Л.В. БОРИСОВА, А.А. МЕТЛЮК

# ТЕОРЕТИЧЕСКАЯ ФОНЕТИКА АНГЛИЙСКОГО ЯЗЫКА

*Допущено Министерством высшего и среднего специального образования БССР в качестве учебного пособия для студентов институтов и факультетов иностранных языков*

Минск
"Вышэйшая школа"
1980

ББК 81.2 Англ-1
Б82

Рецензенты: кафедра романо-германской филологии Азербайджанского педагогического института иностранных языков и кандидат филологических наук *Пилипенко О.Ф.*



**Борисова Л.В., Метлюк А.А.**

**Б82** Теоретическая фонетика английского языка: [ Учеб. пособие для ин-тов и фак. иностр. яз.] — Мн.: Выш.школа, 1980. — 144 с., ил.

В пер.: 60 к.

В пособии рассматриваются основные проблемы современной фонетики английского языка: фонологический анализ звуков, слогообразование и слогоделение, фонетическая природа и характерные особенности английского словесного ударения, интонация, вариативность произносительной нормы английского языка. Введены разделы, освещающие принципы акустической классификации звуков, фоностилистические особенности английской речи, вопросы фонетической интерференции при двуязычии, а также способы графического изображения интонации.

Предназначается студентам факультетов английского языка.

Б $\frac{70104-133}{М304(05)-80}$ 146—80. 4602000000

ББК 81.2 Англ-1
4И (Англ)

## FOREWORD

The book is designed for students of English who have undertaken the study of a theoretical course in English phonetics at Pedagogical Institutes and Faculties of Foreign Languages.

The authors of the book had two objects in view: firstly, to give the prospective teachers of English necessary information on theoretical problems of English phonetics and, secondly, to acquaint them with some of the diverse views of Soviet and foreign linguists. The authors have put forward the opinions of the foremost linguists on the points at issue and expressed their own attitude to them. An endeavour has been made to treat the main problems of theoretical and experimental phonetics with due reference to the latest findings in this field, and to draw the readers' attention to phonetic phenomena which are in the forefront of modern linguistics, such as phonetic styles and prosodic interference. An attempt has also been made to show the relevance of theoretical studies in solving practical problems connected with language teaching.

Every chapter is followed by exercises. Their purpose is not so much to control understanding as to encourage further reading. The exescises contain questions which stimulate independent decisions, contraversial views which are to be refuted or supported, and assignments which require reference to pronouncing dictionaries and analysis of the phonetic features of speech.

Although this book is intended primarily for students, it may be of interest to teachers and to all those concerned with English phonetics.

The authors make grateful acknowledgement to Prof. V.V. Artyomov, Prof. V.A. Vassilyev and the late Prof. K.K.Baryshnicova who have inspired and encouraged this work.

Thanks for valuable suggestions are also expressed to those who have reviewed the MS.

Chapters 1, 2, 4, 6, 7 were written by A.A.Metlyuk; chapters 3, 5, 8, 9, 10 by L.V. Borisova.

## CHAPTER 1. THE ROLE OF SOUND PHENOMENA IN COMMUNICATION

### THE PHONETIC SYSTEM OF A LANGUAGE

Language – *"the immediate actuality of thought"* [46] *and "the most* important means of human intercourse" [47] exists in two main speech forms: oral and written. Speech is a manifestation of language. It is a process of communication by means of language. Both the oral and written speech forms have a material substance. In oral speech the substance is phonic*, it is the *sound substance or the sound matter. In written speech the substance is* graphic.

A substance is not in itself language, but it is what forms patterns of language. The sound substance gives shape to a spoken message in communication, it forms units of the phonetic system of a concrete language.

The phonetic system of a language is a set of phonetic units arranged in an orderly way to replace each other in a given framework. In fact it contains two systems (or *levels*) – *segmental* and *suprasegmental, or prosodic,* each of which is a specially organized language system with a certain number of its units. Segmental units are elementary sounds, vowels and consonants which form the vocalic and the consonantal subsystems. Prosodic units are syllables, accentual (rhythmic) units, intonation groups, utterances, which form the *subsystems of pitch, stress, rhythm, tempo, pauses.*

The sound substance is a medium in which the whole system of language is embodied. Segmental and prosodic units serve to form and differentiate units of other subsystems of language, the lexical and grammatical units. The modifications of words and their combination into utterances (sentences) are first of all sound phenomena. The grammatical form of a verb or a noun can be changed only by changing the sounds which compose them. By changing the prosodic structure (intonation) of an utterance one changes the meaning of the utterance. It is clearly seen from the utterances of identical lexical and grammatical structures. For example, "ˌWell ˊdone?" pronounced with the rising tone, is a question, expressing the speaker's uncertainty and desire to get further information. "ˌWell ˋdone!", pronounced with the falling tone, wide *pitch range and greater loudness, is an exclamation, expressing the speaker's* evaluation. The rising-falling tone in the utterance "ˌWell ˆdone" may express a challenging or quizzical attitude on the part of the speaker.

---

**The terms "phonic", "phonetic", "phonetics" come from the Greek word* φωνη (fo:ne:) – sound.

To underline the importance of the sound medium of language H. Gleason notes that to speak any language a person must know nearly all the 100% of its phonetics, while only 50–90% of its grammar and 1% of the vocabulary may be sufficient [24].

## ASPECTS OF SOUND PHENOMENA

The ability to form language units is not the only property of the sound medium. In addition to it, the sound medium has its own independent properties as a physical phenomenon. Moreover, it is a product of human activity. Being created by the speaker, the sound medium indicates the speaker's personality (sex, age, individual features) and reveals his physiological and emotional state, geographical origin, education, social status and so on.

Every act of speech presupposes the presence of a person who speaks and a person who listens. The speaker produces sounds, the sounds travel through the air to the listener in the form of complex combinations of sound waves, the listener hears and interprets them. Communication is possible only because the speaker and the listener interpret the sounds as units of the same language.

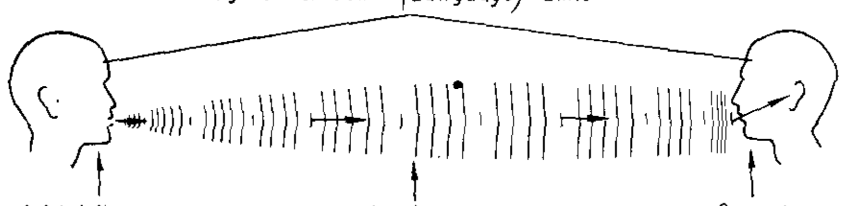


**Fig. 1. Stages in the passing of a spoken message.**

Consequently, sound phenomena have different aspects, which are closely interconnected: the articulatory aspect, the acoustic, the auditory and the linguistic aspect.

The articulatory (sound–production) aspect. Speech sounds are products of human organs of speech. They result from the activities of the diaphragm, the lungs, the bronchi, the trachea, the larynx with the vocal cords in it, the pharynx, the mouth cavity with the speech organs situated in it and the nasal cavity.

Sound production is impossible without respiration, which consists of two alternating phases – inspiration and expiration. Speech sounds are based chiefly on expiration, though in some African languages there are sounds produced by inspiration.

Expiration, during which speech sounds are produced is called phonic expiration as distinct from quiet breathing. In phonic expiration the air co–

mes from the lungs not freely but in spurts, because during speech the air-passage is periodically blocked by the speech organs. Therefore in speech, expiration lasts much longer than inspiration, whereas in quiet breathing inspiration and expiration each take about the same period of time.

The lungs supply the necessary air-pressure and regulate its force, thus producing variations in the intensity of speech sounds. For example, the air pressure is greater on the peak of the syllable and it is less on its margins.

Sound production actually takes place in the larynx, the pharynx and the oral and nasal cavities. The air-stream coming from the lungs undergoes important modifications in them.

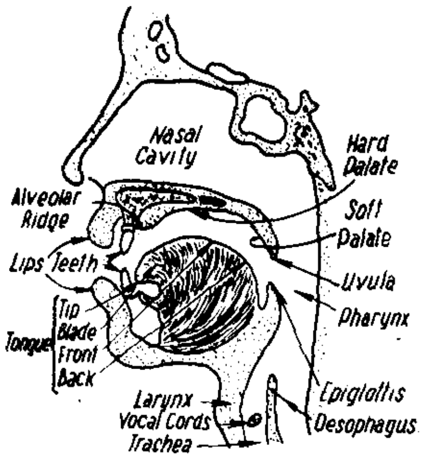


Fig. 2. Organs of speech.

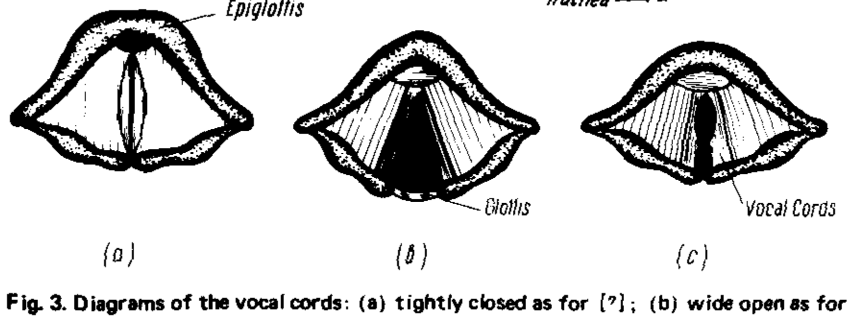


**Fig. 3. Diagrams of the vocal cords: (a) tightly closed as for [ʔ]; (b) wide open as for breath ; (c) loosly together and vibrating as for voice.**

One part of sound production is phonation, or voice-production.

When the vocal cords, situated in the larynx, are tensed and brought loosely together, the air-pressure below the vocal cords becomes very high and the air comes from the lungs in regular puffs making the vocal cords vibrate. Their vibrations are complex, though mainly regular or periodic. The regular vibrations of the vocal cords are transmitted to the air-stream and the acoustic effect perceived by the human ear is that of a vocal tone. This is what we call voice.

The other part of sound–production is *articulation* which comprises all the movements and positions of the speech organs necessary to pronounce a speech sound. The movements of speech organs modify the shape, size and volume of the supralaryngeal cavities (the pharynx, the mouth and the nasal cavity) thus modifying the voice which comes from the lungs. As a result, a *vowel sound* of a certain quality is produced.

When in the supralaryngeal cavities there is an obstruction to the air–stream, a noise is produced. The character of the noise (friction or plosion), depends on the type of obstruction (a constriction or a complete closure) and determines the particular quality of a consonant. When an obstruction is created and the vocal cords vibrate, a voiced consonant is produced. When the vocal cords do not vibrate, the result is a voiceless consonant.

Thus there are two main sources of vibration in the production of speech sounds – the vocal cords and various kinds of obstruction.

The acoustic aspect. Like any other sound of nature speech sounds exist in the form of sound waves and have the same physical properties – frequency, intensity, duration and spectrum.

A sound wave is created by a vibration which may be periodic or non–periodic, simple or complex.

The vocal cords vibrate in such a way that they produce various kinds of waves simultaneously. The basic vibrations of the vocal cords over their whole length produce *the fundamental tone of voice*. The simultaneous vibrations of each part of the vocal cords produce partial tones (overtones, or harmonics).

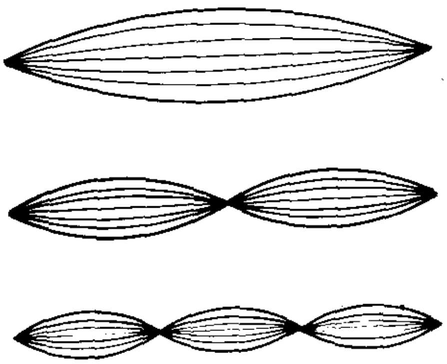

Fig. 4. Vibrations (a) of entire cord; (b) of its parts.

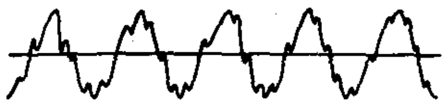

*Fig. 5. Complex vibrations.*

The number of vibrations per second is called frequency. Frequency is measured in hertz or cycles per second (c/s). Frequency of the basic vibrations of the vocal cords is the *fundamental frequency* (marked by Fo). The fundamental frequency determines the pitch of the voice and forms an acoustic basis of speech melody. It is relatively low, about 40–400 c/s in the total range of voice. The frequencies of the overtones are much *higher*.*

The superposition of the fundamental and partial vibrations results in a complex tone.

The complex tone is modified *in the oral and nasal resonators*. Due to the particular changes in their shape, size and volume the oral and nasal cavities function as an acoustic filter: they intensify certain frequencies contained in the complex tone and weaken the others. Thus the specific qualities of vowel sounds are achieved.

The complex range of intensified frequencies which form the quality of a sound is called the acoustic spectrum of the sound. The intensified frequencies in the spectrum which characterize the *quality* of a sound and distinguish it from other sounds of different quality are called formants.

Vowel sounds have at least two formants marked by $F_1$ and $F_2$, which are together responsible for the particular quality of each vowel type. $F_1$ is characterized by lower frequencies, $F_2$ – by higher frequencies. Thus, $F_1$ of the vowel /a:/ is equal to 800 Hz and $F_2$ – to 1100 Hz. The formants of /i:/ are equal to 380 Hz and 2500 Hz, respectively.

The formant of the fundamental tone (Fo) is *irrelevant* to vowel differention. $F_0$ is present in the spectra of vowels, sonants and voiced conconants because these sounds are formed with voice and it is absent in the spectra of voiceless consonants.

Intensity of speech sounds depends on the amplitude of vibration. Changes in intensity are associated with stress in those languages which have dynamic stress. Intensity is measured in decibels (dB**).

Like any other form of matter, sound *exists and moves in time*. Any sound has a certain duration. The duration of a sound is the quantity of time during which the same vibrations continue. For this reason the duration of a sound is often referred to as its quantity. The duration of sounds is measured in milliseconds (ms).

<u>*The auditory (sound–perception) aspect.*</u> Speech sounds may also be analysed from the point of view of perception. The perception of speech sounds involves the activity of our hearing mechanism, which can be viewed in two ways.

---

*Each half of the cord vibrates at a frequency twice as great as that of the cord as a whole. Each third vibrates three times faster, each quarter – four times faster, and so on.

***The decibel* is one tenth of a bel. The bel is so named after Alexander Graham Bell (1847–1922), inventor of the magnetic telephone.

On the one hand, it is a physiological mechanism which reacts to acoustic stimuli: the human ear transforms mechanical vibrations of the air into nervous stimuli and transmits them to the brain. The listener hears the acoustic features of fundamental frequency, formant frequency, intensity and duration in terms of four perceptible categories of pitch, quality, loudness and length.

On the other hand, it is also a psychological mechanism which selects from the great amount of acoustic information only that which is linguistically significant. The human brain interprets acoustic phenomena in terms of a given language system. In this way, different acoustic stimuli may be interpreted as being the same sound unit. Thus for a Russian the open / ɛ / as in "шэст" and the mid–open / e / as in "дети" are one and the same sound unit, as the difference between them is not singificant in distinguishing words or grammatical forms in Russian. An Englishman would consider these sounds as different sound units since in the English language the former resembles the open / æ /as in "had" and the latter is very much like the mid–open / e /as in "head" which serve to differentiate words. So a listener's reactions are conditioned by his experience of handling his own language.

Our physiological capabilities in perception are limited. The human ear does not perceive all the acoustic features present in a sound wave. Thus, the lower limit of frequency which we can perceive is 16 Hz, the hinger limit is 20 000 Hz. In ideal conditions we can perceive even a difference in 3 Hz, but such small distinctions are not usable in speech.

It should be remembered that perceptible features of sounds are not fully conditioned by the related articulatory and acoustic ones. Our perception of the pitch of the voice depends largely (but not solely) on the fundamental frequency generated by vocal cord vibration. The greater is the frequency, the higher is the pitch of the voice and vice versa. But the perceived pitch variation may also be affected by variations of intensity on the same frequency. By reducing the intensity of a sound we can achieve a high pitch in perception.

Changes in intensity are perceived by our ear as variataions in the loudness of a sound. The greater is the intensity of a sound, the louder is the sound. But our perception of loudness does not depend on intensity alone. A sound or a syllable may be perceived as louder, in comparison with neighbouring sounds or syllables, because of a marked pitch change on it or because it is longer than the others.

Moreover, some sounds, owing to their nature, are louder or more sonorous than others. Thus /a:/ is more powerful than /i:/, and vowels generally have more carrying power than consonants.

Our judgements relating to loudness are not as fine as those relating to either quality or pitch.

Different duration of speech sounds is perceived as a difference in their length. But our perception of length does not always correspond to the actual duration of speech sounds or other units. For example, the length of rhythmic units in an English utterance is considered to be approximately the same since it is a characteristic feature of English rhythm that stressed syllables occur at more or less equal intervals of time. But the actual duration of rhythmic units is far from being equal. This is an example of how our brain interprets from the acoustic material only that which is linguistically significant.

Our hearing mechanism acts as a monitor of what we ourselves are saying. The process of communication would be impossible if the speaker himself did not hear the sounds he pronounces. If the link between listening and pronouncing is disturbed, disturbances in the production of speech sounds are likely to appear. The better we hear the differences between the sounds, the better we pronounce them. Therefore in learning to pronounce the sounds of a foreign language one should bear in mind the importance of ear–training.

The linguistic aspect. Segmental sounds and prosodic features are linguistic phenomena. Representing language units in actual speech, they perform certain linguistic functions. They constitute meaningful units – morphemes, words, word–forms, utterances. All the words of a language consist of speech sounds which are grouped and arranged in the way specific for the language and which are unified by stress. All the utterances consist of words, and, consequently, of sounds; they are characterized by certain pitch–and–stress patterns, temporal features, rhythm.

Most of the meaningful distinctions of the language are based on distinctions in sound.

Sounds and prosodic features serve to differentiate the units they form since communication by means of language is possible only because sound phenomena can be opposed to one another for purposes of differentiating words, word–forms and utterances.

Simultaneously, the sound phenomena enable the listener to identify them as concrete words, word–forms or utterances.

Thus, segmental sounds and prosodic features of speech perform constitutive, distinctive and identificatory functions.

The linguistic aspect of speech sounds is also called the functional or social aspect, because of the role which sound matter plays in the functioning of language as a social phenomenon.

The relations between the articulatory, acoustic, auditory and linguistic aspects of speech sounds can be presented roughly in the following way:

| Articulatory characteristics | Acoustic properties | Auditory (perceptible) qualities | Linguistic phenomena |
|---|---|---|---|
| vibrations of the vocal cords | fundamental frequency | pitch | prosody (melody, stress) |
| different positions and movements of speech organs | formant frequencies | quality (timbre) | phoneme, prosody (stress) |
| the amplitude of vibrations | intensity | loudness | prosody (stress) |
| the quantity of time during which the sound is pronounced | duration | length | prosody (tempo, rhythm) |

## PHONETICS AS A SCIENCE

The study of the sound phenomena of language, in all their aspects and varieties, constitutes the subject of the phonetic science.

Phonetics as a branch of linguistics studies sounds in the broad sense, comprising segmental sounds (vowels and consonants) and prosodic phenomena (pitch, stress, tempo, rhythm, pauses). Phonetics occupies itself with the study of the ways in which the sounds are organized into a system of units and the variation of the units in all types and styles of spoken language. It also studies the acoustic properties of sounds, the physiological basis of sound production and the sound phenomena that reveal the individual peculiarities of the speaker.

Thus the sound medium has a special science all to itself because of the exclusive importance of oral speech as compared with written speech. Oral speech is primary, whereas written speech is secondary (it is constructed on the basis of oral speech). The sound medium needs special attention also because of the complex character of its production and perception. Besides, further development of such technical means of sound transmission as the telephone, the radio, recording and speaking machines, speech recognizers, deaf aids raises a great number of primarily phonetic problems.

Phonetics as a science is a branch of linguistics. Being a science in its own right, it is at the same time closely connected with other linguistic sciences —

grammar, lexicology, stylistics and the history of the language, since the phonetic system of a language, its vocabulary and grammar constitute one indivisible whole. It is also closely interconnected with physiology, biology, physics, pedagogy, psychology, mathematics, cybernetics.

Phonetics has a long history. It was known to the ancient Greeks and Hindus. But as a science in its own right it began to develop in Russia and in Western Europe only in the second half of the 19th century.

## BRANCHES OF PHONETICS

Depending on which of sound phenomena is studied, phonetics is subdivided into four main branches.

Articulatory phonetics is concerned with the study of sound as a result of the activities of speech organs. It deals with our voice–producing mechanism and the way we produce sounds, and prosodic phenomena. It studies respiration, phonation (voice–production), articulation and also the mental processes necessary for the mastery of a phonetic system.

Methods employed in articulatory phonetics are experimental. They involve palatography, laryngoscopy, photography, cinematography, X–ray photography, X–ray cinematography, electromyography, and various kinds of technique to study sound–perception.

Besides these objective methods articulatory phonetics uses its oldest, subjective method – the method of direct observation. This method involves observation of the movements of speech organs when pronouncing sounds and analysis of one's muscular sensations during the articulation of speech sounds.

Perceptual (auditory) phonetics occupies itself with the study of man's perception of segmental sounds, pitch variation, loudness and duration. It studies the ways in which sound perception is determined by the phonetic system of a language.

The methods used in perceptual phonetics are also experimental. They include various kinds of auditory tests.

Since sound production and sound perception are physiological processes, articulatory and perceptual phonetics are generally termed physiological phonetics.

Acoustic phonetics is concerned with the acoustic aspect of speech sounds. It studies speech sounds with the help of experimental (instrumental) methods. Various kinds of apparatus are applied for analysing the acoustic structure of segmental sounds and prosodic phenomena. For example, we use a spectrograph to analyse the acoustic spectra of sounds, an oscillograph and an intonograph to analyse frequency, intensity and duration. With the help of an electro–acoustic synthesizer synthetic speech is produced which is a good means of testing the results of the electro–acoustic analysis.

Phonology, or functional phonetics is a purely linguistic branch of phonetics. It deals with the functional aspect of sound phenomena. Phonology sets out to discover those segmental and prosodic features that have a differential value in a language, and it establishes the system of phonemes and prosodemes. It also sets out to determine the frequency of occurrence of these units in syllables, words, rhythmic units and other sequences which form the utterance. The distribution and grouping of phonemes in syllables and words in a particular language are dealt with in an area of phonology which is called phonotactics.

The basis of phonology is the phoneme theory, created in Russia by I. Baudouin de Courtenay and developed by his pupils and followers L.Shcherba, N. Krushevsky and, later, by other Soviet and foreign linguists. Phonology was founded in Prague by a group of linguists (N. Trubetzkoy, R. Jakobson and others).

Phonology of segmental units is often called phonemics, whereas phonology of intonation (prosody) is termed intonology (or prosodemics).

The methods employed by phonology are linguistic.

All the above branches of phonetics are closely connected since the object of their study – speech sounds – is a close unity of acoustic, articulatory, auditory and linguistic aspects. But not all linguists are of the opinion that phonology is an integral part of phonetics.

N. Trubetzkoy [40] claims that phonology should be separated from phonetics. According to the Prague School, established by him, phonetics and phonology are independent sciences: phonetics is a biological science and is concerned with physical and physiological characteristics of speech sounds, phonology is a linguistic science and is concerned with the social function of phonetic phenomena. This point of view is supported by the Danish linguist L.Hjelmslev who advocated total separation between phonetics and phonology. But the vast majority of Soviet phoneticians do not consider it logical to separate function from form and to exclude phonetics from the linguistic sciences [80]. A great number of phoneticians abroad adhere to the same point of view. For instance, B. Malmberg, a Swedish phonetician, writes as follows: "It was a grave error on the part of the Prague School to want to establish a strict separation between phonetics and phonology... The two types of studies are interdependent and condition each other. Consequently it seems preferable to group them together under the traditional general heading of phonetics" [97, p. 97].

Besides the four branches given above, there are other branches of phonetics, such as special, general, historical, descriptive, comparative, applied.

Special phonetics is concerned with the study of the phonetic system of a concrete language. When the phonetic system is studied in its static form at a particular period (synchronically), we deal with descriptive phonetics. When the system is studied in its historical development (diachronically) we speak about historical, or evolutionary phonetics. Historical phonetics uses the philological method of investigation. It studies written documents and compares the spelling and pronunciation of one and the same word in different periods of the history of the language.

General phonetics is concerned with the study of man's sound–producing possibilities and the functioning of his speech mechanism. It establishes the types of speech sounds which exist in various languages, the way they are produced and the role they play when forming and expressing thoughts. General phonetics is based on the extensive material which is provided by the special phonetics of a number of languages and on the material of other sciences. As a result of this, general phonetics has been able to make a number of general conclusions concerning the complex nature of speech sounds and to formulate a number of theories: the phoneme theory, the theory of syllable formation, theories of stress, intonation, etc.

Theoretical phonetics of a particular language applies those theories to the language it analyses.

Comparative phonetics is concerned with the comparative study of the phonetic systems of two or more languages, especially kindred ones.

By practical, or applied phonetics we mean all the practical applications of phonetics. Phonetics is of considerable importance for other fields of language study, which have made use of the approaches and the linguistic methods worked out by phonetics.

All the branches of phonetics are of great use and importance in teaching the pronunciation of foreign languages. Phonetic data is also made use of in teaching children to read and write their mother tongue; in creating orthographies for unwritten languages.

Phonetics is of great practical importance in the teaching of diction to actors, singers, radio–announcers and other public speakers. In order to speak well and to teach others the proper way of speaking one must understand the mechanism of articulation and the mechanism of phonation. Young teachers should bear it in mind that to teach efficiently they must learn to speak efficiently.

Phonetics is applied in logopedics, i.e. in correcting speech defects and in curing pathological phenomena of speech, such as aphasia. To correct deviations from the pronunciation norm, one must have a good knowledge of normal phonetics.

Phonetics is of great importance in surdo–pedagogics, i.e. teaching normal oral speech to deaf–mutes. The deaf people can, in most cases, use only muscular sensation to control and guide their articulation. Therefore, to teach to speak one must know thoroughly the articulatory aspect of speech sounds.

Phonetics is widely used in the field of sound transmission: in telephony, broadcasting, speech recognition. To build a microphone, a tape recorder, a spectrograph, a speech synthesizer or any other apparatus capable of transmitting the spoken language, the sound engineer must know acoustic phonetics and, very often, solve the same problems as the linguist does.

Close interaction and collaboration between phonetics and other sciences has given birth to new scientific branches such as technical acoustics, psychophonetics and other phonetic sciences which contributed considerably to the formation of speechology — the science of speech.

## EXERCISES

A. Think about the following questions for class discussion:

1. What are the relations between language and speech?
2. Define the phonetic system of a language. What units does it include?
3. What are the two main sources of sound?
4. How does the frequency of the vocal cord vibrations change in the production of the falling tone? (the rising tone?)
5. Why do the sounds of the violin, the guitar and other musical instruments differ? What affects their specific qualities? Is there any analogy in the production of speech sounds?
6. In what way does the listener perceive the acoustic properties of speech sounds? What does perception of pitch and loudness depend on?
7. How would you explain the difference between the physiological and psychological aspects of our hearing mechanism? Which of them is responsible for "hearing the message"? Which is responsible for "understanding the message"? Which of them is involved when we listen to a foreign language that we don't know?
8. Why is it important to train one's hearing abilities in foreign language learning?
9. Why is phonetics placed amog linguistic sciences and not among physiological or physical?

B. Select some examples:

1) to illustrate articulatory and functional distinctions between English and Russian (or Byelorussian) sounds;
2) to illustrate the distinctions in sounds which express distinctions in meaning;
3) to illustrate that an utterance, an act of oral speech, is a concrete manifestation of the whole system of language at work.

## CHAPTER 2. ARTICULATORY AND ACOUSTIC ANALYSIS OF ENGLISH SPEECH SOUNDS

### PRINCIPLES OF CLASSIFICATION OF SPEECH SOUNDS

In all languages speech sounds are traditionally divided into two main types – vowels and consonants.

From the articulatory point of view the main principles of the division are as follows: the presence or absence of obstruction; the distribution of muscular tension; the force of the air stream coming from the lungs.

Vowels are speech sounds based on voice which is modified in the supra-laryngeal cavities. There is no obstruction in their articulation. The muscular tension is spread evenly throughout the speech organs. The force of the air stream is rather weak.

Consonants are speech sounds in the articulation of which there is an obstruction, the removal of which causes noise – plosion or friction. The muscular tension is concentrated at the place of obstruction. The air stream is strong.

The articulatory boundary between vowels and consonants is not well marked. There exist speech sounds that occupy an intermediate position between vowels and consonants and have common features with both the vowels and consonants. These are sonorants /m, n, ŋ, j, l, w, r/. There is an obstruction in their articulation and the muscular tension is concentrated at the place of obstruction as in the production of consonants. Like vowels they are largely based on voice. The air passage in their production is rather wide and the force of the air is weak as in the case of vowels. The wide passage for the air stream in the articulation of sonorants means that the oral and nasal cavities are active. It results in greater audibility (sonority, carrying power, or *perceptibility) of the sounds – a feature characteristic of vowels.* Because of their strong vocalic characteristics sonorants /w, j, r/ are often referred to as *semivowels*. Due to their great sonority some sonorants can be syllabic in some particular positions (e.g. teɪ – bl̩; ga: – dn̩). But generally sonorants do not perform the function of syllable formation. That is why they are attributed to consonants. Thus, consonants can be subdivided into sonorants and noise consonants (or contoids, after K.Pike [103]).

From the acoustic point of view vowels are complex periodic vibrations – tones. They are combinations of the main tone and overtones intensified by the supralaryngeal cavities.

Consonants are non-periodic vibrations – noises. Voiceless consonants are pure noises. Voiced consonants are actually a combination of noise and tone. And sonants are predominantly sounds of tone with an admixture of noise.

Thus, the acoustic boundary between vowels and consonants is not well marked either.

Numerous experiments show that the criterion which justifies the division of speech sounds into vowels and consonants is the physiological criterion put forward by V. Bogoroditzky [13]. It is the distinction in the groups of muscles, which operate in vowel and consonant production, and the resulting different articulatory energy in vowel and consonant production.

## THE ARTICULATION BASIS OF ENGLISH

Due to the identical structure of speech organs of people of different races and nationalities, all languages have sounds of identical types (e.g. Russian /Н/ and English /n/). But being identical typologically, the sounds are not identical articulatorily. Their articulatory distinctions are explained by the fact that each language has its own tendencies and modes of articulation. These articulatory habits characteristic of all the native speakers of a language are called the articulation basis of the language.

The articulation basis of English differs from that of Russian: voiced consonants are less energetic, whereas voiceless consonants are much more energetic, the lips do not protrude for /ɔ:, u:/ as they do for /о, у/, the tongue is slightly drawn back.

The following quotation from Webster's New International Dictionary of the English language presents some marked peculiarities of the English articulatory basis: "Every language has certain characteristics of pronunciation which give it distinctive phonetic character. In English, the tongue, when articulating, is characterized by laxity. It is not made so tense as in many other languages, as the French, for instance. In articulating the front vowels it is not pushed so far forward, nor in the back vowels drawn so far back. There is also a constant tendency to lower and flatten the tongue, hollow the front of it more or less, like a spoon, and to draw it away from the teeth, the lips are inactive... English articulation may, in general, be characterized as sluggish or muffled, the French and German, for example, as vigorous and clear" [112, p. xlvii].

The notion of the articulation basis of a language can be broadened if we analyse not only the articulation part in sound production but phonation as well. The phonational habits of the native speakers of different languages may differ depending on the character of sounds (such as clicks and suction which are pronounced during inspiration, checked vowels in English which require a great force of utterance at the end of their articulation). The phonational habits may also be dependent on the amount of consonants that occur in consonantal clusters, and on the character of sound transitions in connected speech.

*The articulation basis, therefore, comprises* both the articulatory and phonational habits of the native speakers of a language.

The peculiarities of the articulation basis of English determine the specific articulatory characteristics of its sound system, the character of sound modifications in connected speech and the physiological mechanism of syllable formation.

## THE ARTICULATORY CLASSIFICATION OF ENGLISH SPEECH SOUNDS

### A. Vowels

The various qualities (timbres) of English vowels are determined by the *oral resonator* – its *size, volume and shape. The resonator is modified* by the most movable speech organs – the tongue and the lips. Moreover, the quality of a vowel depends on whether the speech organs are tense or lax and whether the force of articulation weakens or is stable.

The position of the speech organs in the articulation of vowels may be kept for a variable period of time.

All these factors predetermine the principles according to which vowels *are classified:*

– according to the horizontal movement of the tongue:

– according to the vertical movement of the tongue;

– according to the position of the lips;

– according to the degree of the muscular tension of the articulatory organs;

– according to the force of articulation at the end of a vowel;

– *according to the stability of articulation;*

– according to the length of a vowel.

1. According to the horizontal movement of the tongue, English vowels are classified into front: /i:, e, æ / and the nuclei of the diphthongs /eɪ , ɛə , aɪ /, front-retracted: /I/ and the nucleus of the diphthong /Iə /, mixed: / ɜ:, ə /, back-advanced: / ʊ , ʌ , o:/ and the nuclei of the diphthongs /oʊ , ʊə / and back: /u:, ɔ:, ɒ /.

2. According to the vertical movement of the tongue, English vowels have been traditionally subdivided into close (high), mid and open (low). *It is insufficient, however, to define the articulatory features of vowels* in terms of these 3 degrees of opening of the mouth cavity, since functionally different vowels / i:–ɪ /, /u:–ʊ /, / ɔ:–ɒ / are not described from the point of view of their articulation.

Soviet phoneticians G. Torsuyev [35], A. Trakhterov [39], V. Vassilyev [110] classify these sounds in a more precise manner subdividing each class (close, mid, open) into a narrow and a broad variation. Thus, ac*cording to the height of the tongue, vowels can be classified as* high-narrow ı :, u:/, high–broad / ɪ , ʊ /, mid–narrow /e, ɜ :, o( ʊ )/, mid–broad / ə,ɛ ( ɐ )/, low–narrow / ʌ , ɔ :/, low–broad /æ , ɑ( ɪ , ʊ ), ɑ:, ɒ /.

Their classification reflects the distinctive differencies in the quality of the historically long and historically short vowels.

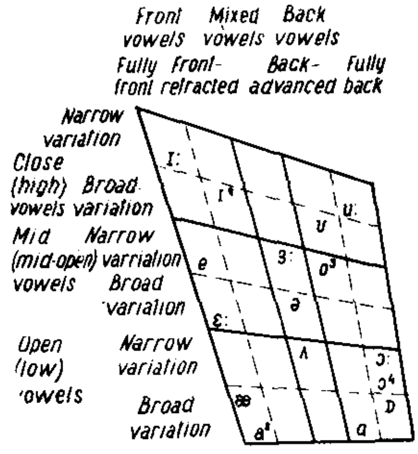


Fig. 6. Table of English vowels:
[1] [ ɛ ] – the nucleus of the diphthong [ɛ ə ], [2] [ ɑ ] – the nucleus of the diphthongs [ɑɪ ] and [ ɑʊ], [3] [o] – the nucleus of the diphthong [o ʊ ], [4] [ ɔ ] – the nucleus of the diphthong [ ɔɪ ].

3. According to the position of the lips, i.e. whether they are rounded, spread or neutral, English vowels are classed into r o u n d e d / ɔ :, ɒ , u: ,ʊ/ and u n r o u n d e d / i :, ɪ , e, æ , ʌ , ɑ :, ɜ :, ə /. The subdivision of vowels into lipspread and lip–neutral is unnecessary for a phonological analysis, but may be useful in describing concreate realizations of the phonemes.

4. According to the degree of muscular tension, English vowels are classified into t e n s e and l a x. Thus, for instance, English /i:/ and /u:/ are characterized as tense, because the speech organs that participate in their formation (the tongue and the lips) are considerably tensed. In the articulation of short / ɪ / and / ʊ / these organs are relatively relaxed, so these vowels are characterized as lax.

All the long vowels are believed to be tense, while short vowels are lax [35]. This is due to the long period of time for which the speech organs are kept in a certain position and this, in its turn, requires greater muscular tension of the speech organs. Not all phoneticians share this opinion. According to D.Jones [84] only the long /i:/ and /u:/ may be considered as tense. D. Jones applies the terms "tense" and "lax" only to close vowels, because in the case of open vowels it is difficult to define whether there is any tenseness or not. This point should be clarified with the help of special electromyographic investigations.

5. According to the force of articulation at the end of the vowel (the character of the end), English vowels are subdivided into f r e e and c h e c k e d. Free vowels are pronounced in an open syllable with a weakening in the force

of articulation towards their end, i.e. they have a fading character. These are all the English *long monophthongs and diphthongs and unstressed short vowels.*

Checked vowels are those in the articulation of which there is no weakening of the force of articulation. They are pronounced abruptly at the end, immediately followed by a consonant that checks them. These are historically short vowels under stress.

6. According to the stability of articulation, English vowels are classed into m o n o p h t h o n g s / iː , ɪ , e , æ , ɑː, ʌ , ɔː , ɒ , uː , ʊ /, d i p h t h o n g s (eɪ , ɑɪ , ɒɪ , ɑʊ, oʊ, ɪə , ɛə , oə , ʊə / and d i p h – t h o n g o i d s , or diphthongized vowels / iː , uː /.

The stability of articulation as in the case of monophthongs or its instability as in the case of diphthongs and diphthongoids is, actually, the stabitity (or instability) of the shape of the oral resonator. When the position of the tongue and the lips during the pronunciation of a vowel is altered to some extent, a new vowel quality is produced. In diphthongs two vowel elements are distinguished – the nucleus and the glide. The nucleus is stronger, more definite in timbre, more prominent and syllabic.

In different languages the nucleus of a diphthong may be either the first or the second element. Diphthongs that consist of a nucleus followed by a glide are falling diphthongs, because the total amount of articulatory energy falls towards the second element. Those consisting of a glide followed by a nucleus are rising diphthongs, since the articulatory energy rises towards second element. English diphthongs are falling. Rising diphthongs are common in Italian.

In some phonetic contexts English diphthongs / ɪə , ʊə / may be pronounced with the second element stronger and more prominent than the first, and are, consequently, rising.

When the diphthong /o ʊ / is pronounced as an exclamation with the high rising tone, the /ʊ/ element in it is as strong and prominent as /o/. So /oʊ / can be called a level diphthong.

D. Jones points out that in unstressed syllables the / ɪ / and / ʊ / elements in / ɪə /, / ʊə / may be weaker than the second element / ə /. E.g. /ˋsɪ ə rɪ̆ ə s/ "serious", /ˋpɪə rɪ̆əd / "period", / ɪ nflŭ ə ns/ "influence", /ˋk ɒ ŋ gr ʊ̆ ənt/ "congruent".

7. Closely connected with the quality of vowels is their quantity, or length. Any speech sound must have certain duration to display its quality, to be perceived as such. According to their length, English vowels are divided into long /iː , ɑː , ɔː , uː , ɜː / and short / ɪ , ʌ , ɒ , e , ʊ , ə , æ /. This length is historical. It differs from the positional length of the same vowels. In connected speech historically long vowels may be of the same length as historically short ones and even shorter. Cf. /biːt/ – /b ɪ d/, /siːt/ – /s ɪ t/. (See Chapter 3).

## B. Consonants

An indispensable constituent of a consonant is noise. The source of noise is an obstruction. There are the following types of obstruction in the production of consonants: 1) complete occlusion (closure), 2) constriction (narrowing) and 3) occlusion–constriction (closure immediately followed by a constriction).

The noise produced by the removal of a closure is that of a plosion, the noise resulting from the movement of the air stream in the narrowing is that of friction. The two effects are combined when closure is followed by a narrowing.

1. According to the type of obstruction and the manner of the production of noise, English consonants are classified in the following way:

- Occlusives
  - stops (plosives) /p, b, t, d, k, g/
  - nasal sonants /m, n, ŋ /
- Constrictives
  - fricatives
    - unicentral /f,v, θ , ð , s, z/
    - bicentral /ʃ ,ʒ /
  - oral sonants
    - medial /j, r, w/
    - lateral /l/
- Occlusive–Constrictives (Affricates) /tʃ , dʒ/

Obstructions may be formed either by two active speech organs or by one active speech organ (articulator) and a passive organ of speech (point or place of articulation).

2. According to the active speech organ which forms an obstruction, English consonants are classed into:

- Labial
  - bilabial /p,b,m,w/
  - labio–dental /v,f/
- Lingual
  - forelingual
    - apical /t,d,n, s,z,θ,ð /
    - cacuminal /r/
  - medio–lingual /j/
  - backlingual /k,g,ŋ /
- Pharyngeal /h/

3. According to the place of obstruction, consonants are classified into d e n t a l (θ,ð /, a l v e o l a r /t, d, n, l, s, z/, p o s t – a l v e o l a r /r/, p a l a t a l /j/, p a l a t o – a l v e o l a r /ʃ ,ʒ , tʃ , dʒ /, v e l a r /ŋ/.

4. According to the presence or absence of voice, English consonants are subdivided into v o i c e d /b, d, g, v, z, ð, ʒ, dʒ/ and v o i c e l e s s /p, t, k, f, s, θ, ʃ, tʃ/.

5. According to the force of articulation, English consonants are classified as l e n i s and f o r t i s.

In the articulation of English voiced consonants the muscular tension is weak – lenis articulation. In the articulation of English voiceless consonants the muscular tension is strong – fortis articulation.

6. According to the position of the soft palate, English consonants are subdivided into o r a l /p, b, t, d, k, g, f, v, ʃ, ʒ, s, z, θ, ð, tʃ, dʒ, w, l, r, j/ and n a s a l /m, n, ŋ/.

In this description of the sounds of the English language we have considered the articulatory characteristics of the sounds, pronounced in isolation. But in connected speech isolated sounds are rather unusual. Sounds are grouped together to form larger units and in the process of grouping they influence one another (the stages of their articulation merge and interpenetrate), and their articulatory features are modified in various ways. Nevertheless, those characteristics of a sound, which are significant for differentiating meaningful units, are preserved in all positions and combinations.

## THE ACOUSTIC CLASSIFICATION OF ENGLISH SPEECH SOUNDS

### A. Vowels

The acoustic classification of speech sounds is based on the analysis of the spectra of the sounds.

The spectra of vowels have a sharply defined formant structure and high total energy. The formants in the spectrum of a vowel are determined by certain articulatory positions and movements of the tongue and the lips. Thus, $F_1$ is conditioned by the vertical position of the tongue. When the tongue is high in the mouth, $F_1$ is low, when the tongue is low, $F_1$ is high. E.g. /i:/ and /u:/ have $F_1$ in the region of 280–300 Hz, whereas /ɑ:/ and /ɒ/ have $F_1$ in the region of 600–800 Hz. The second formant ($F_2$) is conditioned by the horizontal position of the tongue and by the position of the lips. $F_2$ is high in the case of a front vowel and it is low in the case of a back vowel. *Thus, /i:/ has $F_2$ at about 2500 Hz whereas /u:/ has $F_2$ at about 900 Hz.* $F_2$ of rounded vowels is lower than that of unrounded vowels, e.g. $F_2$ of /ʌ/ is 1320 Hz whereas $F_2$ of /ʊ/ is 940 Hz.

If the formants $F_1$ and $F_2$ are in the middle of the spectrum, i.e. close to each other as for /ɑ:, ɒ, æ/, the vowels are classified as compact. If the formants are at each of the extremities of the spectrum as for /u:, ʊ, i:, ɪ/ the vowels are diffuse (see the spectrogram). Open vowels are compact, close vowels are diffuse.

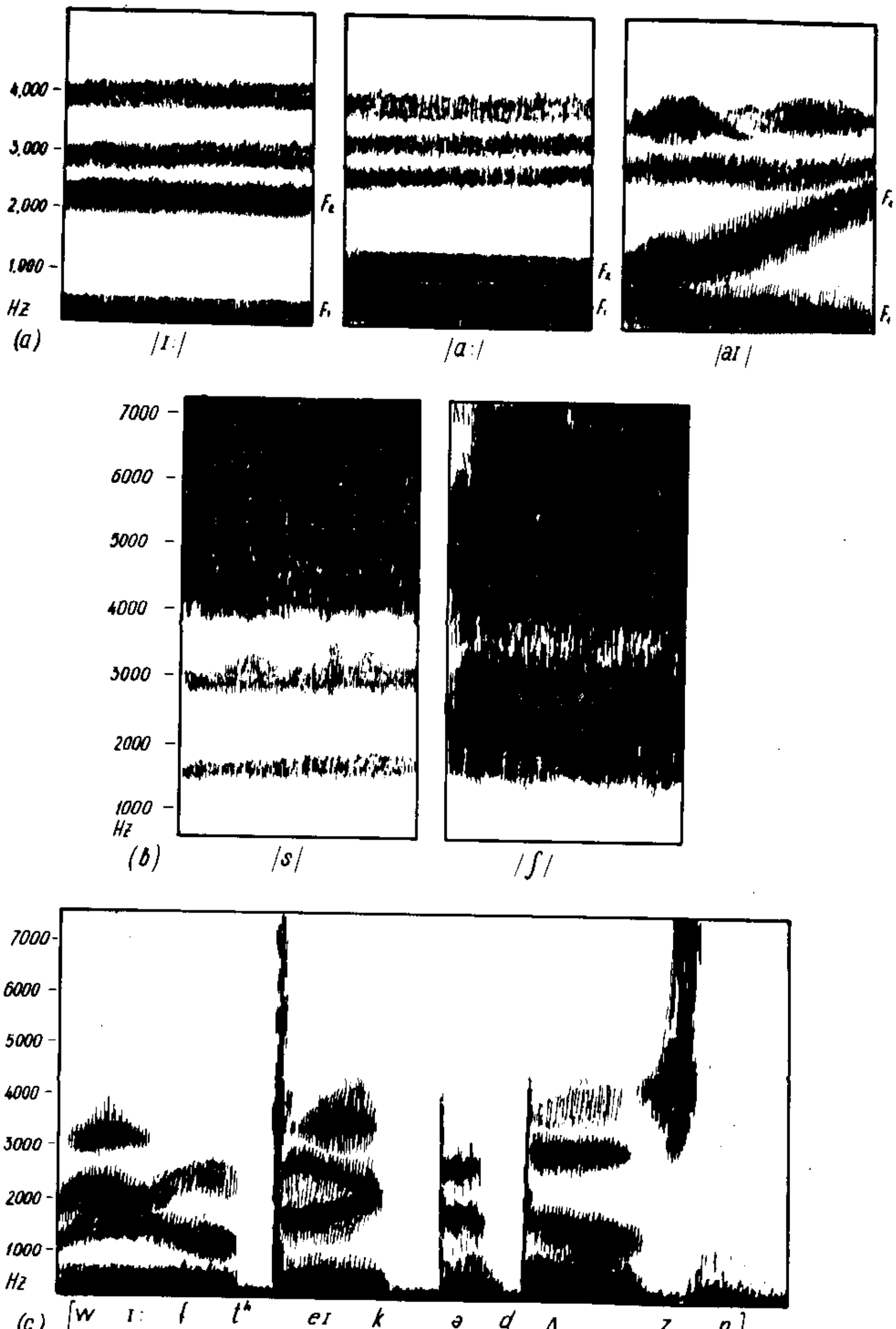


**Fig. 7. Spectrograms of / i :/, /ɑ:/, / ɑɪ /, /s/, / ʃ / and "We'll take a dozen".**

If the second formant is high, as for /i: , e/. the vowels are of a clear or acute timbre. If it is low, so that both $F_1$ and $F_2$ are in the low section of the spectrum (as for /u:, ʊ , ɔ :/)a vowel has a dark or grave timbre. Front vowels are acute, back vowels are grave. $F_2$ is lower in rounded vowels (as ɒ,ɔ :, ʊ, u:) than it is in unrounded (as i:, ɪ , e, æ , ʌ, ɑ :). Acoustically, rounded vowels are opposed to unrounded as flat to plain.

Thus, from the point of view of their acoustic characteristics, the vowel /e/, for instance, is described as acute, compact, plain. The vowel /ɒ/is compact, grave and flat, and /u:/ is diffuse, grave, flat.

## B.Consonants

The formant structure in the spectrum of a consonant is not so sharply defined and the total energy is not so high as in the spectrum of a vowel. The characteristic feature of the spectrum of a consonant is the presence of a formant of noise.

Fricatives /f, v, s, z, θ, ð , ʃ , ʒ/ are characterized by the presence of a noise formant throughout the spectrum.

Plosives and affricates (e.g.t, d, tʃ , dʒ / have a formant of noise only in that part of the spectra which corresponds to plosion and friction. There is no noise in that part of the spectra which corresponds to the articulatory "stop".

Hence fricatives are classed as continuant noises, whereas plosives and affricates are classed as discontinuous.

Voiceless consonants (fortis) are characterized acoustically as tense whereas voiced consonants (lenis) as lax, since the burst of noise in voice-less plosives and the formant of noise in voiceless fricatives are stronger than those in voiced plosives and fricatives.

The noise peculiar to alveolar and dental consonants /t, d, s, z, n, l, θ, ð / is contrasted with that of labial and labio-dental ones /p, b, m, f, v/ because it is sharper in character. This means that in the spectra of /t, d, s, z, n, l, θ, ð/ the formant of noise is higher and in the spectra of /p, b, m,f, v/ the formant of noise is lower.

The fricatives (alveolar and dental) /s, z, θ, ð / have the highest formant of noise in the spectrum. The frequencies of the noise formant in the spectrum of /f, v/ are low. Therefore, /t, d, s, z, θ, ð , n/ are characterized as acute and /p, b, m, v/ as grave. The consonants /k, g, ʃ , ʒ , tʃ , dʒ / are intermediate in this contrast.

The spectrum of velar and palatal consonants /k, g, ŋ , ʃ , ʒ , tʃ, dʒ/ is com-pact while the spectrum of alveolar, labial and dental ones /t, d, n, s, m, p, b, f, v, θ, ð , s, z/ is diffuse. Consequently, the former are classified as compact noises and the latter as diffuse ones.

The sonants /m, n, ŋ / are opposed to all the other consonants as nasal to oral, because in their spectrum there is a special nasal formant.

The sonants /l, r/ are characterized by a sharply defined formant structure (like vowels), but the total energy is low in their spectra (like in consonants). That is why they are classed as vocalic and consonantal.

The sonants /w, j/ are non–vocalic and non–consonantal, because in their spectra the formant structure is not well defined and the total energy is high.

The consonants /s, z/, which have a round narrowing, are opposed to /θ, ð/, which have a flat narrowing, and the affricates /tʃ, dʒ/ are opposed to the plosives /t, d/ as strident to mellow. In the spectrum of strident consonants the intensity of the noise formant is greater than in the spectrum of mellow consonants.

The acoustic characteristics of the sound types of language described above are preserved in the isolated pronunciation of the sounds. But in connected speech the acoustic features of the sounds are considerably modified. The acoustic information about a definite sound is often found not only in the spectrum of this sound but also in the spectrum of the neighbouring sound.

The first attempt to classify speech sounds on the basis of their acoustic distinctions was made by a group of phoneticians R.Jakobson, G.Fant and M.Halle [45]. The scholars established 12 binary distinctive oppositions with the help of which, as they claim, it is possible to classify the phonemes of any language. For the English language, the following 9 binary oppositions are sufficient: 1) vocalic – non–vocalic; 2) consonantal – non–consonantal; 3) compact – diffuse; 4) grave – acute; 5) flat – plain; 6) nasal – oral; 7) tense – lax; 8) discontinuous – continuant; 9) strident – mellow.

The traditional vowel – consonant opposition is divided into two oppositions to define the sounds /r, l, w, j,/. /r, l/ are vocalic and consonantal, because the air passage is free as in vowels and there is an obstruction as in consonants, whereas /w, j/ are non–vocalic and non–consonantal.

In comparison with the articulatory classifications, which have a very long history, the acoustic classification of speech sounds, based on discoveries of modern electro–acoustics, is quite original. The classification worked out by R.Jakobson, G.Fant and M.Halle is of great theoretical importance to linguists. The principle of binary oppositions on which their classification is based is being used not only in phonetics but in other branches of linguistics as well.

Acoustic definitions and classifications of speech sounds are also of great practical importance. They are indispensable in technical acoustics for the solution of the problem of speech synthesis and sound transmission, for the construction of speech recognizers as well as machines capable of putting out information in spoken words.

## EXERCISES

A. Think about the following questions for class discussion:

1. What are the articulatory distionctions between vowels and con-sonants?

2. What are the acoustic distinctions between vowels and consonants? How do their acoustic spectra differ?

3. What articulatory and acoustic features testify to the intermediate position of sonants between vowels and consonants?

4. Does the articulation basis of a language condition the phonetic system of the language? In what way?

5. On what factors does the quality of a vowel sound depend?

6. What are the factors determining the quality of a consonant?

7. How are the formants in the spectrum of a vowel related with its articulation? Can the acoustic engineer do without the knowledge of the articulatory features of speech sounds?

B. Study the table of English vowels:

1) to describe the articulatory characteristics of each vowel;

2) to find out the articulatory distinctions between the pairs of long and short vowels /i: – ɪ , u:–ʊ , ɑ:–ʌ , ɒ:– ɒ /.

C. Study the classification of English consonants:

1) to describe the articulatory characteristics of the sounds /ŋ, n, m, w, h, r, w, j/ ;

2) to find out the articulatory distinctions between the following pairs of consonants /p–b, t–d, k–g, f–v, s–z/, /g–h/.

# CHAPTER 3. PHONOLOGICAL ANALYSIS OF ENGLISH SPEECH SOUNDS

## THE PHONEME

In connected speech a sound is generally modified by its phonetic environment, (i.e. by the neighbouring sounds), by the position it occupies in a word or an utterance; it is also modified by prosodic features, such as stress, speech melody, and tempo of speech.

Compare /p/ in "pill" (i.e. in initial position), in "spill" (i.e. after /s/), in "slip" (i.e. in final position), in "slipper" (i.e. between vowels, the first of which is stressed). These various /p/ sounds differ in manner of articulation and in acoustic qualities. But they do not differ phonologically, if one of the various /p/ sounds is substituted for another, the meaning of the word will not change. That is why for the English speaking people it is of no linguistic importance to discriminate the various /p/ sounds. But it is linguistically important for English speakers to discriminate between /p/ and /b/ (as in "pill" and "bill") or /p/ and /m/ (as in "pill" and "mill"), though the differences in their production might not be much more notable than the differences in the production of the various /p/ sounds. That is why /p, b, m/ are different elements of the English sound system (different English phonemes). The substitution of one for another affects communication.

Every language has a limited number of sound types which are shared by all the speakers of the language and are linguistically important because they distinguish words in the language. In English there are 20 vowel phonemes and 24 consonant phonemes; in Russian there are 6 vowel and 35 consonant phonemes.

All the actual speech sounds are allophones (or variants) of the phonemes that exist in the language. Those that distinguish words, when opposed to one another in the same phonetic position, are realizations of different phonemes. E.g. /v/ and /w/ in English are realizations of two different phonemes because they distinguish such words as "vine" and "wine", "veal" and "wheel" etc.

Those sounds that cannot distinguish words in a definite language and occur only in certain positions or in combination with certain sounds are realizations of one and the same phoneme, its allophones (or variants). In English for example, the "dark" / ɫ / and the "clear" /l/ are variants, or allophones of the same phoneme.

Therefore, the p h o n e m e may be defined as the smallest linguistically relevant unit of the sound structure of a given language which serves to distinguish one word from another.

A l l o p h o n e s (or variants) of a certain phoneme are speech sounds which are realizations of one and the same phoneme and which, therefore, cannot distinguish words. Their articulatory and acoustic distinctions are conditioned by their position and their phonetic environment.

Allophones of a phoneme which never occur in identical positions are said to be in complementary distribution. For example, an RP* speaker pronounces a "dark" allophone of /l/ before consonants in final position, whereas he usually pronounces a "clear" allophone of /l/ only before vowels and /j/. It means that in English both the "dark" and the "clear" allophones of /l/ are in complementary distribution as they are never opposed to each other in identical phonetic positions. In spite of their acoustic and articulatory differences they are not perceived by English speakers as different sounds.

Allophones of a phoneme which do occur in the same phonetic position, but can never distinguish words, are said to be in free variation. For example, /t/ in "Good night" may be either a plosive or a non-plosive sound. Though these sounds differ acoustically, they do not distinguish words in English. Therefore, they are allophones of the same phoneme and are in free variation.

On the one hand, the phoneme is an abstraction and a generalization. It is abstracted from its variants that exist in actual speech and is characterized by features that are common to all its variants (e.g. /b/ is an occlusive, bilabial, lenis consonant, as these features are common to all its allophones).

On the other hand, the phoneme is material, real and objective, because in speech it is represented by concrete material sounds. In other words, the phoneme exists in speech in the material form of speech sounds.

The phoneme correlates with its allophones as the universal correlates with the individual.

"Общее существует лишь в отдельном, через отдельное. Всякое отдельное есть (так или иначе) общее. Всякое общее есть (частичка или сторона или сущность) отдельного. Всякое общее лишь приблизительно охватывает все отдельные предметы. Всякое отдельное неполно входит в общее..." [1, с.318].

The phoneme can therefore be regarded as a dialectical unity of its two aspects: the material and the abstracted aspects. None of these aspects of the phoneme can be neglected or disregarded. That is the materialistic view of the phoneme.

There exist other views of the phoneme. Some linguists consider the phoneme to be but an abstraction and deny its material character. This viewpoint is expressed by linguists of the Prague Phonological School, for whom a phoneme is but an abstract concept.

Other linguists overestimate the material, real and objective character of the phoneme. For example, D.Jones considers a phoneme to be a family of sounds [84], others consider it to be a class of sounds. The /p/ phoneme, for example, is said to 'consist' of its various allophones.

---

*RP stands for "Received Pronunciation", which is generally considered to be the orthoepic norm of British English.

*The linguistic role of the phoneme is clearly seen from the three linguistic* functions of the phoneme, namely, the constitutive, the distinctive, and the identificatory functions [110].

Though the phonemes themselves, in isolation, have no meaning, they are linguistically important, since, in their material form (i.e. in the form of speech sounds) they constitute morphemes, words, all of which are meaningful. Hence, the c o n s t i t u t i v e function of the phoneme.

Besides the constitutive function, the phoneme performs the d i s t i n c t i v e function, because phonemes distinguish one word from another.

On account of the fact that native speakers identify definite combinations of phonemes as meaningful linguistic units (words, word combinations, or phrases), linguists distinguish a third function of the phoneme — the i d e n t i f i c a t o r y (or recognitive) function of the phoneme. It appears that when identifying linguistic units the use of the right phoneme is not the only significant factor, the use of the right allophone is not much less important. Thus, in English it is the aspiration of /p/ rather than its voicelessness, and the non-aspirated character of /b/, that make clear the opposition of /p/ and /b/ in words like "pie" and "buy". This is why an Englishman will often hear "bride" for "pride" when a foreigner uses a non-aspirated /p/.

We may now summarize by saying that the phoneme is a linguistically relevant unit that exists in speech in the material form of its allophones. The phoneme is, therefore, a phonological unit which is represented in speech by phonetic units (the speech sounds). In analysing speech we constantly carry out a phonetic and a phonological analysis. The analysis is primarily phonetic *when we describe the articulatory and acoustic characterictics of particular* sounds and their combinations; but when we determine the role of those sounds in communication, it is mainly phonological analysis. Thus both phonemes and sounds are simply two sides of one and the same phenomenon — the sound substance of language, which can be analysed on either the phonemic (functional) level or the allophonic (variational) level.

## MODIFICATIONS OF PHONEMES IN SPEECH

Every phoneme displays a vast range of variation in connected speech. Among the different types of variation we distinguish idiolectal, diaphonic and allophonic variation.

I d i o l e c t a l *variation embraces the individual peculiarities of* articulating sounds, which are caused by the shape and form of the speaker's speech organs and by his articulatory habits. For instance, a speaker may mumble, or lisp (say "thish ish" for "this is"), or stammer (say "a f—f—f—fine d—d—d—day"). Idiolectal variation may cause a lot of difficulties in communication. At the same time it enables people to identify the speech of definite individuals.

D i a p h o n i c variation affects the quality and quantity of particular phonemes. It is caused by concrete historical tendencies active in certain localities. For example, the diaphonic variation of / æ / involves significant changes in its length, because in some dialects / æ / is much longer than the standard sound. As to its quality, it ranges from a front open /æ / in the southern part of England to /a/ in Northern England. Diaphonic variants do not affect intelligibility of speech, yet they inform the listener about the speaker's origin (i.e. the region he comes from) and his social standing.

The listener easily notices both idiolectal and diaphonic variants, but it does not take him much time to 'tune in' to the speaker's manner of speech and understand him.

The less noticeable variation of phonemes is a l l o p h o n i c variation, which is conditioned by phonetic position and phonetic environment. It has already been mentioned that in connected speech the sounds undergo various modifications under the influence of neighbouring sounds and the intonation patterns they occur in. When viewed linguistically, it means that in a speech continuum there appear a variety of realizations of one and the same phoneme, its allophones. The number of allophones of each phoneme is, therefore, no less than the number of phonetic positions and environments in which the phoneme occurs.

In every language there are positions in which a phoneme can be easily identified because its features are fairly obvious. There are other positions in which the characteristic features of a phoneme are less obvious.

Cf. /oʊ/ in "Oh!" / ^oʊ/, "So—so" /ˈsoʊsoʊ/, "So late" /soˋleɪt/ and in "Not so late" /ˈnɒt sə ˌleɪt/.

Or again, note the differences in /d/ in "Oh, dear" (where its features are obvious), "Good!" (where it generally loses its plosive and voiced character) and in "Good—bye" (where it actually loses all its features).

English vowels are considerably modified in unstressed syllables. The weakening of articulation and shortening of the duration of unstressed vowels results in modification of their quality and quantity. This phonetic phenomenon is known as r e d u c t i o n. Vowels in unstressed syllables are sometimes partially reduced, as in /soˋleɪt/, or reduced to /ə/, as in /ˈnɒt səˌleɪt/.

The obscuring of vowels owing to lack of stress on certain syllables in an utterance is the result of a perfectly normal linguistic development of English. It is the artificial departure from this characteristic of English that is a corruption of the actual pronunciation of the cultivated people.

J. Kenyon remarks that when the schoolboy writes "I wouldn't of gone" he is not making a mistake in grammar, the mistake is merely in spelling. The two expressions "I wouldn't have gone" and "I wouldn't of gone" are exactly alike in standard pronunciation. He says that to pronounce "have" and "of" differently in sentences that require the unstressed forms, would be a worse blunder than to misspell "have". [87]

Apart from that, English vowels are also modified by the neighbouring consonants, mainly by the following consonant. There is always an overlapping of articulatory movements of neighbouring sounds. An illustration of that is the nasalization of /e/ in "men", "ten", the shortening of /i:/ in "cease" (as compared to "seize" or "see"). The process of adapting the articulation of a vowel to a consonant, or a consonant to a vowel, is known as a c c o m o d a – t i o n.

As for the quality of an English consonant, it depends mainly on its position in the intonation group.

Spectrographic analysis makes it evident that the modifications a consonant undergoes in the final position of an intonation group radically differ from those in the initial position. For example, in the final position voiced consonants lose their plosive and voiced character (as in "Good!"), whereas in the initial position they retain those features but are modified in another manner by the following vowel. For example, /ʃ, ʒ, tʃ, dʒ, l/ are slightly palatalized when followed by mid and close vowels or by /j/, as in "shirt", "cheese", "leap" or "June". The palatalization of these sounds is caused by the anticipatory upward movement of the bulk of the tongue for the next sound.

Or again, the height of the central part of the tongue for /j/ fully depends upon the following vowel. Thus before /i:/ (as in "yield") the central part of the tongue is raised higher than for /i:/; in the word "yes" it is as high as when articulating /ʝ/, while in "yard" it reaches but the height of /e/.

When a consonant is a component of a consonant cluster, it is partially or completely assimilated by the neighbouring sounds. A consonant may be voiced (as in /tra:nzˈleɪt/) or devoiced (as in /aɪ ʃt ˋθɪŋk soʊ/), it may lose its plosion (as in "that time"), or the plosion may become restricted (as in "please, "great"), there may even occur coalescent assimilation which results in a new phoneme (as in /ˈhaʊ dʒə ˋdu:/ or /woʊntʃə ˏdu:ɪt/). All English phonemes in various pronouncing conditions undergo a s s i m i l a t i o n.

To state that assimilation, accomodation, reduction depend only on tempo and style, that they occur only in rapid careless speech, would be wrong and would simplify the problem. One should admit that these phenomena occur more obviously in rapid and careless speech, but they are also observed in careful speech with moderate tempo and even when the tempo is slow.

The analysis of the phonetic modifications that occur in the speech continuum reveals the phonetic tendencies of a language. Neglect of these tendencies results in foreign accents and unintelligible speech. This is one of the main reasons why there is a need for further investigation of this problem.

It is generally considered that allophonic modification is caused by "economy of effort". The speaker avoids articulatory movements which are not absolutely necessary for intelligibility of speech. This process is to some extent regulated by the orthoepic norm, by the system of phonemes in the language, and by the system of phonologically relevant features of phonemes in the language.

Thus, English lenis consonants /b, d, g/ in final position can be voiceless, but they cannot be replaced by fortis /p, t, k/, as in English the fortis and lenis consonants distinguish words (e.g. "cab"–"cap", "had"–"hat", "bag"–"back"). It has been noted that /ʃ/ may be "clear" or "dark" in one and the same phonetic position; but the "clear" allophone of /l/ does not occur in the positions in which the "dark" allophone of /l/ is used, if the latter is used instead of the former, it is immediately detected by the native speakers as a local accent.

M.Grammont claims that the different phenomena of combinatory phonetics are regulated by what he calls "the law of the stronger" [74]. According to it, the stronger phoneme influences the weaker one. The stronger phoneme assimilates, or accomodates the neighbouring phoneme because of its articulatory strength and stability, or by its position in the syllable. That can be easily illustrated by an example from the English language: in "of course" / ə f ˋk ɔ :s/ the /v/ is assimilated by /k/ (and not vice versa), because /v/ here is at the end of an unstressed syllable and is therefore weaker than the initial /k/. M. Grammont considers that the analysis of combinatory phenomena in every particular language should aim at revealing the phonemes that resist modifications. But the inventory of the strong and stable phonemes has not as yet been established for any particular language.

Some scholars consider that frequency of occurrence of phonemes and phonemic clusters may be a factor of stability in language in the sense that frequent phonemes resist modifications and modify the rare ones. Analysis of relative frequency of occurrence of English speech sounds shows that the most frequent consonants are /t, n, s, ð, l, d/. [68] Whether they really are the most stable sounds in English requires thorough analysis.

Whatever are the causes, the fact that speech is comprised of a great variety of allophones complicates the identification of phonemes in connected speech. It also accounts for the existence of controversial views on phonological problems. For instance, some phoneticians treat /j/ and /w/ as allophones of /i:/ and /u:/; there are differences in the opinion on the phonological status of /ʊ/, /ʌ/ and /ɒ/, some scholars treat them as allophones of /u:/, /ɑ:/, and /ɔ:/ respectively.

Consequently, the main problems of phonological analysis are as follows:

a) the identification of the phonemic inventory for each individual language;

b) the identification of the inventory of phonologically relevant (distinctive) features of a language;

c) the interrelationships among the phonemes of a language.

## THE PHONEMIC INVENTORY OF ENGLISH

The first problem of phonological analysis is to establish the phonemes in a definite language. This can be carried out only by phonological analysis based on phonological rules. There are two methods to do that: the distributional method and the semantic method.

The d i s t r i b u t i o n a l method is based on the phonological rule that different phonemes can freely occur in one and the same position, while allophones of one and the same phoneme occur in different positions and, therefore, cannot be phonologically opposed to each other. For example, as /p/ and /b/ can freely occur in the same phonetic context (as in "pea" – "bee", "rope" – "robe"), they are consequently different phonemes. But one cannot find [p] aspirated and [p] non–aspirated in the same phonetic position in English. Therefore in English they are allophones of one and the same phoneme, whereas in Chinese the aspirated and non–aspirated stops are regarded as different phonemes, because they occur in the same phonetic positions.

The distributional method of analysis is a purely formal method of identifying the phonemes of a language. This is why the distributional method of identification of the phonemes in a language works even when one does not know the language at all. The method is widely used by the American linguists who study the languages of the Red Indians. But it appears to be complicated and the investigators very often cannot do without native speakers to confirm their conclusions concerning the phonemic status of certain speech sounds.

The s e m a n t i c method, in its turn, is based on the phonological rule that a phoneme can distinguish words when opposed to another phoneme or zero in an identical phonetic position.

The opposition /z/ versus /t/ is called a p h o n o l o g i c a l o p p o s i t i o n. The opposition /z/ versus /–/ is called a z e r o (phonological) opposition. The pairs of words which differ only in one speech sound are called m i n i m a l p a i r s.

The semantic method of identification of the phonemes in a language attaches great significance to meaning. The investigator studies the function of sounds by collecting minimal pairs of words in the language. If two speech sounds distinguish words with different meanings, they form a phonological opposition and are realizations of two different phonemes. If not, they are allophones of one and the same phoneme. Thus, it is clearly evident that in English [s] and [t] are realizations of two different phonemes ("sea"–"tea", "so"–"toe"), while [t] aspirated and [t] non–aspirated are allophones of one and the same phoneme, as they cannot distinguish words. Such analysis is sometimes referred to as "minimal pair test".

The semantic method is widely used by scholars all over the world. It was mainly with the help of the semantic method that Soviet scholars identified the phonemes in the languages of the numerous peoples inhabiting our multi-national country. That made it possible to create written languages for them.

But to identify all the phonemes of a language is not always a simple thing to do. Time and again there emerge difficulties as to the phonemic status of certain sounds. Such difficulties arise when one deals with weakened vowels occurring in unstressed position. It primarily concerns the schwa vowel /ə/ in English which occurs only in unstressed position.

The problem is whether there is a schwa vowel /ə/ phoneme in English. Scholars are not in agreement on this point. Though /ə/ can be opposed only to weakened vowel phonemes, which are partially reduced due to their position in unstressed syllables, it can form phonological oppositions with a number of other phonemes and can distinguish words.

| | | |
|---|---|---|
| E.g., | /ə/ vs. /ɪ/ | accept – except; armour – army |
| | | officers – offices; allusion – illusion |
| | /ə/ vs. /oʊ/ | temper – tempo |
| | | solar – solo |
| | /ə/ vs. /ɜ:/ | forward – foreword |

It is sometimes considered that /ə/ is an allophone of /ʌ/, because /ʌ/ is almost exclusively used in stressed syllables (as in "comfort" /ˈkʌmfət/, "abundant" /əˈbʌndənt/), whereas /ə/ occurs only in unstressed syllables.

There are controversial views on whether /j/ and /w/ in English are allophones of /ɪ/ and /u/ or they are separate phonemes. R.Jakobson and other American linguists treat them as allophones of /ɪ/ and /u/ on account of their weakness and unstable articulatory features. [82] Whereas other scholars treat /j/ and /w/ as phonemes, because as they say /j/ and /w/ can form phonological oppositions with each other and with other phonemes (e.g. "yell" – "well", "yet" – "met", "wheat" – "meat"). We support the second viewpoint not only for the reason mentioned above. One ought to mention the fact that /j/ and /w/ in English occur in phonetic positions that are generally occupied by consonant phonemes; consequently, they cannot be considered to be allophones of vowel phonemes.

There are cases when the establishment of phonological oppositions is not sufficient to determine the phonemic status of a sound, especially when the sound is of a complex nature.

In the English language the sounds /tʃ/, /dʒ/, /tr/, /dr/, /ts/, /dz/ form phonological oppositions and distinguish such words as "eat – each", "head – hedge", "tie – try", "die – dry", "hat – hats", "buzz – buds". But does that mean that all of them are monophonemic and should be included into the phonemic inventory?

N. Trubetzkoy worked out a number of rules which help to determine whether a sound of a complex nature is monophonemic. The main rules state that, firstly, a phoneme is indivisible, as no syllabic division can occur within a phoneme. Secondly, a phoneme is produced by one articulatory effort.

Thirdly, the duration of a phoneme should not exceed that of other phonemes in the language. [40]

Consequently, /tʃ/ and /dʒ/ in words like "cheese, each, jail, hedge" are monophonemic, because both acoustic and physiological analysis provide sufficient evidence that these sounds are produced by one articulatory effort, their duration does not exceed the duration of either /t/ (as in "tear"), or /ʃ/ (as in "share"), or /d/ (as in "dare"). Besides that, in words like "cheese, chair, each, hedge, John, jail", no syllabic division occurs within the sounds /tʃ/ and /dʒ/.

/ts/, /tz/ are obviously biphonemic combinations (i.e. combinations of two phonemes), because their duration exceeds the average duration of either /t/, /d/, /s/ or /z/.

As for /tr/, /dr/ (as in "tree, dream") their phonemic status will remain undecided until special acoustic and physiological analysis is made. As /t/ and /r/, on the one hand, and /d/ and /r/, on the other, are so closely linked in the pronunciation of Englishmen, D.Jones calls them affricates alongside of /tʃ/ and /dʒ/. [102] (He distinguished them from sequences /t/ + /r/ and /d/ + /r/ as in "rest–room, hand–writing".) Most phoneticians regard /tr/ and /dr/ as biphonemic clusters.

There appears to be another analogical problem. It concerns the phonemic status of the English diphthongs and the so–called "triphthongs". Are they monophonemic or biphonemic clusters in English?

The syllabic and articulatory indivisibility of English diphthongs* and their duration which does not exceed the duration of English historically long vowels /i:, u:, ɔ:, ɑ:, ɜ:/, clearly determine their monophonemic character in English.

As for /aɪə/, /aʊə/, it has been proved acoustically and physiologically that in English they cannot be considered monophonemic. They are not produced by a single articulatory effort, as there is an increase in the force of articulation and intensity not only for the first element, but for the last element as well. The syllabic division generally occurs in between the diphthong and the schwa vowel (/aɪ–ə/, /aʊ–ə/ as in "flier", "flower"). On account of that they should be regarded as biphonemic clusters of a diphthong with the schwa vowel.

In such a way it has been established that in RP there are 12 vowel phonemes: /i:, ɪ, e, æ, ɒ, ɑ:, ʊ, u:, ɔ:, ʌ, ɜ:, ə/, 8 diphthongs:

3 glides to /ɪ/ – (eɪ, aɪ, ɒɪ/,

2 glides to /ʊ/ – /oʊ, aʊ/,

3 glides to /ə/ – /ɪə, ɛə, ʊə/;

---

*Cf. /bloʊ–ɪŋ/ "blowing", /pleɪ–ə/ "player" and [пó–йу] "пою", [ус–пэ–йу] "успею".

24 consonant phonemes: /n, t, d, s, l, ð, v, m, k, w, z, r, b, f, p, h, ŋ, g, ʃ, j, dʒ, tʃ, θ, ʒ/.

As the diphthong /ɒə/ and the labiovelar fricative /ʍ/ are not used by all RP speakers (only some RP speakers differentiate words as "pour – paw", "which – witch") they are generally called facultative phonemes. But since they are used by many RP speakers as phonemes, V.A. Vassilyev includes them into the phonemic inventory and states that in RP there are 21 vowel phonemes and 25 consonant phonemes. [110]

## THE SYSTEM OF PHONOLOGICAL OPPOSITIONS IN ENGLISH

The second problem of phonological analysis is the identification of the inventory of distinctive features on which all the phonological oppositions in the language are based.

Every sound is characterized by a number of features, not all of which are equally important for communication. If one compares some of the allophones of /p/, it appears that all of them have common features and features which characterize only a few of them. The problem is to decide which of the features of a group of common sounds in a certain language are phonologically relevant and which of them are irrelevant, or incidental. This is important not only for a detailed characterization of the phonemic systems of languages and the identification of their typology. It is also most essential for teaching purposes, since the phonologically relevant features require special attention in pronunciation teaching.

We primarily need to explain the following:

Each phoneme is characterized by a certain number of **phonologically relevant features**, which are its constant distinctive features (as they distinguish the phoneme from all the other phonemes of the language).

Each allophone of a certain phoneme is characterized by definite phonologically relevant features (which are common to all its allophones) plus a number of **irrelevent**, or incidental, **features** (which distinguish the allophone from all the other allophones of the phoneme).

Phoneme: /p/

Allophones:

| $p_1$ | $p_2$ | $p_3$ |
|---|---|---|
| pi: | pɔ: | ʌp |
| bilabial | bilabial | bilabial |
| occlusive | occlusive | occlusive |
| fortis | fortis | fortis |
| aspirated | aspirated | non–aspirated |
| plosive | plosive | non–plosive |
| unrounded | rounded | unrounded |

The phonologically relevant features that characterize the phoneme /p/ are, therefore, bilabial, occlusive and fortis. Aspiration, plosiveness, labialization, etc. are phonologically irrelevant features.

Phonologically irrelevant does not necessarily mean useless for communication. It has already been mentioned that the aspiration of /p/ helps the listener to distinguish it from /b/ (as in "pride" – "bride", "pie" – "buy").

The point is that if the speaker substitutes one phonologically relevant feature (say, bilabial) for any other relevant feature (say, forelingual), the phoneme ceases to be the phoneme it was and becomes a different phoneme (in this case /p/ is replaced by /t/). Such a substitution is easily perceived by any native speaker whether he had been trained in phonetics or not. (Cf. "pie" – "tie", "cap" – "cat").

The substitution of one irrelevant feature for another (say, aspirated for non–aspirated) results in a different allophone of one and the same phoneme ([p ] aspirated and [p] non–aspirated). Such a substitution does not affect communication.

When interpreted phonologically, it means that not all the articulatory features of actual sounds are equally important for the identification of the phonemes of a language and the words they constitute. Some of the features are phonologically relevant, others are irrelevant.

The phonologically relevant features are normally identified by opposing one phoneme to every other phoneme in the language. But there often occur difficulties, which can be overcome with the aid of physiological and acoustic analyses.

Thus, until recently it was considered that the oppositions /p–b/, /t–d/, /k–g/, /f–v/, /θ – ð/, /s–z/, /ʃ – ʒ/, /tʃ–dʒ/ were based on the presence and absence of voice. But it has been proved that the presence or absence of voice in these oppositions is not a constant distinctive feature, because the so–called "voiced" consonants in many phonetic positions are actually not voiced throughout. (/d/ in "<u>Do</u> that" is voiceless in its initial stage, while /d/ in final position is voiceless either in its last phase or throughout, as in "Yes, I did". Ch. Barber states that the so–called "voiced"/b, d, g/ are actually voiceless after diphthongs and historically long vowels, as in "rogue", "feed"). [54] Consequently, their voiced character cannot be considered to be a phonologically relevant feature. Yet, the oppositions /p–b/, /t–d/, etc. exist in the English language, and "cob" is never pronounced as "cop", "had" can never be substituted for "hat", and so on. So there must be at least one phonologically relevant feature on which such oppositions are based.

Acoustic and physiological analyses have proved that the so–called "voiced" consonants in English are always l e n i s (lax, weak) irrespective of their phonetic environment and position, as compared to the so–called "voiceless" consonants in English which are always f o r t i s (energetic, strong) in all phonetic positions. Compare /p ɪ k/ and /b ɪ g/. /p/ and /k/ are pronounced with the lips quite tense, with aspiration (especially in the initial /p/) and with the vocal cords *not vibrating: the /b/ and /g/ are pronounced with the lips* more relaxed, with no aspiration, and with the vocal cords vibrating only in the initial /b/.

Therefore, the phonological oppositions /p–b/, /t–d/, /k–g/, /f–v/, /θ–ð/, /s–z/, /ʃ–ʒ/, /tʃ–dʒ/ are primarily based on fortis vs. lenis articulation, which are their phonologically relevant features. Besides the energy difference, the fortis/lenis correlation also implies that the lenis sounds are regularly shorter than the fortis ones. Thus /z/ is a good deal shorter and much less energetic than /s/ (cf. "lose" – "loose", "as" – "ass"). The presence of voice in /b, d, g, v, z, ʒ, dʒ/ is an incidental feature that can be neutralized in certain phonetic positions. That is why, when teaching English pronunciation, the importance of the voiced character of these sounds should not be overestimated, whereas special attention should be drawn to their lenis character.

Or again, until recently duration in the English language was regarded as a phonologically relevant feature capable of distinguishing /i:– ɪ /, /u:– ʊ /, /ɔ:–ɒ/, /ɑ:–ʌ/, /ɜ:–ə/. But in English the quality of the historically long and short vowels differs, so length is not the only feature that distinguishes them. Besides, length differences are conditioned, they therefore cannot be distinctive. Acoustic analysis shows that the length of vowels varies in different phonetic environment and in different positions.

/si: – si:d – si:t/

It is a well known fact that /i:/ in /si:/ is longer than the same sound in /si:d/, and that /i:/ in /si:t/ is the shortest: it is almost as short as /ɪ/ in /sɪd/.

It has also been established that a vowel is longer in front of a fricative than in front of a stop. Thus, /æ/ is longer in /æs/ than in /æt/. Besides that, vowel length depends on whether the syllable it occurs in is stressed. Stressed vowels are generally longer than the unstressed vowels. Vowel length also depends on the number of syllables in a word. Compare the duration of /ɑ:/ in "arm" – "disarm" – " disarmament".

As length varies and does not characterize all the allophones of a historically long vowel, length cannot be considered its phonologically relevant feature. But there are perceptual features which constantly distinguish all the English vowels: these are distinctions in their quality, which are based on the slight differences in the tongue positions when producing these vowels, i.e. their articulatory differences.

Thus, the opposition /i:– ɪ/ is based on the following phonologically relevant features: high–narrow vs. high–broad, fully–front vs. front–retracted. The opposition /ɔ :– ɒ / is based on the following phonologically relevant features: back–advanced vs. fully back, low–narrow vs. low–broad.

Duration, though it is an incidental feature and therefore cannot be considered phonologically relevant, is nevertheless a very important feature that serves as an additional means of identifying English sounds. For example, the shortening of a vowel generally signals that the following consonant is fortis and voiceless (cf. "seed"–"seat", "had–"hat").

When a monophthong is opposed to a diphthong, there is one extra distinctive feature that differentiates them: it is absence vs. presence of a glide.

Therefore, in the system of English vowel phonemes there are oppositions of

| | |
|---|---|
| monophthongs between them-selves and diphthongs between themselves | these oppositions are based on differences in the movements and positions of the tongue and the lips. |
| monophthongs vs. diphthongs | these oppositions are based on both the above mentioned differences and the absence or presence of a glide. |

There are no phonological oppositions in the system of English vowel phonemes based on length alone.

In the system of English consonant phonemes there are oppositions based on the force of articulation:

fortis vs. lenis (two – do, back – bag, etc.)

There are oppositions based on the active organ of speech:

bilabial vs. forelingual (pen – ten, mat – sat, wet – let)
bilabial vs. mediolingual (wail – Yale)
bilabial vs. backlingual (bay – gay)
forelingual vs. mediolingual (less – yes)
forelingual vs. backlingual (take – cake)
forelingual vs. pharyngal (they – hay, see – he)
labiodental vs. forelingual (fat – sat, van – than)

There are oppositions based on the type of obstruction:

plosive vs. fricative (pen – when, berry – very)
plosive vs. affricate (till – chill, do – Jew)
plosive vs. sonorant (bad – mad, deck – neck)
fricative vs. sonorant (vine – wine)
fricative vs. affricate (share – chair)
constrictive sonorant vs. occlusive sonorant (we – me, low – no)
flat narrowing vs. round narrowing (thing – sing)
unicentral vs. bicentral (see – she, thigh – shy)

## THE INTERRELATIONSHIPS AMONG THE PHONEMES OF A LANGUAGE

Once the phonemes of a language are established and their phonologically relevant features are determined, there arises another phonological problem: to describe the interrelationships among the phonemes of a language.

Can different phonemes have common allophones? Can allophones of a phoneme lose any of their phonologically relevant features in certain phonetic positions?

There are three views on the problem.

Scholars who support the morphonological viewpoint (R. Avanesov [20], P. Kuznetsov [29], A. Reformatsky [33] and others) claim that a phoneme in a "weak" position may lose one of its distinctive (phonologically relevant) features and, therefore, lose its distinctive function. For example, Russian voiced consonants lose their voiced character and are pronounced as voiceless in final position (as in "луг" /k/ "глаз" /c/, etc.). This leads to the loss of the distinction between /к/ and /г/, /с/ and /з/. Therefore, in word final position the phonological oppositions based on the phonologically relevant features "voiced vs. voiceless" are neutralized in Russian. Scholars term this phenomenon n e u t r a l i z a t i o n of phonological oppositions.

*Neutralization of phonological oppositions is the loss of a distinctive* (phonologically relevant) feature by one of the phonemes of an opposition.

Those who support this view consider that a phoneme is morphemically bound and, therefore, in all the derivatives of "луга" (лугов, луг) we deal with the allophones of one and the same phoneme /г/, and in all the derivatives of "лука" (луком, лук) we deal with the allophones of the phoneme /к/.

Consequently, different phonemes may have common allophones and sometimes a sound may be assigned to either of two phonemes. In the case of [k], it may either be considered an allophone of the phoneme /к/ (as in "лук") or an allophone of the phoneme /г/ (as in "луг").

But the Russian language is the only language in which the phenomenon of neutralization has been examined more or less in depth.

The supporters of the phonological viewpoint (L. Shcherba [44], D.Jones [86], K.Pike [104] and others) reject the notion of "neutralization of phonological oppositions". They consider that an allophone cannot lose any of *its distinctive features. If it does, it becomes an allophone of the phoneme the* distinctive features of which it acquires. Thus, [к] in "луг" is an allophone of /к/, [ə ] in "addition" is an allophone of the schwa vowel phoneme /ə / (and is not an allophone of /æ/, as in "add"); [t] in "walked" is an allophone of /t/.

The third viewpoint is that of N. Trubetzkoy [40], R. Jakobson [82], and some other linguists who consider that there are phonological units higher than a phoneme — the archiphonemes.

The archiphoneme is an abstraction which combines the distinctive features common to two phonemes. According to this viewpoint both [k] and [г] in "лук" and "луг" are assigned to the archiphoneme /K/ which is neither voiced, nor voiceless.

We assume that for teaching purposes the most suitable viewpoint is that of L. Shcherba and his followers. Accordingly, the phoneme is characterized by definite articulatory and acoustic characteristics and can be easily described as a separate unit of the sound system of language. Whereas the other viewpoints treat the phoneme as a phonological unit which is actually devoid of articulatory and acoustic characteristics, because even its phonologically relevant features appear to be unstable (they can be neutralized). Moreover, the phoneme in that sense embraces sounds that can be assigned to other phonemes as well (the so–called "common" allophones). Such an approach hinders the practical application of phonology to teaching pronunciation.

The existence of a number of viewpoints on phonological problems can be explained by the well–known fact that language is too complicated for all its features to be described in terms of any one theory.

## TYPES OF TRANSCRIPTION

Besides the problems of phonological analysis of speech sounds discussed above, phonology deals with the problem of representing speech visually. This problem is closely connected with the problems of phonological analysis already discussed, because any system of writing is not a simple record of speech utterances, it is always a generalization about them. Be it ideographic writing (with a different symbol for each word, as in Chinese writing), or syllabic writing (with a symbol for each syllable, as in Japanese writing), or alphabetic writing (with a symbol for each phoneme or combination of phonemes, as in English), writing systematizes and provides a distinctive symbol for each class of sounds it represents.

A t r a n s c r i p t i o n, which is a visual system of notation of the sound structure of speech, is also a generalization of a great variety of sounds that are uttered by speakers of a given language.

The extent of the generalization may vary. One can classify the sounds into phonemes disregarding the different degrees of aspiration, labialization, length, palatalization and other phonologically irrelevant features of the sounds. On the other hand, one can differentiate between all those features and classify them as well. Consequently, there may be different types of transcriptions depending upon the degree of exactness required.

If it is accuracy only in the representation of the phonemes of the language that is required, the transcription should provide each phoneme with a distinctive symbol to avoid ambiguity. Such a transcription is generally called p h o n e m i c, or b r o a d, transcription. It contains as many symbols as there are phonemes in the language. The phonemic data is usually enclosed between virgules (also called diagonals): /t/.

If it is exactness in the differentiation of the allophones of each phoneme that is required, the transcription should provide either different symbols for each allophone, or introduce special marks to represent the different features of the allophones. The former would increase the number of symbols considerably, and that would cause great difficulties for those who use it. The latter is more economical, although it makes the notation rather complicated. Scholars usually make use of both ways: they provide some of the typical allophones with distinctive symbols and introduce special marks (called "diacritic marks") to denote the different features the allophones are characterized by. Such a transcription is called a p h o n e t i c, or n a r r o w transcription. The phonetic data is customarily enclosed in square brackets: [t].

It has always been one of the main concerns of the phonetic science to work out a transcription. A phonetic transcription is essential for scientific and practical use. One can hardly do without it in foreign language teaching, in studying spoken languages, dialects, accents. The first attempts to represent speech sounds visually by means of special symbols were made as far back as the 16th century.

The modern phonetic transcription that is most widely used now is the International Phonetic Transcription devised by the International Phonetic Association in 1904. This transcription is a phonetic alphabet which may be applied to most of the languages. That is why it contains symbols that stand for phonemes in different languages. E.g. /æ / (as in "bag"), /y/ (close lip rounded/ i/in German "u"), /o/ (close lip rounded/e/in French "peu"), etc. For this reason the transcription is often referred to as the "u n i v e r s a l t r a n s c r i p t i o n" of the IPA (International Phonetic Association).

One of the principles of this transcription is to use the fewest possible symbols of the simplest possible shape. Most of the symbols it uses are letters of the Latin alphabet. Besides, it contains a series of diacritic marks.

The broad type of the International Phonetic Transcription was first used by D.Jones in his "English Pronouncing Dictionary", published in 1917. These are the symbols that he selected for English: /i:, i, e, æ , ɑ:, ɔ :, ɔ , u: ,u, ʌ , ə :,ə , ei, ou, ai, au, ɔ i, iə , ɛə , u ə /.

The narrow type of the transcription makes use of extra symbols /ɒ , ʊ, ɜ , ɪ / and diacritic marks, such as

~ nasalization; ɛ̃ = nasalized ɛ .
• devoicing; n̥, l̥, ʒ̊ = unvoiced n, l, ʒ .
ˇ voicing; s̬ = z , t̬ = American "voiced" t.
+ advanced variety; u + or u̟ = sound between u and ʉ.
– retracted variety; a– or a̠ = sound between a and ɑ .
˔ raised variety; a˔ or a̝ = ạ = æ .
˕ lowered variety; e˕ or e̞ = ẹ .
ʿ slight aspiration after p, t, etc.
ˌ under a letter (or over it if the letter has a tail below) means that the sound is syllabic; n̩ = syllabic n.
: length mark.
• half length. [86]

The American linguists use what is often called the "linguistic alphabet". L. Bloomfield [ 55 ] was the first to use it, later it was expanded by B.Bloch, G.Trager [56] and other American linguists. The "linguistic alphabet" includes new symbols:

/ɨ/ for /ʌ/ (e.g. "just" /ǰɨst/),
/ih/ for /ɪə/ (e.g. "near" /nih/),
/iy/ for /i:/ (e.g. "mean" /muyn/).
/š/ for /ʃ/,
/ž/ for /ʒ/,
/č/ for /tʃ/,
/ǰ/ for /dʒ/.

If the symbols / š , ž , č , ǰ / suggest that the sounds are really monophonemic consonants, it is doubtful whether the postvocalic glides /w, h/ in /aw, ow, uw, ih/ are really allophones of the initial /w, h/.

The "linguistic alphabet" is widely used by those who study American English pronunciation.

Phonology has developed rapidly and made a profound study of the functions of sounds in most of the living languages. But so far it has not as yet fully examined the functions of most of the prosodic features, such as speech melody, tempo of speech, rhythm and others. Scholars are beginning to tackle these and similar problems of intonology.

## EXERCISES

A. Think about the following questions for class discussion:

1. Every language contains an immense amount of actual speech sounds. What helps the listener to cope with what he hears? Why is the phoneme concept essential to the phonetic science?

2. Why can't allophones of one and the same phoneme be in contrast with each other? Can they occur in exactly the same phonetic environment? Can they distinguish differences of meaning?

3. Why isn't it possible to make any rigid separation of phonetics and phonology? Does phonetic analysis necessarily involve phonological considerations? Are phonological considerations always based on the analysis of the relevance of phonetic features?

4. All the similarities and differences between the phonemes of a language can be reduced to a comparatively small number of distinctive features. These features have been presented in articulatory terms. Can they also be presented as oppositions between acoustic features, e.g. grave and acute, compact and diffuse, etc? Which of the two presentations is preferable for making comparisons between the phonemes of

different languages? Which of the two presentations is preferable for teaching purposes?

5. Why isn't there unanimous agreement among the phonologists as to the system of English phonemes? Why can /tr/ and /ju/ be treated as single–unit phonemes? Are there any other problems which have no single solution?

6. It is a characteristic feature of many RP speakers and non–RP speakers to pronounce the glottal stop [ʔ] before /p/, particularly when /p/ is followed by a pause. For instance, in "What a hope!" the vowel /oʊ/ is cut off by a complete closure of the vocal cords, the /p/ is exploded only after the glottal stop is released. The glottal stop is used instead of /t/ when the following consonant is a stop, as in "Not good enough". The glottal stop is also used in other phonetic positions. Why isn't the glottal stop included into the system of English phonemes in spite of the fact that it is so widely used? Can it be regarded as an allophone of /t/?

7. In every language there are certain regularities in the modifications of sounds in connected speech. Is it important to be well aware of these regularities to identify easily the words and phonemes of the language? What peculiarities of English combinatory phenomena should the learner's attention be directed at?

8. Why isn't it possible to develop an orthography for an unwritten language without an adequate phonemic analysis of the language?

B. Provide some evidence to prove that in English

(a) aspirated and non–aspirated /t/ are allophones of one and the same phoneme;

(b) /ŋ/ and /n/ are realizations of two different phonemes;

(c) /aɪ/, /aʊ/, /ɔɪ/ are single–unit phonemes.

C. English spelling is by no means a reliable guide to correct pronunciation, as it uses the same letter or letters for different sounds and it gives the same sound different spellings. Compare, for example, the pronunciation of the "ough" combination in "cough", "'tough", "thought", "through", "though", "thorough". Or again, compare the spelling of /ɪ/ in "little", "busy", "enough", "women". Find more examples to show how essential a phonetic transcription is for practical purposes.

D. Compare /si:d/ and /si:t/. The phonemic transcription suggests that the only perceptual difference between the two words lies in the final consonants. What else contributes a great deal to our recognition of one word or the other? Is there any difference in the length of /i:/ which is conditioned by the lenis /d/ vs. fortis /t/ consonants?

Find other examples to illustrate that a phonemic transcription does not symbolize conditioned differences.

E. Transcribe the words "spend", "peace", "life", "choose", "joy" using
   (a) the broad (phonemic) transcription of the IPA;
   (b) the narrow transcription of the IPA;
   (c) the American transcription (the "linguistic alphabet").

# CHAPTER 4. THE SYLLABLE AS A PHONETIC AND PHONOLOGICAL UNIT

## GENERAL NOTES ON THE SYLLABLE

Sounds (phonemes) are the smallest segments into which the speech continuum is generally divided for purposes of analysis, because these units serve to differentiate words. But in connected speech sounds are not pronounced separatelv, by "themselves". It is practically impossible to draw articulatory boundaries between them. If we slow down the tempo of utterance and articulate the sounds distinctly we shall see that the smallest units, into which the speech continuum is divided, are s y l l a b l e s [27].

When we pronounce a syllable, the speech organs, while producing a consonant, take all the positions necessary for the following vowel, for example note the movements of the tongue and the lips in /su:n/ "soon", /lu:z/ "lose". That is why the boundaries between the consonant and the vowel are not clearly marked. On the contrary, boundaries between syllables are marked by the alternation of openings and closings in sound production and, as a result, by the alternation of increases and decreases in articulatory tension. So the smallest pronunciation (articulatory) unit is the syllable.

It has been proved experimentally that the syllable is also the smallest perceptible unit. A number of experiments, carried out by Soviet linguists L.Chistovitch, V. Kozhevnikov [41], Z. Dzhaparidze [25], show that the listener can recognize the preceding sound only after he has analysed the whole syllable. And it takes less time to identify a syllable than the isolated sounds contained in it.

The syllable can be considered as both a phonetic and a phonological unit. As a phonetic unit the syllable is defined in articulatory, auditory (perceptual) and acoustic terms with universal application for all languages.

As a phonological unit the syllable can be defined and described only with reference to the structure of one particular language. The very term "syllable" denotes particular ways in which phonemes are combined in a language. (Cf. the Greek syllabḗ, "something taken together", from syn–, "together", and labein, "take").

Each language has its own rules of combining its phonemes into syllables. Some combinations are permissible in a language, others are not. Therefore, without any reference to morphology (to the meaning), it is possible to say that such nonsense words as "bulling", "mimsy", "slithy", "wabe", "toves" etc. from Lewis Carol's "Alice in Wonderland", are English and "kpo", "fsple" cannot be English as far as the combinations of phonemes are concerned [64]. And because of the specific grouping and distribution of phonemes in different languages one and the same word may, with certainty, be interpreted as bisyllabic by a speaker of one language, and as trisyllabic by a speaker of another language. For example, a German pronounces the word "Knabe" as bisyllabic, whereas an Englishman would make it trisyllabic – /k ə n ɑ:bə /, because the English language does not permit /kn/ as an initial sound combination [99].

The ancient Greak scholars noticed that the two main phonological types of sounds – vowels and consonants fulfil different functions in speech. The function of a vowel is to occupy the central position in certain combinations of sounds, whereas consonants serve as the margins of the sound combinations. (Hence the term "consonant", which means "sounding with something" – con+sonant.) In other words, vowels are always syllabic and consonants are incapable of forming syllables without vowels.

But in a number of languages some sonorous consonants, such as /n, l, r, m/, can also be syllabic because of their strong vocalic features, for instance, in Czeck – "krk" (neck) and in English "garden" /gɑ: –dn̩/, "needn't" /ni: –dn̩t/, "castle" /kɑ: –sl̩/, "lighten" /laɪ –tn̩/.

So, phonologically, the syllable is a structural unit, which consists of a vowel alone or of a vowel (or a syllabic sonorant) surrounded by consonants in the numbers and arrangements permitted by a given language.

## THE PHONETIC ASPECT OF THE SYLLABLE

Phoneticians are not always in agreement in their definition of the syllable because in their analysis they proceed from either articulatory or acoustic aspects of the unit.

One of the phonetic theories – t h e  e x p i r a t o r y, or c h e s t  p u l s e  t h e o r y – defines the syllable as a sound or a group of sounds that are pronounced in one chest pulse, accompanied by increases in air pressure. According to this definition, there are as many syllables in a word as there are chest pulses (expirations) made during the utterance of the word. Each vowel sound is pronounced with increased expiration. Consequently, vowels are always syllabic. Boundaries between syllables are in the place where there occur changes in the air pressure. But it is impossible to explain all cases of syllable formation on the basis of the expiratory theory, and therefore, to determine boundaries between syllables. A.Gimson notes that it is doubtful whether a double chest pulse will be evident in the pronunciation of juxtaposed vowels as, for instance, in "seeing" / sɪ: – ɪŋ/, though such words consist of two syllables [71].

The r e l a t i v e  s o n o r i t y  t h e o r y, or the prominence theory, created by the Danish phonetician O. Jespersen, considers that sounds tend to group themselves according to their sonority.

Pronounced with uniform force, length and pitch, speech sounds differ in sonority (prominence, audibility or carrying power). The most sonorous sounds are vowels, less sonorous are sonorants /w, j, r, m, n, ŋ / and the least sonorous are noise consonants. O. Jespersen classifies sounds according to the degree of sonority in the following way (beginning with the most sonorous):

1. Open vowels /æ, ɒ, ɑ:, ɔ:/
2. Mid-open vowels /e, ɜ:, ə, ʌ/
3. Close vowels /i:, ɪ, ʊ, u:/
4. Sonants /r, l, m, n, ŋ, w, j/
5. Voiced fricatives /v, ð, z, ʒ/
6. Voiced stops /b, d, g/
7. Voiceless fricatives /f, θ, s, ʃ/
8. Voiceless stops (p, t, k/

Sounds are grouped around the most sonorous ones, i.e. vowels (and sometimes sonants) which form the peak of sonority in a syllable. One peak of sonority is separated from another peak by sounds of lower sonority, i.e. consonants. The distance between the two points of lower sonority is a syllable, e.g. / ɒk–ˈtoʊ –bə / "october". The number of syllables is determined by the number of peaks of prominence. Thus in the word / melt / "melt" there is one peak of sonority /e/ and the word is monosyllabic. In the word / metl /"metal" there are two peaks of sonority – /e/ and /l/, separated by the least sonorous /t/, and consequently, there are two syllables.

Here are some more examples to illustrate the sonority theory: /wʊdn/ "wooden" /fild/ "filled", /fidl/ "fiddle".

In Czeck words like "vlk", "krk" and in English "pst" the sounds /l, r, s/ are sonorous peaks.

But there are cases that contradict Jespersen's theory.

e.g. /stɑ:/ "star", / skeɪt/ "skate", /nekst/ "next".

In these words the sound /s/ is more sonorous than /t/ and /k/ and forms the second peak of sonority. Yet, the words are monosyllabic.

It is evident that the relative sonority theory does not explain the mechanism of syllable formation. It only makes an attempt at explaining our perception of a syllable. Neither does it explain syllable division, as it does not say to which syllable the less sonorous sounds belong, e.g. / ən ˈaɪsmæn/ "an iceman" and /əˈnaɪsˈmæn/ "a nice man", / ən ˈeɪm/ "an aim" and /əˈneɪm/ "a name", /sʌm əˈdresɪz/ "some addresses" and /sʌmə ˈdresɪz/ "summer dresses"

Nevertheless, the relative sonority theory has been accepted by D.Jones and some other phoneticians.

The most widespread among Soviet linguists is the m u s c u l a r t e n s i o n (or the articulatory effort) t h e o r y which is known as Shcherba's theory.

According to this theory a syllable is characterized by variations in muscular tension. The energy of articulation increases at the beginning of a syllable, reaches its maximum with the vowel (or the sonant) and decreases towards the end of the syllable. So, a syllable is an arc of muscular tension. The boundaries between syllables are determined by the occurrence of the lowest articulatory energy.

There are as many syllables in a word as there are maxima of muscular tension in it. Cf. /tɑ:/ "tar" and /tɑ:/ "tower" (a reduced variant of /taʊə/). The sound / ɑ:/ in the second example is pronounced with two articulatory efforts, so there are two arcs of muscular tension and, therefore, two syllables.

Consonants within a syllable are characterized by different distribution of muscular tension. In accordance with this, L. Schherba distinguished the following three types of consonants.

1) Initially strong consonants, in the articulation of which the beginning is stronger while the end is weaker. They occur at the end of a closed syllable.

E.g. i¦t, ʌ¦s, p i¦n, sæ¦d, pɑ:¦t*

2) Finally strong consonants, in the articulation of which the beginning is weak while the end is more energetic. They occur at the beginning of a syllable.

E.g. m¦i , t¦ai , p¦ɑ:t , s¦æd

3) Double–peaked consonants, in the articulation of which both the beginning and the end are energetic whereas the middle is weak. Acoustically they produce the impression of two consonants. These consonants occur at the junction of words or morphemes.

E.g. pe¦nn¦aɪf , ðæ¦t t¦aɪm, mɪ¦dd¦eɪ .

The type of consonant is therefore a cue for syllable division. If in /ənaɪs haʊs/ the sound /n/ is initially strong, the syllabic boundary is after the /n/ – / ən ais'haʊs/. If the sound /n/ is finally strong, the boundary is before it – /ə 'nais'haʊs/. In other words, if there is a new onset of muscular tension on the sound /n/, the latter belongs to the second syllable, and if the new onset of muscular tension is on /aɪ/ , the sound /n/ belongs to the first syllable.

The above theories define the syllable on either the production or perception level.

N. Zhinkin has worked out the so–called loudness theory [26], which takes into account both the levels. On the perception level the syllable is defined as an arc of actual loudness. The experiments carried out by N.Zhinkin, showed that the organ immediately responsible for the variations in loudness of a syllable is the pharynx. The narrowing of the pharyngeal passage and the resulting increase in muscular tension of its walls reinforce the actual loudness of the vowel thus forming the peak of the syllable, while the loudness of the

---

* / ¦ / is used to isolate the consonant out of the syllable.

marginal consonants is weakened. In the production of loudness variations of all the speech mechanisms are involved. So on the speech production level the correlate of "the arc of loudness" is "the arc of articulatory effort".*

The acoustic aspect of the syllable has been studied by E.Zwirner, R.Jakobson and M.Halle. According to the results obtained, the peak of the syllable (a vowel or a sonant) has a higher intensity than its consonants, and in many cases a higher fundamental frequency. Perceptually, the peak is louder and higher in pitch. These acoustic features easily agree with the physiological definition of the syllable as an arc of articulatory effort (or muscular tension).

In analysing the above theories of the syllable, we cannot but agree with the scholars who point out that each of the existing theories is correct to a certain extent, but none of them is able to explain reliably all the cases of syllable boundaries.

It seems that the phonetic definition of the syllable should also take into account the peculiarties of the articulatory basis of a concrete language, the characteristic tendencies in articulatory transitions from a consonant to a vowel (CV transition), from a vowel, to a consonant (VC transition) and from a consonant to a consonant (CC transition).

In English CV transitions are loose, and, therefore the voiceless plosives /p, t, k/ before stressed vowels are aspirated. So the presence of aspiration can indicate that the consonant and the vowel belong to the same syllable, e.g. "plum pie" / plʌm paɪ /. The absence of aspiration shows that they belong to different syllables, i.e. there is a syllable boundary after the consonant, e.g. in "plump eye" /plʌmp aɪ /.

VC transitions are often close in English, because of the checked character of short vowels under stress. Such a VC combination forms one syllable. In unstressed position a short English vowel is not checked and because of the loose VC transition it may form one syllable and the following consonant may belong to another one, e.g. "positive" / pɒz–ɪ–tɪv/.

However the syllabic boundaries are not always well marked in connected speech (e.g./ ə n ɑ is hɑʊ s/) which leads to controversial views on syllbble division.

## THE STRUCTURAL ASPECT OF THE ENGLISH SYLLABLE

To study the syllable as a phonological unit of English is to describe the structure**of the syllable by stating the functions of the phonemes in it and the relations between the phonemes [99].

*The term is suggested by V.A.Vassilyev [110].

**The term "structure" denotes the relations between elements of a system. Applied to the syllable, it refers to the relationship between the phonemes constituting the syllable.

Syllable formation in English, as in other languages, is based on the phonological opposition of vowels and consonants. Vowels are always syllabic, they occupy a central position in the syllable. Consonants are non–syllabic and marginal. The sounds /w, r, j/, despite their strong vocalic features, function as consonants, occurring only before the vowel, e.g. /ˈwɪn– tə /, /ˈrɪ :–də /, /ja:d/. The sounds /l, m, n/ normally function as consonants, in various sound combinations before the vowel. But in unstressed final position, when preceded by a noise consonant, they are syllabic, e.g. /petl̩/ "petal", /blɒ sm̩/ "blossom", /laɪtn̩/ "lighten".

It should be noted specially, that historically short English vowels /ɪ, e,æ, ʌ , ɒ , ʊ, ə / never occur in stressed final position without the following consonant. In unstressed position the vowels /ɪ, ə / can occur as final.

Consonants present particular interest in the study of the syllable, because it is due to the number and arrangement of consonants that the structure of the syllable varies. And it is largely due to consonants that we understand the utterance.

Depending on the position of consonants (C) in relation to the vowel (V), there are 4 types of syllables: 1) open syllables (CV), when there is no consonant after the vowel, e.g. /fa:/ "far", /taɪ/ "tie", /si:/ "sea"; 2) closed syllables (VC), when the vowel is followed by a consonant, e.g. /ɑ:t/ "art", /sɪt/ "sit", /laɪf/ "life"; 3) covered syllables (CV (C), when the vowel is preceded by a consonant, e.g. /seɪ/ "say", /laɪk/ "like", / ʃɔ:/ "shore"; 4) uncovered syllables (V (C), then there is no consonant before the vowel, e.g. / æpt/ "apt", /ɪ :t/ "eat", /eɪt/ "eight". The fundamental syllable type in English is the closed syllable, whereas in Russian it is the open syllable.

As to the presence, number and arrangement of consonants there are 23 syllable patterns in English [36], such as V, VC, CVC, CV, CCVC, CCVCC, CCCVC, CCCVCC, etc. The vowel may occur alone in a syllable or it may have up to 3 consonants before it and up to 4 consonants after it. The most frequent and fundamental pattern in English is CVC.

It is a feature of English that in initial position, i.e. before the vowel, there can be any consonant except / ŋ /; no consonant combinations are possible with / ð , z, tʃ, dʒ /, and such consonant clusters as /mh, sr, sʃ, spw, fs, hr, stl/ cannot occur initially either.

J. O'Connor notes that final clusters are much more complex in English than initial ones. This is due to the fact that final clusters are used to express grammatical meanings of plurality, tense, ordinal number, e.g. "texts" /–ksts/, "mixed" /–kst/, "glimpsed"/–mpst/. In Russian initial clusters are more complex and more numerous than the final ones, because they represent grammatical prefixes, e.g. /fskr–, fspl–, vzr–, kst–/. [99]

Phonotactic possibilities of English phonemes predetermine the rules of syllable division.

English historically short vowels under stress (checked vowels) occur only in a closed syllable. Checked vowels are always followed by a consonant. So the syllabic boundary never occurs after these vowels.

It lies after the following consonant, as in /ˈlʌv-lɪ/ "lovely", /ˈtwen–tɪ/ "twenty", /ˈkwɪk–lɪ/ "quickly", /ˈhɒt–lɪ/ "hotly", /gʊd–nɪs/ "goodness" or within it, if it is the only consonant between the checked vowel and the succeeding vowel. E.g. /letə/ "letter", /bɪgə/ "bigger", /ˈhɒtə/ "hotter", /ˈʃɪlɪŋ/ "shilling".

The preceding and following vowels attract this consonant and the consonant is split into two. In speech the consonant forms a close link between the two syllables. If a checked vowel is separated by one consonant from a syllabic sonant the boundary between the two syllables is also within the consonant. E.g. lɪtl, kɒtn, fætn, bɪtn, rɪðm.

Historically long monophthongs, diphthongs and unstressed short monophthongs (free vowels) can occur both in the open and in the closed syllable. E.g. /kɑ:/ "car", /kɑ:t/ "cart", /ˈtaɪ–nɪ/ "tiny".

When a free vowel is separated from a succeeding vowel by only one consonant sound, the syllable in which such a vowel occurs , is always open. E.g. /aɪ–ˈdɪə/ "idea", /ka:–ˈtu:n/ "cartoon", /ɪ–ˈrekt/ "erect".

When a post–stressed short vowel is separated from a succeeding vowel by a single consonant, the boundary is most probably, before the consonant, because the short vowel is free in unstressed position. E.g. /ˈfæmɪlɪ/ "family", /ˈpɒl–ɪ–sɪ/ "policy", /ˌpɒs–ɪ–ˈbɪl–ɪ–tɪ/ "possibility".

When there is a cluster of consonants between two vowels, the place of the syllabic boundary is conditioned by whether this cluster is permitted at the beginning of words or not. If it does occur in initial position in English the syllabic boundary is before it. If it doesn't, the boundary is between the consonants. For instance the cluster /gr/ is used word–initially in English, therefore it can occur at the beginning of a syllable and the syllabic boundary is before the cluster. E.g. /ə–ˈgri:/ "agree", /rɪ–ˈgret/ "regret".

The clusters /dm/, /dv/ do not occur word–initially and cannot occur at the beginning of a syllable. The syllabic boundary is therefore between the consonants constituting the clusters. E.g. /ədˈmɪt/ "admit", /əd–ˈmaɪə/ "admire", /ədˈvaɪs/ "advice", /ədˈmɪʃn/ "admission".

When two vowels are separated by more than two consonants as for example in /ˈekstrə/ the boundary may be both before /s/ and /t/ because both /str/ and /tr/ occur at the beginning of words and /ks/ can occur in final position.

The so–called triphthongs in English are disyllabic combinations, because they contain two vowel phonemes. E.g. /ˈsaɪ– əns/ "science", /ˈflaʊ – ə/ "flower", /ˈvaʊ–əl/ "vowel"

The structure of an English syllable depends on whether it is stressed or not. The peak of the stressed syllable is always a vowel. In the unstressed syl–

lable the peak may be a vowel or a sonant. When the peak of the stressed syl– lable is a short vowel, the syllable must be "closed" by a consonant.

## FUNCTIONS OF THE SYLLABLE

As a phonological unit the syllable performs several functions, that may be combined into the main three: constitutive, distinctive and identificatory.

The c o n s t i t u t i v e function of the syllable manifests itself in the fact that the syllable forms higher–level units – words, accentual or rhythmic groups, utterances. Two aspects of this function can be emphasized. On the one hand, the syllable is a unit in which segmental phonemes are realized. L. Bondarko has proved experimentally that the relations between the distinctive features of the phonemes and their acoustic correlates can be revealed only within the syllable [14]. On the other hand, within a syllable or a sequence of syllables prosodic (or suprasegmental) features of speech are also realized. These are distinctive variations in loudness (stress) in pitch (tone), and in duration (tempo, lenght). Thus syllables may be stressed and unstressed, high, mid or low, rising or falling, long or short. All these prosodic features are significant for constituting the stress–pattern of a word and the tonal and rhythmic structures of an utterance.

So we can say, that the syllable is not a mere sum of sounds; it is a specific minimal structure of both phonemic and prosodic features. In the taxanomical scale of language units the syllable occupies the position between the phoneme and the word [64].

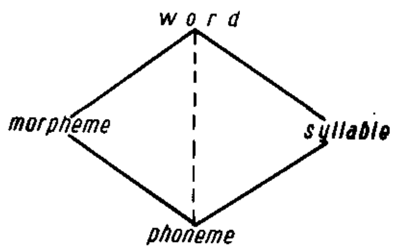


Fig. 8. Relations of language units.

In forming words and utterances the syllable performs the delimitative function which is inseparable from the constitutive function. Some syllables can occur only word–initially (/gr/, /str/) and others only word–finally; /tn/, /dn/, /stl/ thus making the boundaries between words.

The d i s t i n c t i v e function of the syllable is to differentiate words and word combinations. It has been mentioned that phonemes exist and function within the syllable. Therefore words are actually differentiated by the syllable as one articulatory and perceptible unit. For instance, the mono– syllabic words /bi:t/ "beat" and /bi:d/ "bead" differ not only in their conso– nant phonemes /t/ and /d/, but also in the length of /i:/, which is conditioned by the neighbouring fortis and lenis consonants. Such words as /ga:dn/ "garden" – /ga:dz/ "guards", /bi:tn/ "beaten" – /bi:ts/ "beats" are distinguished not only by the phonemes /n/ versus /z/ and /n/ versus /s/ but also by their syllables as bisyllabic and monosyllabic words.

There are some words in English where syllabicity alone is responsible for the differenation of the words: /laɪtnɪŋ/ "lightening" (освещение) and /laɪtnɪŋ/ "lightening" (молния).

V.Vassilyev notes that the existence of such pairs of words makes it possible to consider syllabicity the only distinctive feature of the words and, on this account he distinguishes a separate phonological unit – the syllabeme [110].

Variations in the syllabic structure of one and the same word or a word combination may serve to differentiate styles of pronunciation. For example the word "national" may be pronounced with 3 and 2 syllables – /næʃnəl/ and /n æ ʃ n ə l/; the word combination "little and nice" may have 4 or 3 syllables– /lɪtlənnaɪs/ and /lɪtl ə nn a ɪs/. The second variants are characteristic of the colloquial style.

Syllable division (syllabification) is very important too in distinguishing words and utterances.

The distinctive role of syllabification is illustrated by examples like

/naɪtreɪt/ "nitrate" – /naɪtreɪt/ night-rate,

/əneɪm/ "a name" – /əneɪm/ "an aim",

/wɪloun/ "we'll own" – /wɪloʊn/ "we loan",

/aɪskrɪ:m/ "I scream" – /aɪskrɪ:m/ "ice cream",

/aɪs ɔ:hər aɪz/ "I saw her eyes"–/aɪ s ɔ:hər aɪz/"I saw her rise".

Due to the distinctive importance of syllable division, the syllabic boundary is often regarded by the American descriptivists as a separate phonological unit – the juncture phoneme.

There are two types of juncture: open and close. Open juncture (or open transition) occurs between syllables, i.e. between two articulatory units. It may also be called intersyllabic juncture. Thus, in "we'll own" /wɪl oʊn/ the open juncture is between /l/ and /o/, and in "we loan" /wɪ loʊn/ it is between /ɪ/ and /l/;

Close juncture (or close transition) occurs between sounds within one syllable, i.e. within one articulatory unit. Therefore, the transitions from one sound to another are closer within a syllable than between syllables. Thus in "we loan" /wɪ loʊn/ the close juncture is between /l/ and /oʊ/, /oʊ/ and /n/. This juncture may also be called intrasyllabic juncture.

The latest acoustic investigations of juncture show that the factors determining an open or a close juncture are the duration of the sounds, their intensity and formant transitions. Thus, according to the data obtained by

I.Lehiste [94], the initial /n/ in "a nice man" is longer than the final /n/ in "an iceman". The pre–junctural /n/ has falling intensity, while the post–junctural /n/ has rising intensity. Formant transitions of /n/ and /ai/ are different in the contrasted pairs.

While the phonetic realization of open juncture is described by different phoneticians in approximately the same terms, there is less uniformity in their phonological interpretation of the phenomena. Some phoneticians consider the open juncture to be a segmental phoneme [77], others consider it a suprasegmental phoneme [78] or a phoneme in its own right [79, 109].

K.Pike and I.Lehiste regard the juncture to be a contrastive feature of high–level units but not a phonological unit in its own right [94, 104].

The i d e n t i f i c a t o r y function of the syllable is conditioned by the hearer's perception of syllables as entire phonetic units with their concrete allophones and syllabic boundaries.

The listener identifies two syllables in "plum pie" and "plump eye" with the corresponding boundaries before /p/ and after /p/, because in the first example /p/ is aspirated and /m/ is as long as if it were final, whereas in the second example /p/ is unaspirated and /m/ is shorter on account of the following fortis /p/.

That is why learners of English should take care not to mispronounce English sounds and not to shift the syllabic boundary as it may cause not only a strong foreign accent, but also misunderstanding on the part of the listener.

## EXERCISES

A. Think about the following questions for class discussion:

1. Why is the syllable considered to be the smallest articulatory unit? Why can't we consider the sound to be the smallest articulatory unit?
2. What makes us assert that the syllable is also the smallest perceptible unit? Don't we perceive separate sounds in connected speech?
3. Is the syllable a linguistic unit? Is it the smallest linguistic unit?
4. What are the main problems of the phonetic aspect of the syllable? Which of the theories of syllable formation do you think is the most consistent and helpful?
5. Could we describe the structure of the syllable with reference to all the languages? What universal features of the syllabic structure could you name?
6. What are typically English structural characteristics of the syllable?

B. Choose one of the nonsense words given below which can be recognized as either English or Russian due to the structural characteristics of the syllables:

vzol, mimsy, fole, strem, kpi, rzhest, pernit.

C. Transcribe the following words and divide them into syllables:

1) river, runner, petal, solid, cricket, level, sonority, specific, syllabic;

2) towel, bowl, shower, science, require;

3) express, strengthen, extraordinary, cluster, succeed, function, obstruction.

E. Collect pairs of words and word combinations which are differentiated by the position of the syllabic boundary as are "plum pie" and "plump eye".

## CHAPTER 5. WORD STRESS

### THE NATURE OF ENGLISH WORD STRESS

A word, as a meaningful language unit, has a definite phonetic structure. The phonetic structure of a word comprises not only the sounds that the word is composed of and not only the syllabic structure that these sounds form, it also has a definite stress pattern.

The auditory impression of stress is that of prominence. And if a word contains more than one syllable, the relative prominence of those syllables differs. There may be one prominent syllable in a word as compared to the rest of the syllables of the same word (as in "im'portant"), there may be two equally prominent syllables (as in "'mis'behave"), two unequally prominent syllables (as in "ˌexami'nation") or more prominent syllables (as in "'unreˌlia'bility" And this correlation of degrees of prominence of the syllables in a word forms the stress pattern of the word, which is often called the accentual structure of a word.

The stress patterns of different words may coincide. Thus the words "mother", "table", "happy", "after" have an identical stress pattern ⊥ — , though their sound structures have nothing in common. The stress pattern of these words differs from that of "analyse", "prominent", "syllable", "character", which is ⊥ — — .

Monosyllabic words have no stress pattern, because there can be established no correlation of prominence within it. Yet as lexical units monosyllables are regarded as stressed.

The stress patterns of words are generally perceived without difficulty. People easily distinguish between "'subject" and "sub'ject".

Actual speech does not consist of isolated words. And the stress pattern of a word is deduced from how the word is accented in connected speech. On the other hand, the stress pattern of a word is only its potential pattern in an utterance. Though English words generally retain their stress patterns in connected speech, there occur numerous instances when the stress pattern of a word is altered.

Cf. 'un'happy — He was 'so un'happy. — He 'remembered those 'unhappy 'days.

Thus, word stress may be said to be a word-level concept [61], which should not be confused with utterance stress. Word stress belongs to the word when said in isolation. Whereas utterance stress belongs to the utterance.

The placement of utterance stress is primarily conditioned by the situational and linguistic context. It is also conditioned by subjective factors: by the speaker's intention to bring out words which are considered by him to be semantically important in the situational context. As for the stress pattern of a word, it is conditioned only by objective factors: pronunciation tendencies and the orthoepic norm. One cannot distort the stress pattern of a word on one's own, because such a distortion will make speech unintelligible.

As stated above, the auditory impression of stress is that of prominence. So a stressed syllable on the auditory level is a syllable that has special prominence. The effect of prominence may be produced by a greater degree of loudness, greater length of the stressed syllable, some modifications in its pitch and quality.

Acoustic analysis shows that the perception of prominence may be due to definite variations of the following acoustic parameters: intensity, duration, frequency, formant structure*. All these parameters generally interact to produce the effect of prominence.

In different languages stress may be achieved by various combinations of these parameters. Depending upon which parameter is the principal one in producing the effect of stress, word stress in languages may be of different types.

There are languages with dynamic word stress. Stress in such languages is mainly achieved by a greater force of articulation which results in greater loudness, on the auditory level, and greater intensity on the acoustic level. The stressed syllables are louder than the unstressed ones. All the other parameters play a less important role in producing the effect of stress in such languages.

In languages with musical word stress prominence is mainly achieved by variations in pitch level, the main acoustic parameter being fundamental frequency. Chinese, Japanese, Vietnamese are languages with musical word stress (or tonic word stress). The meaning of the words in those languages depends on the pitch levels of their syllables.

Swedish word stress is characterized as dynamic and musical, because both loudness and pitch variations are relevant factors in producing prominence. For instance, the Swedish word "Anden" with falls in pitch on both syllables means "soul", but when pronounced with a fall in pitch on the first syllable and low pitch on the second syllable means "duck" [99].

In languages with quantitative word stress the effect of stress is mainly based on the quantity of the sound, i.e. its duration. In such languages vowels in the stressed syllables are always longer than vowels in unstressed syllables. Russian word stress is considered to be mainly quantitative though it has been proved that duration is not the only parameter that produces the effect of stress in Russian [28].

Besides those types of word stress, linguists distinguish qualitative word stress,as in many languages the quality of vowels in stressed syllables is unobscur—

---

*It is generally accepted that intensity, duration, frequency and formant structure are the physical correlates of loudness, length, pitch and quality correspondingly.