles are rarely found in their pure form; they depend on the purpose and situation of communication. Each style is determined by a combination of characteristics: form of speech (oral / written), type (monologue / dialogue), preparedness (prepared / spontaneous), number of participants (public / non-public), nature of relations (official / unofficial).

Revision: Section 9 (Повторение: Раздел 9)

1. What has been the recent trend in the style of public speaking and BBC news reading, moving away from the traditional formal standard?
2. What is accent in the meaning of the type of articulation?
3. How can the style of pronunciation be briefly described?
4. How is colloquial speech best described?
5. What main pronunciational styles are recognized?
6. What is the main difference between speech culture and speech etiquette?
7. What are the subject matter and the parameters of speech culture?
8. What are the components of speech etiquette?
9. What is the definition of the phonetic style?
10. What are phonetic styles classified into and how can these subtypes be described in short?

10. Variation / Variants in English Pronunciation

(Вариативность / варианты английского произношения)

!

Pronunciation: Non-English Words

When English borrows a word from another language, speakers rarely aim for a perfect imitation of its original sound. More often, they haven't even heard it spoken in its native context. The standard approach is **nativisation** – pronouncing the loanword using the familiar sounds and patterns of English, guided by its spelling, cf. (Lindsey, 2011). This is how English has absorbed most of its vocabulary for centuries; words like 'choice' from French or 'status' from Latin no longer retain their foreign sounds.

However, practices shift over time, and **denativisation** – reverting to a pronunciation closer to the source – can occur. For instance, the British pronunciation of 'Lyon' and 'Don Quixote' has moved towards the original in recent decades, and many younger Brits now adopt the American-influenced 'oh-mage' for 'homage'. A key divergence lies between British and American English. While both nativise, they often do so differently. Americans frequently apply a **'one-size-fits-all' strategy** for stressed vowels in loanwords, using a set of five tense vowels (as in *father, day, see, go, too*). This leads to pronunciations like 'pAHsta', 'NEEsa', and 'BOOdha'. This strategy, partly born from America's 'melting pot' need to handle diverse names and words, and heavily influenced by the five-vowel system of Spanish, aims to signal foreignness by avoiding native English vowel patterns.

British English, less influenced by a single prototypical foreign language and more familiar with the complex vowel system of French, tends to **nativise more consistently**. Brits are more likely to use the same vowel in 'pasta' as in native 'plastic', and to nativise stress patterns (e.g., 'Salon' vs. American 'saLON'). This can sometimes preserve phonetic distinctions that the American strategy merges, as with the German vowels in 'Schubert' versus 'Mahler'. Ultimately, pronouncing a loanword 'correctly' is less about accuracy to the source and more about conforming to the internal, evolving rules of the borrowing language. The chosen strategy – whether leaning towards nativisation or a marked 'foreign' pronunciation – reveals deep-seated phonetic habits and cultural attitudes towards language and identity.

10.1. Pronunciation norm: orthoepy & the orthoepic norm; literary pronunciation, the concept of; the emergence of pronunciation norms

(Произносительная норма: орфоэпия и орфоэпическая норма;
литературное произношение, возникновение нормы произношения)

When describing the English language for either research or teaching purposes, one normally considers and mentions the contrast between British English and American English.

In British English, Received Pronunciation (BBC English) was generally accepted throughout the UK as a suitable accent for such occasions as the reading of news bulletins and the making of announcements of nationwide importance (Соколова, 2010 : 159).

In American English, General American is the form of speech used by radio and TV, though its existence is questioned, see the discussion in the sections below (Соколова, 2010 : 182). It is mostly used in cultural, scientific and business intercourse. General American is the language of broadcasts and publications and thus it is familiar to almost all Americans. It is the language without any flavour, the language they use to sound educated, professional and businesslike.

!

Accents: British by Americans

The curious task of an American actor adopting a British accent for a film – a far rarer feat than the reverse – provides a fascinating lens through which to examine the peculiarities of the British sound. A case study is William Hurt's performance in the 1983 film *Gorky Park* carried out by Dr. G. Lindsey (Lindsey, 2011). Playing a Russian policeman but directed by Brits, Hurt aimed not for a caricature of posh aristocracy but for a credible, contemporary **Standard Southern British (SSB)** accent.

He displayed mastering several key British phonetic hallmarks. He adopted **non-rhoticity**, dropping his /r/ sounds unless they preceded a vowel. He employed **BATH-broadening**, using the broad /ɑ:/ vowel in words like "bath" rather than the American flat /æ/, cf. (Lindsey, 2019 : 41-43). He consistently used the short, rounded **LOT vowel** (/ɒ/), which doesn't exist in General American, and generally avoided **American T-flapping**, pronouncing his /t/ sounds more clearly. Crucially, his **GOAT vowel** was modern, avoiding the old-fashioned, very fronted quality stereotyped by Americans, cf. (Lindsey, 2011).

However, traces of his native accent subtly persisted, revealing where the boundaries blur or where strict "rules" are myths. While he avoided the optional **intrusive R**, he occasionally used **T-voicing** (a light, almost /d/-like sound between vowels), a feature common in SSB. His articulation of /l/ sounds was sometimes slightly **dark and dental**, and his **THOUGHT vowel** was somewhat more open than typical, cf. (Lindsey, 2019 : 29-30). A more noticeable tell was a tendency to **avoid weak forms and contractions**, making his speech sound unusually deliberate and full compared to the more rhythmic, reduced flow of native casual speech.

This performance underscores that the modern British "accent" targeted by learners and actors is not the plummy **Received Pronunciation (RP)** of the past, but the more neutral and geographically rooted **SSB**. It also highlights that accent adoption is an exercise in selective phonetic filtering – where an actor must consciously suppress some native habits while learning new ones, all while navigating features that exist on a spectrum within the target accent itself.

10.2. Regional & Social Variantology

(Региональная и социальная вариантология)

Variantology is the study of how a language becomes stratified and functions when used by distinct national speech communities (Скреб, 2017). This falls within the domain of

sociolinguistics, which can be divided into: **regional variantology** (the study of geographically based language forms, such as national varieties and dialects) and **social variantology** (the study of language forms linked to social groups, contexts, and communicative functions).

!

Pronunciation: Language “Mistakes”

The common frustration – why do linguists so often defend what seems like a blatant language mistake? – stems from a fundamental difference in perspective. Where the public may see error, linguists see a complex, living system governed by deep principles.

At the heart of their view is a concept articulated by Ferdinand de Saussure: the **arbitrariness of the linguistic sign**. The connection between a word’s sound and its meaning is mostly random – this is why languages differ wildly and why words can change sound over time without changing sense. This same arbitrariness applies to **standards** of grammar, spelling, and pronunciation. These conventions are human-made, often dating from the 18th or 19th centuries, and are not eternal truths but historical decisions that can shift.

Linguists, therefore, aim to be **descriptive**, not prescriptive. They act like scientists, analysing how language actually works, not how it *should* work. A key insight is that most language is an **unconscious iceberg**; speakers master immensely complex grammatical rules in childhood without explicit teaching. The conventions that must be painstakingly taught in school are often the artificial standards, not the innate rules of the language itself.

When confronted with a new usage or an apparent mistake, a linguist’s instinct isn’t to condemn but to **analyse**. They look for **regularities and precedents**, asking: does this fit a pattern seen elsewhere in the language or in its history?, cf. (Lindsey 2011). What might have caused this reanalysis? For example, the development of plural pronouns like “y’all” or “yous” fills a gap the standard language lacks, following a logical linguistic pattern, regardless of social stigma.

This leads to the **power of association**. People swiftly link speech forms to stereotypes – judging a construction as “uneducated” or “quaint” based on who uses it, not on its linguistic merit. While linguists are not immune to these associations, their training helps them see past them to the historical accidents and systematic processes that truly shape how we speak.

So, defending a “mistake” is rarely about endorsing carelessness. It is about recognising the **rule-governed nature** of all language variation, understanding the fluidity of standards, and separating the deep, unconscious logic of speech from the superficial judgments we attach to it.

10.3. National Pronunciation Variants of English: British, American, Australian (Национальные произносительные варианты английского языка: британский, американский, австралийский)

Over the past three centuries, the English language has expanded across all continents, and the number of its speakers has increased significantly (Петроченко, 2001 : 91-93). During the Old English and Early Middle English periods, English dialects were limited to parts of the British Isles, primarily what is now England. From the 13th to the 17th century, English spread

throughout the British Isles, except for some mountainous areas of Wales, northern Scotland, and parts of Ireland. The number of English speakers grew steadily: from an estimated 1.5 to 2 million at the end of the 11th century to over 8 million by 1700. During two centuries of British colonial expansion and emigration overseas, the number of speakers increased sharply, reaching around 123 million by 1900.

England's colonial expansion into the New World began in the late 16th century. These early colonists came mainly from London, East Anglia, and Yorkshire, with later arrivals from other regions, including Scotland and Ireland. Their various dialects gradually blended in North America, forming what is now known as American English.

Australia was used as a penal colony from the late 18th century, later attracting immigrants through land grants and gold discoveries. Most of the population in Australia and New Zealand originated from Great Britain.

A pidgin is a communication system that develops among groups without a common language. It has a limited vocabulary, simplified grammar, and narrower functions than its source language. Widely used English-based pidgins include Krio (Sierra Leone), Nigerian Pidgin English, and Tok Pisin (Papua New Guinea).

!

Pronunciation: CUBE Dictionary

CUBE is a freely accessible online pronunciation dictionary specialising in contemporary British English, cf. (Lindsey, 2011). Its core functionality is built around two search windows: one for **spelling** and one for **sound**, allowing users to find words by their written form or phonetic transcription.

The dictionary employs multiple transcription systems, including a traditional and a more contemporary scheme, with the latter representing a modern standard while the former may reflect somewhat older pronunciations.

While CUBE already includes a vast number of entries, including many proper names, should a search for a specific word yield no results, users can opt to be notified by email. Typically, such missing words are added to the database within 24 hours. Searches can also be refined and sorted by web frequency.

Beyond searching for individual phonemes, the sound window enables queries based on broader **sound categories**, all of which are detailed on the dictionary's symbols page. An additional useful feature is the ability to generate and display **minimal pairs**.

!

Accents: Brits Sing in American

The phenomenon of adopting an American accent while singing is not unique to British performers, but common among singers from Ireland, South Africa, Australia, New Zealand, and other English-speaking regions. What does "singing in an American accent" actually involve?, cf. (Lindsey, 2011). It often includes pronouncing **R sounds** before consonants and at word endings – a feature absent in many non-rhotic speaking accents – and using a **tap** sound (similar to a light /d/) for /t/ and /d/ between vowels, typical of North American speech.

The trend can be traced back to early British acts like The Beatles, whose adoption of American vocal styles was not always widely welcomed. So why do so many non-Americans shift toward an American sound when singing? Importantly, this shift doesn't mean singers lose their accents entirely – it's specifically linked to performing **American genres of music**.

The United States, particularly through the profound influence of **African American communities**, has played a dominant role in shaping global popular music throughout the 20th century. This explains the frequent presence of Southern American speech features in many singing styles. There is a natural connection between music and pronunciation: the vocal placement that suits genres like jazz, blues, R&B, rock and roll, and pop aligns closely with the articulatory patterns of American accents. This makes the music not only sound more authentic but also easier to sing with those phonetic qualities.

Interestingly, even American singers often adjust their accents when performing. Features such as a drawn-out “ah” vowel or non-rhotic pronunciation – originating from Southern American speech – are frequently adopted by singers whose everyday speech is General American. In many singing styles, dropping the /r/ before consonants and at word ends is common practice.

Ultimately, singing makes it remarkably easy to mimic a different pronunciation, blending vocal style with linguistic pattern in a way that feels almost instinctive.

10.4. Regional Pronunciation Types of British English: Southern and Northern English Pronunciation, Standard Scottish Pronunciation, Pronunciation in Wales & (Northern) Ireland (Региональные произносительные типы британского варианта английского языка: южноанглийское и североанглийское произношение, общепринятый шотландский тип произношения, произношение в Уэльсе и (Северной) Ирландии)

In Great Britain, three principal types of English pronunciation exist: Southern, Northern, and Scottish.

Southern English Pronunciation, also termed Received Pronunciation (RP), Public School Pronunciation, or BBC English, originated in southeast England and became the prestige standard from the 16th century. It spread through public schools, broadcasting, and is often taught internationally. At present, its modern form, Standard Southern British, SSB, has been developing, cf. (Lindsey, 2019). Key trends include vowel neutralisation, consonant elision, stress shifts, and influence from American usage (which is debatable, though).

Northern English Pronunciation is spoken from Birmingham to the Scottish border and retains features closer to older forms of London English, some of which influenced American dialects. Characteristics include distinct vowel realisations, such as /a/ instead of /ɑ:/ in words like *dance*.

Scottish Pronunciation differs more markedly from Southern English, with features such as a trilled /r/, dark /l/ in all positions, and the retention of the phoneme /x/ (as in *loch*). It also uses monophthongs where Southern English uses diphthongs before /r/.

Of the three, Southern English remains the least geographically restricted and is widely associated with educated speech across Britain and beyond.

The English spoken in Ireland diverges more sharply from British English than Scottish dialects, having developed in relative isolation over centuries. Despite efforts to revive the Irish language, by 1900 a strongly accented variety known as the “brogue” had become dominant. Urban accents such as those of London, Birmingham, and Liverpool are often perceived as ‘harsh’, while rural accents are considered ‘quaint’.

10.5. American English Pronunciation (Американское произношение)

About 300 million people speak American English.

It is generally assumed that American English in its spoken form is essentially Southern English Standard of the 17th and 18th centuries as modified locally in the course of the last century or two. Basically 3 types of American English are generally recognized (Shakhbagova, 1982 : 15-16; Швейцер, 2008):

1) the Eastern American – spoken in New England, North-East of the Connecticut River, New Hampshire, Massachusetts, Vermont, New York;

2) the Southern American – Virginia, North Carolina, South Carolina, Tennessee, Florida, Georgia, Mississippi, Arkansas, Louisiana, Texas, West Virginia, Oklahoma, Kentucky, cf. (Скреб, 2017), which has a specific way of diphthongization of vowels (pure ones) and monophthongization of some diphthongs at the expense of “drawling” (prolonging their nuclei and dropping the glide);

3) General American – the rest the USA from Ohio via Midwest to the Pacific Coast, New Jersey, Wisconsin and other states.

General American is also known as North American and Western American spoken in the central states. It is known to be the pronunciation Standard of the USA and is the form of speech used by radio and TV presenters. It is mostly used in cultural, scientific and business intercourse.

Pronunciation variants assigned to particular regions are regarded as the ones prevailing there. Many American linguists specify the fact that the diversities in pronunciation between the standards in the USA (with the exception of Southern American) are not so marked and they tend to be levelled out.

The neutral speech of the Midwest and West has become norm for almost all educated persons, for businessmen, professionals in consulting companies as well as the vast suburban middle class. But the lower-status local and regional accents have stayed on. The well-educated people have made a point of avoiding local ways of speaking as too provincial. That is why they say: “we speak a version of General American at work, at home we enjoy putting on our local ways of speaking (even showing off our local ways)” (Metcalf, 2000).

Thanks to modern mobility and broadcast media, fewer and fewer Americans are stuck with only one way of speaking. As they also say: “we enjoy our local speech more nowadays because we have that choice”. But it is necessary to state the fact that General American is the language of broadcasts and publications and thus it is familiar to almost all Americans. It is the language without any flavor, the language they use to sound educated, professional and

businesslike. Of course, you will find residents, even natives, of Atlanta with hardly a trace of a Southern accent; Bostonians who sound like Californians; Minnesotans who could be from Ohio. Southern American differs from General American.

But the 20th and the 21st centuries brought many changes to the South. Descendants of slaves migrated North and West, bringing southern speech patterns to the rest of the USA. Vacationers and retirees from the cold North flocked South, bringing northern speech to Florida. National Corporations established offices in southern cities like Atlanta and Dallas. CNN headquarters in Atlanta, Disney & Universal Studios established national attractions in Orlando bringing in Northern speech.

Thus, we may say that American English shows a lesser degree of dialect than British English due to high mobility of population, and internal migrations of different communities.

!

American English features

Producing a clear and consistent pronunciation guide for American English presents a unique set of linguistic challenges, centred on the difficulty of **phonemicization** – the process of defining the core set of distinctive sounds, or phonemes, that differentiate words within a dialect.

This task is notably more complex for **General American** than for, say, Standard Southern British English, cf. (Lindsey, 2011). A key sign of this complexity is the striking lack of agreement among dictionaries on how to transcribe even basic words, differing in both the choice and number of phonemes they assign.

The core principles of phonemicization rely on evidence like **minimal pairs** (e.g., *sax* vs. *six*, proving /æ/ and /ɪ/ are distinct phonemes), **complementary distribution** (where sounds appear in mutually exclusive environments, suggesting they are variants of one phoneme), and **neutralization** (where a normally robust contrast disappears in a specific context).

American English throws up significant obstacles to applying these principles neatly:

1) **the flap/tap:** the widespread use of a flapped consonant (sounding like a light /d/) for both /t/ and /d/ between vowels (e.g., in *latter* and *ladder*) creates **neutralization**; dictionaries disagree profoundly on how to represent this, variously ignoring it, using a /d/ symbol, or inventing a new symbol, none of which is fully satisfactory;

2) **vowel mergers:** several vowel contrasts are absent for large groups of speakers, creating divergent phonemic inventories; the **cot-caught** merger and the **pin-pen** merger mean millions of Americans have one fewer vowel phoneme than those who maintain the distinction; dictionaries must choose which system(s) to represent;

3) **the STRUT-commA question:** there is fundamental disagreement over whether words like *above* use one vowel phoneme (schwa) in both syllables or two different ones (the STRUT vowel and schwa), cf. (Lindsey, 2011); this challenges the very phonemic inventory;

4) **vowels before R:** the powerful influence of /r/ on preceding vowels in rhotic American accents complicates analysis. Well-known mergers (like *merry-marry-Mary*) represent neutralization. Furthermore, there is little consensus on how to phonemicize the vowel in words like *near* or *spirit* – is it the KIT vowel, the FLEECE vowel, or a unique R-influenced vowel?

These challenges are compounded by the nature of **General American** itself. Unlike the historically monolithic prestige of Received Pronunciation in Britain, American English is a **family of similar regional accents** with no single, universally accepted standard. This inherent diversity resists a one-size-fits-all phonemic description.

Ultimately, this “phonemic variability” does not hinder everyday communication, as native speakers process speech fluidly, not as strings of dictionary phonemes. The problem is intrinsic to the analytical task of phonemicization. When creating a pronunciation guide, lexicographers must navigate a landscape of variable mergers, contextual neutralizations, and regional diversity, leading to the divergent transcription systems we see across different dictionaries.

10.5.1. The Eastern, Southern, and Western American Pronunciation

(Восточный, южный и западный тип американского произношения)

!

The Northern Cities Vowel Shift

The distinctive accent of Chicago, part of America’s Inland North, offers a fascinating window into a major linguistic phenomenon: the **Northern Cities Vowel Shift (NCVS)**. This is a **chain shift**, where the pronunciation of one vowel sound changes, prompting neighbouring vowels to move in a connected, domino-like fashion within the auditory space, much like colours shifting on a spectrum, cf. (Lindsey. 2011).

The shift is broadly **clockwise** when compared to General American vowels. The most recognisable elements involve the **TRAP** vowel (as in ‘cat’), which is raised and often becomes a diphthong (sounding like ‘ceah’t’), and the **LOT** vowel (as in ‘cot’), which is strongly fronted (towards a sound like ‘caht’). This movement is thought to create a **push or drag chain**: as LOT moves forward, it may pull the **THOUGHT** vowel (as in ‘caught’) down and open, helping to maintain the distinction between words for the majority in this region who do **not** have the *cot-caught* merger.

The chain continues to affect other vowels, though less uniformly. The **DRESS** vowel (as in ‘bed’) often centralises, sounding more like the General American **STRUT** vowel, especially before an /l/ sound. Similarly, the **KIT** vowel (as in ‘bid’) can also centralise. The behaviour of the **STRUT** vowel itself is complex; while the classic description says it shifts back to sound like ‘boss’, in reality its movement is inconsistent, leading to theoretical puzzles where shifted DRESS and unshifted STRUT can sound identical.

This core shift also influences diphthongs. The **PRICE** vowel (as in ‘ride’) fronts, while the **MOUTH** vowel (as in ‘house’) does the opposite, backing – possibly to keep the two clearly distinct. Notably, this pattern is the **exact converse** of the anti-clockwise shift that occurred in Southern British English over the 20th century, where PRICE moved back and MOUTH moved forward.

In essence, the Chicago accent exemplifies how vowel systems are dynamic, interconnected structures. A change in one area of the vowel space creates ripple effects, reshaping the entire soundscape of a dialect in a systematic yet often untidy evolutionary process.

10.5.2. Features of the Consonant & Vowel Systems in American English

(Особенности консонантной и вокалической системы в американском варианте английского языка)

!

Accents: American by Brits

The journey of a British actor mastering an American accent is fraught with a particular phonetic pitfall: the misplaced “R” sound. This challenge stems from a fundamental divide between the two dialects: **rhoticity**.

American English is broadly **rhotic**, meaning speakers pronounce the consonant /r/ wherever it appears in spelling. British English, however, is largely **non-rhotic**; speakers like myself only pronounce /r/ when it is immediately followed by a vowel sound. For us, the /r/ in “car” is silent, but it appears in “car engine”.

This system leads to the phenomenon of **linking R** (often pejoratively called “intrusive R”), cf. (Lindsey, 2019 : 87-89). In British English, to avoid two vowels clashing awkwardly – a sequence the language disfavours – an /r/ sound is naturally inserted between them. This happens whether the first word is spelled with an ‘r’ (“far out”) or without one (“vanilla ice” becomes “vanilla /r/ ice”). For British speakers, this linking /r/ is a single, rule-governed process of sound connection, not a historical or spelling-based one.

The peril for the actor lies in **overgeneralising** this rule when switching to a rhotic American accent. The critical mistake is assuming that certain British vowel sounds can *always* be converted to an American vowel-plus-/r/ combination. This hypercorrection creates the truly “intrusive R”, cf. (Lindsey, 2011). The danger zones are three British vowels that commonly occur at word endings: the /ɔ:/ in words like **law**, **saw**, and **draw**; the /ɑ:/ in words like **spa**, **bra**, and **grandma**; the schwa /ə/ in hundreds of words ending in ‘-a’, like **idea**, **pasta**, and **Sarah**.

When an actor carries over the linking habit, they might mistakenly add an /r/ to these words before a following vowel (e.g., “I saw/r/ it”), creating an “R” that exists in neither the American target accent nor their own British speech. An even graver error is adding a non-linking, final /r/ where there is no following vowel at all (e.g., “See-saw/r/”).

Thus, the peculiarity for the British actor is not merely learning to pronounce every written ‘r’. It is the more nuanced task of **suppressing a deeply ingrained, subconscious phonetic rule** of their native tongue – the automatic insertion of a linking consonant – while simultaneously activating a new rule where /r/ presence is dictated strictly by orthography. The battle is against a phonological reflex, making the American “r” a deceptively tricky sound to master authentically.

10.5.2.1. Features of the Consonant System

(Особенности консонантной системы)

Let us consider the peculiar features of General American. In the sphere of consonants General American has the following peculiarities:

- 1) [r] is retroflex and pronounced in all positions. It is pronounced in the throat;

- 2) dark [ɫ] is used in all positions. Clear [ɪ] is sometimes used in Southern American;
- 3) [t] is voiced in -VCV- position. Held [t̪] before [n] is used, e.g.: *certain*;
- 4) [hw] in wh-words (see also the comment in the previous essay about the Southern pronunciation of British English). It is formed by rounding the lips and raising the back of the tongue to the soft palate (the velum);
- 5) [h] is lost in pronouns *him, her, his, he*, and the verb *have*, e.g.: *Where (h)ave you put it? I told (h)im find* – [h] is not pronounced;
- 6) [ju:] sequence is retained in American English & Received Pronunciation in word-initial position, e.g. in: *unit, unite, uniform*. It is also pronounced after [p], [b], [m], [f], [v], [k], [g], [j], e.g.: *few, fume, beauty, hew, cube*. But after [t], [d], [n], [s], or [z] the sound [j] is dropped, e.g.: *news, suit, student, assume, due, dew, Zurich*;
- 7) [ʃ] in General American is vocalized in final unstressed syllables, e.g.: *version* ['vɜ:ʒn], *Asia* ['eɪʒə], *Persia* ['pɜ:ʒə], *excursion* [ɪks'kɜ:ʒn]

10.5.2.2. Features of the Vowel System

(Особенности вокалической системы)

In the sphere of vowels the peculiarities of General American are as follows:

- 1) There is no strong opposition between short and long vowels, [i:] and [u:] are diphthongized.
 - 2) There are only 5 diphthongs [eɪ], [aɪ], [ɔɪ], [aʊ], [əʊ] / [oʊ]. In American English they pronounce [dɪr] for *dear*, [deɪr] for *dare*, [tʊr] for *tour*.
 - 3) The symbol [ɑ] is used to denote the pronunciation of the stressed sound in such words as *father*.
 - 4) [æ] is longer in General American, and is nasalized before [d], [m], [n], e.g.: [bæ:d], [æ:nsə].
 - 5) The 3 Received Pronunciation vowels [ɒ], [æ], [ɑ:] correspond to only two vowels in General American [ɑ], [æ], e.g.: Received Pronunciation *dad* [dæd] / [dɑd], *path* [pa:θ], *dance* [da:ns], *dog* [dɒg] correspond to General American [dæd], [pæθ], *dance* [dæns], *dog* [dag].
 - 6) [r] is treated as a semi-vowel.
- As for the intonation General American has its own peculiarities:
- 1) American intonation more dynamic, confident, and persuasive.
 - 2) Phoneticians suggest 3 ways to make intonation sound more like American – to pronounce louder (raise the volume); to stretch the word (to command attention); to change the pitch at greater rates und more often.
 - 3) Use liaisons, i.e. connect words to form sound groups, e.g.: *My name is Ann* [maɪ neɪ mɪ zæn], *Bob is on the phone* [bɒbɪzən ðə fəʊn] (American English [ɑ] – is [ɒ] or [ɔ:] in Received Pronunciation).

10.6. Australian English Pronunciation (Австралийское произношение)

Australian English features

The story of the Australian accent, particularly its vowels, is tangled in a legacy of linguistic tradition and colonial influence, cf. (Lindsey, 2011). It begins with a peculiarity: why are Australian English vowel charts often misleading? They sometimes list steady, single-sound monophthongs alongside gliding diphthongs, creating a confusing picture. The central controversy revolves around vowels like the one in **FLEECE** (as in 'bead') and **GOOSE** (as in 'who'd').

Australian English vowel charts often incorrectly list the FLEECE and GOOSE vowels as steady monophthongs. Phonetic evidence demonstrates they are, in fact, **closing diphthongs** (vowels with a glide), patterning with other diphthongs like FACE and GOAT.

Since the 19th century, Southern British English possessed a set of seven closing diphthongs (e.g., FACE, FLEECE, GOOSE). Esteemed phoneticians like Henry Sweet and later American linguists consistently classified and transcribed them as diphthongs, using symbols like /iy/ and /uw/.

In the 20th century, a separate tradition emerged, centred on **Received Pronunciation (RP)**, the prestige accent of the British Empire, and propagated through dictionaries. This tradition **rejected the diphthongal transcription** for FLEECE and GOOSE, instead representing them as long monophthongs /i:/ and /u:/, aligning them with vowels like /ɑ:/ and /ɔ:/. This was despite phonetic descriptions of RP acknowledging their diphthongal nature.

The root of the confusion is traced back to an RP-centric tradition of transcription, championed by phonetician **Daniel Jones**, took hold. Jones, obsessed with vowel length, devised a system that treated FLEECE and GOOSE as long monophthongs (like /i:/ and /u:/), ignoring their diphthongal nature. He also believed vowels like KIT and FLEECE shared the same phoneme, with length distinction marked by a "chroneme". His symbols persisted even after his theory was abandoned. This was partly influenced by the class snobbery of Victorian England, where minimizing the glide in these vowels was seen as more "posh" and distinct from working-class Cockney speech. Maximising the diphthongal quality of FLEECE was associated with Cockney and lower-class speech. Affecting a pure, monophthongal pronunciation (like German *i:*) became a marker of upper-class RP "purity".

This RP-centric system was exported to Australia. Australian English (AusE) developed from contact between southeastern English dialects, heavily influenced by both RP and Cockney. Early Australian linguists, considering RP a standard to aspire to, imported its transcription methods. Thus, Australian English pronunciation was often described using symbols that didn't fit its reality.

Even a revised, influential 1997 acoustic study of AusE vowels retained the RP symbol /i:/ for FLEECE, treating it as a monophthong with an optional "onglide" [ɪi:], while correctly transcribing FACE and PRICE as diphthongs.

This **methodologically flawed** the study, as it measured FLEECE and GOOSE in closed syllables (where they are more monophthongal) rather than in open syllables (where their full diphthongal form is clear). Acoustically, Australian FLEECE behaves similarly to FACE, both being diphthongs. The RP-centric transcription forces a misleading separation: FLEECE is stuck on monophthong

charts while its phonetic siblings, FACE and PRICE, are on diphthong charts. This masks the fact that in accents worldwide, including Scottish English, FLEECE and FACE pattern together. In Australian English, all seven closing diphthongs (including FLEECE and GOOSE) clearly reveal their gliding nature when followed by another vowel, as in words like *seeing*, *going*, or *doing*.

Clinging to these inappropriate RP symbols is a piece of “colonial baggage”. It creates bizarre analytical outcomes and obscures the true, systematic pattern of Australian vowels – a pattern where FLEECE and GOOSE belong firmly with the family of closing diphthongs, their behaviour echoing their historical origins in the speech of early settlers from Southeast England, rather than in the artificial standards of a distant elite. The real peculiarity, then, isn’t the Australian accent itself, but the outdated linguistic lens through which it has often been forced to be seen.

Using glide symbols (e.g., /j/ for front glides, /w/ for back glides) in broad transcription is obviously more accurate and economical than using multiple vowel symbols. It correctly shows the syllable structure and falling sonority of a diphthong.

Revision: Section 10 (Повторение: Раздел 10)

1. When describing the English language for research or educational purposes, which norms does one normally imply?
2. What is variantology and what are its sections?
3. What are the principal types of modern British English?
4. What pronunciation norms are generally accepted in the UK and the USA for broadcasts?
5. What are the three traditionally recognized major types of American English?
6. What is the defining phonetic characteristic of Southern American English?
7. What is the current social dynamic regarding the use of General American versus local accents in the USA?
8. List the key consonantal features that distinguish General American pronunciation from Received Pronunciation (RP).
9. What are the main differences between the vowel systems of Received Pronunciation (RP) and General American (GA) as outlined in the text?
10. How was the Australian English accent formed?

11. Present Day Changes in Modern English Pronunciation

(Изменения в произношении, происходящие в современном английском языке)

!

Accents: Shakespeare's Original Pronunciation

At the Globe Theatre in the UK, in 2004, for the first time on that stage since Shakespeare's own day, the words of a play were spoken not in modern received pronunciation, but in the authentic accent of the early 17th century – what scholars term Original Pronunciation, or OP. Against initial fears that audiences would find it incomprehensible, the experiment proved a resounding success, unlocking the plays in unexpected ways.

This OP was not a single, fixed accent, but the living, shifting soundscape of Early Modern English, a period of remarkably rapid phonetic change. We hear it evolve even within Shakespeare's lifetime, in words like 'musician', which journeyed through several distinct pronunciations.

But how can we know how they truly spoke? Scholars piece together the lost accent like detectives, using three crucial forms of evidence. First, they scour the writings of contemporary commentators, like the playwright Ben Jonson, who explicitly described the sounds of speech. Second, they study the often erratic spellings of the time, which faithfully mirrored pronunciation. The third and most telling clue, however, lies in Shakespeare's own verse: **nearly two-thirds of his sonnets contain rhymes that fall flat in modern English but spring to life in OP**, proving its patterns beyond doubt.

Hearing it is transformative. In a passage from *Henry V*, "the swelling scene" becomes "the swelling **sane**"; "his heels" shortens to "**is 'eels**"; and "warlike Harry" loses a letter to become "**warlike 'Arry**". The effect is startlingly direct – faster, earthier, and physically grounded. Actors report that performing in OP connects them more viscerally to the text, stripping away centuries of reverential delivery. An OP *Romeo and Juliet*, for instance, raced along ten minutes faster than its modern-pronunciation counterpart. Most powerfully, OP restores lost meaning and humour. The famous, lofty Sonnet 116 ("Let me not to the marriage of true minds...") takes on a new character when its original rhymes are heard. And in *As You Like It*, a seemingly innocuous line – "from hour to hour we ripe and ripe... and thereby hangs a tale" – conceals a bawdy pun. In OP, 'hour' was pronounced '**ore**' (with the 'h' dropped), creating a clear pun on 'whore,' a joke entirely lost to the modern ear, cf. (Lindsey, 2019 : 73-74).

Thus, Original Pronunciation does not obscure Shakespeare; it illuminates him. It returns the vital pulse of rhyme, reveals hidden layers of wit, and reconnects the words to the robust, energetic world for which they were written.

11.1. Causes for Changes in Modern English

(Причины изменений в современном английском языке)

All languages are constantly changing. There are two reasons for these changes, two main factors:

1) internal factors which do with the nature of the language itself, caused by changes in the culture, society of the speakers and the environment in which this culture is set, and etc.;

2) external factors caused by other languages of the world and different varieties of English (American, African, West, Australian) spoken today – called extra-linguistic factors.

Historically English changed a good deal between 1350-1550 and 1700-1900. English is changing because it is becoming a world language. During the Reign of Queen Elizabeth I (1558-1603) only 5-7 million spoke English. Nowadays during the Reign of Queen Elizabeth II – 400 million speak English, 420 million use English as a second language (ESL) and 500 million speak English as a foreign language (EFL) and these figures continue to grow due to the following reasons: 1) increase of population in the USA (about 300 mln nowadays), Canada, and Australia, and exodus of people to English speaking countries; 2) mother tongues become inadequate for world communications (limited spheres); 3) if you want to get on in life you have to learn English; 4) after the fall of the Berlin wall, Russian has been replaced by English in East European countries; 5) 2/3 of world qualified scientists speak and write English; 6) 3/4 of world mail is written in English; 7) 85% of information is stored in computers in English.

Modern English is influenced to a greater extent by American English (AmE). Many linguists say that the world is a global village as a result of electronic interdependence in today's business world. Today business, commerce, films, radio and TV, advertising sections of many French, Spanish, Russian mass media are full of "American linguistic invasion". American English pushes British English a bit further along the way it is already going – the increasing uniformity of language inside Britain. With the great increase of world-wide communications, it is probable that various forms of English spoken today (American, African, Australian, Canadian, and etc.) are now influencing one another and a trend to greater mixing is taking place in English on a world scale. AmE has a rare forcefulness and flexibility and a greater permissiveness on points of pronunciation. BBC English was generally accepted throughout the UK as a suitable accent for such occasions as the reading of news bulletins and the making of announcements of nationwide importance. But fashions have changed and the movement in recent years has been towards a more relaxed and more personal style on the air. Most public speakers and BBC news readers adopt an informal "chatty style" to project a friendly image to their audience. American pronunciation is now completely accepted in Britain though American pronunciation 30-40 years ago was considered strange and difficult to understand. Common features in BBC English and AmE prevail over differences.

!

Vocabulary: “Americanization” of BrE

In the ever-evolving tapestry of the English language, American influence has long woven its threads through British lexicon, gifting us words like ‘movie’ and phrases like ‘do the math’, cf. (Lindsey, 2019 : 11-12). Yet, while vocabulary and pronunciation have drifted across the Atlantic for years, it is far rarer for the very structure of grammar to follow suit. Now, however, in the digital age, two subtle but significant grammatical shifts are taking root on British shores, propelled by a powerful new conduit: the internet, and online platforms.

The first of these shifts centres on the verb ‘get’. For many Britons of a certain generation, the past participle is unequivocally ‘got’. The American ‘gotten’ feels distinctly foreign. However, a simple poll among British viewers revealed a telling divide: while 71% still preferred “has got”, a substantial 29% – nearly a third – opted for “has gotten”, cf. (Lindsey, 2011). For these users, ‘gotten’ introduces a useful distinction, specifically meaning ‘has received’, whereas ‘has got’ remains pleasingly ambiguous, covering both possession and reception. This adoption of ‘gotten’ represents a conscious grammatical change that actually decreases ambiguity, a reversal of a centuries-old trend where the form had fallen dormant in Britain since the 1700s, only surviving across the Atlantic.

The second shift moves in the opposite direction, trading distinction for simplicity. This concerns the use of the simple past tense where traditional British grammar demands the present perfect. For instance, where a Brit might say “I’ve just eaten” to connect a recent action to the present moment, an American might say “I just ate”. This ‘present-relevant past’ is increasingly heard from Brits, particularly the young. Ironically, this change *increases* ambiguity, blurring the clear line British grammar draws between the recent, relevant past and the historical past.

The catalyst for both these changes is unmistakably the digital era. The pervasive language of American tech giants – from online notifications (“people followed me”) to advertising memes (“just got better”) – bombards young Brits daily. This exposure is intimate and unmarked; there is no American accent to flag it as ‘other’. It simply becomes part of the linguistic fabric of online life, cf. (Lindsey, 2011). Thus, the internet acts as a Trojan horse, quietly importing these grammatical patterns, with ‘gotten’ making a historic comeback and the simple past tense subtly rewriting the rules of recency. The story is one of grammatical frontiers quietly redrawn, not by decree, but by the relentless, personal tide of globalised digital communication.

!

Changes in pronunciation (exemplified by the word *covert*)

The tale of the word ‘covert’ /ˈkʌvərt, ˈkʊvɜːrt/ (a French borrowing into Middle English, related to the word *cover*) is one of mistaken identity, Cold War intrigue, and a quiet revolution in pronunciation. Contrary to popular belief, it is not merely the negated form of ‘overt’ /ˈoʊvɜːrt/ or /oʊˈvɜːrt/. While both have distant French roots, ‘covert’ truly stems from the verb ‘to cover’, giving it its original meaning of ‘covered’ or ‘concealed’.

For centuries, it was a somewhat marginal word, but its fortunes – and its sound – were dramatically transformed in the mid-20th century. As America’s clandestine Cold War activities surged, so did the need for a term to describe them. ‘Covert’ was plucked from obscurity and thrust into daily use by governments, spies, and journalists.

This sudden popularity among speakers unfamiliar with the word triggered a profound change. The traditional British pronunciation, **CO-vert** (with a weak second syllable), began to be remodelled by many to rhyme with 'overt', creating a neat pair of opposites: **co-VERT** and **o-VERT**. This 'spelling pronunciation' was likely helped by the word's new, technical aura and its familiar-looking parts (/co-/ and -vert), cf. (Lindsey, 2019 : 9-10).

The shift was led by the United States, becoming dominant there after President Lyndon B. Johnson – a noted user of the old form – left office. Britain followed, but with a lag of several years and a key phonetic difference: where Americans often stressed the second syllable but kept the first vowel short (co-VERT), most Brits and Southern Hemisphere speakers adopted a long vowel sound in that second syllable (co-**VERT**).

Despite the change, the old pronunciation proved resilient. Polls for the *Longman Pronunciation Dictionary* in the late 1980s showed over half of British respondents still preferred **CO-vert**. Even today, dictionaries often list both variants, a testament to a linguistic Cold War where two pronunciations have, for the moment, reached an uneasy stalemate.

11.2. Modern Types of RP (Received Pronunciation) (Современные виды RP, нормативного произношения)

Received Pronunciation, RP, has been long considered (at least in our country) the standard (i.e. orthoepic norm) and the only option for learning by foreigners. The main courses in practical phonetics of English, the classic ones probably being the books by the Soviet linguist V. D. Arakin (Аракин, 2017) and others were mostly prepared in the second half of the 20th century beyond the Iron Curtain presumably based on the relatively open materials 'at hand' which would have been the speeches by the head of the UK, and the BBC broadcast. As it was described above, the language both Elizabeth II and the BBC presenters spoke in the second half of the 20th century was RP which itself has been undergoing significant changes, is not perceived as the ultimate standard anymore and is spoken by a decreasing number of the native English speakers (giving way to the modern Standard Southern British, SSB).

!

Accents: The King's Speech

King Charles's accent is essentially the classic Received Pronunciation (RP) (Lindsey, 2011). This was the upper-class accent throughout much of the final century of the British Empire. Not many people still speak it, because the profound social shifts in Britain – beginning in the 1960s – swiftly rendered it rather posh and dated.

The modern standard accent in southern Britain is known as Standard Southern British (SSB). Charles was just old enough for his accent to have settled before the Swinging Sixties. Also fixed at that time was the set of vowel symbols used in most British English dictionaries. Consequently, King Charles is one of the few individuals who actually sounds like those symbols.

The King's vowels

In present-day SSB, the PRICE vowel starts quite far back in the mouth: /ɑj/, cf. (Lindsey, 2011). However, the King pronounces it with the more traditional symbol /aɪ/, beginning with a fronter quality. He also pronounces the DRESS and FACE vowels as the familiar /e/ and /eɪ/. In SSB, these are more open: /ɛ/ and /ɛj/.

On the other hand, in contemporary SSB the CHOICE vowel is rather close: /oj/, whereas the King uses the more open RP vowel: /ɔɪ/. When he says 'joy', it starts rather like 'job'. The King even retains a few vowels that are actually posher and more old-fashioned than the dictionary symbols. One is the GOAT vowel: the familiar symbol begins with a schwa (as in SSB): /ə/, /əʊ/, but the King begins it further forward in the mouth: /ɜʊ/. His MOUTH vowel is quite distinct from today's /aw/, starting somewhat further back and with no lip-rounding. Britons occasionally mimic this posh, old-fashioned sound by spelling 'house' as 'hice'. A difference in weak syllables involves the so-called *happy* vowel. SSB speakers today use the same vowel in both syllables of 'deeply', but in the King's RP the word ends with the lax /ɪ/ of KIT: *deep*/ɪ/.

Turning to **consonants**, one rather posh, old-fashioned feature is the weakening of /t/ to a sort of /s/ – heard, for example, in the second part of 'fighting' (almost like 'sing'). In SSB, /t/ is typically affricated, pronounced with a slight /s/-like release: /tʃ/. In the King's speech, when /t/ occurs in a weak position it often becomes a full fricative – a process called spirantisation, which can also be heard in Liverpool and Irish speech. Occasionally the King uses American-style voiced /t/, as in 'forty'. Yet one form of /t/-weakening that the upper class, including the King, never adopt is using a glottal stop before a vowel – as exemplified by the new Prime Minister, cf. (Lindsey, 2019 : 67-69).

Weakening is something the King does quite broadly; he has sometimes been described as **mumbling**. A somewhat reduced style of speech was typical of many RP speakers. There are words where, in the King's speech, sounds are noticeably reduced or absent altogether. The King does often stress words or syllables, but rather than articulating them more clearly, he frequently prefers to emphasise them with a gesture of his head or whole body.

!

Accents: the Princes' Speech

King Charles III and his sons, Prince William and Prince Harry, actually speak different accents. The fundamental contrast is between King Charles's Received Pronunciation (RP) accent – the traditional, upper-class pronunciation of the mid-20th century – and the contemporary Standard Southern British (SSB) accent used by William and Harry. This reflects significant social changes in Britain from the 1960s onwards, which made RP sound increasingly "posh" and led even the upper classes to modify their speech.

The key **vowel** distinctions between the accents may be summarized as follows. The King Charles uses /aɪ/ (starting with a fronter sound) in **PRICE**, while his sons use the SSB /ɑj/ (starting further back in the mouth). In words like 'while', SSB can smooth this to a monophthong, cf. (Lindsey, 2019 : 45-46). Charles uses closer vowels /e/ and /eɪ/ in words like **DRESS** & **FACE** respectively, whereas SSB uses more open versions /ɛ/ and /ɛj/. In the **CHOICE** vowel the King has a more open /ɔɪ/ (making 'joy' sound like 'job'), while Harry uses a closer SSB /oj/ ('joy' sounding like 'jaw'). The

SQUARE vowel is represented in RP by a diphthong /εə/, but SSB uses a steady monophthong /ε:/ . The final vowel in **happy** is the lax /ɪ/ of ‘kit’ in RP, and in SSB, it is the tense vowel of ‘tea’.

The **consonant** system is displayed as follows. William and Harry exhibit **L-vocalisation** (e.g., merging the PRICE vowel with a following ‘L’), which is less common in Charles’s speech, cf. (Lindsey, 2019 : 71-72). /t/ pronunciations also differ. King Charles *does not* use **glottaling**, i.e. glottal stops (e.g., for final /t/ before a vowel). William and Harry use them frequently, cf. (Lindsey, 2019 : 67-69). King Charles occasionally uses American-style /t/ **voicing** (or “flapping”). Harry does this more often, likely due to his time in the United States. **Affrication**, in this case releasing /t/ with an /s/-like sound is common in SSB. Charles sometimes does this, but William uses it extensively, to a degree that resembles features of Liverpool English (‘Scouse’).

Charles’s occasionally very relaxed, clipped **speech style** has been described as “mumbling”. William also speeds through or omits syllables, but his delivery sounds more rushed than relaxed. The King does not use several traits common in modern SSB, such as: **Uptalk** (a rising intonation on statements), cf. (Lindsey, 2019 : 99-108); **TH-fronting** (pronouncing ‘th’ as /f/ or /v/); **ejective /k/** (a glottalised ‘k’ sound), cf. (Lindsey 2019 : 75-76).

!

Accents: the Decline of RP

The dramatic and rapid decline of Received Pronunciation (RP) began in the 1960s, identifying 1962 as the pivotal year that marked the beginning of its end, cf. (Lindsey 2011).

The story begins with the publication in 1962 of A. C. Gimson’s classic description of RP, which established the vowel symbols still familiar today. However, the prestige of the accent underwent an astonishingly swift collapse almost immediately thereafter. This is starkly illustrated by comparing the first edition of Gimson’s book with the second edition just eight years later in 1970. The first edition presented RP as the firmly established speech of the ruling class. By 1970, the preface noted that RP had “lost ground”, especially among a younger generation who rejected it due to its association with the ‘Establishment’, preferring regional accents instead.

This seismic shift is framed against the backdrop of the 1960s, a period of profound social change. Key events included the end of National Service, the rise of working-class culture in media like *Saturday Night and Sunday Morning* and *Coronation Street*, and the launch of culturally defining phenomena like The Beatles’ *Love Me Do* and the film *Dr. No*. Crucially, these new icons flouted RP: The Beatles had Liverpool accents, and James Bond was played by a working-class Scot, not an RP speaker.

The narrative then explores Gimson’s own 1962 categorization of RP, which included a forward-looking variant called **Advanced RP**. This was the speech of fashionable young people, featuring traits like diphthongising the DRESS vowel, using a very back PALM vowel, and monophthongising diphthongs (making ‘loud’ sound almost like ‘lard’, for more details on the broadening of vowels see Lindsey, 2019 : 41-43). Ironically, within eight years, this very “advanced” accent became the most ridiculed. Its overt prestige vanished, replaced by the “covert prestige” of working-class speech. The powerful force was no longer attraction to RP, but **repulsion** from it, as the accent became stigmatised for symbolising unfair privilege.

Linguistically, the collapse of RP's prestige acted like removing a corset, allowing the tightly compressed short front vowels to relax into a more natural, open distribution – a shift even adopted to a degree by the Queen. This was separate from the later, more deliberate adoption of London features like T-glottalling.

Still, there is a paradoxical legacy of 1962. While the RP vowel system was doomed, **Gimson's 1962 phonetic symbols were fossilised**, continuing to be taught worldwide for decades as the definitive guide to British pronunciation, even as the accent they described ceased to be the standard. Furthermore, the term 'RP' itself became problematic; within Britain, it is now often used only to describe markedly posh speech, while linguists have largely replaced it with **Standard Southern British (SSB)**. The narrative posits that classic RP, as the speech of a specific privileged class, is now a historical phenomenon, cf. (Lindsey, 2011).

!

“Villainous” Accents

There has long been an enduring Hollywood trope of casting villains with British accents, examining its origins and nuances beyond a simple explanation, cf. (Lindsey, 2011).

It begins by noting the common cinematic device where a British-accented antagonist corrects an American hero, a detail audiences readily perceive even when directors claim it is minor. The observation is made that in Hollywood films, villains are frequently portrayed either by British actors or by Americans adopting British accents.

While many might first attribute this trend to **colonial history** – a lingering American resentment from the war of independence – this is not a complete explanation. It points out that Hollywood's founding studio heads were not from colonised nations and were historically very Brit-friendly. Furthermore, not all British accents carry the same villainous weight.

The narrative clarifies that the archetypal “villainous accent” people imagine is, or is based upon, **Received Pronunciation (RP)**. This accent solidified alongside the industrial revolution and the Second British Empire, becoming the hallmark of a new, powerful social elite. Within British storytelling itself, RP has long functioned as a **domestic accent of evil**, often spoken by powerful antagonists while sympathetic underdogs use more local or working-class accents.

This hierarchical dynamic extends beyond the UK. In countries like Australia, accents are still socially stratified in implicit reference to RP. Even in the USA, traces remain in the older upper-class accents of New England and New York, which share more features with RP than General American does, and RP historically influenced the **American Theater Standard**.

There is also a shift: today's **Standard Southern British (SSB)** accent is less extreme than the classic RP of the past. Therefore, in modern films where all characters are English, filmmakers often “crank up the posh” to an exaggerated degree to signal that a character is unsympathetic or evil.

11.3. The Main Changes in the Systems of the English Vowels & Consonants (Основные изменения в системе английских гласных и согласных фонем)

Below we quote some changes traced in Modern English notwithstanding its type or region of use:

1) in the system of vowels:

a) historically long vowels [i:], [u:] become diphthongized [iⁱ], [u^u];

b) there is a tendency for some diphthongs to become shorter and they sound like pure vowels – fire – [fa:], tire – [ta:], poor – [po:] – especially among young speakers *careful* – ['kɜ:ful] becomes ['kɜ:ful];

c) the vowel phoneme [æ] is replaced by [a] – [hæv] becomes [hav];

d) centralization of vowels – [jes] tends to be [jəs], conservative speakers say [nou], so [sou], [nout], whereas young speakers say [nəu], [səu], [nəut] for *no*, *so*, *note*, correspondingly, also ['ɒnist] tends to be ['ɒnəst] for *honest*, also [bət] for *but* and [gəd] for *good*,

e) change of [ju:] into [u:] after [t], [d], [n], [s], or [z], e.g.: *news*, *suit*, *student*, *assume*, *due*, *dew*, but after [p], [b], [m], [f], [v], [k], [g], [j] it is retained, cf.: *few*, *fume*, *beauty*, *kew*,

f) a tendency to lengthen short vowels in monosyllabic words, e.g.: *bad* [bæ:d], *Dad* [dæ:d], *jam* [dʒæ:m], *bag* [bæg], *big* [bi:g], *his* [hi:z], *is* [i:z];

g) young people use [æ] instead of [a:], e.g.: *can't* [kænt], *dance* [dæns], *chance* [tʃæns], and etc.

2) in the system of consonants:

a) voicing of [t] in intervocal position ['bridɪf], ['sɪdi];

b) word-final [b], [d], [g] are termed as lenis voiceless;

c) loss of [h] in rapid speech, e.g.: *have*, *him*, *his*, *her* – [aɪ'təʊld əɾ];

d) loss of plosives, e.g.: *sit down* [si dæun];

e) sound combinations [tj], [dj], [sj] are pronounced as [tʃ], [dʒ], [ʃ], e.g.: *graduate* ['grædʒuɪt], *actual* ['æktʃuəl], *issue* ['ɪʃu:];

f) spread of dark [ɪ] in all positions;

g) initial [hw] in wh-words, e.g.: *why* [hwai] and [hwæi], *where* [hwɜə] and ['hwere], *when* [hwen], *what* [hwɒt], *which* [hwɪtʃ], and etc.;

h) intrusive [r] on analogy with linking [r], e.g.: *Asia* [r] and *Africa*, *idea* [r] of it, *law* [r] of the sea. NB: two last cases at least – g) and h) – can be traced back to the Southern pronunciation of BrE, see, for instance, (Поливанов, 2004).

Revision: Section 11 (Повторение: Раздел 11)

1. What are the two main categories of factors causing language change, and what is the primary external factor influencing English today?
2. What is the primary external factor influencing English today?
3. What are the main reasons for the number of the English language speakers to grow?
4. What has been the recent trend in the style of pronunciation, marking a shift from traditional RP?

5. What are the specific changes in the vowel system noted in Modern English?
6. What are the specific changes in the consonant system noted in Modern English?
7. Based on the comparison of King Charles III's speech and that of his sons, what are the key phonetic differences between classic RP and modern Standard Southern British (SSB)?
8. When and why did RP start to decline?
9. How is the modern American influence reshaping the British English grammar?
10. Explain the phonetic changes in a language caused by the so-called 'spelling pronunciation'.

12. Revision: Final Test

(Повторение: ИТОГОВЫЙ ТЕСТ)

1. Phonetics is defined as the science that studies ...
 - A. speech sounds only
 - B. the sound system of language, phonemes, stress, syllable structure, and intonation
 - C. written symbols of language
2. The Greek word *phōnē*, from which the term *phonetics* originates, means ...
 - A. language
 - B. sound or voice
 - C. articulation
3. Why is the definition of phonetics considered complicated?
 - A. Because speech sounds are abstract
 - B. Because speech sounds are studied only in linguistics
 - C. Because speech sounds are also studied in acoustics and physiology
4. Which objects are fundamental for a phonetic study?
 - A. Phonemes and morphemes
 - B. Phonemes and speech sounds
 - C. Words and sentences
5. Articulatory phonetics focuses on ...
 - A. perception of speech sounds
 - B. physical properties of sound waves
 - C. production of speech sounds by the speech organs
6. Acoustic phonetics studies speech sounds as ...
 - A. linguistic signs
 - B. psychological phenomena
 - C. physical vibrations in the air
7. Auditory phonetics is mainly connected with ...
 - A. anatomy
 - B. experimental psychology
 - C. historical linguistics
8. The term *anthropophonics* was introduced by ...
 - A. R. Jakobson
 - B. Baudouin de Courtenay
 - C. N. Trubetzkoy
9. General phonetics studies ...
 - A. sound-producing possibilities of the human speech apparatus
 - B. one national language
 - C. spelling systems
10. Special phonetics focuses on ...
 - A. universal sound laws
 - B. phonetic systems of individual languages
 - C. speech disorders
11. Historical phonetics is concerned with ...
 - A. present-day pronunciation norms
 - B. future language change
 - C. sound changes through different historical stages
12. Descriptive phonetics studies a language ...
 - A. at one stage of its development
 - B. across several centuries
 - C. only in dialectal forms
13. Comparative phonetics establishes ...
 - A. stylistic norms
 - B. correspondences between related languages
 - C. speech perception mechanisms
14. Theoretical phonetics ...
 - A. deals with theoretical issues and debatable questions
 - B. solves practical pronunciation problems
 - C. studies only laboratory data

- 15.** Applied phonetics is primarily concerned with
 A. practical use of phonetic knowledge
 B. abstract theory
 C. language history
- 16.** Experimental phonetics is characterized by
 A. observation only
 B. laboratory conditions and instruments
 C. purely theoretical reasoning
- 17.** Phonetics is connected with grammar mainly through
 A. sentence structure
 B. rules of reading and pronunciation
 C. word meaning
- 18.** The pronunciation difference in *leaf* – *leaves* demonstrates the link between phonetics and
 A. syntax
 B. stylistics
 C. grammar
- 19.** Stress differences in /'object/ – /ob'ject/ relate phonetics to
 A. lexicology
 B. morphology
 C. phonotactics
- 20.** Intonation primarily connects phonetics with
 A. stylistics
 B. lexicography
 C. etymology
- 21.** Speech sounds as sound waves can be analysed from the perspectives of both physiology and
 A. psychology
 B. psychiatry
 C. acoustics
- 22.** The articulatory stage as a coding activity in phonetics is described as
 A. psychological
 B. physiological
 C. physical
- 23.** Phonetics is related to psychology because it studies
 A. social identity
 B. encoding and decoding of speech
 C. dialect classification
- 24.** Sociological relevance of phonetics lies in
 A. biological constraints
 B. writing systems
 C. social variation and identity
- 25.** Logical stress serves to
 A. mark grammatical tense
 B. create rhyme
 C. express logical relations and focus
- 26.** The direct observation method includes observation
 A. by ear, sight, and muscular sensation
 B. only by ear
 C. only by instruments
- 27.** The experimental method differs from observation because it
 A. creates artificial conditions
 B. excludes listening
 C. avoids hypotheses
- 28.** The laryngoscope allows observation of
 A. tongue movement
 B. vocal cords
 C. nasal airflow
- 29.** A palatogram shows
 A. sound frequency
 B. tongue–palate contact
 C. air pressure
- 30.** Oscillography represents
 A. lexical meaning
 B. grammatical categories
 C. acoustic characteristics of sound
- 31.** Loudness corresponds to
 A. frequency
 B. amplitude
 C. spectrum

- 32.** Pitch is correlated with
 A. frequency
 B. amplitude
 C. timbre
- 33.** Timbre corresponds to
 A. spectrum
 B. amplitude
 C. duration
- 34.** Formants are defined as
 A. speech organs
 B. frequency maxima in the spectrum
 C. phonological units
- 35.** The fundamental frequency is denoted as
 A. F1
 B. F0
 C. F2
- 36.** Supra-glottal cavities primarily influence
 A. formant structure
 B. spelling
 C. grammatical meaning
- 37.** According to Helmholtz's theory, supra-glottal cavities function as
 A. generators
 B. filters
 C. amplifiers
- 38.** The auditory aspect of speech is described as
 A. fully understood
 B. least studied
 C. irrelevant
- 39.** The cochlea belongs to
 A. outer ear
 B. middle ear
 C. inner ear
- 40.** The articulation basis of a language is
 A. a set of phonemes
 B. habitual articulatory tendencies
 C. spelling tradition
- 41.** The articulation basis is transmitted
 A. socially from generation to generation
 B. genetically
 C. through writing
- 42.** English articulation is characterized by
 A. advanced tongue tip
 B. retracted tongue tip and alveolar tapping
 C. inactive tongue
- 43.** The most important defining feature of English pronunciation is
 A. vowel length
 B. consonant clusters
 C. stress
- 44.** Phonetic interference in bilingualism refers to
 A. grammatical errors only
 B. influence of the native sound system
 C. spelling mistakes
- 45.** Phonemes relate to speech sounds as
 A. form and content
 B. cause and effect
 C. sound and noise
- 46.** Speech sounds have
 A. one aspect
 B. two aspects
 C. three aspects
- 47.** The distributional method studies
 A. meaning contrasts
 B. acoustic spectra
 C. sound distribution in positions
- 48.** The method of minimal pairs belongs to
 A. articulatory analysis
 B. phonological analysis
 C. acoustic analysis
- 49.** The ultimate aim of phonological analysis is to
 A. describe intonation
 B. establish the phoneme inventory
 C. explain spelling rules

- 50.** The speaking–hearing process model consists of
 A. three stages
 B. four stages
 C. five stages
- 51.** The first stage of the process is
 A. psychological (coding)
 B. articulatory
 C. acoustic
- 52.** The final stage of the process is
 A. acoustic
 B. auditory
 C. psychological (decoding)
- 53.** Literary pronunciation is
 A. a standardized norm
 B. a dialect
 C. an individual variant
- 54.** Literary pronunciation is associated with
 A. regional identity
 B. education and public speech
 C. informal communication
- 55.** General American pronunciation is mainly used in
 A. rural speech
 B. broadcasting and public communication
 C. historical linguistics
- 56.** The Great Vowel Shift primarily affected
 A. long vowels
 B. consonants
 C. stress patterns
- 57.** The Great Vowel Shift involved
 A. lowering only
 B. consonant palatalization
 C. raising and diphthongization
- 58.** It is rejected that Modern English has
 A. diphthongs
 B. paired short and long vowels
 C. stress contrasts
- 59.** English checked vowels occur
 A. before voiceless consonants
 B. in open syllables only
 C. only in unstressed positions
- 60.** Apical articulation dominates in
 A. English
 B. Russian
 C. both equally
- 61.** Russian consonant articulation is mostly
 A. apical
 B. dorsal
 C. glottal
- 62.** How many zones of consonant articulation does English have?
 A. four
 B. five
 C. six
- 63.** Aspiration is
 A. absent in English
 B. absent in Russian
 C. identical in both
- 64.** Phonetic styles are studied in
 A. phonology
 B. phonostylistics
 C. lexicology
- 65.** Culture of speech is discussed within
 A. phonostylistics
 B. historical phonetics
 C. acoustic phonetics

Keys (Ключи)

Revision: Section 1 Answers (Повторение: ответы, Раздел 1)

1. Phonetics is defined as a branch of linguistics which studies the sound system of language and is concerned with speech sounds, word stress, syllabic structure, and intonation, as well as the relation between written and spoken language.

2. The three branches are: articulatory phonetics, which studies speech sounds as regards their production by the human speech apparatus; acoustic phonetics, which studies the physical properties of sound waves; auditory phonetics, which investigates the hearing process and perception of speech sounds.

3. General phonetics studies all the sound-producing possibilities of the human speech apparatus and the laws governing sound changes in various languages, whereas special phonetics analyses and describes the phonetic system of a particular language.

4. Descriptive phonetics analyses the phonetic system of a language at a single point in time; historical phonetics traces successive changes in the phonetic system of a language at different stages of its development; comparative phonetics studies correspondences between the phonetic systems of two or more languages, especially kindred ones.

5. Theoretical phonetics deals with theoretical problems and views on phonetic issues; applied phonetics is concerned with the practical utilization of phonetic knowledge; experimental phonetics studies phonetic phenomena under laboratory conditions using instruments and special equipment.

6. Phonetics is connected with grammar through rules of reading and pronunciation of grammatical forms (e.g. -ed endings, plural nouns); with lexicology through word stress, which helps to distinguish parts of speech and differentiate homographs; with stylistics through intonation, speech melody, rhythm, word stress, tempo, and loudness, which help to differentiate styles of speech, express emotions, attitudes, and distinguish types of utterances.

7. It is difficult to give an exact definition because speech sounds are studied not only in phonetics but also in acoustics and physiology, and because phonetics encompasses both material (physical) and ideal (functional) phenomena of language.

8. The main methods of research in phonetics are the direct observation method, the experimental method, instrumental methods, and methods of phonological analysis, including distributional, statistical, and semantic (commutation) methods aimed at establishing the inventory of phonemes.

9. Phonetics is connected with acoustics, anatomy, physiology, psychology, sociology, logic, neuroscience, cognitive science, mathematics, statistics, computer technology, telecommunications, and semiotics.

10. The speaking–hearing process includes: 1) the psychological (coding) stage, 2) the physiological (articulatory) stage, 3) the physical (acoustic) stage, 4) the physiological (auditory) stage, and 5) the psychological (decoding) stage.

Revision: Section 2 Answers (Повторение: ответы, Раздел 2)

1. The articulation basis is defined as the unity of habitual movements and positions of the articulatory organs for the given language, or as “a total of articulatory tendencies and speech-motor of the speakers of a given language”.

2. The articulation basis comprises differences in tension, tongue shape, pressure of the articulators, and lip, cheek, and jaw posture and movement, which run through the whole articulatory process.

3. The articulation basis may be considered as a language tradition and should be understood as communication in the given language from one generation to another.

4. Although many languages have identical articulatory organs and similar sounds, these sounds differ from the articulatory point of view because each language has its own articulatory tendencies and peculiarities that are part and parcel of native speakers.

5. The articulatory setting of English is characterised by a retracted tongue position, the tongue spread at the back and lightly touching the rear molars, with the tip of the tongue tapping the alveolar ridge to produce sounds such as /t/, /d/, /n/, and /l/.

6. The articulatory approach treats pronunciation as a physical motor skill and focuses on developing articulatory awareness, mastering stress, understanding aspiration, and adopting the English articulatory setting.

7. In English and Russian only the tongue position and lip position coincide, while vowel length, degree of tenseness, and stability of articulation are characteristic of English only; Russian vowels show stronger lip rounding and protrusion.

8. English consonants have six zones of articulation compared to four in Russian; apical articulation dominates in English, dorsal articulation in Russian; aspiration is present in English and absent in Russian.

9. Literary pronunciation is a standardised, normative form of pronunciation accepted as the norm for public speaking and educated speech and opposed to dialects, accents, and regional variants.

10. Phonetic interference refers to the influence of the native language’s sound system on the pronunciation of a second language, including sound substitution, stress and intonation patterns, and difficulties with sounds absent in the native language.

Revision: Section 3 Answers (Повторение: ответы, Раздел 3)

1. The phonetic system of a language is understood as the totality of its phonemes and the rules governing their functioning, including their classification, distribution, and combination in the flow of speech.

2. It is the smallest (i.e., further indivisible into smaller consecutive segments) language unit (sound type) which are capable of distinguishing one word from another word of the same language or one grammatical form of a word from another grammatical form of the same word.

3. Allophones are positional and combinatory variants of the same phoneme which do not differentiate meaning and occur in different phonetic environments.

4. Allophones are divided into: (1) principal allophones, and (2) subsidiary allophones.

5. The three aspects of the phoneme are: (1) material; (2) abstractional (generalized); (3) functional.

6. The phoneme performs in speech the following functions: (1) constitutive (because sounds constitute words, phrases and sentences which are meaningful); (2) distinctive (because sounds help to distinguish one word from another); (3) recognitive – because sounds (or allophones) help to recognize words.

7. The difficulty of giving a definition of the phoneme lies in the fact that the phoneme has several aspects and functions that must be combined in one definition.

8. A speech sound is a concrete phonetic unit produced in speech, whereas a phoneme is a language unit capable of differentiating meaning and existing as an abstract sound type.

9. J. Baudouin de Courtenay gave a mentalistic view on the phoneme defining it as ‘mental image of a sound’.

10. Initially, L. V. Scherba defined it as an abstract element associated with meaning, but later in his book he gave a materialistic definition of the phoneme, describing it as sound types represented by actual varied sounds (allophones).

11. The most general and the old distinction in the phonemic composition of a language is between vowels and consonants.

12. The Great Vowel Shift was a major series of sound changes that altered the pronunciation of long vowels in English through raising and diphthongization and destroyed phonemically paired short and long vowels, replacing them with qualitatively different vowels and diphthongs.

13. English vowels classified as tense and lax according to the degree of tenseness (muscular tension); as historically long and historically short according to the length; and as monophthongs, diphthongs, and diphthongoids according to the stability of articulation.

14. The orthographic form of a word does not always correspond to the sequence of sounds produced in speech because pronunciation reflects the phonetic and phonological system, whereas spelling preserves historical forms.

15. The speech continuum (i.e. sound stream) consists of phonetic units such as the phonetic word, the syntagm, the phrase, the phonetic paragraph (phono-paragraph), and the text.

Revision: Section 4 Answers (Повторение: ответы, Раздел 4)

1. Phonetics studies the material, physical, and physiological aspects of speech sounds, whereas phonology studies the purely linguistic, functional aspect of speech sounds.

2. The Prague phonological school is characterised by the functional approach to phonology and the theory of phonological oppositions developed by N. S. Trubetzkoy as well as the notion of the archiphoneme for a unit higher than the phoneme.

3. The Moscow phonological school is characterised by its emphasis on the morphemic principle and its methods of phonological analysis based on morphological alternations, whereas the St.-Petersburg (Leningrad) phonological school treats phonemes as sound types closely connected with concrete phonetic realizations and gives priority to phonetic criteria.

4. The method of phonological analysis comprises a system of principles and procedures aimed at establishing the phonemic inventory of a language.

5. It establishes the distribution (all positions and combinations) of speech sounds in the words of a language.

6. It establishes the frequency, probability, and predictability of occurrence of phonemes and their allophones.

7. It establishes the ability of sound segments to differentiate meaning by substituting sounds in a constant phonetic environment.

8. Phonological features are properties of speech sounds that are capable of differentiating meaning and are used to establish phonemic oppositions in a language.

9. A phonological opposition is a relation between phonemes based on the presence or absence of distinctive features and forming a system of oppositions in a language.

10. Two main types of transcription are distinguished: phonemic (connected with the abstractional aspect of phonemes), and phonetic (or allophonic; connected with the material aspect of phonemes); whereas the International Phonetic Alphabet (IPA) is used as a system of symbols for the consistent representation of speech sounds.

Revision: Section 5 Answers (Повторение: ответы, Раздел 5)

1. Modification of phonemes in speech refers to changes in the realization of phonemes caused by their position and combinability in the flow of speech.

2. Allophonic variations are different realizations of the same phoneme which do not differentiate meaning and are determined by phonetic environment.

3. The main causes of allophonic variation are the influence of neighbouring sounds, the position of a phoneme in a word or utterance, and the conditions of speech.

4. Positional changes depend on the position of a phoneme in a word or syllable, while combinatory changes depend on the influence of adjacent sounds.

5. "Assimilation; Accommodation; Elision; Reduction".

6. Assimilation is a type of combinatory phoneme modification in which one sound becomes more like a neighbouring sound under the influence of that sound; Assimilation may be progressive, regressive, or reciprocal, and may also be complete or incomplete

7. Accommodation is a type of phoneme modification resulting from the mutual influence of vowels and consonants within a syllable.

8. Elision is the loss of a sound or sounds in connected speech as a result of phonetic context or speech economy.

9. Reduction is the weakening of vowels in unstressed positions, resulting in qualitative and quantitative changes.

10. Phoneme modifications reveal regular phonetic laws of connected speech and explain the difference between isolated pronunciation and natural fluent speech.

Revision: Section 6 Answers (Повторение: ответы, Раздел 6)

1. A syllable is a minimal unit of speech consisting of a vowel sound or a vowel with surrounding consonants and forming a single articulatory and acoustic unit; the main role of the syllable in speech is in that it serves as a basic unit of rhythm and articulation in speech.

2. The articulatory, acoustic, auditory, and functional approaches to the definition of the syllable are distinguished.

3. According to the articulatory theory, a syllable is formed by one muscular contraction of the respiratory muscles.

4. The sonority theory states that a syllable is formed by sounds arranged according to their degree of sonority, with the most sonorous sound forming the syllable peak.

5. The key features are pitch (speech melody), rhythm, stress, loudness, tempo, and timbre.

6. Syllables may be open or closed, covered or uncovered, and simple or complex.

7. Syllable division is the process of dividing a word into syllables according to phonetic and phonological principles; the phonetic, phonological, and morphological principles of syllable division are applied in English.

8. Vowels and sonorants are syllabic in English.

9. A syllabic consonant is formed when a weak vowel is followed by [n], [l], or [r]: V(weak) + [n]/[l]/[r] = C(syllabic). Example: garden ['gɑ:dn].

10. It proves that neither a morpheme boundary nor a syllable boundary can pass within a diphthong; the nucleus and glide cannot be separated.

Revision: Section 7 Answers (Повторение: ответы, Раздел 7)

1. Word stress is the greater prominence of one syllable in a word as compared to the others.

2. Word stress may be fixed or free.

3. Word stress performs constitutive (organising), distinctive (meaning-distinguishing), and cognitive functions.

4. Word stress differentiates words and grammatical forms, e.g. English 'import – noun vs im'port – verb; Russian пеку́ (genitive singular) – пе́ки (nominative or accusative plural).

5. If the first word is a noun, stress is on the first word (e.g., BATHroom, GEOGRAPHY professor). If the first word is an adjective, stress is on the second word (e.g., small ROOM, nutty PROFESSOR).

6. When a noun is formed by adding some suffixes (i.e. -ion) to a verb, the primary stress shifts to the syllable immediately preceding the suffix. The original verbal stress is retained as a secondary stress

7. The dynamic stress is achieved through physical force, while the musical stress is achieved by the specific pitch contour that carries meaning.

8. When determining the stressed syllable, grammatical endings (i.e. inflection) and productive suffixes are not taken into account.

9. In the standard literary Russian language, a qualitative reduction occurs where distinct unstressed vowels (like /a, o, ə/) in the first pre-stressed syllable merge into a single sound, typically an /и/-like vowel after palatalised consonants.

10. In disyllabic words, stress typically falls on the first syllable, unless the first syllable is a semantically weak prefix, and in polysyllabic words, the primary stress most commonly falls on the third syllable from the end.

Revision: Section 8 Answers (Повторение: ответы, Раздел 8)

1. Intonation is a complex unity of phonetic features that constitute a sentence, like speech melody, sentence stress, rhythm, tempo, and timbre.

2. Intonation can change the total meaning of the utterance by the use of pitch.

3. The two parameters are pitch level (determined by the pitch of the first stressed syllable) and pitch range (the interval between the highest and lowest pitched syllables). They are interdependent: the higher the pitch level, the wider the range.

4. In English speech, the rises and falls of pitch (intonation) are structured, language-specific patterns that are consciously controlled by the speaker to carry linguistic and communicative meaning, as opposed to being random or physiologically inevitable fluctuations.

5. English has a stress-timed rhythm, where the time between stressed syllables tends to be equal. Consequently, all unstressed syllables between stresses are compressed as far as possible to fit this regular beat.

6. Excitement is usually accompanied by an increase in loudness and tempo, using the upper pitch range. Depression correlates with lower pitch, slow tempo, and quietness. Semantically important parts sound slower, while unimportant parts are pronounced faster.

7. Vocal fry is characterized by very low-frequency, often irregular vibrations of the vocal cords. Its most common place in English is as the end part (the aftermath) of a falling intonation pattern.

8. Pausation is a controlled, functional mechanism—not just random silence—that is critical for organizing, producing, and comprehending spoken language.

9. Duration can be measured acoustically in milliseconds (ms), though many speakers perceive sounds as long or short.

10. Intonation has the following functions: (1) communication function (conveys meaning and shows which words carry new information), 2) sentence-forming (constitutive) function (joins words into sense-groups and sentences using pitch and stress), 3) sentence-delimiting function (divides longer sentences into shorter sense-groups), 4) distinctive function (distinguishes communicatively different types of sentences), 5) attitudinal function (shows the speaker's emotional attitude to the situation).

Revision: Section 9 Answers (Повторение: ответы, Раздел 9)

1. The recent movement has been towards a more relaxed and personal style. Most public speakers and BBC news readers now adopt an informal “chatty style” to project a friendly image.

2. Accent in the meaning of the type of articulation is a phonetic feature of a person’s speech which refers to specific regional or foreign phonetic traits (e.g., a Russian accent in English).

3. The style of pronunciation implies the speaker’s adoption of different registers, from formal to highly informal colloquial speech.

4. Colloquial speech is the informal, spontaneous way native speakers talk in everyday situations. It is characterized by unplanned delivery, the use of contractions, and reliance on gestures and facial expressions.

5. Pronunciation has two primary styles: a conversational style, which is relaxed, everyday speech that tolerates variations, and a public (or bookish) style, used in formal settings like lectures, which requires precise, norm-following pronunciation and avoids deviations.

6. Speech culture is a discipline concerned with the strict adherence to the standards of the literary language, investigating the refinement of language as an instrument of culture, while speech etiquette is the nationally specific and socially conditioned conventions governing behaviour during interaction, constituting a system of established communicative formulas, prescribed by society as rules of verbal conduct.

7. The subject matter of speech culture includes the standardisation of the literary language, the principles and rules of verbal communication, and the ethical norms of interaction; whereas its parameters are logicity, lexical richness, purity, grammatical correctness, vocabulary range, intonation, adherence to speech conventions, good manners, and facial expression.

8. The components of speech etiquette components are: (1) politeness formulas, (2) forms of address, (3) adherence to standards of speech, (4) observance of linguistic and value-based taboos, (5) the use of non-verbal means of communication.

9. A phonetic style, or phonetic style is a set of segmental and suprasegmental features applied in a communicative context or extra-linguistic situation to fulfil specific functions.

10. Phonetic styles are classified into: informational (neutral, factual), scientific (intellectual, persuasive), publicistic (persuasive, influential), artistic (emotional, expressive), and conversational (informal, diverse).

Revision: Section 10 Answers (Повторение: ответы, Раздел 10)

1. The two main norms considered when studying English are pronunciation norms exemplified by standards like Received Pronunciation (BBC English) in British English and General American in American English.

2. The study of variation, or variantology, examines how language is stratified. This includes regional variantology, which studies geographically based language forms, and social variantology whose scope is research of language forms linked to social groups, etc.

3. Within British English, three principal types exist: Southern English Pronunciation (Received Pronunciation or its modern successor, Standard Southern British), Northern English Pronunciation, and Scottish Pronunciation. They differ in features like vowel realization and rhoticity.

4. BBC English (or Received Pronunciation) was generally accepted in the UK for national broadcasts and announcements, and General American, which is the form of speech used by radio and TV in the USA for cultural, scientific, and business intercourse, and is used to sound educated and professional.

5. The three types are: (1) Eastern American, (2) Southern American, and (3) General American.

6. The Southern type is characterized by a specific way of diphthongization of pure vowels and monophthongization of some diphthongs at the expense of “drawling” (prolonging their nuclei and dropping the glide).

7. General American serves as the professional norm, yet local accents remain. Educated individuals often switch between the standard at work and local speech at home. However, American English shows a lesser degree of dialect than British English due to high mobility.

8. Key consonantal features of General American include: (1) [r] is retroflex and pronounced in all positions (rhotic), (2) [t] is voiced in intervocalic positions, (3) [l] is dark in all positions, (4) [h] is often lost in pronouns, (5) after [t], [d], [n], [s], [z], the sound [j] is dropped, (6) [ʃ] can be vocalized to [ɜ] in final unstressed syllables.

9. The main differences in General American vowels include: (1) no strong opposition between short and long vowels; [i:] and [u:] are diphthongized, (2) 5 diphthongs: [eɪ], [aɪ], [ɔɪ], [aʊ], [əʊ]/[ɔʊ], (3) the three RP vowels [ɒ], [æ], [ɑ:] correspond to only two vowels in GA: [a] and [æ], (4) the vowel [æ] is longer and often nasalized before [d], [m], [n], (5) [r] is treated as a semi-vowel and influences the preceding vowels.

10. The Australian English accent emerged from the mixing of southeastern English dialects brought by early settlers, which were heavily influenced by both the prestige of Received Pronunciation (RP) and the working-class tones of Cockney speech.

Revision: Section 11 Answers (Повторение: ответы, Раздел 11)

1. Changes are caused by internal factors (related to the nature of the language, culture, and society) and external factors (influence from other languages and varieties of English).

2. The primary external factor is the influence of American English (AmE), due to global communications, business, and media.

3. The number of English speakers continues to grow primarily due to demographic expansion in core English-speaking nations and ongoing immigration to them, as well as by the role of English as a global lingua franca, essential for socioeconomic advancement and

international communication, it being the primary language for most of the world's scientists, international mail, and digitally stored information.

4. A new style of pronunciation, Standard Southern British, SSB, has emerged replacing the previously accepted RP standard.

5. Modern English is witnessing notable shifts in its vowel system. Historically long vowels are becoming diphthongized, while a counter-trend sees some diphthongs shortening to sound like pure vowels, particularly among younger speakers. The system is also marked by a widespread centralization of vowels and specific phoneme replacements, such as [æ] being replaced by [a]. Furthermore, there is a loss of the [j] glide after certain consonants and a tendency to lengthen short vowels in monosyllabic words.

6. Several key changes are occurring in the consonant system of Modern English. There is a noticeable voicing of [t] in intervocalic positions and a tendency for word-final voiced plosives to become lenis voiceless. In rapid speech, the [h] sound is often lost, and plosives may be dropped altogether. Furthermore, specific sound combinations are simplifying through assimilation, such as [tj], [dj], [sj] becoming [tʃ], [dʒ], [ʃ]. Other widespread features include the spread of dark [ɫ] to all positions, the use of an initial [hw] in wh-words, and the insertion of an intrusive [r] between vowels. Many of these features, like the intrusive [r] and [hw] pronunciation, originate from Southern British English.

7. Some of the key phonetic differences between classic RP and modern SSB include: (1) The PRICE vowel: RP begins with a more front [aɪ] (Charles), while SSB begins with a backer [ɑj] (William / Harry); (2) T-glottaling: SSB speakers (William / Harry) frequently use a glottal stop for /t/ before a vowel, a feature typically avoided in classic RP (Charles).

8. The decline of Received Pronunciation (RP) began dramatically in the 1960s driven by the profound social changes of the 1960s, with 1962 serving as a pivotal year. The era's new cultural icons, like The Beatles with their Liverpool accents and a working-class Scottish James Bond, explicitly flouted RP, giving "covert prestige" to working-class speech. The accent became stigmatized as a symbol of unfair privilege.

9. There is a resurgence of the American past participle "gotten" for the verb "get", which is employed by approximately one-third of Britons to explicitly denote the meaning of "has received" and the American simple past tense, exemplified by phrases such as "I just ate", is increasingly supplanting the British present perfect tense, as in "I've just eaten".

10. Phonetic changes in a language can be caused by the so-called 'spelling pronunciation'. This can be best exemplified by the word covert and its development of standard pronunciation. The reason for the original pronunciation /'kʌvərt/ to change into /'koʊvɜ:t/ was its sudden increase in popular usage among speakers unfamiliar with its traditional pronunciation and remodelled it based on its spelling to create a logical, analogical pair with its apparent opposite, overt /'oʊvɜ:t/.

Final Test Answers (*Ответы итогового теста*)

1. B	23. B	45. A
2. B	24. C	46. C
3. C	25. C	47. C
4. B	26. A	48. B
5. C	27. A	49. B
6. C	28. B	50. C
7. B	29. B	51. A
8. B	30. C	52. C
9. A	31. B	53. A
10. B	32. A	54. B
11. C	33. A	55. B
12. A	34. B	56. A
13. B	35. B	57. C
14. A	36. A	58. B
15. A	37. C	59. A
16. B	38. B	60. A
17. B	39. C	61. B
18. C	40. B	62. C
19. A	41. A	63. B
20. A	42. B	64. B
21. C	43. C	65. A
22. B	44. B	

References

(Литература)

- Аракин, В. Д. Практический курс английского языка. 1 курс : учебник для студентов высших учебных заведений / В. Д. Аракин, Л. И. Селянина, К. П. Гинтовт [и др.] ; под ред. В. Д. Аракина. 6-е изд., доп. и испр. Москва : ВЛАДОС, 2017. 535 с. ISBN 978-5-906992-35-2.
- Ахманова, О. С. Словарь лингвистических терминов : [лингвостилистика, синтаксис, морфология, фономорфология, фонетика, лексикология : около 7 000 терминов] / О. С. Ахманова. 5-е изд. Москва : Либроком : URSS, 2009. 569 с. ISBN 978-5-397-00901-0.
- Беляева, М. Ю. Фонетика : учебное пособие для студентов бакалавриата, обучающихся по направлению 44.03.01 Педагогическое образования (профиль подготовки – Филологическое образования) очной формы обучения / М. Ю. Беляева ; Министерство науки и высшего образования Российской Федерации, Филиал ФГБОУ ВО «Кубанский государственный университет» в г. Славянске-на-Кубани, кафедра русской и зарубежной филологии. Краснодар : Экоинфест, 2019. 106 с. ISBN 978-5-94215-491-2.
- Васильева, В. В. Русский прозаический ритм. Динамический аспект / В. В. Васильева. 2-е изд., испр. Санкт-Петербург : Геликон Плюс, 2017. 142 с. ISBN 978-5-0098-135-1. EDN XPEVNZ.
- Великая, Е. В. Просодия в стилевой дифференциации языка : монография / Е. В. Великая. Москва : Прометей, 2009. 256 с. ISBN 978-5-7042-2201-9.
- Воронцова, Ю. А. Русский язык в деловой документации. Культура речи : учебное пособие / Ю. А. Воронцова, Е. Ю. Хорошко ; Белгородский юридический институт МВД России имени И.Д. Путилина. Белгород : Белгородский юридический институт МВД России имени И. Д. Путилина, 2021. 216 с. ISBN 978-5-91776-400-9.
- Дикушина, О. И. Фонетика английского языка : теоретический курс : учебное пособие для студентов педагогических институтов / О. И. Дикушина. Москва ; Ленинград : Просвещение, 1965. 201 с.
- Евстифеева, М. В. Теоретическая фонетика английского языка : лекции, семинары, упражнения : учебное пособие для студентов высших учебных заведений, обучающихся по специальности 050303 – иностранный язык / М. В. Евстифеева. 4-е изд., стер. Москва : Флинта, 2022. 165 с. ISBN 978-5-9765-1115-6.
- Зиндер, Л. Р. Общая фонетика и избранные статьи : [учебное пособие] / Л. Р. Зиндер ; [сост., вступ. ст. Л. В. Бондаренко]. 2-е изд., испр. и доп. Москва : Academia ; Санкт-Петербург : Санкт-Петербургский гос. ун-т, Филологический факультет , 2007. 574 с. ISBN 978-5-8465-0623-7.
- Зияданова, К. А. Экспериментальное сопоставительное ультразвуковое исследование вокализма языка казахов Алтая и Монголии / К. А. Зияданова, Я. М. Рабинович, В. М. Лемская, Д. М. Токмашев // Экспериментальные исследования языка : материалы конференции, Москва, 19–21 июня 2025 года. Москва : ООО Издательство «Наш мир», 2025. С. 92–94. EDN PILAFR.
- Златоустова, Л. В. Общая и прикладная фонетика : [учебник для вузов по направлению и специальности «Лингвистика»] / Л. В. Златоустова, Р. К. Потапова, В. В. Потапов, В. Н. Трунин-Донской ; под общ. ред. Р. К. Потаповой ; Московский гос. ун-т им.

- М. В. Ломоносова, Филологический факультет. Москва : Изд. Московского ун-та, 1997. 414 с. ISBN 5-211-03041-9.
- Калинина, Г. С. Речевой этикет в разноструктурных языках : диссертация ... кандидата филологических наук : 10.02.19 / Галина Сергеевна Калинина ; [Место защиты: ФГАОУ ВО «Казанский (Приволжский) федеральный университет»]. Казань, 2022. 244 с.
- Ким-Малони, А. А. Преодоление языковой интерференции у студентов из стран центральной Азии с целью повышения эффективности их культурно-языковой адаптации к условиям обучения и проживания на территории России / А. А. Ким-Малони, А. С. Ксенц, В. М. Лемская [и др.] // Вестник Томского государственного педагогического университета (TSPU Bulletin). 2014. Вып. 4 (145). С. 129–135. EDN SCKTUF.
- Лайонз, Дж. Введение в теоретическую лингвистику = Introduction to theoretical linguistics / Дж. Лайонз ; пер. с англ. под ред. и с предисл. В. А. Звегинцева. 2-е изд. Москва : URSS, 2010. 541 с. ISBN 978-5-354-01287-9.
- Лемская, В. М. Применение данных ультразвуковых исследований разносистемных идиолектов в практике изучения языков коренных народов Сибири и преподавания иностранных языков / В. М. Лемская, Д. М. Токмашев, А. И. Кузьмина [и др.] // Родной и иностранные языки в поликультурном пространстве : сборник статей Международной научно-практической конференции, Сургут, 01–02 ноября 2024 года. Сургут : Сургутский государственный университет, 2025. С. 27–31. EDN DCNGHZ.
- Леонтьева, С. Ф. Теоретическая фонетика английского языка : [учеб. для веч. и заоч. отд-ний педвузов] / С. Ф. Леонтьева. 2-е изд., испр. и доп. Москва : Высшая школа, 1988. 270 с.
- Маркова, М. А. Русская и чувашская библионимия в контрастивном аспекте : диссертация ... кандидата филологических наук : 5.9.8. / Мария Александровна Маркова ; [Место защиты : «Чувашский государственный университет имени И. Н. Ульянова» ; Диссовет 24.2.434.02 (Д 212.301.XX)]. Чебоксары, 2025. 158 с.
- Петроченко, Л. А. Введение в историю английского языка : учебное пособие / Л. А. Петроченко. Томск : Изд-во ТГПУ, 2001. 120 с.
- Плоткин, В. Я. Эволюция фонологических систем : на материале германских языков / В. Я. Плоткин. Москва : Наука, 1982. 129 с.
- Поливанов, Е. Д. Лекции по введению в языкознание и общей фонетике / Е. Д. Поливанов ; послесл. В. К. Журавлева и И. В. Журавлева. 2-е изд., стер. Москва : УРСС, 2004. 110 с. ISBN 5-354-01055-1.
- Сафронова, Н. А. Фонетика и фонология : курс лекций / Н. А. Сафронова ; Министерство науки и высшего образования Российской Федерации, ФГБОУ ВО «Владимирский государственный университет им. А. Г. и Н. Г. Столетовых». Владимир : Изд-во Владимирского гос. ун-та, 2023. 87 с. ISBN 978-5-9984-1861-7.
- Потапова, Р. К. Функционально-речевая специфика просодии и семантики / Р. К. Потапова // Речевые технологии. 2014. № 1-2. С. 3–21. EDN XYMSWD.
- Современный русский язык : в 3 ч. Ч. 1: Фонетика. Лексикология : сборник элементов учебно-методического комплекса по дисциплине для студентов, обучающихся по направлению подготовки специальности 031000 «Филология. Русский язык и литература» / сост. И. А. Измestьева [и др.]. Тольятти : Изд-во ТГУ, 2012. 251 с.
- Скроб, Т. В. Особенности билингвизма и диглоссии в территориальном варианте национального языка : на материале техасской разновидности американского варианта английского языка : диссертация ... кандидата филологических наук : 10.02.19 / Татьяна Владимировна Скроб ; [Место защиты: Сарат. гос. ун-т им. Н. Г. Чернышевского]. Саратов, 2017. 235 с.

- Соколова, М. А. Теоретическая фонетика английского языка = Theoretical phonetics of English : учебник для студентов высших учебных заведений, обучающихся по специальности 050303.65 (033200) – иностранный язык / М. А. Соколова [и др.]. Дубна : Феникс+, 2010. 192 с. ISBN 978-5-9279-0153-1.
- Тарева, Е. Г. Система культуросообразных подходов к обучению иностранному языку / Е. Г. Тарева // Язык и культура. 2017. № 40. С. 302–320. doi: 10.17223/19996195/40/22. EDN KZXYAH.
- Токмашев, Д. М. Артикуляционные параметры дорсального смычного /к/ в тюркских языках Южной Сибири по данным ультразвукового исследования (на материале телеутского языка и диалектов татар Сибири) / Д. М. Токмашев, В. М. Лемская, А. А. Хакимова, Н. В. Субботина // Урало-алтайские исследования. 2023. № 4 (51). С. 106–119. doi: 10.37892/2500-2902-2023-51-4-106-119. EDN XDMPTG.
- Трубецкой, Н. С. Основы фонологии / Н. С. Трубецкой ; пер. с нем. А. А. Холодовича ; под ред. С. Д. Кацнельсона. 4-е изд. Москва : URSS : ЛЕНАНД, 2012. 335 с. ISBN 978-5-382-01374-9.
- Фант, Г. Акустическая теория речеобразования / Г. Фант ; пер. с англ. Л. А. Варшавского и В. И. Медведева ; под ред. В.С. Григорьева. Москва : Наука, 1964. 284 с.
- Швейцер, А. Д. Литературный английский язык в США и Англии / А. Д. Швейцер. 3-е изд. Москва : URSS, 2008. 199 с. ISBN 978-5-382-00774-8.
- Щерба, Л. В. Фонетика французского языка : очерк французского произношения в сравнении с русским : [учебное пособие для ин-тов и фак. иностр. яз.] / Акад. Л. В. Щерба. 7-е изд. Москва : Высшая школа, 1963. 309 с.
- Якобсон, Р. Фонология и её отношение к фонетике / Р. Якобсон, М. Халле // Новое в лингвистике. Вып. 2. Москва : Изд-во иностранной литературы, 1962. С. 231–278.
- Brown, G. Listening to Spoken English / G. Brown. 2nd ed. Routledge, 2017. 192 p. ISBN 1134961057, 9781134961054.
- Cook, A. American Accent Training / A. Cook. 4th ed. Barrons Educational Services, 2017. 228 p. ISBN 1438065035, 9781438065038.
- Crystal, D. A Dictionary of Linguistics and Phonetics / D. Crystal. 7th ed. Wiley-Blackwell, 2024. 512 p. ISBN 978-1-119-18459-1.
- Fuster, J. M. Cortex and Mind: Unifying Cognition / J. M. Fuster. Oxford : Oxford University Press, 2003. 294 p. ISBN 9780195300840.
- Gimson, A. C. An Introduction to the Pronunciation of English / A. C. Gimson. 4th ed. London : Edward Arnold, 1989. 364 p. ISBN 0713165944, 9780713165944.
- Grant, L. Well Said (Well Said, New Edition) / L. Grant. 4th ed. Boston : Heinle ELT, 2016. 241 p. ISBN 9781305641365.
- Harris, J. Vowel Patterns in Mind and Sound / J. Harris, G. Lindsey // Phonological Knowledge: Conceptual and Empirical Issues / ed. by N. Burton-Roberts, P. Carr, G. Docherty. Oxford : Oxford University Press, 2000. P. 185–205. ISBN 9780198241270.
- Hock, H. H. Principles of Historical Linguistics / H. H. Hock. 3rd ed. Walter de Gruyter GmbH & Co KG, 2021. 1101 p. ISBN 3110746565, 9783110746563.
- Jones, D. An English Pronouncing Dictionary (on Strictly Phonetic Principles) / D. Jones. London : J. M. Dent & Sons, Ltd. ; New York : E. P. Dutton, 1917. xxviii, 419 p.
- Lindsey, G. Intonation and pragmatics / G. Lindsey // Journal of the International Phonetic Association. 1981. Vol. 11, № 1. P. 2–21. doi: 10.1017/S0025100300002176.
- Lindsey, G. English Speech Services / G. Lindsey. 2011. URL: <https://www.englishspeech-services.com/> (accessed: 01.12.2025).

- Lindsey, G. English after RP: Standard British Pronunciation Today / G. Lindsey. Cham : Springer, 2019. XVI, 153 p. ISBN 3030043576, 9783030043575.
- Messum, P. Understanding and Teaching the English Articulatory Setting / P. Messum // Speak Out! (IATEFL Pronunciation Special Interest Group Newsletter). 2010. Issue 43 (September). P. 20–24.
- Messum, P. Teaching the Basis of Articulation / P. Messum // Speak Out! (IATEFL Pronunciation Special Interest Group Newsletter). 2012. Issue 47 (September). P. 28.
- Messum, P. Bringing the English Articulatory Setting into the Classroom: (1) the Tongue / P. Messum, R. Young // Speak Out! Journal of the IATEFL Pronunciation Special Interest Group. 2017. Issue 57 (September). P. 29–38.
- Messum, P. Teaching the Underlying Systems of English Pronunciation as Motor Skills / P. Messum, R. Young // Speak Out! Journal of the IATEFL Pronunciation Special Interest Group. 2019. Issue 61 (September). P. 18–31.
- Metcalf, A. How We Talk, American Regional English Today / A. Metcalf. Boston ; New York : Houghton Mifflin Company, 2000. 207 p.
- O'Connor, J. D. Phonetics / J. D. O'Connor. Harmondsworth : Penguin Books Ltd., 1978. 320 p.
- Roach, P. English Phonetics and Phonology: A Practical Course / P. Roach. 2nd ed. Cambridge : Cambridge University Press, 1990. 212 p.
- Shakhbagova, D. A. Varieties of English Pronunciation / D. A. Shakhbagova. Moscow : Vyssaja Shkola, 1982. 128 p.
- Shillcock, R. C. Phonologically Motivated Input Representation for the Modelling of Auditory Word Perception in Continuous Speech / R. C. Shillcock, G. Lindsey, J. Levy, N. Chater // Proceedings of the 14th Annual Conference of the Cognitive Science Society. Hillsdale, N.J. : Lawrence Erlbaum, 1992. P. 408-413.
- Vassilyev, V. A. English Phonetics. A Normative Course / V. A. Vassilyev, O. V. Burenkova, A. R. Katanskaya, N. D. Lukina, L. P. Maslova, E. I. Torsuyeva. Leningrad : Uchpedgiz, 1962. 384 p.
- Vassilyev, V. A. English Phonetics. A Theoretical Course / V. A. Vassilyev. Moscow : Vyssaja Shkola, 1970. 323 p.
- Washington, J. An ultrasound study of dorsal stops in Tatar and Qumuq / J. Washington, D. Tokmashev, A. Badrtdinova, Z. Tarkovskaya, M. Zimin, V. Maltseva, N. Urtegeshev, T. Ryzhikova, V. Lemskaya, A. Dybo // UltraFest IX: Ultrasound Imaging for Speech and Language Virtual Conference. Indiana University, October 21–24, 2020. URL: https://ultrafest2020.indiana.edu/abstracts/UltraFest_IX_Washington_EtAl_TatarQumuq.pdf (accessed: 01.04.2025).

Учебное издание

Валерия Михайловна ЛЕМСКАЯ, Сергей Викторович ГЛУШКОВ,
Ярослав Андреевич ГЛУХИЙ

**ТЕОРЕТИЧЕСКАЯ ФОНЕТИКА
АНГЛИЙСКОГО ЯЗЫКА**

Учебно-методическое пособие

Электронное издание локального распространения

Ответственный за выпуск: Е. В. Гребенникова
Корректор: Ю. П. Готфрид
Технический редактор: А. И. Лелоюр
Дизайн обложки Н. В. Удлер

Изображение на обложке создано с помощью нейросети Seedream

Подписано к использованию: 15.10.2026.
Гарнитура Times. Объем издания: 3,24 Мб.
Комплектация издания – 1 CD.
Тираж 100 CD. Заказ № 109/ЭУ.

Издательство
Томского государственного педагогического университета
634061, г. Томск, ул. Киевская, 60
тел. 8(3822)311-484
E-mail: izdatel@tspu.ru

© Лемская В. М., Глушков С. В., Глухий В. Я., 2026

© Томский государственный педагогический университет, 2026

ISBN 978-5-907791-81-7