AN INTRODUCTION TO PHONETICS

Parviz Birjandi (PhD)
Allameh Tabatabaii University

Mohammad Ali Salmani-Nodoushan (PhD)
University of Zanjan

ZABANKADEH PUBLICATIONS
8 Bazarcheh Ketaab, Enghelab Avenue, Tehran, Iran
Phone: +98 21 66402367
Fax: +98 21 66492961

No 8, Bazarcheh Ketaab, Enghelab Sq., P.O. Box: 13145-564 Tehran, Iran
E-Mail: info@zabankadeh.net Tel.: 02166402367 Fax: 02166492961

An Introduction to Phonetics

Authors: Parviz Birjandi (PhD) Mohammad Ali Salmani-Nodoushan (PhD)

Printed in Iran

سلمانی ـ نُدوشن، محمدعلی، ۱۳۴۸ ـ Salmani-Nodoushan, Mohammad Ali
(ان اینترو داکشن تو فُونِتیکس)
An Introduction to Phonetics / Mohammad Ali Salmani-Nodoushan, Parviz Birjandi.-
تهران : زبانکده ، ۱۳۸۴ = ۲۰۰۵ م.
xii ، ۱۸۰ ص.
ISBN: 964-6117-57-0
انگلیسی.
فهرستنویسی بر اساس اطلاعات فیپا.
کتابنامه: ص. ۱۷۷ ـ ۱۸۰.
۱. زبان انگلیسی ــ آواشناسی. ۲. زبان انگلیسی ــ واج شناسی.
الف. بیرجندی، پرویز، ۱۳۱۹- ، Birjandi, Parviz. ب. عنوان: An Introduction to Phonetics.
۹الف۸س / ۱۱۳۵ PE ۴۲۱/ ۵
۱۳۸۴
کتابخانه ملی ایران ۶۵۵۰ ـ ۸۴م

An Introduction to Phonetics

مولفان: دکتر پرویز بیرجندی / دکتر محمدعلی سلمانی نُدوشن

انتشارات زبانکده

چاپ اول ۱۳۸۴، تیراژ ۳۰۰۰ نسخه، چاپ دیبا

تهران، روبروی دبیرخانه دانشگاه تهران، بازارچه کتاب، شماره ۸
تلفن: ۶۶۴۰۲۳۶۷ (۰۲۱) فاکس: ۶۶۴۹۲۹۶۱ (۰۲۱)
صندوق پستی: ۵۶۴ ـ ۱۳۱۴۵

قیمت: ۲۳۰۰۰ ریال

ISBN: 964-6117-57-0
شابک: ۰-۵۷-۶۱۱۷-۹۶۴

ACKNOWLEDGMENTS

Many people have played an important role in the development of this book. Colleague **Richard W. Sorfleet**, of Ontario College of Teachers, Canada, was extremely supportive of this book, offering numerous suggestions for our consideration. He read the manuscript and provided us with useful comments which enabled us to convert turgid prose into readable English.

Special acknowledgment and thanks are also extended to colleagues **Dr Mehdi Nowroozi Khiabani** and **Dr Kurosh Lachini** for their insightful reviews of the manuscript. Their comments and suggestions were very valuable to us. Our special thanks also go to **Mrs. Arezou Pashapur** for typing the manuscript. We would also like to thank the people at **Zabankadeh Publications** for publishing and marketing the book.

Parviz Birjandi

Mohammad Ali Salmani-Nodoushan

August 2005

CONTENTS

PREFACE

A major difficulty facing almost any foreign language learner is the achievement of acceptable pronunciation which marks his success in mastering the language. Many EFL learners master such aspects of language as syntax, semantics, morphology, and even pragmatics to the point of native-like competence, but fail to master phonology. This is partly because of the physiological constraints that make the pronunciation of a foreign or second language sound different from that of the native language of the speakers, and partly due to the lack of appropriate training in phonology courses.

An Introduction to phonetics is designed to support EFL learners in achieving native-like pronunciation:

- Chapter one deals with the history of phonology and phonetics and provides a brief overview of the impact of philosophy and psychology on the emergence of phonology.
- Chapter two defines the notion of phoneme, describes IPA phonetic alphabet, and distinguishes between broad and narrow transcriptions.
- Chapters three and four provide an in-depth account of traditional and systematic articulatory phonetics respectively.
- Chapter five discusses the place of suprasegmentals in phonology.
- Chapter six seeks to explain phonemics.
- Chapter seven provides a brief introduction to the rudiments of acoustic or physical phonetics.
- Chapter eight introduces the reader to the notion of auditory phonetics.

- The tables, figures, and photos that are presented throughout the book are designed to give the reader an instant reference for the precise articulation of English phonemes.

The book is designed for use in undergraduate classes of phonology and phonetics. The fifth chapter is also useful for students of conversation classes. Teachers at high school level may also find the fifth chapter valuable. The step by step approach of the book towards its subject matter makes it is easy for the reader to follow the line of discussion without the help of a phonology teacher.

Parviz Birjandi
Mohammad Ali Salmani-Nodoushan

August 2005

CHAPTER ONE
PRELIMINARIES

1. INTRODUCTION

Phonetics is a branch of linguistics which is concerned with the production, physical nature, and perception of speech sounds. The main fields of study are experimental phonetics, articulatory phonetics, phonemics, acoustic phonetics, and auditory phonetics.

2. EARLY CONTRIBUTORS

The earliest contributions to phonetics were made more than 2000 years ago by Sanskrit scholars such as the grammarian Panini in the 400s who dealt with articulation to keep the pronunciation of ancient rituals unchanged. The first phonetician of the modern world was Dane J. Matthias, author of *De Litteris* (1586). English mathematician John Wallis, who instructed deaf-mutes, was the first to classify vowels, in 1653, according to their place of articulation. The vowel triangle was invented in 1781 by C. F. Hellwag from Germany. Ten years later, Austrian mechanician Wolfgang von Kempelen invented a machine that produced speech sounds. German physicist Hermann Helmholtz, who wrote *Sensations of Tone* (1863), inaugurated the study of acoustic phonetics. Frenchman Abbé Jean Pierre Rousselot pioneered in experimental phonetics.

Late in the 19th century, the theory of the phoneme was advanced by Jan Baudouin de Courtenay from Poland and Ferdinand de Saussure from Switzerland. In the United States, linguist Leonard Bloomfield and anthropologist and linguist Edward Sapir contributed greatly to the phonetic theory. Linguist Roman Jakobson developed a theory of the universal characteristics of all phonemic systems. Perhaps Ferdinand de Saussure

(1857-1913), a Swiss linguist, has provided the clearest ideas about phonetics. His book *Memoir on the Original Vowel System in the Indo-European Languages* (1879) is an important work on the vowel system of Proto-Indo-European—considered the parent language from which the Indo-European languages descended.

3. THE ROLE OF PHILOSOPHY

Recent scientific endeavors all get their insights from the developments in philosophy. In fact, schools of philosophy have shed light on the paths in which different branches of science have stepped. Phonetics, like other fields, has been deeply influenced—at least in its origins—by philosophy. Specifically, the ideas of Kant have had the greatest impact on the early steps of phonetics.

Perhaps the greatest influence is due to the Scottish philosopher and historian David Hume who argued that no observable evidence is available for the existence of a mind substance, spirit, or God. In other words, all metaphysical assertions about things that cannot be directly perceived (that is, through the five sense modalities of sight, taste, touch, hearing, and/or smell) are equally meaningless and non-existent, he claimed, and should be “committed to the flames.” Following Hume's line of thought, Immanuel Kant later categorized knowledge to two major branches: (a) the “phenomenal world” of experience, and the noumenal world of faith. He maintained that metaphysical beliefs about the soul, the cosmos, and God are matters of faith rather than of scientific knowledge.

Along the same lines, early linguists, and most notably among them Leonard Bloomfield maintained that language only consists of elements which are perceivable through the five sense modalities. That is, the only scientifically acceptable parts of language are those elements which we can experience through our five senses. As such, meaning was out of consideration in linguistic theories. The most readily tangible part of language was its phonological part. Therefore, linguists were expected to focus on the study of sounds.

4. THE ROLE OF PHYSIOLOGY AND PSYCHOLOGY

The origins of phonetics are also traceable in physiology and psychology. Perhaps the most famous physiologist whose ideas were widely incorporated into the developments of phonetics was Ivan Pavlov. He is noted for his pioneer work in the physiology of the heart, nervous system, and digestive system. His most famous experiments, begun in 1889, demonstrated the conditioned and unconditioned reflexes in dogs, and they had an influence on the development of physiologically oriented behaviorist theories of psychology during the early years of the 20th century. His work on the physiology of the digestive glands won him the 1904 Nobel Prize in physiology or medicine.

Pavlov carried out a series of experiments in which he provided what he called stimuli for his dog and measured what he called responses. In these experiments, Pavlov accompanied the call of a bell with dog-food and measured the secretion of saliva by the dog. He repeated this stimulus several times. When he predicted that the dog could associate the ring of the bell with the presence of food, Pavlov omitted the food but kept the bell. To his surprise, Pavlov observed that the dog's secretion of saliva in the absence of food increased as if food was present. That is, the dog had learnt to react to the call of the bell. Pavlov concluded that organisms could learn new behavior upon the repetition of appropriate stimuli to the point of automaticity. Pavlov's experiments led to the notion of Behaviorism in psychology which is, in turn, at the heart of structuralism in linguistics.

Behaviorism is a movement in psychology that advocates the use of strict experimental procedures to study observable behavior (or responses) in relation to the environment (or stimuli). The behavioristic view of psychology has its roots in the writings of the British associationist philosophers, as well as in the American functionalist school of psychology and the Darwinian theory of evolution, both of which emphasize the way that individuals adapt and adjust to the environment.

Behaviorism was first developed in the early 20th century by the American psychologist John B. Watson. The dominant view of that time was that psychology is the study of inner experiences or feelings by subjective, introspective methods. Following Kantian philosophy, Watson did not deny the existence of inner experiences, but he insisted that these experiences could not be studied because they were not observable. He was greatly influenced by the pioneering investigations of the Russian physiologists Ivan Pavlov and Vladimir Bekhterev on conditioning of animals (i.e., classical conditioning). Watson proposed to make the study of psychology scientific by using only objective procedures such as laboratory experiments designed to establish statistically significant results. The behavioristic view led him to formulate a stimulus-response (S-R) theory of psychology. In this theory all complex forms of behavior—emotions, habits, and such—are seen as composed of simple muscular and glandular elements that can be observed and measured. He claimed that emotional reactions are learned in much the same way as other skills.

Watson's stimulus-response theory resulted in a tremendous increase in research activity on learning in animals and in humans, from infancy to early adulthood. Between 1920 and midcentury, behaviorism dominated psychology in the United States and also had wide international influence. By the 1950s, the new behavioral movement had produced a mass of data on learning that led such American experimental psychologists as Edward C. Tolman, Clark L. Hull, and B. F. Skinner to formulate their own theories of learning and behavior based on laboratory experiments instead of introspective observations.

5. THE ROLE OF AMERICAN STRUCTURALISM

Early linguistic studies were concerned with historical issues such as language families. As linguists in the United States became involved in the study of American Indian languages during the late 19th and early 20th centuries, it became increasingly clear that the historical orientation of nineteenth-century

European linguistics was not very practical for work with languages that lacked extensive written materials from the past. Furthermore, many American linguists of this period were also anthropologists—scientists who study human beings.

These factors led to a speech-oriented focus in the linguistic studies in the United States. Scholars turned their attention to the form of languages, emphasizing the description of the phonological and other surface forms of these languages. American linguists frequently approached American Indian languages without the advantage of knowing even the basic sound system, let alone the principles of sentence formation or meaning of these languages.

Thus, they had to start their studies with what was most immediately observable in these languages—the sounds. Since few linguists ever achieved this degree of fluency in American Indian languages, American linguistics of the first half of the twentieth century was characterized by an intensive investigation of sounds and the principles of word formation. Little (if at all) attention was paid to syntax (i.e., word-order) or semantics (i.e., meaning).

Because of its attention to the form, or structure of language, American linguistics of the early 20^{th} century came to be known as Structural Linguistics. Attempting to describe languages that they themselves could not speak, American linguists were forced to concentrate on the directly observable aspects of these languages—their sounds. In their work, they gradually evolved a set of procedures considered useful in determining the sound system of languages. These procedures and techniques were later known as Discovery Procedures.

Despite the limitations of this early work on American Indian languages, the detailed, objective investigations of linguists during the first half of the twentieth century provided concrete evidence about the diversity that exists among human languages.

6. THE EMERGENCE OF PHONETICS

All these developments resulted in the emergence of a new branch within linguistics which was concerned with the study of the phonological component of language—phonology. Phonology is the study of all aspects of the sounds and sound system of a language. It includes two major sub-branches: (a) **phonetics**, and (b) **phonemics**.

Phonetics is the field of language study concerned with the physical properties of sounds, and it has three subfields:

(a) articulatory phonetics (i.e., the study of how the human vocal organ produces sound)
(b) acoustic phonetics (i.e., the study of the sound waves produced by the human vocal apparatus)
(c) auditory phonetics (i.e., the examination of how speech sounds are perceived by the human ear)

Phonemics, in contrast, is not concerned with the physical properties of sounds. Rather, it focuses on how sounds function in a particular language. The following example illustrates the difference between phonetics and phonology. In the English language, when the sound /k/ (usually spelled c) occurs at the beginning of a word, as in the word cut, it is pronounced with aspiration (a puff of breath). However, when this sound occurs at the end of a word, as in tuck, there is no aspiration. Phonetically, the aspirated [k^h] and unaspirated [k] are different sounds, but in English these different sounds never distinguish one word from another or bring about differences in meaning, and English speakers are usually unaware of the phonetic difference until it is pointed out to them. Thus English makes no phonological distinction between the aspirated and unaspirated /k/. The Hindi language, however, uses this sound difference to distinguish words such as kal (meaning time), which has an unaspirated /k/, and khal (meaning skin), in which /kh/ represents the aspirated [k^h]. Therefore, in Hindi the distinction between the aspirated and unaspirated

/k/ is both phonetic and phonological—phonemic (i.e., any difference in pronunciations which brings about a difference in meaning is said to be phonemic). The following chapters will provide an in-depth discussion of both phonetics (i.e., auditory, acoustic, and articulatory phonetics) and phonemics.

EXERCISE

Decide whether the following statements are true or false. Mark Ⓕ for false and Ⓣ for true statements.

01 Phonetics is concerned with the production, physical nature, and perception of speech sounds. Ⓣ Ⓕ

02 The main fields of phonetics and phonology are experimental phonetics, articulatory phonetics, acoustic phonetics, auditory phonetics, and phonemics. Ⓣ Ⓕ

03 The theory of the universal characteristics of all phonemic systems was developed by Roman Jacobson. Ⓣ Ⓕ

04 Because of its emphasis on the five senses, Kantian philosophy helped the emergence of phonetics. Ⓣ Ⓕ

05 Phonology is the study of all aspects of the sounds and sound system of a language. Ⓣ Ⓕ

06 The major difference between phonetics and phonology is that phonetics is concerned with the description of speech sounds while phonology is concerned with the description of the sound system of language. Ⓣ Ⓕ

07 The branch of phonetics that has to do with the study of how the human vocal organ produces speech sounds is called auditory phonetics. Ⓣ Ⓕ

08 Phonetics and phonemics are the sub-branches of phonology. Ⓣ Ⓕ

09 Phonemics focuses on how sounds function in a particular language. Ⓣ Ⓕ

10 American structuralism is the school of linguistics which had the most influence on the emergence and development of phonology. Ⓣ Ⓕ

CHAPTER TWO
ARTICULATORY PHONETICS: RUDIMENTS

1. INTRODUCTION

Articulatory phonetics describes speech sounds genetically—that is, with respect to the ways by which the vocal organs modify the air stream in the mouth, nose, and throat in order to produce a sound. To date, articulatory phonetics has witnessed two major movements: (a) traditional phonetics, and (b) modern or systematic phonetics. These movements will be discussed in detail in the following chapters. This chapter will focus on the rudimentary topics which are crucial to the understanding of articulatory phonetics.

To show how a speech sound is articulated, all the vocal activities involved in the production of a sound need not be described. Only a selection of them such as the place and manner of articulation is enough. Phonetic symbols and their articulatory definitions are abbreviated descriptions of these selected activities. The symbols most commonly used are those adopted by the International Phonetic Association (IPA) and are written in brackets or between slant lines. An understanding of articulatory phonetics requires an understanding of a number of basic concepts usually employed by phoneticians in their discussion of phonetics.

2. PHONEME

The basic building block of any discussion of articulatory phonetics is phoneme. The technical term phoneme is usually used to refer to sound segments. Linguists define phoneme as the minimal unit of sound (or sometimes syntax). The study of phonemes is the study of the sounds of speech in their primary function, which is to make vocal signs that refer to the fact that different things sound different. The phonemes of a particular language are those minimal distinct units of sound that can distinguish meaning in that language. In English, the /p/ sound is a phoneme because it is the smallest unit of sound that can make a difference of meaning if, for example, it replaces the initial sound of such words as **bill, till, or dill,** making the word **pill**. The vowel sound /ɪ/ of pill is also a phoneme because its distinctness in sound makes **pill**, which means one thing, sound different from **pal**, which means another. Two different sounds, reflecting distinct articulatory activities, may represent two phonemes in one language but only a single phoneme in another. Thus phonetic /r/ and /l/ are distinct phonemes in English, whereas these sounds represent a single phoneme in Japanese, just as [p^h] and [p] in **pie** and **spy**, respectively, represent a single phoneme in English although these sounds are phonetically distinct.

Phonemes are not letters; they refer to the sound of a spoken utterance. For example, **flocks** and **phlox** have exactly the same five phonemes. Similarly, **bill** and **Bill** are identical phonemically, regardless of the difference in meaning. Each language has its own inventory of phonetic differences that it treats as phonemic (that is, as necessary to distinguish meaning). For practical purposes, the total number of phonemes for a language is the least number of different symbols adequate to make an unambiguous graphic representation of its speech that any native speaker could read if given a sound value for each symbol, and that any foreigner could pronounce correctly if given additional rules covering nondistinctive phonetic variations that the native speaker makes automatically. For convenience, each phoneme of any language may be given a symbol.

3. THE NEED FOR PHONETIC WRITING SYSTEM

It is quite clear that ordinary writing systems (i.e., alphabet or orthography) cannot illustrate pronunciation differences. In fact, there are many words that are written with the same set of letters but pronounced differently. There are also some words which are written with different sets of letters but pronounced the same. As such, phonologists and phoneticians felt the need for a new writing system, one in which the symbols fully represent the sounds of any language.

Standard English orthography (the writing system) cannot capture all the sounds of English pronunciation. The same is true for writing systems of many of the world's languages. Even if the writing system does a good job of capturing all the sounds of a language, what happens when the pronunciation of a word changes over time? Or when there are multiple pronunciations for a word? To overcome this, linguists use the phonetic alphabet, designed to represent all the possible sounds of the world's languages in a standard way. The most commonly-used alphabet is known as the **Phonetic Alphabet** designed by the International Phonetic Association (IPA) in the late 19th century.

The anomaly between writing and pronunciation was the main drive behind the development of the phonetic alphabet. Perhaps you have already noticed that some similar sounds have different representations in English orthography. Take the following examples:

[f]	as in	trough, fun
[n]	as in	knife, night
[i]	as in	tiny, ceiling, tin

In addition, different sounds in many languages may be used in the writing system with the same representation. Take the following English examples:

gh	Pronounced as /f/	as in	cough
	Pronounced as /aʊ/	as in	bough
ou	Pronounced as /ɔ:/	as in	bought
	Pronounced as /ʌ/	as in	tough

Language (all languages) changes over time. Spelling (orthography) is slower to change than pronunciation. As a consequence, the discrepancy between spelling and sounds gave impetus to a number of scholars to want to revise the alphabet so that each sound would be represented by one and only one symbol and each symbol would represent one and only one sound. Robert Robinson (1617), Cave Beck (1657), Bishop John Wilkins (1668), Francis Lodwick (1686) are some of the scholars who developed their own phonetic writing systems. In 1888, the International Phonetic Association (IPA) developed the most comprehensive phonetic chart which could be used to symbolize the sounds that appear in all languages of the world. The main characteristic of the phonetic alphabet is the one-to-one correspondence between sound and symbol. In other words, each sound is represented by one and only one symbol and each symbol represents one and only one sound. Today, the phonetic alphabet is widely used to transcribe or write sounds in all languages of the world. Take the following example:

EXAMPLE

English Orthography	The man wanted to know the truth.
IPA Transcription	/ðə mæn 'wɒntɪd tə nəʊ ðə tru:θ/

Each symbol in the ordinary orthography (alphabet) of a language is called a letter. Each symbol in the phonetic alphabet is called a **sound segment**. As such, the word *truth* is composed of five letters (t, r, u, t, h) but four phonemes /t/, /r/, /u:/, /θ/. The ordinary writing in a language (i.e., use of letters) is called orthography or writing while phonetic writing (i.e., use of phonemes) is usually called **transcription**. The inventory of phonemes in the IPA phonetic alphabet is so rich that it can be used to represent all sounds that appear in all languages of the world. However, some but not all of these sounds appear in the English language. The following table summarizes all the basic sounds that are employed by English speakers in their speech (i.e., English vowels, diphthongs, triphthongs, and consonants).

	SYMBOL	EXAMPLE	SYMBOL	EXAMPLE
VOWELS	/iː/	beat	/ə/	an
	/ɪ/	bit	/uː/	boot
	/e/	hen	/ʊ/	book
	/ɜː/	bird	/ɔː/	four
	/æ/	hat	/ɑː/	car
	/ʌ/	cut	/ɒ/	hot
DIPHTHONGS	/ʊə/	poor	/eɪ/	day
	/ɔɪ/	boy	/aʊ/	now
	/ɪə/	here	/aɪ/	dry
	/əʊ/	go	/eə/	hair
TRIPHTHONGS	/aʊə/	power	/ɔɪə/	loyal
	/aɪə/	liar	/əʊə/	mower
	/eɪə/	layer		
CONSONANTS	/p/	pot	/ʃ/	she
	/b/	bat	/ʒ/	vision
	/t/	top	/ʧ/	chat
	/d/	desk	/ʤ/	judge
	/k/	cut	/m/	man
	/g/	good	/n/	not
	/f/	five	/ŋ/	thing
	/v/	van	/r/	rat
	/θ/	think	/l/	look
	/ð/	this	/w/	one
	/s/	sit	/j/	yard
	/z/	zip	/h/	hat

Table 2.1: IPA symbols for British English vowels, diphthongs, triphthongs, and consonants.

	SYMBOL	EXAMPLE	SYMBOL	EXAMPLE
VOWELS	/i/	beat	/ə/	an
	/ɪ/	bit	/u/	boot
	/e/	hen	/ʊ/	book
	/ɚ/	bird	/ɔ/	four
	/æ/	hat	/ɑ/	car
	/ʌ/	cut	/a/	hot
DIPHTHONGS	/ʊɚ/	poor	/eɪ/	day
	/ɔɪ/	boy	/aʊ/	now
	/ɪɚ/	here	/aɪ/	dry
	/əʊ/	go	/eɚ/	hair
CONSONANTS	/p/	pot	/ž/	vision
	/b/	bat	/č/	chat
	/t/	top	/ǰ/	judge
	/d/	desk	/m/	man
	/k/	cut	/n/	not
	/g/	good	/ŋ/	thing
	/f/	five	/r/	rat
	/v/	van	/l/	look
	/θ/	think	/w/	one
	/ð/	this	/ʍ /	why
	/s/	sit	/j/	yard
	/z/	zip	/h/	hat
	/š/	she		

Table 2. 2: IPA symbols for American English vowels, diphthongs, and consonants.

In addition to phonemes, a number of other symbols are used in the phonetic alphabet to represent the intricacies and nuances of sounds such as length, aspiration, etc. These symbols are normally put either above (superscript) or below (subscript) phonemes and indicate the features of phonemes. These symbols are called **diacritics** or **sporadic features.** Arabic is a language which uses sporadic features even in its ordinary orthography. In ordinary English orthography, too, a limited number of diacritics are used (e.g., ℃).

Table 2.2 summarizes the most frequent diacritic symbols which are used in IPA phonetic alphabet:

SYMBOL	MEANING	LOCATION
:	long	right of phoneme
̥	voiceless	under phoneme
•	syllabic	under phoneme
h	aspirated	right superscript
w	rounded	right superscript
~	nasalized	above phoneme
̩	syllabic	under phoneme

Table 2.3: Sporadic symbols used in phonetic transcription in IPA

In addition to these, phonologists often use cover symbols to refer to classes of sounds. They use a capital **C** to refer to the class of consonants, a capital **V** to represent vowels, a capital **L** for liquids, a capital **N** for nasals, and a capital **G** for glides:

SYMBOL	MEANING	SYMBOL	MEANING
C:	long C	Ṽ	nasalized V
V:	long V	V́	stressed V
L̥	voiceless L	Ṇ	syllabic N
G̥	voiceless G	L̩	syllabic L
V̥	voiceless V		

Phoneticians also repeat these symbols when they want to show a cluster of phonemes. For example, to show a consonant cluster, the symbol **C** is repeated.

CC (cluster of 2 consonants) **CCC** (cluster of 3 consonants)

4. BROAD AND NARROW TRANSCRIPTION

An interesting point about phonemes is that they are not always pronounced with the same pronunciation. In fact, the co-text and environment of the occurrence of a phoneme plays a crucial role in the way the phoneme is pronounced. All of us have the experience of changing the sound /n/, unconsciously of course, into the sound /m/ in such Persian words as /ʃænbe/ (meaning Saturday) so that the word is pronounced as /ʃæmbe/. In fact, the /b/ sound following the /n/ sound causes this pronunciation difference. For the most part, these pronunciation differences are surface phenomena. In other words, our brains form the exact pronunciations of words (i.e., similar to those found in standard dictionaries). When the brain orders the vocal organ to vocalize these words, the physiological shortcomings of the human vocal organ cause these pronunciation differences. Many phonologists use the phrase **'Ease of Pronunciation'** to refer to this physiological phenomenon.

We can now conceptualize two types of pronunciations: phonetic and phonemic. Phonemic representation refers to the pronunciation of words as they exist in our minds; phonetic representation refers to the pronunciation of words as they are actually pronounced by our tongues. We should, however, be aware that only a very limited number of phonemes have different phonemic and phonetic representations. Take the phoneme /p/ in English as an example. When this phoneme appears in word-initial contexts, it is pronounced with a puff of air. This phenomenon is known as aspiration. In non-word-initial contexts, however, the phoneme /p/ is reduced to a phoneme which stands between the phonemes /p/ and /b/.

Phonologists have developed two types of phonetic writing system to capture these differences: (a) one in which only the mental (phonemic) representation of phonemes is shown, and (b) one in which the actual-speech (phonetic) representation of phonemes is shown. The former is called **broad transcription** while the latter is called **narrow transcription**. Broad transcription only utilizes a basic set of symbols. Narrow transcription utilizes the same set of symbols with the addition of diacritics and other symbols. The second difference between broad and narrow transcriptions is that phonemes represented in broad transcription are put between two slant lines // whereas phonemes represented in narrow transcription are put inside square brackets []. Take the following examples:

SPELLING	BROAD SYMBOL	NARROW SYMBOLS
p	/p/	[p] and [pʰ]
l	/l/	[l] and [ɫ]

The different representations of a phoneme in narrow transcription are called the **allophones** of that phoneme. Take the following examples:

(a) The phoneme /p/ has two allophones: unaspirated [p] and aspirated [pʰ].

EXAMPLE	BROAD	NARROW
spit	/spɪt/	[spɪt]
Peter	/'pi:tə/	['pʰi:tə]
slap	/slæp/	[slæp]

(b) The phoneme /l/ has two allophones: light [l] and dark [ɫ].

(c) The phoneme /k/ has two allophones: aspirated [kʰ] and unaspirated [k].

EXAMPLE	BROAD	NARROW
look	/lʊk/	[lʊk]
black	/blæk/	[blæk]
kill	/kɪl/	[kʰɪɫ]

Allophones of a phoneme are in **complimentary distribution**. That is, they cannot occur in the same context. For example, [ph] comes at the beginning of a word while [p] occurs in other contexts.

Another point to notice about allophones is that the differences between them are phonetic rather than phonemic. A phonetic difference does not cause a change in meaning. A phonemic difference, however, brings about a change in meaning. For instance, the difference in words like ship /ʃɪp/ and sheep /ʃi:p/ is phonemic because these two words have two different meanings. Therefore, we cannot consider /ɪ/ and /i:/ to be allophones of a basic phoneme. Words like sheep /ʃi:p/ and ship /ʃɪp/ are called **minimal pairs**. Traditionally, minimal pairs were defined as pairs of words that differ in one and only one phoneme. Take the following examples:

thy	/ðaɪ/	thigh	/θaɪ/
ship	/ʃɪp/	sheep	/ʃi:p/
bow	/baʊ/	wow	/waʊ/
kill	/kɪl/	keel	/ki:l/

If one of the words in a minimal pair is repeated, a **minimal set** will result. Take the following examples:

thy	thigh	thigh
ship	ship	sheep
bow	wow	bow
keel	kill	kill

One classic book with a good number of minimal pairs in it is the *American PDs* (or *American Pronunciation Drills*). The *American PDs* is still widely used in phonetics classes in a good number of language schools throughout the world.

It should be noted that some phonologists are inclined to use the diacritic symbols that represent primary and secondary stress in ordinary writing too. In this case, the symbol ´ is used to represent primary stress and the symbol ` to represent secondary stress. In phonetic transcription, however, the symbol ' is used for primary stress and the symbol ˌ for secondary stress. The other difference is that in ordinary writing, the symbols appear over the vowels that carry them. In phonetic transcription, on the other hand, the ˈ symbol is put at the top-left corner and the symbol ˌ at the bottom-left corner of the syllables that carry them in British English. In American English, the ˈ symbol appears at the top-right corner and the symbol ˌ at the bottom-right corner of the syllables that carry them. Take the following examples:

WRITING	BRITISH METHOD	AMERICAN METHOD
dèmonstrátion	/ˌdemənˈstreɪʃən/	/demˌənstreɪˈʃən/
àristócracy	/ˌærɪˈstɒkrəsɪ/	/ærˌɪstɒkˈrəsɪ/
hèrmeneútics	/ˌhɜːmɪˈnjuːtɪks/	/hɚˌmɪnuːˈtɪks/

(The /ɚ/ sound in American English stands for the /ɜː/ sound of British English, and is pronounced as a combination of /ə/ and /r/ sounds).

5. SYLLABLE STRUCTURE

The sounds that result from one chest pulse form a **syllable**. In its minimal form, a syllable consists of a vowel. In addition to the vowel a syllable may consist of one or more consonants that appear on either or both sides of the vowel. In some languages like Japanese, most often the syllable is composed of one consonant followed by one vowel. These languages are called **syllabic** languages. In syllabic languages, each syllable is represented by a symbol (called **syllabary**) in the writing system. The word TOYOTA from the Japanese language for example includes three syllables: TO, YO, and TA. Therefore, the syllable structure of most Japanese syllables is very simple: Consonant + Vowel (CV). Most languages are, however, **alphabetic** in that symbols (called characters or letters) in their orthography represent sound segments or phonemes rather than syllables. In these languages, the consonants and vowels are arrayed in a linear fashion to represent the syllables, words, sentences, etc. Arabic and Hebrew, however, tend to arrange their consonants in a linear fashion, and superscribe or subscribe their vowels as diacritics or sporadic features above or under their consonants. As such, Arabic and Hebrew can ironically be called betagamic rather than alphabetic languages.

Many of the most famous languages of the world, including English, are, however, alphabetic in the sense that they represent both their vowels and consonants in the form of letters in their orthography. In such languages, words are composed of one or more syllables. A **syllable** is a phonological unit composed of one or more phonemes. Every syllable has a nucleus, which is usually a vowel (but which may be a syllabic liquid or nasal). The nucleus may be preceded by one or more phonemes called the syllable **onset** and followed by one or more segments called the **coda**. English is an alphabetic language which has a complex syllable structure. The syllable structure of English has been presented in table 2.4 below.

SYLLABLE STRUCTURE	EXAMPLE
V	I /aɪ/
CV	key /ki:/
CCV	tree /tri:/
CCCV	spree /spri:/
CVC	seek /si:k/
CCVC	speak /spi:k/
CCCVC	scram /skræm/
CCCVCC	striped /straɪpt/
VC	an /æn/
VCC	ant /ænt/
VCCC	ants /ænts/
CVCC	pant /pænt/
CVCCC	pants /pænts/
CCCVCCC	splints /splɪnts/
CCVCC	stamp /stæmp/

Table 2.4: Syllable structure of English

Table 2.4 has used the symbol **C** to represent consonants and the symbol **V** to represent vowels. Notice that the syllable structure of English includes at least fifteen different types of syllables:

In this connection, it is interesting to notice that in alphabetic languages, the number of vowels that appear in a word can be used as an index for determining the number of syllables that make that word. A close look at the syllable structures presented in table 2.4 above reveals that, in English, consonant clusters can occur in both syllable-initial and syllable-final positions (i.e., as onset or coda). Moreover, consonant clusters are not limited to two consonants in English. In a word like street three consonants cluster together at the beginning of the syllable to produce a **CCCVC** syllable. Another interesting observation is that vowels can initiate syllables in English.

The syllable structure of Persian is, however, different. On the one hand, Persian syllables cannot be initiated with vowels; even words that seem to start with a vowel include the glottal stop /ʔ/ as the syllable onset. On the other hand, syllable-initial consonant clusters are impossible in Persian. In addition, syllable-final consonant clusters in Persian normally take no more than two consonants in their structure. As such, most Persian syllables belong in one of the three syllable structures (i.e., **CV**, **CVC**, or **CVCC**) presented in table 2.5 below. Take the following Persian examples:

PERSIAN SYLLABLES	EXAMPLE
CV	ba /ba:/ meaning with
CVC	toop /tu:p/ meaning ball
CVCC	satr /sætr/ meaning line
	abr /ʔæbr/ meaning cloud

Table 2.5: Persian syllable structure

The differences between the syllable structure of Persian and English are responsible for a good portion of Iranian EFL learners' pronunciation problems.

In fact, many Iranian EFL learners tend to insert the vowel /e/ in many mono-syllabic English words to make them readily pronounceable. In addition, since Persian syllables cannot be initiated by vowels, many Iranian EFL learners start pronouncing vowel-initial English syllables with the consonant /ʔ/. The term Penglish is sometimes used to refer to Persian pronunciation of English words. The result of such mispronunciations is that many monosyllabic English words are rendered as bi- or tri-syllabic by some Iranian EFL students. Take the following examples:

EXAMPLE	PRONUNCIATION	MISPRONUNCIATION
out	/aʊt/	/ʔaʊt/
tree	/tri:/	/teri:/
dress	/dres/	/deres/
street	/stri:t/	/ʔesteri:t/
sky	/skaɪ/	/ʔeskaɪ/

6. RECEIVED PRONUNCIATION

Why should anyone want to learn the speech sounds for a British accent that is spoken by less than 3% of the population of that country? And, Britain itself provides only a minority of the English speaking peoples of this world. The reason is mainly to do with a legacy of history. Throughout the nineteenth century and throughout the early part of the twentieth century, **Received Pronunciation (RP)** was very much the language of the ruling and educated classes. A vicious circle was then instituted: those who could afford an education went to the private schools and to university where they learnt RP; the teachers for the next generation were then drawn from this class of people to teach the next generation of the ruling elite. Thus, the educators were instructed in RP to teach RP. Therefore, most of the early phonetics work was carried out by RP speakers using their own accent as the "standard" from which all other varieties were measured.

RP does have the advantage of being a regionless accent although many of these speakers are concentrated in the southeast of England. This is partly to do with its circulation amongst the educated and elite, but it was also the variety that was exported through the colonies during the time of the British Empire. Consequently, the ruling elites of many of these countries also adopted RP as their "standard."

One of the problems of RP is that its association with the ruling elite has meant that it is regarded as being a classist accent. Aloofness and snobbishness are characteristics often associated with RP speakers. The legacy of history has meant that RP is the closest that British English has towards a standard variety. It has been subjected to a great deal of academic scrutiny, it is the choice of many broadcasters and it is an accent that many parents aspire for their children to emulate. The main advantage of learning and using RP is that it is an accent that all English speakers can understand, wherever they are in the world. It may have many problems attached to it, but it is still the most widely understood and respected of the British English varieties. The American counterpart of RP is usually referred to as **General American English (GAE)**. Foreign language learners should make no value judgments on the advantage of learning one (RP or GAE) over the other.

EXERCISE

1. Write the phonetic symbol for the first sound in each of the following words according to the way you pronounce it.

Examples:	ooze [u:]			psycho	[s]
write	[ ]	advantage	[ ]	counter	[ ]
mnemonic	[ ]	variety	[ ]	accent	[ ]
philosophy	[ ]	standard	[ ]	post	[ ]
read	[ ]	still	[ ]	cheat	[ ]

2. Write the following words in both broad and narrow phonetic transcriptions, according to your pronunciation.

WORD	BROAD TRANSCRIPTION	NARROW TRANSCRIPTION
rumor		
pacific		
presumptuous		
characteristic		
redemption		
soliloquy		
catastrophe		
posthumous		
realization		
rupture		

3. The following lines of poetry include mispronunciations. Rewrite them in the correct narrow phonetic transcription.

MISPRONUNCIATIONS	CORRECT NARROW TRANSCRIPTIONS
[ðə tʰaɪm hæz cʌm]	
[ði: vɔlrəs seɪd]	
[tʰu: tɔ:lk əv mænɪ θɪŋz]	
[əv ʃu:z ænd ʃɪpz]	
[ænd si:lɪŋ wæx]	

4. The following sentences are written in normal English orthography. Write them in broad transcription.

ORTHOGRAPHY	BROAD TRANSCRIPTIONS
How was your weekend?	
Have you ever seen him?	
He does some incredible things.	
I've never been there.	
How many rooms are there?	
How do you like it there?	
You're in great shape, Keith.	
I won't play too hard.	
I hardly ever do yoga at noon.	
How often do you exercise?	

5. Write the onset of each of the following syllables in the spaces provided. If the syllable has no onset, write Ø. The first one has been done as an example for you.

SYLLABLE	ONSET	SYLLABLE	ONSET
e.g.: street	str	pant	
speak		ant	
scram		striped	
an		splints	
ants		pants	

6. Write the coda of each of the following syllables in the spaces provided. If the syllable has no coda, write Ø. The first one has been done as an example for you.

SYLLABLE	CODA	SYLLABLE	CODA
e.g.: street	t	pant	
speak		ant	
scram		striped	
an		splints	
ants		pants	

7. Define the following terms briefly.

complimentary distribution
ease of Pronunciation
diacritics
sporadic features
transcription
sound segment
phonetic alphabet
phoneme
allophone
broad transcription
narrow transcription
coda
onset
syllabary
syllable
minimal pair

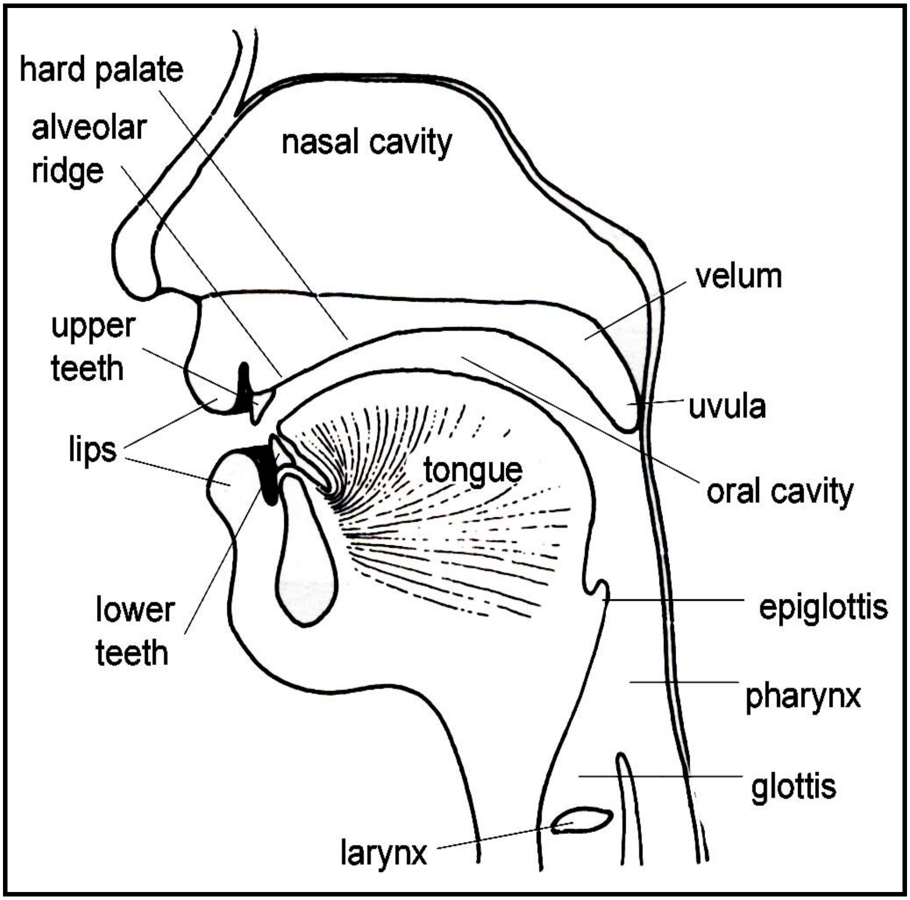
CHAPTER THREE
TRADITIONAL ARTICULATORY PHONETICS

1. INTRODUCTION

Early phoneticians used phonetic symbols as abbreviations for full descriptions of the sounds of language. In order to afford a precise account of the sounds of language, they used the **vocal tract activities** during **sound articulation** as their classification criterion. In other words, each sound in any language is considered to be a complex pattern of overlapping waves moving through air. These patterns are thought to be caused by different configurations of the various organs of speech as air travels from the lungs out through the nose or mouth or both.

So, phoneticians argued that a sound can be best described if we describe it in terms of air modifications that occur in the vocal organ of speakers. In fact, some phoneticians turned to anatomists for help in this connection. This gained so much importance that a good number of books on phonetics included chapters devoted partially or completely to a discussion of the functions of human respiratory system. A discussion of the human vocal organ is, therefore, necessary here.

The primary function of the human vocal organ is not speech production. In fact, this organ has been developed for purposes of breathing. In other words, the primary function of this organ is respiration. We take oxygen into our lungs (inhalation) and send carbon dioxide out of our lungs (exhalation) in order to remain alive. Speech is only the peripheral or subsidiary function of human vocal organ also called the human vocal tract. Speech sounds are not the only sounds produced by human vocal tract. All of us sneeze, snore, grunt, groan, scream, hiccup, cough, etc. It is, however, impossible to find any language in which these noises are part of its speech articulation system.



Sagittal section of human vocal tract (Reprinted from Lieberman and Blumstein, 1988, P. 43)

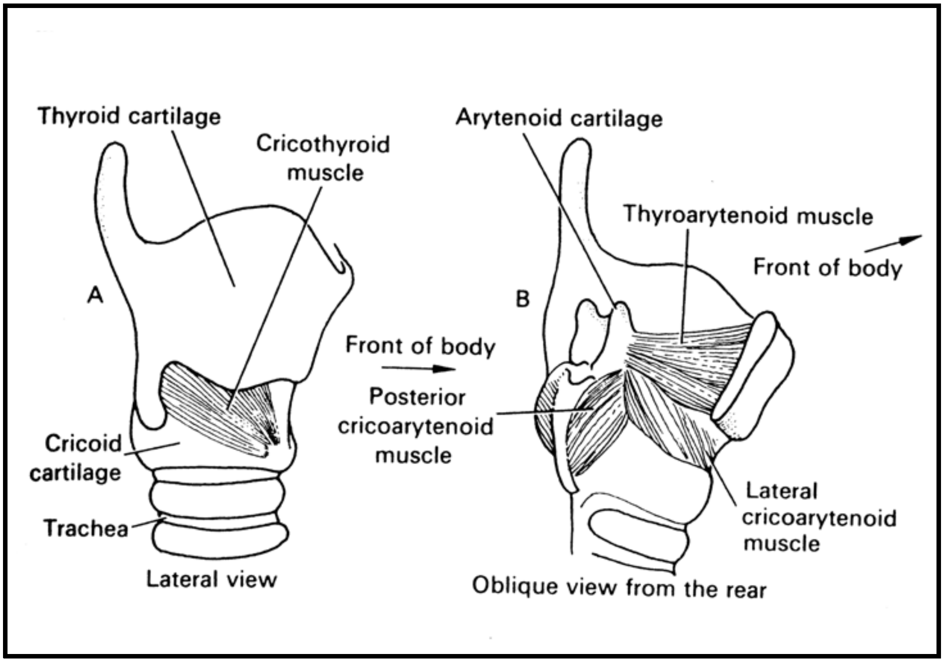
In order to produce speech, we modify the flow of air that passes between our lungs and the outside atmosphere. These modifications are normally applied to the flow of carbon dioxide during exhalation to produce a class of sounds technically called **egressive sounds**. In some African languages, however, people make some speech sounds when they are inhaling air **(ingressive**

sounds). It can, therefore, be concluded that **air stream mechanism** is at the heart of speech.

Normal name	Fancy name	Adjective
lips	labia	labial
teeth	----	dental
alveolar ridge	----	alveolar
(hard) palate	----	palatal
soft palate	velum	velar
uvula	----	uvular
upper throat	pharynx	pharyngeal
voicebox	larynx	laryngeal
tongue tip	apex	apical
tongue blade	lamina	laminal
tongue body	dorsum (back)	dorsal
tongue root	----	radical
glottis	----	glottal
nose	nasal cavity	nasal
mouth	oral cavity	oral

In addition to their normal names, many of the parts of the vocal tract have Latin or Greek fancy names. The adjectives we use to describe sounds made with each part are usually based on the Latin/Greek name.

In phonetics, the terms *velum, pharynx, larynx,* and *dorsum* are used as often, or more often, than the simpler names. Many of the names that appear in the sagittal section of the human vocal tract are already known to you. There are, however, some names that you may find new. It is, therefore, necessary to provide a definition for each of these new terms.



Cartilages and intrinsic muscles of the larynx (Reprinted from Lieberman and Blumstein, 1988, P. 98)

The **alveolar ridge** is a short distance behind the upper teeth. It is a change in the angle of the roof of the mouth. (In some people it's quite abrupt, in others very slight.) Sounds which involve the area between the upper teeth and this ridge are called **alveolars**. **Hard palate** is the hard portion of the roof of the mouth. The term "palate" by itself usually refers to the hard palate. **Soft palate/velum** is the soft portion of the roof of the mouth, lying behind the hard palate. The velum can also move: if it lowers, it creates an opening that allows air to flow out through the nose; if it stays raised, the opening is blocked, and no air can flow through the nose. **Uvula** is the small, dangly thing at the back of the soft palate. The uvula vibrates during the /r/ sound in many French dialects. **Pharynx** is the cavity between the root of the tongue and the walls of the upper throat. **Tongue blade** is the flat surface of the tongue just behind the tip. **Tongue body/dorsum** refers to the main part of the tongue, lying below the hard and soft palate. The body, specifically the back part of the body (hence "dorsum", Latin for "back"), moves to make vowels and many consonants. **Tongue root** is the lowest part of the tongue in the throat. **Epiglottis** refers to the fold of tissue below the root of the tongue. The epiglottis helps cover the larynx during swallowing, making sure that food goes into the stomach (through the esophagus) and not into the lungs. A few languages use the epiglottis in making sounds. English is fortunately not one of them. **Vocal folds/vocal cords** are folds of tissue (i.e., membranes) stretched across the airway to the lungs. They can vibrate against each other, providing much of the sound during speech. **Glottis** refers to the opening between the vocal cords. During a glottal stop, the vocal cords are held together and there is no opening between them. **Larynx** is the structure that holds and manipulates the vocal cords. The "Adam's apple" in males is the bump formed by the front part of the larynx.

As it was mentioned in the previous chapter, human beings can not only produce sounds but also noise by means of their vocal tracts. All of us have the experience of coughing after catching cold or flu. We also sneeze, snore, grunt, groan, scream, etc. You, however, cannot find any one language in which words are formed as sequences of these noises. There is a quite simple reason for this. Noises interfere with the primary functions of our vocal organs, that of respiration. In addition, they cause pain and bruise in different parts of the vocal tract. Now, you can imagine what happens to the health of an individual if he had to use these noises in his speech. In fact all natural languages employ sounds (i.e., consonants and vowels). These are well-organized sounds which do not cause problems for respiration.

In order to produce speech sounds, the flow of pulmonary air in the vocal tract must somehow be modified. These modifications are of four main types: (1) complete blockage followed by sudden release, (2) complete blockage followed by gradual release, (3) constriction or narrowing, and (4) no blockage at all. In fact, these modification types have been used in traditional phonetics as the criteria for the classification of speech sounds. These modifications are technically referred to as **manner of articulation**. These modifications do not occur at one single part of the vocal tract. Rather, they may happen at any part of the super-glottal (from the glottis upwards) section of the vocal tract. These places (sometimes called articulators) include the lips, the teeth, the alveolar ridge, the palate, the velum, and the glottis. **Place of articulation** is the technical term which is used to refer to the places along the vocal tract at which air modifications take place. In addition to place and manner of articulation, speech sounds can be **voiced** or **unvoiced** (voiceless) depending on whether the vocal cords are set into vibration by the impact of pulmonary air. This phenomenon is technically referred to as **voicing**.

2. SPEECH ARTICULATION

Voicing, manners, and places of articulation are responsible for the production (articulation) of different speech sounds. Sound is produced by the interference of the flow of air through the mouth (and nose). **Consonants** are created when the airflow is directly restricted, or obstructed. As such, pulmonary air cannot escape from the oral cavity without creating audible friction. By way of contrast, **vowels** are created when the airflow is not crucially restricted or obstructed. Therefore, pulmonary air can escape the oral cavity without creating audible friction. Defining characteristics of consonants include: (a) voicing (b) place of articulation (c) manner of articulation. Defining characteristics of vowels are: (a) lip rounding, and (b) relative tongue position.

2.1. MANNER OF ARTICULATION

Manner of articulation refers to the nature of the obstruction of pulmonary air flow. In order to fully appreciate the differences among speech sounds, as well as indicating the place of articulation, it is necessary to determine the nature and extent of the obstruction of airflow involved in their articulation. The type of airflow obstruction is known as the **manner of articulation**. The manner of articulation is particularly defined by four major factors: (a) whether there is vibration of the vocal cords (voiced vs. voiceless), (b) whether there is obstruction of the airstream at any point above the glottis (consonant vs. vowel), (c) whether the airstream passes through the nasal cavity in addition to the oral cavity (nasal vs. oral), and (d) whether the airstream passes through the middle of the oral cavity or along the side(s) (non-lateral vs. lateral). An example of this can be found by looking at the following words:

nine	/naɪn/	dine	/daɪn/	line	/laɪn/

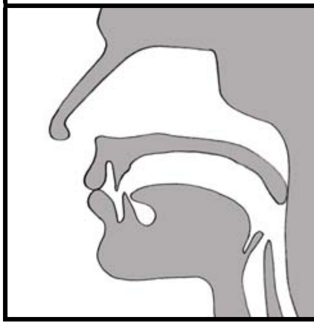
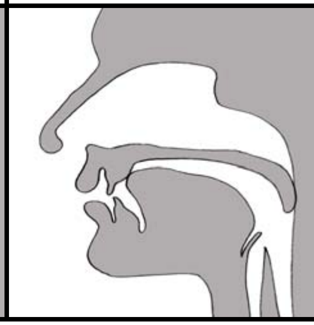
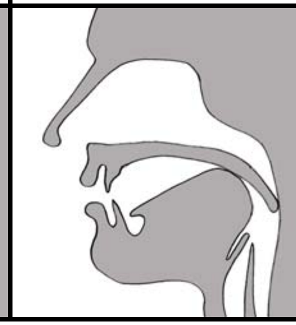
They all begin with voiced, alveolar consonants /n/, /d/, and /l/. Yet, they are all clearly different in both sound and meaning. The kinds of constriction made by the articulators are what make up this further dimension of classification. There are two common kinds of constriction that often occur in English: **plosive** and **fricative**. Also, there are other less common constrictions: **nasal**, **affricate**, **lateral**, and **approximant**. Traditional phonetics, however, used three cover terms to refer to all kinds of constriction: plosive, fricative, and affricate. In the following sections, these manners of articulation will be discussed with greater detail so that the reader can fully understand what they mean.

2.1.1. OCCLUSIVE AND PLOSIVE

Occlusives require a complete closure of the speech canal, not just a restriction. This distinguishes them from the continuants. The occlusion is twofold: (a) the airstream is halted by a sudden closure in the oral cavity; (b) the trapped air is freed by abruptly releasing the closure. If the trapped air is gradually released, an **affricate** consonant is articulated. Occlusives in English include /p/, /b/, /m/, /t/, /d/, /n/, /k/, /g/, and /ŋ/.

[p] is a voiceless bilabial stop consonant. The lips are pressed tightly together. The air is trapped behind the lips. The vocal cords are kept far apart, and the nasal cavity is closed by the velum. Then the trapped air is suddenly released. [b] is the voiced counterpart of [p]. The only difference is that the vocal cords

are close to each other and vibrate during the articulation of [b]. In the case of /m/, the nasal cavity is open.

/b/ and /p/	/t/ and /d/	/k/ and /g/
		

[t] is a voiceless dental or alveolar stop. The tongue makes contact with the front teeth or with the alveolar ridge directly above them. There is no vocal cord vibration and the nasal cavity is blocked. [d] is a voiced dental or alveolar stop. It is produced in the same way as [t] but with vibration of the vocal cords. In the case of /n/, the nasal cavity is open to let the air pass through it.

[k] is a voiceless velar stop. With the tongue tip resting against the lower teeth, the back of the tongue makes contact with the soft palate. [g] is its voiced counterpart. Its articulation is the same as [k], but with vibration of the vocal cords. The corresponding velar nasal [ŋ] is usually voiced as well.

Some languages, including Persian, have a glottal occlusive [ʔ] too. The glottal stop can be produced in either of the two ways: (a) by the sudden opening of the glottis under pressure from the air below, or (b) by the abrupt closure of the glottis to block the airstream. The glottal stop is always voiceless, as the complete closure of the vocal cords precludes their vibration.

Occlusives can be categorized into two major types: **stops** and **plosives**. The two categories are in inclusional distribution. That is, all plosives are stops but all stops are not necessarily plosive. This relationship can be schematically represented as:

(PLOSIVE ➙ STOP).

Plosive sounds are made by forming a complete obstruction to the flow of air through the mouth and nose. The first stage is that a closure occurs. Then the flow of air builds up and finally the closure is released, making an explosion of air that causes a sharp noise. Try to slowly say /p/ to yourself. You should be able to feel the build up of air that bursts into the /p/ sound when you open your lips. It should be noted that a plosive cannot be prolonged or maintained so that once the air has been released, the sound has escaped. As such, plosive sounds lack the length feature. Contrast this quality of plosives with a fricative in which you can lengthen the sound. The plosive sounds in RP are: /b/, /p/, /t/, /d/, /k/, and /g/. As it was mentioned earlier, plosive sounds belong to a more general class of sounds called stops. A stop sound is one in which the flow of air is completely blocked only in the oral cavity. Stops also include such sounds as /m/, /n/, and /ŋ/. Take the following examples:

moon /mu:n/ night /naɪt/ thing /θɪŋ/

You can feel that in the production of such sounds as /m/, /n/, and /ŋ/ the flow of air is completely blocked in the mouth. However, air can flow through the nose. As such, the air cannot burst into these sounds because they can be lengthened. In addition to these sounds, /ʧ/ and /ʤ/ are also marked by a complete blockage of pulmonary air in the oral cavity. Here, again, the blockage is not followed by an abrupt release. Rather, the blocked air is gradually released to create friction. Some phoneticians rank these two sounds among the stop consonants while many others classify them as affricates. Take the following examples:

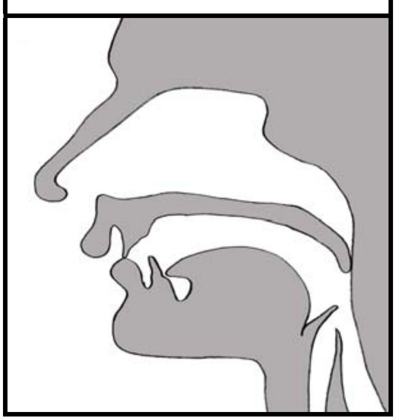
Jack /ʤæk/ chat /ʧæt/

2.1.2. FRICATIVE

A fricative is the type of consonant that is formed by forcing air through a narrow gap in the oral cavity so that a hissing sound is created. Typically air is forced between the tongue and the place of articulation for the particular sound. Try it yourself. Say the /f/ in fin /fɪn/, the /θ/ in thin /θɪn/ and the /ʃ/ in shin /ʃɪn/. You should be able to feel the turbulence created by the sounds. It is possible to maintain a fricative sound for as long as your breath holds out. This is very different from a plosive sound. Other fricatives include the /v/ in van, the /s/ in sin, the /h/ in hat /hæt/, the /ð/ in that /ðæt/, the /z/ in zoo /zu:/ and the /Ʒ/ sound in genre /'ʒɑ:nrə/.

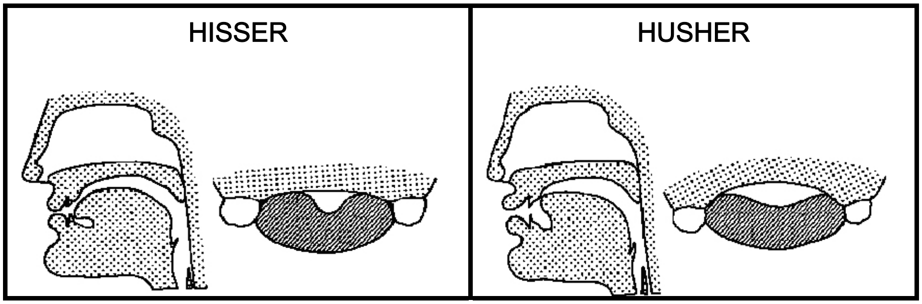
Fricative consonants result from a narrowing of the speech canal that does not achieve the full closure characteristic of the occlusives. The shape and position of the lips and/or tongue determine the type of fricative produced. Phoneticians usually distinguish between so-called true fricatives and the related class of spirants. During the production of a fricative, the airstream can be directed in several ways. First, in the case of true fricatives, the tongue channels the air through the center of the mouth (like in the case of the dorsal fricatives). Second, the tongue can also channel the air down the side(s) of the mouth (like in the case of the lateral fricatives). Finally, in the case of labial and dental fricatives, the shape and position of the tongue is not important. This makes sense because the place of articulation is not, strictly speaking, in the oral cavity at all.

/f/ and /v/



[f] is a voiceless labiodental fricative. The lower lip is brought close to the upper teeth, occasionally even grazing the teeth with its outer surface, or with its inner surface, imparting in this case a slight hushing sound. Considering its place of articulation, it is unimportant to class this sound as dorsal or lateral fricative. [v] is a voiced labiodental fricative. Its articulation is the same as [f], but with vibration of the vocal cords. Like [f], considering its place of articulation, it is unimportant to class [v] as dorsal or lateral fricative.

Among the fricatives are ones described as hissers and hushers. The realization of a hisser requires a high degree of tension in the tongue: a groove is formed along the whole length of the tongue, in particular at the place of articulation where the air passes through a little round opening. English hissers are /s/, /z/, /ʧ/, and /ʤ/. The hushers are produced similarly, but with a shallower groove in the tongue, and a little opening more oval than round. The lips are often rounded or projected outwards during the realization of a husher. English hushers are /ʃ/, and /ʒ/.



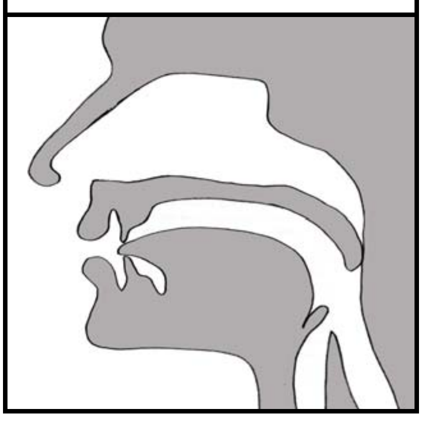
Spirants involve the same restriction of the speech canal as fricatives, but the speech organs are substantially less tense during the articulation of a spirant. Rather than friction, a resonant sound is produced at the place of articulation.

[s] is a voiceless alveolar fricative (also a hisser). This apico-alveolar hisser is produced by bringing the end of the tongue close to the alveolar ridge. Hissers like /s/ can be divided into three categories, according to the precise part of the tongue that comes into play: (a) **coronal hissers** which involve the front margin

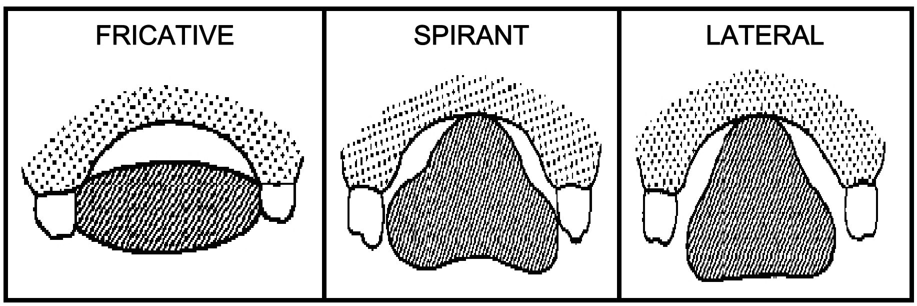
of the tongue (as in English), (b) **apical hissers** which involve the very tip or apex of the tongue (as in Castilian Spanish), and (c) **post-dental** hissers where the front part of the tongue body is involved (as in French). The quality of the sound is noticeably altered in these three types of hissers. The IPA uses diacritical marks to indicate distinctions of this magnitude. [z] is a voiced alveolar fricative (and a hisser). The same mechanism that produces [s] also produces [z] but with vibration of the vocal cords. In general, the remarks made for the voiceless sound are equally valid for the voiced variant. Other hissers are [ʧ] and [ʤ]. [ʧ] is a voiceless palatal fricative. The tongue body forms a groove and approaches the hard palate. In terms of general tongue shape, this articulation qualifies as a hisser. [ʤ] is a voiced palatal fricative. It is articulated in the same way as [ʧ] but with vibration of the vocal cords. [ʃ] is a voiceless alveolar fricative (and a husher). The tip of the tongue touches the alveolar ridge. The groove in the center of the tongue is shallower than is the case for hissers. [ʒ] is its voiced counterpart.

Basically, friction and fricatives develop from tense articulations. When the articulation is lax, resonance, and thus a spirant, occurs. Also realize that many spirants can be thought of as the lax counterparts of stop consonants. English interdentals /ð/ and /θ/ are spirants.

/θ/ and /ð/



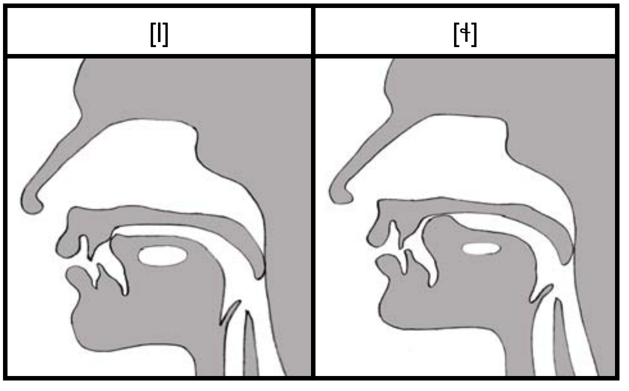
[θ] is a voiceless dental spirant. The tongue tip is held close to the upper teeth, either behind them (dental) or just underneath them (the interdental articulation). This spirant is the lax counterpart of the stop [t]. Considering its place of articulation, it is unimportant to class this sound as a dorsal or lateral fricative. [ð] is the voiced counterpart of /θ/. Its articulation is the same as /θ/ but with vibration of the vocal cords. This spirant is the lax counterpart of the stop [d].



Laterals are generally considered to be a special case, since physically speaking they could be grouped among the fricatives and spirants. They are called **laterals** since, during their production, the back of the tongue makes contact with the hard palate while the front of the tongue sinks down, channeling the air laterally around the tongue, down the side (or sometimes both sides) of the mouth. On the other hand, for non-lateral articulations, the back of the tongue rests against the top molars, and the air flows over the tongue down the center of the mouth. There are two distinct types of lateral: (a) lateral fricatives, where the articulation, requiring a great deal of muscular tension, resembles that of the fricatives (except for the position of the tongue), and (b) non-fricative lateral, often called **liquids**, whose articulation is very close to the spirants. It is interesting to note that the location of the lateral channel through which the air flows is unimportant. That is, whether it is on the left, the right, or both sides of the mouth, the nature of the sound produced is unchanged. English laterals are usually non-fricative. In addition, no distinction is made between voiceless and voiced variants because it is very rare for a language to distinguish laterals according to voice. English liquids are /r/ and /l/.

2.1.3. LATERAL

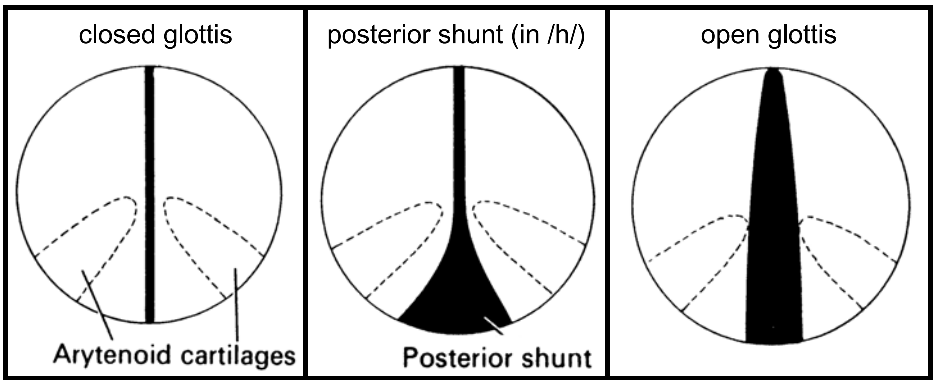
To produce a lateral sound, air is obstructed by the tongue at a point along the centre of the mouth but the sides of the tongue are left low so that air can escape over its sides. In fact, the tongue is strongly flexed and the air is forced through a narrow oval cavity, producing a hushing sound. /l/ is the clearest example of a lateral sound in English. Both the clear [l] and the word-final dark [ɫ] allophones (i.e., the variants of the same phoneme) of /l/ are lateral sounds. When an alveolar plosive is followed by the lateral /l/, then what happens is that we simply lower the sides of the tongue to release the compressed air, rather than raising and lowering the blade of the tongue. If you say the word 'bottle' to yourself you can feel the sides of the tongue lower to let out the air.



2.1.4. APPROXIMANT

An approximant is a consonant that makes very little obstruction to the airflow. Approximants are divided into two main groups: **semivowels (also called glides)** and **liquids**. The semivowels are /h/ as in hat /hæt/, /j/ as in yellow /'jeləʊ/, and

/w/ as in one /wʌn/. They are very similar to the vowels /ɜː/, /uː/ and /iː/, respectively. However, semivowels are produced as a rapid glide. The liquids include the lateral /l/ and /r/ sounds in that these sounds have an identifiable constriction of the airflow, but not one sufficiently obstructive enough to produce a fricative sound. Approximants are never fricative and never completely block the flow of air.



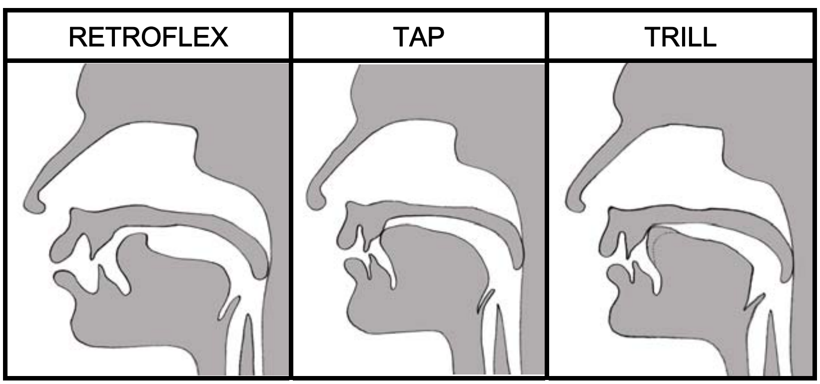
Adjustments of the larynx for phonation (Reprinted—with some modifications—from Lieberman and Blumstein, 1988, P. 111)

[h] is an interesting case. It is a voiceless glottal fricative. The glottis is almost completely closed, except for a narrow opening (i.e., the posterior shunt) in its upper part at the level of the arythenoidal cartilage. A strong friction develops when air flows through this opening.

One point of caution is, however, necessary here. English and Persian make use of different forms of the approximant /r/. On the whole, the approximant /r/ is of three types: (a) flap, (b) trill, and (c) retroflex. English draws on trill and retroflex types while Persian employs trill and flap types. In Persian, the /r/ sound becomes flap whenever it appears between two vowels. A flap /r/ is articulated when the blade of the tongue touches the roof of the mouth only once. Try saying the following Persian words to yourselves:

/bɑːˈrɑːn/ (meaning rain)   /ˈmibɑːræd/ (meaning precipitates)

You should be able to feel the /r/ sound in these Persian words. It is somewhat similar to the pronunciation of the intervocalic (between two vowels) /t/ sound in American pronunciation of such words as 'water', 'butter', 'better', etc. Notice, however, that English lacks flap /r/. The trill /r/ is articulated when the blade of the tongue repeatedly touches the roof of the mouth for several times (similar to the noise made by a chain gun). The /r/ sound in the Russian language is an example of the extreme case of trill /r/. Both English and Persian employ trill /r/. The retroflex /r/ is specific to English. Whenever the /r/ sound starts a word in English, it becomes retroflex: (a) the tip of the tongue curls back, (b) the upper and lower teeth approach each other, and (c) the lips are rounded and sent forward. Hence, the retroflex /r/. try saying the English words 'write', 'red', etc. to yourselves. They start with retroflex /r/.

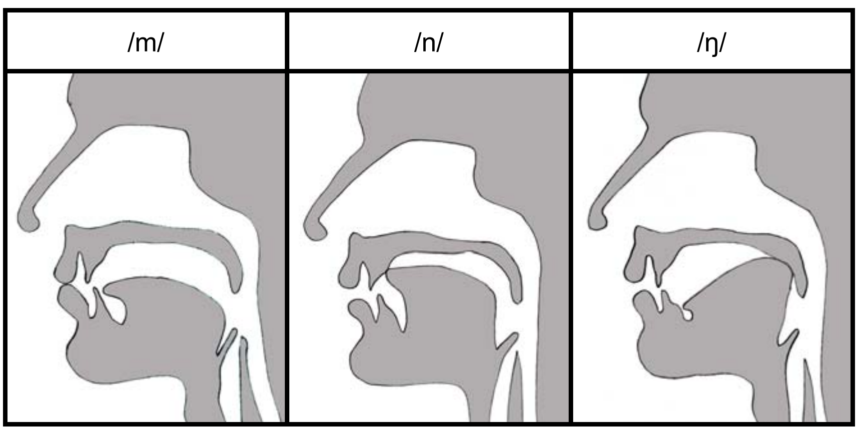


Many phoneticians use the term **vibrant** to refer to these and similar consonants. Vibrant consonants involve one or more tapping or flapping vibrations of the speech organs under pressure from the airstream. Part of the tongue makes contact with the palate, most commonly at the alveolar ridge, the soft palate, or (in some languages) the uvula. One or more very brief occlusions (or air blockage) occur successively, accompanied by short resonances. Vibrants are generally voiced. In narrow transcription, a small subscript circle may be added to any IPA symbol to indicate a voiceless variant. There are two distinct classes of vibrants: (a) those with only one vibration, called **taps**, and (b) those with multiple vibrations, called **trills**. [ɾ] is an alveolar tap (sometimes called flap). The alveolar region serves as the target for the tongue tip, which vibrates there briefly before falling back to rest against the lower teeth. [ɽ] is a retroflex tap. The tip of the tongue curves up and back, and its underside vibrates briefly against the roof of the mouth, before falling back to rest against the lower teeth. [r] is an alveolar trill. The alveolar region serves as the target for the tongue tip, which vibrates there under pressure from the airstream behind. The vibration produces occlusive sounds and vocalic-type resonances in rapid alternation. This is the famous rolled /r/ of Spanish, Russian, and some other languages.

2.1.5. NASAL

You have already noted in the discussion of plosives that, in the articulation of some sounds, pulmonary air flows throw the nasal cavity to produce a nasal consonant. A nasal consonant is a consonant in which air escapes only through the nose. For this to happen, the soft dorsal part of the soft palate is lowered to allow air to pass it, whilst a closure is made somewhere in the oral cavity to stop air escaping through the mouth. You can feel if a sound is a nasal sound or not by placing your hand in front of your mouth and feeling if any air is escaping or not. There are three nasal sounds in English. The /m/ in mat /mæt/, the/n/ in not /nɒt/ and the /ŋ/ in sing /sɪŋ/ or think /θɪŋk/. The velar /ŋ/ does not occur in Persian. The Iranian learner of English, therefore, replaces the /ŋ/ phoneme by a sequence of two phonemes /n/, and /g/. In English, whenever the letter **n** appears before the letters **g** and **k**, it is pronounced as /ŋ/. Nasal sounds are considered to belong to the stop category along with plosives and affricates.

The nasal “occlusives” of the vast majority of the world's languages are voiced. Very few not-so-famous languages have voiceless nasals too. During the production of these nasal “occlusives”, the soft palate is lowered to a greater or lesser extent, allowing a portion of the airstream to pass through the nasal cavity. Occlusion occurs in the mouth only; the nasal resonance is continuous. Indeed, many linguists rank the nasals among the continuants. /m/ is a bilabial nasal. The mouth is configured just as for the corresponding bilabial stop /p/ and /b/. The lips are pressed tightly together. The air builds up and is suddenly released. /n/ is a dental or alveolar nasal. The mouth is configured just as for the corresponding dental or alveolar stop /t/ and /d/. The tongue makes contact either with the front teeth, or with the alveolar ridge directly above them. /ŋ/ is a velar nasal. The configuration of the mouth is very close to that of the corresponding velar stop /k/ and /g/. With the tongue tip resting against the lower teeth, the back of the tongue makes contact with the soft palate. But as the soft palate is lowered (to allow air to flow through the nasal cavity), the tongue's movement is more important for the nasal than for the oral sound.

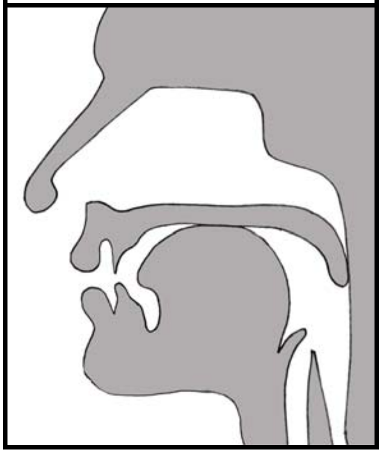


2.1.6. AFFRICATE

An affricate is a plosive immediately followed by a fricative in the same place of articulation.

The /ʧ/ in chap /ʧæp/ and the /ʤ/ in jeep /ʤi:p/ are the two clear affricates in English. If you think about it, the /ʧ/ sound is made up from the plosive /t/ and the fricative /ʃ/ sounds. Likewise, the /ʤ/ sound is made up from the plosive /d/ immediately followed by the fricative /z/.

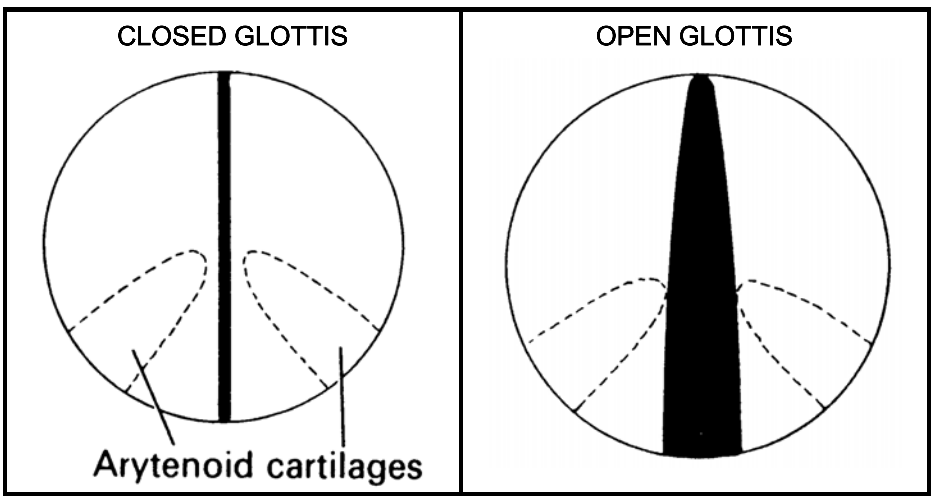
/ʧ/ and /ʤ/



2.1.7. VOICING (LEVEL OF VIBRATION)

All of us inhale oxygen and exhale carbon dioxide. Pulmonary air gets out the lungs and enters the bronchi. The two bronchi meet each other and form the trachea. The trachea is intercepted by the voicebox or larynx on its way out. Inside the voice box, there are two membranes that are hinged together at the back. These membranes are called the vocal cords or the vocal folds. The vocal cords can make a wedge-shaped opening when they are far apart. This opening is called the glottis.

The level of vibration of the vocal cords determines whether a sound is voiced or unvoiced. When the glottis is open, pulmonary air passes through it easily without causing any friction. That is, if the vocal cords are apart, then air can escape unimpeded. Sounds produced in this way are said to be voiceless. The easiest example of this is to whisper. When you whisper, your glottis is wide open and, therefore, all the sounds produced are voiceless. However, when the glottis is closed, the vocal cords are set into vibration by the impact of the pulmonary air. When the vocal cords vibrate, voice sounds are produced. When they do not vibrate, voiceless sounds result. That is, if the vocal cords are very close together, the air will blow them apart as it forces its way through. This makes the cords vibrate, producing a voiced sound. Vocal cord vibration is technically referred to as voicing. In English, only a limited number of consonants are voiceless. /ʧ/, /s/, /p/, /k/, /f/, /ʃ/, /t/, /θ/, and /h/ are voiceless. All other phonemes are voiced.



Adjustments of the larynx for phonation (Reprinted—with some modifications—from Lieberman and Blumstein, 1988, P. 111)

To feel the distinction between voiced and voiceless sounds is very easy. Place your finger and thumb lightly on your throat. Say sssssss to yourself. Then say zzzzzzz. Repeat these a few times. Then substitute fffffff and vvvvvvv sounds. You should be able to feel the vibration of the cords when you say zzzzzz and vvvvvv, but not when you say sssssss and fffffff. It is also possible to hear the vibration through bone conduction. Instead of putting your fingers on your throat, put your index fingers in your ears and repeat the above sounds. You should hear a low buzzing sound when you articulate zzzzzz and vvvvvv, but hear almost nothing for the other two sounds.

Voicing is important in a language like English because the meaning of a sound often depends on whether that sound is voiced or not. For example, 'big' carries a very different meaning from 'pig'. English has many sounds that are paired up in this way. In fact, there are a number of cases in which the place and manner of articulation of phonemes are the same, but the meaning is dependant upon whether the sound is voiced or not.

2.1.8. FORTIS AND LENIS

Fortis and Lenis are closely related to level of vibration. Fortis sounds are those that are made with strong muscular effort, originating in the lungs. Lenis is the opposite. To fully understand what is meant by fortis and lenis, your attention is drawn to the fact that all of us have the ability to whisper. In fact, during a whisper, all voiced sounds are devoiced and yet still remain distinct from voiceless sounds. Try whispering /pig/ and then /big/. You can still differentiate the two because the intensity of the burst of air is greater for the fortis /p/ than for the lenis /b/. Other fortis sounds include /t/ and /k/. Other Lenis sounds include /d/ and /g/. In English, the fortis/lenis distinction happens to duplicate the voiced/voiceless one found in minimal pairs. So all the voiced sounds are lenis and all the voiceless sounds are fortis sounds.

2.1.9. LIP ROPUNDING

Another important form of air modification in which the lips are involved is lip rounding. In the production of some vowels, the glide /w/, and the retroflex /r/, the lips are rounded so that their shape becomes circular. This phenomenon is called lip-rounding. It is a useful defining feature in the classification of vowels.

2.2. PLACE OF ARTICULATION

As it was noted earlier, the distinction between manner of articulation and place of articulation is particularly important for the classification of consonants. The **place of articulation** is the point where the airstream is obstructed. In general, the place of articulation is simply that point on the palate where the tongue is placed to block the stream of air. However, the palate is not the only place of articulation.

The place of articulation can be any of these points: (a) the lips (labials and bilabials), (b) the teeth (dentals), (c) the lips and teeth (labio-dentals—here the tongue is not directly involved), (d) the alveolar ridge (that part of the gums behind the upper front teeth—alveolar articulations), (e) the hard palate (given its large size, one can distinguish between palato-alveolars, palatals and palato-velars), (f) the soft palate (or velum—velar articulations), (g) the uvula (uvulars), (h) the pharynx (pharyngeals), and (i) the glottis (glottals).

After the air has left the larynx, it passes into the vocal tract. Consonants are produced by obstructing the air flow through the vocal tract. There are a number of places where these obstructions can take place. These places are known as the **articulators**. They include the lips, the teeth, the alveolar ridge, the hard palate, the soft palate, and the throat. Some phoneticians define articulators as the movable parts of the vocal tract.

2.2.1. LIPS

Some sounds are produced when the pulmonary air is modified (blocked or constricted) by the lips. If both of the lips are used to articulate a sound, then it is said to be a **bilabial** sound. Examples of bilabial sounds include /p/, /b/ and /m/. Sometimes, the upper teeth and the lower lip collaborate to produce a sound. In this case, the sound is said to be a labiodental sound. In English /f/ and /v/ are labiodental consonants.

2.2.2. TEETH

The two 'th' sounds of English (ð and θ) are formed by forcing air through the teeth. The apex of the tongue is normally inserted between the upper and the lower teeth and air is forced out. This causes a friction which is realized in the form of two different consonant sounds: /ð/ as in this /ðɪs/, and /θ/ as in thin /θɪn/. If you say the soft /θ/ in /thin/ and then the hard /ð/ sound in /then/, you can feel the air being forced through the teeth. The tongue tip and rims are articulating with the upper teeth. These sounds are called dental or interdental sounds. The upper teeth are also used when you say /f/ and /v/. In this case however, air is being forced through the upper teeth and lower lip.

2.2.3. ALVEOLAR RIDGE

An alveolar sound is produced when the tongue tip, or blade, touches the bony prominence behind the top teeth. This prominence is in fact that part of the gum which lies behind the upper teeth. In English, the following phonemes are alveolar: /t/ as in tin /tɪn/, /d/ as in din /dɪn/, /s/ as in sin /sɪn/, /z/ as in zip /zɪp/, /l/ as in look /lʊk/, /r/ as in roof /ru:f/, and /n/ as in /night /naɪt/. Try saying all of these phonemes to yourself. They do not all touch in exactly the same way due to the **manner of articulation.** Some are **plosives** while the others are **fricatives** or **laterals**. But the place of articulation is clearly the alveolar ridge for all of them. The Iranian EFL learners' attention is specifically drawn to the difference between the Persian and English /t/ and /d/ phonemes. Whereas in English these sounds are purely alveolar stops, in Persian they are dentoalveolar phonemes. The blade of the tongue touches the back of upper teeth to form these sounds in Persian whereas, in English, the tongue lies solely behind the alveolar ridge. The difference can be felt if English /t/ and /d/ phonemes are inserted in Persian words.

PHONEME	EXAMPLE	PRONUNCIATION
/ʃ/	sheep	/ʃi:p/
/ʒ/	genre	/ˈʒɑ:nrə/
/ʧ/	cheap	/ʧi:p/
/ʤ/	jeep	/ʤi:p/

Four sounds are said to be palatoalveolar. This is partly because the blade of the tongue straddles both the alveolar ridge and the front of the hard palate as air is forced through to make the sounds. These sounds include /ʃ/, /ʒ/, /ʧ/, and /ʤ/. Take the following examples:

2.2.4. HARD PALATE

This is the hard bit of the arched bony structure that forms the roof of the mouth. The /j/ sound in yes is the clearest example of a palatal sound in RP. You can feel the fricative sound being forced between the tongue and the very top of your mouth. The blade of the tongue touches the hard palate to articulate palatal sounds. Earliest accounts of phonetics also classified affricates as being palatal. Today, however, they are considered to be palatoalveolar rather than palatal *per se*. Still, there are some phoneticians who prefer the older classification.

2.2.5. THE SOFT PALATE (VELUM)

The soft palate is toward the back of the mouth. It is where the roof of the mouth gives way to the soft area behind it. It can just be felt with your tongue if you curl it as far back and as high as you can so that the apex of your tongue can feel the soft area of the back-roof of your mouth. The soft palate is technically called the velum and sounds which are produced by a constriction or blockage at this part of the vocal tract are called velar sounds. Thus, velar sounds are usually made when the back of the tongue is pressed against the soft palate. They include the /k/ in cat, the /g/ in girl and the /ŋ/ in hang. The glide /w/ is also regarded as a labiovelar sound, because it simultaneously uses both lips whilst raising the back of the tongue towards the velum. Try saying /wheel/ and /win/ and feel the position of your tongue.

2.2.6. GLOTTIS

Glottal sounds are those sounds that are made in the larynx through the closure or narrowing of the glottis. It was mentioned earlier that there is an opening between the vocal cords which is called the glottis. If pulmonary air is constricted or blocked at this place, glottal sounds are articulated. The glide /h/

as in Helen is an example of a glottal sound. It is physically impossible to feel the process using your tongue. It is as far back as you can get in your mouth. Unlike English, Persian syllables cannot start with a vowel. As such, Iranian EFL learners are inclined to insert the glottal stop consonant /ʔ/ at the beginning of English syllables that start with a vowel. The /æ/ sound, as it appears in Arabic, is an extreme example of a glottal stop. The glottal stop is becoming a more widespread part of British English, but is still uncommon in RP. You also use your glottis for speech when you whisper or speak in a creaky voice. Take the following examples:

WORD	ENGLISH	IRANIAN MISPRONUNCIATION
ago	/ə'gəʊ/	/ʔə'gəʊ/
out	/aʊt/	/ʔaʊt/
apple	/æpl/	/ʔæpl/

3. CONSONANTS

Consonants in English are distinguished from vowels on the basis of the modifications of pulmonary air in the oral cavity. Consonants are distinguished from one another on the basis of their differences in three respects: (a) manner of articulations, (b) place of articulation, and (c) voicing. By way of contrast, vowels are distinguished from one another on the basis of two criteria: (a) relative position of the tongue in the mouth, and (b) lip rounding.

Consonant is the general term that refers to a class of sounds where there is obstruction of some kind (i.e., complete blockage, or constriction) to the flow of pulmonary air. As it was mentioned earlier, there are six different degrees of obstruction. Therefore, consonants can be classified into six different categories on the basis of their manner of articulation:

TYPE	PHONEME								
PLOSIVES	/p/	/b/	/t/	/d/	/k/	/g/			
FRICATIVES	/f/	/v/	/θ/	/ð/	/h/	/s/	/z/	/ʃ/	/ʒ/
AFFRICATES	/ʧ/	/ʤ/							
NASALS	/m/	/n/	/ŋ/						
APPROXIMANT	/r/	/w/	/j/						
LATERAL	/l/								

However, as the table shows, more than one consonants fall within almost all of these categories. Therefore, other criteria are needed to distinguish one consonant from the other. For example, /p/ and /b/ cannot solely be distinguished on the basis of their manner of articulation. Moreover, they are articulated at the same place of articulation. Yet they are different since they assign different meanings to the two English words 'pat' /pæt/ and 'bat' /bæt/.

Consonants that share the same manner of articulation may be different in terms of place of articulation. Consonants are classified into nine different classes according to their place of articulation:

TYPE	PHONEME
BILABIAL	/m/, /p/, /b/
LABIODENTAL	/f/, /v/
INTERDENTAL	/θ/, /ð/
ALVEOLAR	/t/, /d/, /n/, /s/, /z/, /r/, /l/
PALATOALVEOLAR	/ʃ/, /ʒ/, /ʧ/, /ʤ/
PALATAL	/j/
LABIOVELAR	/w/
VELAR	/k/, /g/, /ŋ/
GLOTTAL	/h/

Even those consonant that share the place and manner of articulation may be different in terms of voicing and nasality. According to the level of vibration of the vocal cords, consonants are classified into two groups: voiced, and voiceless:

TYPE	PHONEME
VOICELESS	/ʧ/, /s/, /p/, /k/, /f/, /ʃ/, /t/, /θ/, /h/
VOICED	/ʤ/, /z/, /b/, /g/, /v/, /ʒ/, /d/, /ð/, /w/, /j/, /l/, /r/, /m/, /n/, /ŋ/

On the basis of nasality, consonants are divided into two groups: nasal, and non-nasal. As you have already noticed, nasality is identified by a free flow of air through the nose.

TYPE	PHONEME
NASAL	/m/, /n/, /ŋ/
NON-NASAL	/ʤ/, /z/, /b/, /g/, /v/, /ʒ/, /d/, /ð/, /w/, /j/, /l/, /r/, /ʧ/, /s/, /p/, /k/, /f/, /ʃ/, /t/, /θ/, /h/

These differences in place of articulation, manner of articulation, nasality, and voicing led traditional phoneticians to assign different names to different consonants. The name that was given to any given consonant was based on the air stream mechanism which led to the articulation of that consonant:

Consonant name = Place of articulation + voicing + manner of articulation

For example the consonant /p/ would be identified as **bilabial voiceless stop.** By way of contrast, the consonant /b/ was defined as **bilabial voiced stop.** As such /b/ and /p/ were distinguished on the basis of the level of vibration of the vocal cords (i.e., voicing). /m/ was considered to be a **bilabial nasal.**

In traditional phonetics, consonants were named after their particular characteristics:

PHONEME	TRADITIONAL NAME
/p/	Bilabial voiceless stop
/b/	Bilabial voiced stop
/m/	Bilabial nasal
/f/	Labiodental voiceless fricative
/v/	Labiodental voiced fricative
/θ/	Interdental voiceless fricative
/ð/	Interdental voiced fricative
/t/	Alveolar voiceless stop
/d/	Alveolar voiced stop
/n/	Alveolar nasal
/s/	Alveolar voiceless fricative
/z/	Alveolar voiced fricative
/l/	lateral
/r/	Alveolar approximant
/ʃ/	Palatoalveolar voiceless fricative
/ʒ/	Palatoalveolar voiced fricative
/ʧ/	Voiceless affricate
/ʤ/	Voiced affricate
/j/	Palatal approximant
/w/	Labiovelar approximant
/k/	Velar voiceless stop
/g/	Velar voiced stop
/ŋ/	Velar nasal
/h/	Glottal fricative

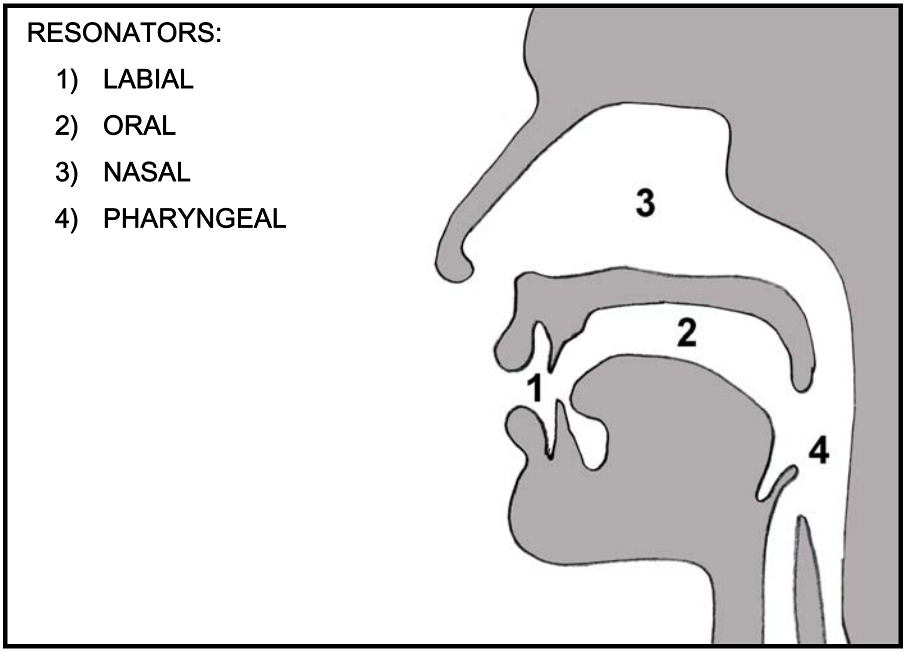
Also notice that earlier phoneticians had a tendency to distinguish approximants, glides, and laterals as fricatives. In addition, they also classified /s/, /z/, /ʃ/, /ʒ/, /ʧ/, /ʤ/, /r/, and /l/ as palatals.

A closer look at the names given to consonants reveals that only the minimum number of criteria needed to distinguish one consonant from the rest should be used to name that consonant. These features are called **distinctive features**. In addition to distinctive features, there are a number of characteristics of phonemes that can be guessed on the basis of their distinctive features. These predictable characteristics are called **redundant features**. Redundant features should not be mentioned in the naming of phonemes. The phoneme /ŋ/, for example, is voiced. However, the voicing feature has not been used in its name. This is because it is possible to guess that the phoneme /ŋ/ is voiced due to the fact that it is a nasal consonant. In other words, nasality can predict voicing. Therefore, voicing is said to be a **redundant feature** for nasal consonants, and is not used in their names. A feature which is redundant for one phoneme may be distinctive for another. For example, whereas voicing is considered to be a redundant feature for the phoneme /ŋ/, it is certainly a distinctive feature for the phonemes /g/, and /k/ because voicing is the only difference between them.

Phoneticians usually express redundant features in terms of statements which they call **redundancy rules**. Thus, the fact that voicing can be predicted by nasality is expressed in the following redundancy rule: whenever a phoneme is nasal, it is deemed to be voiced. Like other scientific fields, phonetics draws on a system of notations to simplify its redundancy rules. Take the following example: [nasal] $\rightarrow$ [voiced] (where $\rightarrow$ is read as 'rewrites as'). Redundancy rules will be fully discussed in the following chapters.

4. VOWELS

It was mentioned earlier that most sounds in speech are produced by passing a stream of air from the lungs through one or more resonators belonging to the phonetic apparatus. There are four principle resonators: (1) the pharyngeal cavity, (2) the oral cavity, (3) the labial cavity, and (4) the nasal cavity. Air flows through these resonators.



The presence or absence of obstructions in the course of the airstream modifies the nature of the sound produced. By classifying the different types of obstructions that are possible, articulatory phonetics distinguishes four different major sound classes: consonants, vowels, glides or semivowels, and liquids.

The distinction between consonants and vowels is quite simple. If the air, once out of the glottis, is allowed to pass freely through the oral cavity, the sound is a vowel. If the air, once out of the glottis, is partially or totally obstructed in one or more places in the oral cavity, the sound is a consonant. It should be noted that the line between vowels and consonants cannot be clearly drawn. A continuum exists between the two extremes. In English, there are also intermediate instances: (a) liquids, and (b) glides.

In order for a phoneme to be a vowel, it should meet certain vowel-hood criteria. These criteria include: (a) the degree of openness of the oral cavity also known as the **degree of aperture**, (b) the degree of **tension** or **laxity** of the vocal tract muscles, and (c) amount of **duration** or **length** of articulation.

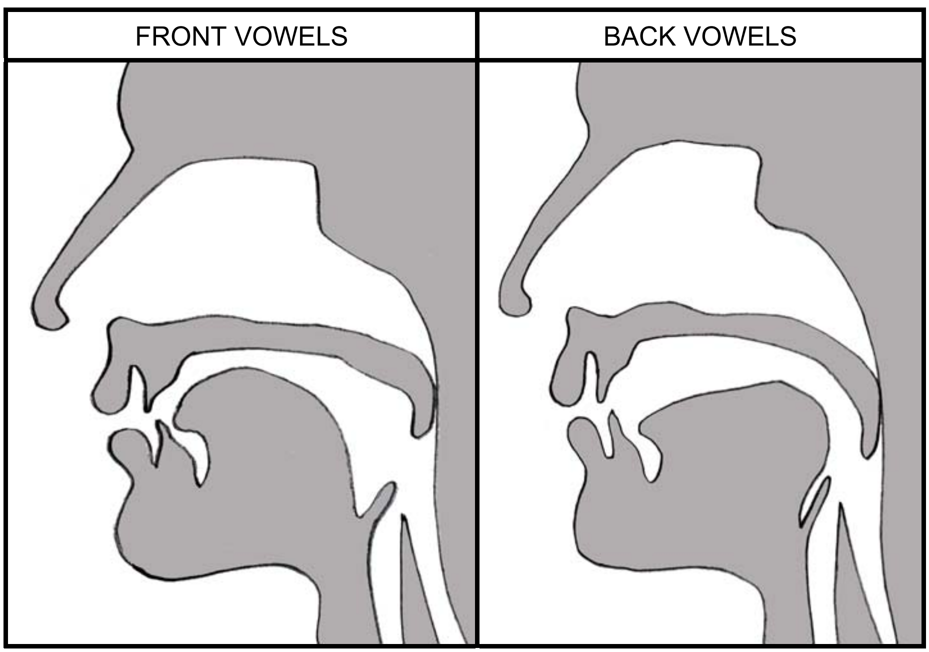
There is a large degree of freedom in the articulation of open vowels. However, this freedom is not endless. On the one hand, no vowel can be more 'open' than the standard open vowels (fourth degree of aperture). On the other hand, a vowel could not be much more 'close' than the standard close vowels (first degree of aperture). Additionally, the close vowels must have a certain minimum duration in order to be perceived as a vowel rather than a consonant.

The chief characteristic of vowels is the freedom with which the airstream, once out of the glottis, passes through the speech organs. The supra-glottal resonators do not cut off or constrict the air. They only cause resonance, that is to say, the reinforcement of certain frequency ranges.

A vowel's timbre (or quality) depends on the following elements: (a) the number of active resonators, (b) the shape of the oral cavity, and (c) the size of the oral cavity. There are **three possible resonators** involved in the articulation of a vowel: (a) the **oral cavity**, (b) the **labial** cavity, and (c) the **nasal** cavity. If the soft palate is raised, the air does not enter the nasal cavity, and passes exclusively through the oral cavity; if the soft palate is lowered, however, air can pass through nose and mouth simultaneously. If the lips are pushed forward and **rounded**, a third, labial resonator is formed; if, on the other hand, the lips are spread sideways or pressed against the teeth, no labial resonator is formed. It is thus the number of resonators at stake in distinguishing **nasal vowels** (nasal resonator active) from **oral vowels** (no nasal resonance), and **rounded vowels** (labial resonator active) from **unrounded vowels** (no labial resonator/no labial resonance).

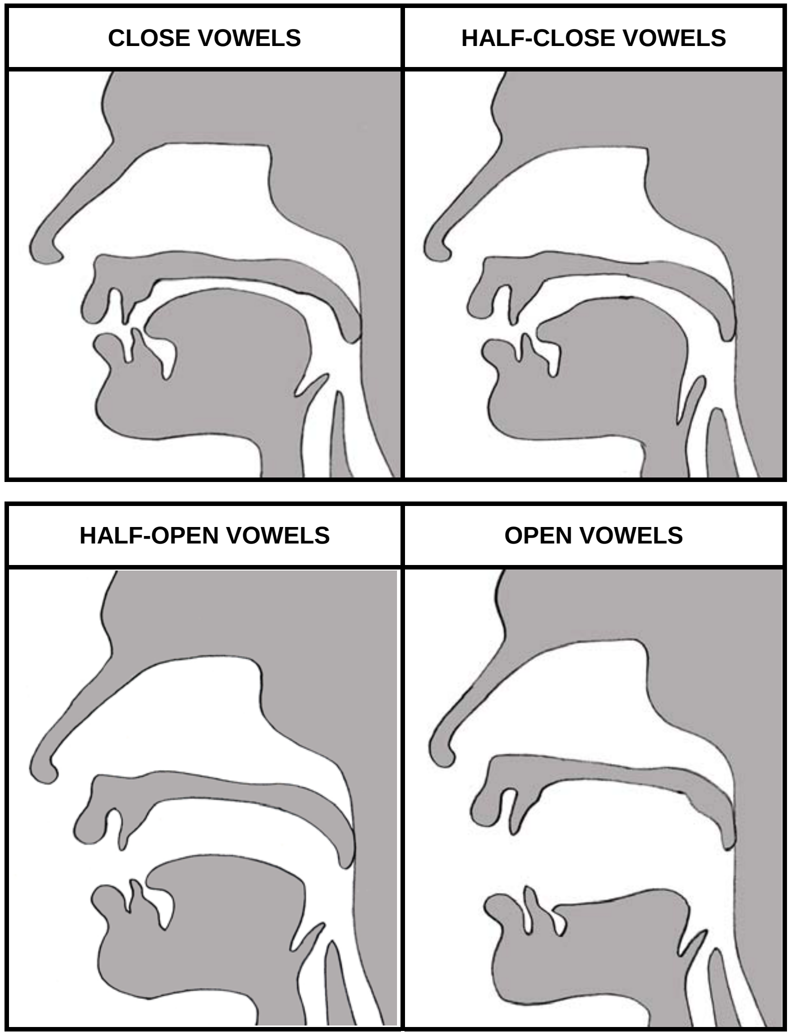
Ordinarily, English vowels do not involve the nasal cavity. They can, however, become nasalized in certain contexts (e.g., when they follow nasal consonants). The **shape of the oral cavity** is determined by the general position of the tongue in the mouth. This divides the vowels into three great classes: (a)

front vowels, (b) **back vowels**, and (c) **central vowels.** In the articulation of front vowels, the tongue body is in the pre-palatal region. In the production of back vowels, the tongue body is in the post-palatal or velar region. Finally, in the articulation of central vowels, the tongue body is in the medio-palatal region.

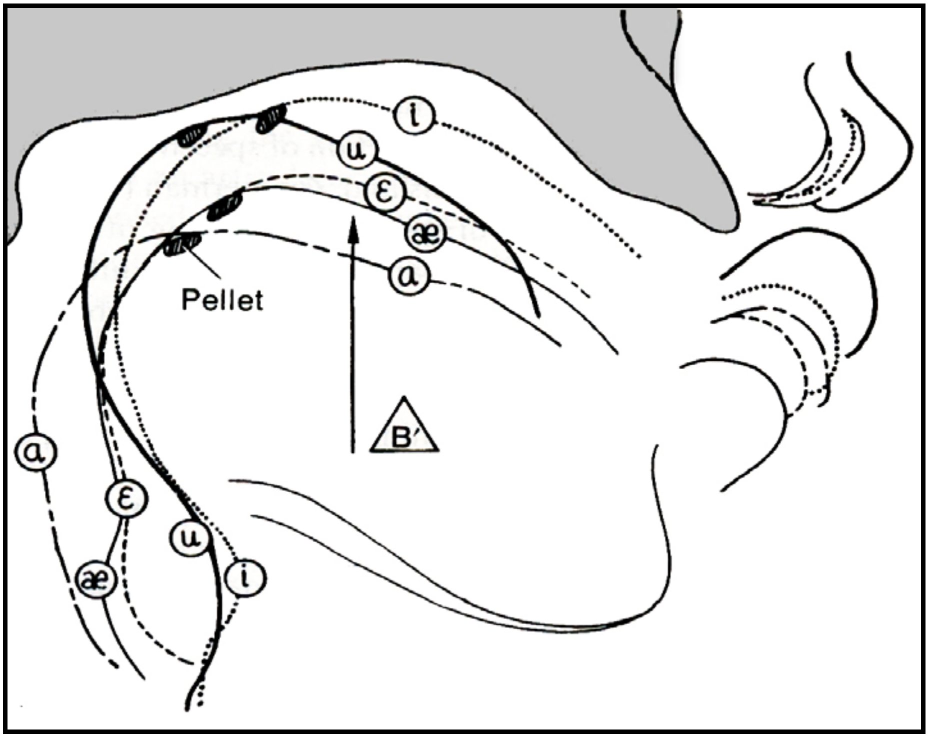


The **size of the oral cavity** is the last factor considered in the articulatory classification of a vowel. The size of the oral cavity depends directly upon the **degree of aperture (of the mouth)**, that is, upon the distance between the palate and the tongue's highest point. Arbitrarily, four degrees of aperture are distinguished, from the most closed (first degree) to the most open (fourth degree). The descriptions of the vocalic articulations (i.e., names given to vowels) should be grouped according to the following principle:

An initial classification is made based on degree of aperture; within each group, the vowels are then divided according to mouth shape, and then as rounded or unrounded.



Vowels differ from consonants in that there is no noticeable obstruction in the vocal tract during their production. Air escapes in a relatively unimpeded way through the mouth and/or nose. If you try saying aaaaa, iiiii, uuuuu, eeeeee, oooooo to yourself you should be able to feel that, although your tongue moves about your mouth, it never actually obstructs the airflow.



Tongue positions for vowel production (Reprinted from Lieberman and Blumstein, 1988, P. 126)

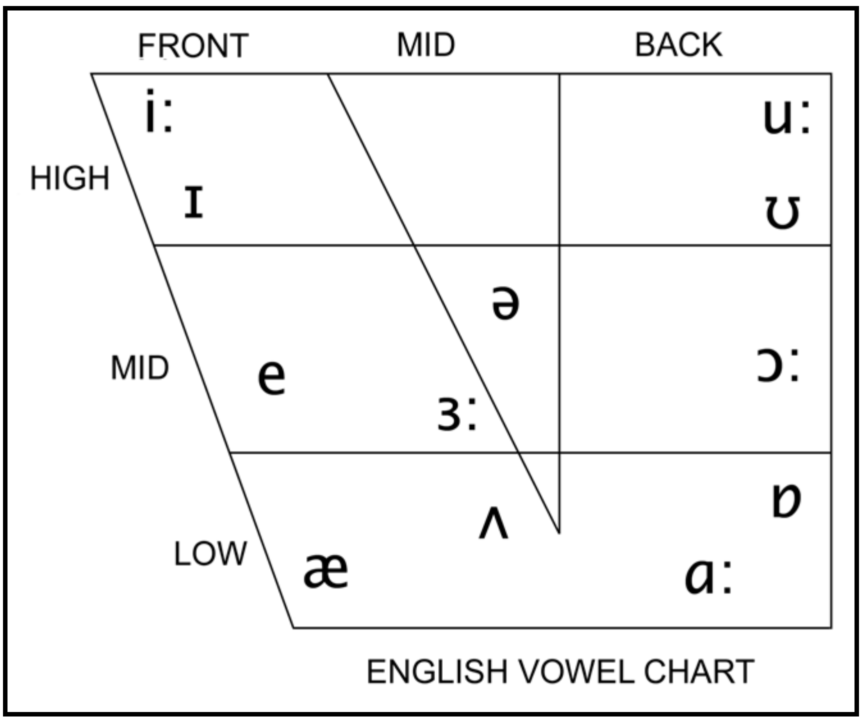
You should also be able to feel that the position of the tongue changes for each of those vowels. Vowels are far more difficult to transcribe than consonants and are also an extremely important area of English phonology as they make up the greatest difference between English varieties. In other words, the differences among varieties of English are mostly due to the speakers' articulation of vowels. In addition, vowels are the stress carrying phonemes in stressed syllables.

The relative position of the tongue in the mouth is the most crucial criterion that determines the differences between vowels. The changes of the tongue position can be very slight and difficult to detect. Therefore, phoneticians produced what they called the vowel chart to describe these tongue positions.

As such, the vowel chart attempts to map the positions of the tongue and jaw in articulating vowels. It is worth mentioning that, as it is so difficult to determine the exact position of the lips, tongue and palate, there is no single agreed upon vowel chart.

4.1. THE VOWEL CHART

The vowel chart is quite a complicated diagram.



All that the vowel chart is trying to do is to represent where the tongue lies in relation to the openness of the mouth when you sound a vowel. So the front closed vowel /i:/ means that your tongue is in a forward position in the mouth which is in a relatively closed position. Try saying it to yourself and then contrast it with the open back sound in the diagram.

As such, the English vowels can be represented in the form of a table or chart that can show the relative position of the tongue in the mouth. This chart can tell the location in the mouth where a vowel is articulated. It can also indicate the degree to which the mouth is open (technically known as **degree of aperture**) when a given vowel is produced. For example, you may have the experience of going to a doctor because of a cold or flu. The doctor wants you to say æææææææ to be able to see the back of your mouth. This shows that the /æ/ sound is articulated with a lowered jaw which adds to the openness of the mouth. Look at the following figure to see how English vowels are articulated.

4.2. TYPOLOGY OF VOWELS

There are two major types of vowels on the basis of their length: **long** vowels, and **short** vowels. Long vowels are usually distinguished from short vowels in that the duration of time that speakers spend in articulating them is somewhat longer than the time spent for the articulation of short vowels. This characteristic of long vowels is known as **length** and is schematically represented by putting a colon (:) after the phoneme which represents the vowel in question. In British English, there are five long vowels:

/i:/	as in	sheep	/ʃi:p/
/ɜ:/	as in	fur	/fɜ:/
/ɔ:/	as in	four	/fɔ:/
/ɑ:/	as in	car	/kɑ:/
/u:/	as in	boot	/bu:t/

As it was mentioned above, short vowels are formed in a much shorter time than long vowels. In other words, short vowels lack the **length** feature. In British English, there are seven short vowels:

/ɪ/	as in	hit	/hɪt/
/e/	as in	hen	/hen/

/æ/	as in	hat	/hæt/
/ə/	as in	ago	/ə'gəʊ/
/ʊ/	as in	book	/bʊk/
/ʌ/	as in	bus	/bʌs/
/ɒ/	as in	hot	/hɒt/

In English, vowels can also glide into one another to form diphthongs (i.e., sequence of two vowels realized as one sound in pronunciation) and even triphthongs (i.e., sequence of three vowels realized as one sound in pronunciation).

Diphthongs are those sounds that consist of a movement or glide from one vowel to another. The first part of a diphthong is always longer and stronger than the second part; as the sound glides into the second part of the diphthong the loudness of the sound decreases. One of the most frequent errors made by language learners is to use pure vowels instead of diphthongs. English diphthongs include the following:

/ʊə/	as in	poor	/pʊə/
/ɔɪ/	as in	boy	/bɔɪ/
/ɪə/	as in	here	/hɪə/
/əʊ/	as in	know	/nəʊ/
/eɪ/	as in	day	/deɪ/
/aʊ/	as in	cow	/kaʊ/
/aɪ/ or /ʌɪ/	as in	my	/maɪ/ or /mʌɪ/
/eə/ or /ɛə/	as in	hair	/heə/ or /hɛə/

Triphthongs are those sounds that consist of a movement or glide from one vowel to another and then onto a third. They are very similar to diphthongs, but have an extra schwa on the end of the diphthongs. There are said to be only five triphthongs, but there are a number of occasions when diphthongs meet other vowels over word edge boundaries. The five clear examples of triphthongs are as follows:

/aʊə/	as in	power	/paʊə/
/aɪə/	as in	liar	/laɪə/
/eɪə/	as in	layer	/leɪə/
/ɔɪə/	as in	loyal	/lɔɪəl/
/əʊə/	as in	mower	/məʊə/

4.3. WEAK FORMS OF VOWELS

Vowels in English can appear in stressed and unstressed syllables. When a vowel appears in a stressed syllable, it is pronounced with its **strong** or **full form**. However, when vowels appear in unstressed syllables, they are normally neutralized/centralized and adopt a rather weak identity known as the **weak form**. The weak form of all vowels is /ə/ except for /iː/ and /ɪ/ of which the weak form is /ɪ/. This can be schematically represented as the following:

Strong	/iː/	/ɪ/	/uː/	/ʊ/	/ɔː/	/ɒ/	/ɑː/	/ʌ/	/ɜː/	/æ/	/e/
Weak	/ɪ/	/ɪ/	/ə/	/ə/	/ə/	/ə/	/ə/	/ə/	/ə/	/ə/	/ə/

In morphology, weak forms are those words that are pronounced in an unstressed manner. Many of the most common words in English can come in either a strong or a weak form. The most crucial factors in determining whether a word should be pronounced with a weak vowel or a strong one are (a) the type of the word (i.e., function words versus content words), and (b) the number of syllables that compose the word (i.e., whether the word is mono-syllabic, bi-syllabic, or poly-syllabic). A function word is a word that has a grammatical rather than semantic function in the language. In other words, the words which are used to express grammatical relationships among words in a sentence are called function words. Functions words include: conjunctions, articles, pronouns, prepositions and some auxiliary and modal verbs. Content words, on the other hand, are rather encyclopedic than linguistic. That is, those words that have independent encyclopedic meanings of their own, and that can be used independently, are called content words. Content words include: main verbs, adjectives, adverbs, and nouns.

ARTICLES	THE	/ðə/ before consonants (the man /ðəmæn/)
		/ði/ before vowels (the orange /ði'ɒrɪnʤ/)
	A, AN	/ə/ before consonants (a bike /əbaɪk/)
		/ən/ before vowels (an apple /ənæpl/)
CONJUNCTIONS	AND	/ən(d)/ (come and see /kɒmənsi:/)
		Fish and chips /fiʃənʧɪps/
	BUT	/bət/ (cheap but reliable /ʧI:pbətrɪlaɪəb(ə)l/)
		(Last but one /læstbətwʌn/)
	THAN	/ðən/ (better than expected /betəðənɪk'spektɪd/)
		(More than one /mɔ:ðənwʌn/)
PRONOUNS	HIS	/hɪz/, /həz/, /əz/ before a noun
		Take his coat /teɪkəzkəʊt/
	HER	/ə/ before consonants (take her home /teɪkəhəʊm/)
		/ə(r)/ Before vowels (take her out /teɪkə(r)aʊt/)
PREPOSITIONS		un-emphatically or in non-sentence-final positions
	FROM	/frəm/ (he is from london /hi:ɪzfrəmlɒnd(ə)n/)
	AT	/ət/ (at the table /ətðəteɪbl/)
	OF	/əv/ (the leg of the table /ðəlegəvðəteɪbl/)
	ON	/ən/ (and on thy cheeks /ənənðaɪʧi:ks/)
	etc.	

Weak forms are all function words. Generally the strong forms of these words are used (a) when they are being directly quoted, (b) when they are being contrasted or (c) if they appear at the end of a sentence. The pronunciation of a weak form can be so different from the strong form that it is barely recognizable as being the same word. If said in isolation, it would be all but unintelligible. Usually, it is the context that makes it understandable.

It is possible to use only strong forms of vowels in English, and some non-native speakers do exactly this. However, it sounds very unnatural to a native speaker and it will also mean that a person who only uses the strong forms of English vowels will have trouble understanding native speakers of English who use the weak form all of the time. The most common weak form examples are introduced as follows:

In traditional phonetics, vowels were named on the basis of such criteria as the relative position of the tongue in the mouth, lip rounding, and length.

VOWEL	NAME	VOWEL	NAME
/iː/	High front vowel	/ʌ/	Low mid vowel
/ɪ/	High-mid front vowel	/uː/	High back vowel
/e/	Mid front vowel	/ʊ/	High-mid back vowel
/æ/	Low front vowel	/ɔː/	Mid back vowel
/ə/	Mid mid vowel	/ɒ/	Mid-low back vowel
/ɜː/	Mid-low mid vowel	/ɑː/	Low back vowel

As such, the vowel /iː/ would, for example, be considered a **high front vowel**. Other vowels also received their names in this way:

vertical position + horizontal position + vowel.

5. SEMIVOWELS OR GLIDES

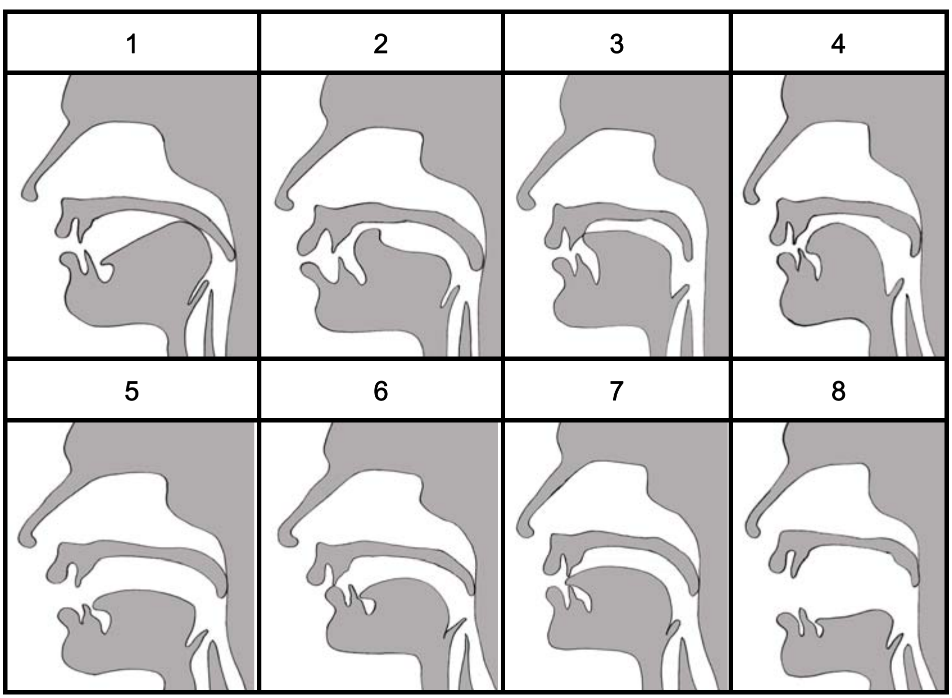
Semivowels or **glides** are those sounds which, though articulated quite variously, in one way or another fail to meet the above criteria for full vowel-hood. They also fail to meet the **obstruction** or **blockage** criterion for full consonant-hood. They are, therefore, considered to be an intermediate state between vowels and consonants. They involve the same restriction of the speech canal as fricative consonants, but the speech organs are substantially less tense during the articulation of semivowels. Rather than friction, a resonant sound is produced at the place of articulation. The English language includes three semivowels: /h/, /w/, and /j/.

6. LIQUIDS

Liquids meet not only the vowel-hood criteria but also the consonant-hood criterion. They are, like semivowels, considered to be an intermediate state between vowels and consonants. Liquids are generally considered to be a special case. During their production, the oral cavity is constricted while the tongue sinks down in such a way as to channel the air out of the mouth as if a vowel is being articulated. English includes two liquid sounds: /l/, and /r/.

EXERCISE

1. Look at the following illustrations:



Each of these illustrations represents the shape of the human vocal organ for the articulation of a sound or a class of sounds. Complete the following table by indicating the sound/class of sounds represented by each illustration.

ILLUSTRATION	YOUR ANSWER
1	
2	
3	
e.g., 4	Front vowels
5	
6	
7	
8	

2. Write the symbol that corresponds to each of the following phonetic descriptions, then give an English word that contains this sound.

	Example: voiced alveolar stop	[d]	dough
A	voiceless bilabial unaspirated stop	[ ]	
B	low front vowel	[ ]	
C	lateral liquid	[ ]	
D	velar nasal	[ ]	
E	voiced interdental fricative	[ ]	
F	voiceless affricate	[ ]	
G	palatal glide	[ ]	
H	mid lax front vowel	[ ]	
I	high back tense vowel	[ ]	
J	voiceless aspirated alveolar stop	[ ]	

3. For each group of sounds listed, state the phonetic feature(s) they all share.

Example: [p] [b] [m] *Features:* bilabial, stop, consonant

	GROUP OF SOUNDS	SHARED FEATURES
A	[g] [p] [t] [d] [k] [b]	
B	[u:] [ʊ] [ɔ:] [ɒ]	
C	[i:] [ɪ] [ɜ:] [e] [æ]	
D	[t] [s] [ʃ] [p] [k] [ʧ] [f] [h] [θ]	
E	[v] [z] [ʒ] [ʤ] [n] [g] [d] [b] [l] [r] [w] [j]	
F	[θ] [ð]	
G	[t] [d] [n] [s] [z] [r] [l]	
H	[ʃ] [ʒ] [ʧ] [ʤ]	
I	[k] [g] [ŋ]	
J	[j] [w] [h]	
K	[t] [d] [s] [ʃ] [n] [ʧ] [ʤ]	

4. What phonetic property or feature distinguishes the sets of sounds in column A from those in column B?

	COLUMN A	COLUMN B	YOUR ANSWER
A	[i:] [ɪ]	[u:] [ʊ]	
B	[p] [t] [k] [s] [f]	[b] [d] [g] [z] [v]	
C	[p] [b] [m]	[t] [d] [n] [k] [g] [ŋ]	
D	[i:] [ɪ] [u:] [ʊ]	[e] [ɜ:] [ɔ:] [ɒ] [æ] [ʌ]	
E	[f] [v] [s] [z] [ʃ] [ʒ]	[ʧ] [ʤ]	
F	[i:] [ɪ] [e] [ə] [ɜ:] [æ]	[u:] [ʊ] [ɔ:] [ɒ] [a:]	

5. Which of the following sound pairs have the same manner of articulation? Mark Ⓨ if the sound pairs have the same manner of articulation, and Ⓝ if they have different manners of articulation.

A	[h] [ʔ]	Ⓨ	Ⓝ	F	[f] [ʃ]	Ⓨ	Ⓝ
B	[r] [w]	Ⓨ	Ⓝ	G	[k] [θ]	Ⓨ	Ⓝ
C	[m] [ŋ]	Ⓨ	Ⓝ	H	[s] [g]	Ⓨ	Ⓝ
D	[ð] [v]	Ⓨ	Ⓝ	I	[j] [w]	Ⓨ	Ⓝ
E	[t] [r]	Ⓨ	Ⓝ	J	[j] [ʤ]	Ⓨ	Ⓝ

6. Identify the manner of articulation for each of the following sounds.

	YOUR ANSWER		YOUR ANSWER
[h]	( )	[f]	( )
[ʔ]	( )	[ʃ]	( )
[w]	( )	[k]	( )
[ŋ]	( )	[ʤ]	( )
[ð]	( )	[ʧ]	( )
[θ]	( )	[ʒ]	( )
[v]	( )	[j]	( )
[r]	( )	[s]	( )
[t]	( )	[z]	( )
[m]	( )	[n]	( )

CHAPTER FOUR
SYSTEMATIC ARTICULATORY PHONETICS

1. INTRODUCTION

Despite its potentials, traditional phonetics was inadequate in a number of cases. For example, the phoneme /p/, as it was mentioned earlier, has two allophonic realizations [p], and [p^h]. Vowels, too, can mix to form diphthongs or triphthongs. However, the naming convention used in traditional phonetics could not easily capture these nuances. Therefore, phoneticians sought to develop a new comprehensive and exhaustive system. This new perspective on phonetics was called **systematic** or **modern phonetics**.

2. WHAT IS SYSTEMATIC/MODERN PHONETICS

Modern phonetics is based on the idea that each phonetic segment or phoneme is composed of a bundle of features. These features provide a good description of the phoneme. As such, phonemes are distinguished from one another on the basis of their phonetic features. The presence of a feature in the bundle of features that identify a phoneme is marked by the plus [+] symbol and the absence of the feature by the minus [-] symbol. Therefore, the phonetic features used in modern phonetics are called **binary features** due to their dual [±] values. For example, the allophone [p^h] is said to be [+aspiration] whereas its allophone [p] is identified as [-aspiration]. The symbol [α] is used to show both [±] values. The symbol [-α] is used to show the [∓] value.

Each phoneme is, therefore, adequately represented as a matrix of binary phonetic features. All sounds marked by the same value [+ or -] for a feature belong to a natural class named after that feature. For instance, all sounds that are marked [+voice] belong in the voiced class and all sounds that are marked

[-voice] belong in the voiceless class. Unlike traditional phonetics, phonemes are not given individual names in modern or systematic phonetics. Rather, the bundle of distinctive features that distinguish one phoneme from the rest is used as a description (or name) for that phoneme. For example, the nasal consonants /ŋ/, /n/, and /m/ are described as:

$$\begin{array}{ccc} \text{/ŋ/} & \text{/n/} & \text{/m/} \\ \begin{bmatrix} \text{-vocalic} \\ \text{+consonantal} \\ \text{+nasal} \\ \text{-anterior} \end{bmatrix} & \begin{bmatrix} \text{-vocalic} \\ \text{+consonantal} \\ \text{+nasal} \\ \text{+coronal} \end{bmatrix} & \begin{bmatrix} \text{-vocalic} \\ \text{+consonantal} \\ \text{-coronal} \\ \text{+anterior} \end{bmatrix} \end{array}$$

Each of these features describes one of the discrete activities in (or parts of) the vocal tract that are crucial to the articulation of the phoneme. It should be noted that, like in traditional phonetics, in modern phonetics, too, each phoneme may be composed of more features than are actually in its description. However, many features are predictable on the basis of other features. In our example here, the phoneme /ŋ/ is a [+voice] phoneme. However, we did not use this feature in the bundle of features that describe the phoneme because the feature [+nasal] can predict the feature [+voice]. Therefore, only the minimum inevitable number of features that differentiate between a given phoneme and the rest of sound segments should be used in the description of that phoneme. These are, thus, called the **distinctive features** of that phoneme. The predictable features of a phoneme are called the **redundant features** of that phoneme.

Phonetic features not only differentiate between phonemes but also between classes of phonemes. In fact the notion of natural classes of sounds derives from this ability of binary features. For instance, /m/, /n/, and /ŋ/ are all [+nasal] while other phonemes are [-nasal]. So, these three phonemes compose a natural class of sounds in English. A **natural class** is defined as: A class of phonemes in which the number of features that must be specified to define that class is smaller than the number of features required to distinguish any member of that class.

In our example, only the feature [+nasal] is required to distinguish the nasal class whereas, as you have noticed, each of the nasal phonemes /ŋ/, /n/, and /m/ needs more than three distinctive features to be identified. Nasals do not comprise the only natural class of sounds. There are, in fact, many different natural classes of sounds in the English language. The major natural classes of sounds of English are consonants, vowels, liquids, and glides. These classes can easily be distinguished from each other on the basis of values (±) they assign to the phonetic features [consonantal] and [vocalic]. Compare the following:

$$\begin{array}{cc} \text{VOWELS} & \text{CONSONANTS} \\ \begin{bmatrix} \text{+vocalic} \\ \text{-consonantal} \end{bmatrix} & \begin{bmatrix} \text{-vocalic} \\ \text{+consonantal} \end{bmatrix} \\ \text{LIQUIDS} & \text{GLIDES} \\ \begin{bmatrix} \text{+vocalic} \\ \text{+consonantal} \end{bmatrix} & \begin{bmatrix} \text{-vocalic} \\ \text{-consonantal} \end{bmatrix} \end{array}$$

Before we embark on any discussion of phonetic features, we had better identify English phonemes in terms of their distinctive features.

1) English liquids are identified by the following features:

$$\begin{array}{cc} \text{/r/} & \text{/l/} \\ \begin{bmatrix} \text{+consonantal} \\ \text{+vocalic} \\ \text{-lateral} \end{bmatrix} & \begin{bmatrix} \text{+consonantal} \\ \text{+vocalic} \\ \text{+lateral} \end{bmatrix} \end{array}$$

2) English glides are identified by the following features:

$$\underset{}{\overset{/j/}{\begin{bmatrix}\text{-vocalic}\\ \text{-consonantal}\\ \text{+voice}\\ \text{-round}\end{bmatrix}}}\quad \overset{/w/}{\begin{bmatrix}\text{-vocalic}\\ \text{-consonantal}\\ \text{+voice}\\ \text{+round}\end{bmatrix}}\quad \overset{/h/}{\begin{bmatrix}\text{-vocalic}\\ \text{-consonantal}\\ \text{-voice}\\ \\ \end{bmatrix}}$$

3) English voiceless consonants are identified by the following features:

$$\overset{/p/}{\begin{bmatrix}\text{-vocalic}\\ \text{+consonantal}\\ \text{-voice}\\ \text{+abrupt release}\\ \text{+anterior}\\ \text{-coronal}\end{bmatrix}}\quad \overset{/f/}{\begin{bmatrix}\text{-vocalic}\\ \text{+consonantal}\\ \text{-voice}\\ \text{+strident}\\ \text{-coronal}\\ \\ \end{bmatrix}}\quad \overset{/\theta/}{\begin{bmatrix}\text{-vocalic}\\ \text{+consonantal}\\ \text{-voice}\\ \text{+continuant}\\ \text{-strident}\\ \text{+coronal}\end{bmatrix}}$$

$$\overset{/t/}{\begin{bmatrix}\text{+consonantal}\\ \text{-vocalic}\\ \text{-voice}\\ \text{+abrupt release}\\ \text{+coronal}\\ \\ \\ \end{bmatrix}}\quad \overset{/s/}{\begin{bmatrix}\text{+consonantal}\\ \text{-vocalic}\\ \text{-voice}\\ \text{+strident}\\ \text{+anterior}\\ \text{+coronal}\\ \\ \end{bmatrix}}\quad \overset{/\text{ʃ}/}{\begin{bmatrix}\text{+consonantal}\\ \text{-vocalic}\\ \text{-voice}\\ \text{+strident}\\ \text{+continuant}\\ \text{+coronal}\\ \text{-anterior}\end{bmatrix}}$$

$$\overset{/\text{ʧ}/}{\begin{bmatrix}\text{+consonantal}\\ \text{-vocalic}\\ \text{-voice}\\ \text{-abrupt release}\\ \text{-continuant}\end{bmatrix}}\quad \overset{/k/}{\begin{bmatrix}\text{+consonantal}\\ \text{-vocalic}\\ \text{-voice}\\ \text{+abrupt release}\\ \text{-anterior}\end{bmatrix}}$$

4) English voiced consonants are identified by the following features:

/b/

$$\begin{bmatrix}\text{-vocalic}\\ \text{+consonantal}\\ \text{+voice}\\ \text{+abrupt release}\\ \text{+anterior}\\ \text{-coronal}\end{bmatrix}$$

/v/

$$\begin{bmatrix}\text{-vocalic}\\ \text{+consonantal}\\ \text{+voice}\\ \text{+strident}\\ \text{-coronal}\end{bmatrix}$$

/ð/

$$\begin{bmatrix}\text{-vocalic}\\ \text{+consonantal}\\ \text{+voice}\\ \text{+continuant}\\ \text{-strident}\\ \text{+coronal}\end{bmatrix}$$

/d/

$$\begin{bmatrix}\text{+consonantal}\\ \text{-vocalic}\\ \text{+voice}\\ \text{-nasal}\\ \text{+abrupt release}\\ \text{+coronal}\end{bmatrix}$$

/z/

$$\begin{bmatrix}\text{+consonantal}\\ \text{-vocalic}\\ \text{+voice}\\ \text{+strident}\\ \text{+anterior}\\ \text{+coronal}\end{bmatrix}$$

/ʒ/

$$\begin{bmatrix}\text{+consonantal}\\ \text{-vocalic}\\ \text{+voice}\\ \text{+strident}\\ \text{+continuant}\\ \text{+coronal}\\ \text{-anterior}\end{bmatrix}$$

/ʤ/

$$\begin{bmatrix}\text{+consonantal}\\ \text{-vocalic}\\ \text{+voice}\\ \text{-abrupt release}\\ \text{-continuant}\end{bmatrix}$$

/g/

$$\begin{bmatrix}\text{+consonantal}\\ \text{-vocalic}\\ \text{+voice}\\ \text{-nasal}\\ \text{+abrupt release}\\ \text{-anterior}\end{bmatrix}$$

5) English nasals are identified by the following features:

/ŋ/

$$\begin{bmatrix}\text{-vocalic}\\ \text{+consonantal}\\ \text{+nasal}\\ \text{-anterior}\end{bmatrix}$$

/n/

$$\begin{bmatrix}\text{-vocalic}\\ \text{+consonantal}\\ \text{+nasal}\\ \text{+coronal}\end{bmatrix}$$

/m/

$$\begin{bmatrix}\text{-vocalic}\\ \text{+consonantal}\\ \text{-coronal}\\ \text{+anterior}\end{bmatrix}$$

Like other English phonemes, English vowels, too, have been identified in terms of distinctive features. However, the set of distinctive features used to identify vowels is somewhat different from the set of features used for the identification of other phonemes. English vowels are identified in terms of the following features:

/iː/	/ɪ/	/e/
-consonantal	-consonantal	-consonantal
+vocalic	+vocalic	+vocalic
+high	+high	-high
-back	-back	-low
+tense	-tense	-back
		+tense

/ɜː/	/æ/	/ə/
-consonantal	-consonantal	-consonantal
+vocalic	+vocalic	+vocalic
-high	+low	-low
-low	-back	-high
-back		-back
-tense		-tense

/ʌ/	/ɑː/	/ɒ/
-consonantal	-consonantal	-consonantal
+vocalic	+vocalic	+vocalic
-high	+low	-high
-low	+back	-low
-back		+back
-rounded		-tense
		+rounded

/ɔː/ /uː/ /ʊ/

$$\begin{bmatrix} \text{-consonantal} \\ \text{+vocalic} \\ \text{-high} \\ \text{-low} \\ \text{+back} \\ \text{+tense} \\ \text{+rounded} \end{bmatrix} \quad \begin{bmatrix} \text{-consonantal} \\ \text{+vocalic} \\ \text{+high} \\ \text{+back} \\ \text{+rounded} \\ \text{+tense} \\ \end{bmatrix} \quad \begin{bmatrix} \text{-consonantal} \\ \text{+vocalic} \\ \text{+high} \\ \text{+back} \\ \text{+rounded} \\ \text{-tense} \\ \end{bmatrix}$$

The features used to distinguish English vowels can be summed up in the form of a matrix (table 4.1. below).

	/ iː/	/ɪ/	/e/	/ɜː/	/æ/	/ʌ/	/uː/	/ʊ/	/ɔː/	/ɒ/	/ɑː/
high	+	+	-	-	-	-	+	+	-	-	-
Mid	-	-	+	+	-	+	-	-	+	+	-
low	-	-	-	-	+	-	-	-	-	-	+
back	-	-	-	-	-	-	+	+	+	+	+
Central	-	-	-	-	-	+	-	-	-	-	-
rounded	-	-	-	-	-	-	+	+	+	+	-
tense	+	-	+	-	-	-	+	-	+	-	-

Table 4.1.: Phonetic features of English vowels (Reprinted from Fromkin, Rodman, and Hyams (2003) p. 299—with some modifications).

Notice that all English vowels are [+syllabic] and [+voiced]. Also notice that the identification of diphthongs and triphthongs requires more than one column of features. The diphthongs would be represented by a two-column feature matrix of the vowel followed by the glide; the triphthongs would be represented by a three-column feature matrix of the vowel followed by the glide followed by the vowel.

		PHONEME	Vocalic	Consonantal	Voice	Nasal	Abrupt release	Continuant	Aspiration	Anterior	Coronal	Strident	Lateral	Sonorant	Syllabic
STOPS	**PLOSIVES**	/p/	-	+	-	-	+	-	-	+	-	-	-	-	-
		/b/	-	+	+	-	+	-	-	+	-	-	-	-	-
		/t/	-	+	-	-	+	-	-	+	+	-	-	-	-
		/d/	-	+	+	-	+	-	-	+	+	-	-	-	-
		/k/	-	+	-	-	+	-	-	-	-	-	-	-	-
		/g/	-	+	+	-	+	-	-	-	-	-	-	-	-
	NASALS	/m/	-	+	+	+	+	-	-	+	-	-	-	+	±
		/n/	-	+	+	+	+	-	-	+	+	-	-	+	±
		/ŋ/	-	+	+	+	+	-	-	-	-	-	-	+	±
	AFFRICATES	/ʧ/	-	+	-	-	-	-	-	-	+	+	-	-	-
		/ʤ/	-	+	+	-	-	-	-	-	+	+	-	-	-
FRICATIVES		/f/	-	+	-	-	-	+	-	+	-	+	-	-	-
		/v/	-	+	+	-	-	+	-	+	-	+	-	-	-
		/θ/	-	+	-	-	-	+	-	+	+	-	-	-	-
		/ð/	-	+	+	-	-	+	-	+	+	-	-	-	-
		/s/	-	+	-	-	-	+	-	+	+	+	-	-	-
		/z/	-	+	+	-	-	+	-	+	+	+	-	-	-
		/ʃ/	-	+	-	-	-	+	-	-	+	+	-	-	-
		/ʒ/	-	+	+	-	-	+	-	-	+	+	-	-	-

TABLE 4.2.: Features of English consonants.

The phonetic features that can be used for the identification of English consonants have been summarized in table 4.2. above.

The phonetic features of English glides and liquids are identified in table 4.3. below.

	PHONEME	Vocalic	Consonantal	Voice	Nasal	Abrupt release	Continuant	Aspiration	Anterior	Coronal	Strident	Lateral	Sonorant	Syllabic
LIQUIDS	/l/	+	+	+	-	-	+	-	+	-	-	+	+	±
	/r/	+	+	+	-	-	+	-	+	-	-	-	+	±
GLIDES	/h/	-	-	-	-	-	+	-	-	-	-	-	+	-
	/w/	-	-	+	-	-	+	-	-	-	-	-	+	-
	/j/	-	-	+	-	-	+	-	-	-	-	-	+	-

TABLE 4.3.: Features of English liquids and glides.

3. AIRSTREAM MECHANISMS

It was pointed out before that most sounds in speech are produced by passing a stream of air from the lungs through one or more resonators belonging to the phonetic apparatus. In a small number of languages, the airstream does not originate in the lungs, but rather from outside. This **velaric ingressive airstream mechanism** produces implosive and click sounds through inhalation. A speech sound can also be generated from a difference in pressure of the air inside and outside a resonator. In the case of pharyngeal cavity, this pressure difference can be created without using the lungs at all whereby producing ejective sounds. This is called the **glottalic egressive airstream mechanism**. The main airstream mechanism in most of the world's languages, however, is the **pulmonic egressive airstream mechanism** which produces almost all the phonemes of the world's most important languages. In this case, the flow of air originates in the lungs. As such, it is called pulmonic airstream (or simply pulmonic air). It then moves outward through the resonators; hence, the name egressive. When articulators (i.e., parts of the vocal tract that modify the airstream) influence the flow of pulmonic air, sounds are produced. The nature of the modifications that produce one phoneme differs from those of any other phoneme. As it was mentioned earlier, modern or systematic phonetics uses binary phonetic features to describe the nature of these modifications.

4. PHONETIC FEATURES

You have already noticed that sound segments, in systematic phonetics, are said to be composed of bundles of features. In fact, the identification of sound segments on preceding features is based on the binary features of those sounds that we have put within square brackets. You have also noted that the set of features used to identify vowels differs from the set used to identify other types of phonemes. Now, it is important to know what we mean by these features.

4.1. VOCALIC VERSUS CONSONANTAL

There are a good number of phonetic features. However, two of them are considered to be the most crucial ones: (a) [±consonantal], and (b) [±vocalic]. Their importance is in that they classify all sounds (in the case of English and Farsi at least) into the four major sound classes: (a) consonants, (b) vowels, (c) liquids, and (d) glides:

VOWELS

$$\begin{bmatrix} \text{+vocalic} \\ \text{-consonantal} \end{bmatrix}$$

CONSONANTS

$$\begin{bmatrix} \text{-vocalic} \\ \text{+consonantal} \end{bmatrix}$$

LIQUIDS

$$\begin{bmatrix} \text{+vocalic} \\ \text{+consonantal} \end{bmatrix}$$

GLIDES

$$\begin{bmatrix} \text{-vocalic} \\ \text{-consonantal} \end{bmatrix}$$

The distinction between [+consonantal] and [+vocalic] is quite simple. If the air, once out of the glottis, is allowed to pass freely through the oral cavity, the sound is a **[+vocalic]** phoneme. If the air, once out of the glottis, is partially or totally obstructed in one or more places in the oral cavity, the sound is a

[+consonantal] phoneme. It should be noted that the line between [+vocalic] and [+consonantal] phonemes cannot be clearly drawn. A continuum exists between the two extremes. In English, there are also intermediate instances: (a) liquids, and (b) glides. Liquids share features of both [+consonantal] and [+vocalic] sounds. Glides lack features of both [+consonantal] and [+vocalic] sounds.

In order for a phoneme to be [+vocalic], it should meet certain vowel-hood criteria. These criteria include: (a) the degree of openness of the oral cavity also known as the **degree of aperture**, (b) the degree of **tension** or **laxity** of the vocal tract muscles, and (c) amount of **duration** or **length** of articulation. There is a large degree of freedom in the articulation of open vowels. However, this freedom is not endless. On the one hand, no vowel can be more open than the standard open vowels (fourth degree of aperture). On the other hand, a vowel could not be much more "close" than the standard close vowels (first degree of aperture). Additionally, the close vowels must have a certain minimum duration in order to be perceived as [+vocalic] rather than [+consonantal].

4.2. VOICE

A sound is described as [-voice] when the vocal cords do not vibrate during its articulation. If the vocal cords do vibrate, the sound is considered [+voice]. It was mentioned earlier that the vocal cords are folds of muscle (membranes) located at the level of the glottis (in fact, the glottis is nothing other than the space between the vocal cords). The vocal cords vibrate when they are closed to obstruct the airflow through the glottis. They vibrate under the pressure of the air being forced through them by the lungs. The [±voice] opposition is mainly useful for the classification of consonants. Because all vowels are [+voice], it is considered a redundant feature for vowels that can be captured by a redundancy rule. In English, the phonemes /ʧ/, /s/, /p/, /k/, /f/, /ʃ/, /t/, /θ/, and /h/ are [-voice]. Other phonemes are [+voice].

4.3. NASAL AND NASALIZATION

The top of the pharynx is like a crossroads. The airstream can exit the pharynx in either of two ways, depending on the position of the soft palate. If the soft palate is lowered, a portion of the air will pass through the nasal cavity (the remainder finding its way through the oral cavity). If the soft palate is raised, access to the nasal cavity is cut off, and the air can only pass through the oral cavity. The sounds produced via the first method are [+nasal]. Those produced the other way are [-nasal].

Nasality primarily concerns consonants. However, vowels can be nasalized in specific contexts (i.e., when they are followed by nasal consonants). Therefore, the term **nasality** is used for consonants and **nasalization** for vowels. Nasalization is fully explained in the following chapters.

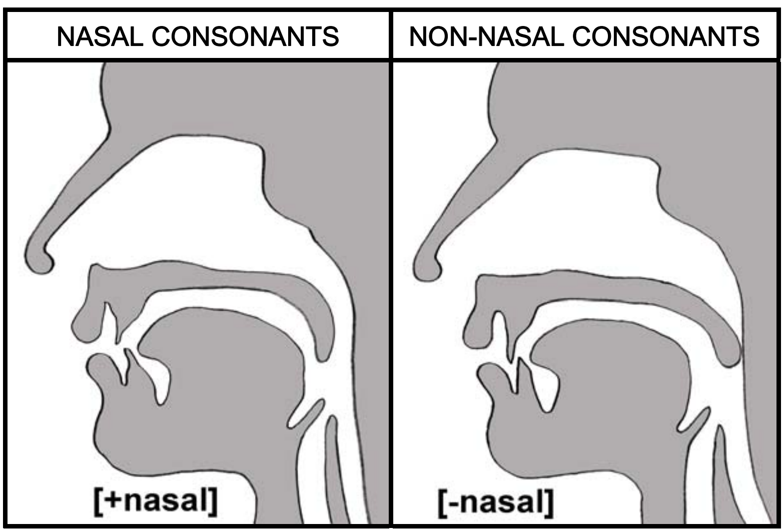
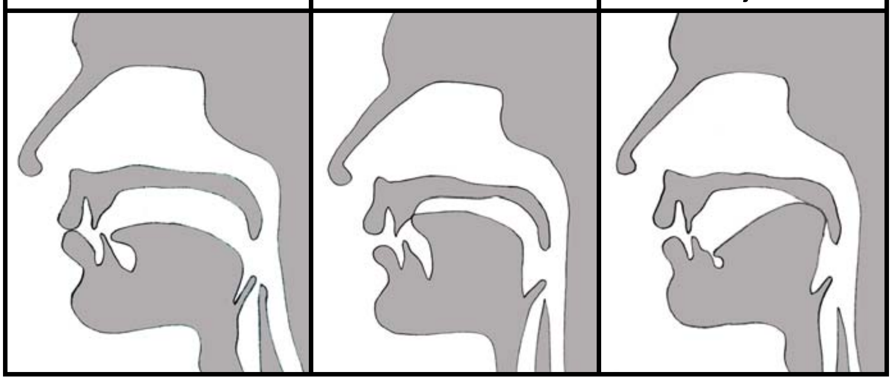


Figure 4.1: Nasal versus non-nasal phonemes

[+nasal] phonemes of the vast majority of the world's languages are [+voice]. So, the value [±] for the voicing feature is predictable through a redundancy rule when a phoneme is [+nasal]. Therefore, [+voice] is not included in the bundle of distinctive features of a [+nasal] sound segment.

During the production of these nasal stops, the soft palate is lowered to a greater or lesser extent, allowing a portion of the airstream to pass through the nasal cavity. Occlusion (i.e., air blockage) occurs in the mouth only. Many linguists, therefore, consider [+nasal] phonemes to be stop sounds. However, [+nasal] phonemes can endure through the nasal cavity. In other words, the nasal resonance is continuous. Although a few linguists rank the nasals among the [+continuant] phonemes, the majority of them categorize [+nasal] sounds as [-continuant].

/m/	/n/	/ŋ/



The English language has three [+nasal] sounds in its inventory of phonemes: /m/, /n/, and /ŋ/. Farsi only includes the first two: /m/, and /n/. The phoneme /m/ is a bilabial nasal. The mouth is configured just as for the articulation of the corresponding [-nasal] bilabial stops /p/, and /b/. The lips are pressed tightly together, but the velum is somewhat lowered to let pulmonary air escape through the nasal cavity. The vocal cords are kept close to each other to create audible vibration. The phoneme /n/ is a dental or alveolar nasal. The mouth is configured just as for the corresponding dental or alveolar stops /d/ and /t/. The tongue makes contact either with the front teeth, or with the alveolar ridge directly above them. The phoneme /ŋ/ is a velar nasal. The configuration of the mouth is very close to that of the corresponding velar stops /g/ and /k/. The tongue tip rests against the lower teeth while the back of the tongue makes contact with the soft palate. However, as the soft palate is lowered (to allow air to flow through the nasal cavity), the tongue's back is detached from the soft palate to let a burst of air flow through the mouth.

4.4. ABRUPT RELEASE

Pulmonic egressive air is modified in the oral cavity to produce the phonemes of certain languages. In fact, there are three major types of air modifications: (a) complete blockage followed by sudden release, (b) complete blockage followed by gradual release, and (c) constriction or narrowing of the passage of air. The first results in the production of stops and plosives, the second in the production of affricates, and the last in the production of fricatives.

In systematic phonetics, the feature [±abrupt release] indicates whether the first of these types of air modifications is present in the articulation of a phoneme or not. In other words, when a phoneme is articulated by a complete blockage followed by a sudden release of the air in the oral cavity, that phoneme is [+abrupt release]. Otherwise, the phoneme is [-abrupt release]. Therefore, all stop consonants (including nasals /m/, /n/, /ŋ/, and plosives /p/, /b/, /t/, /d/, /k/, /g/) have the feature [+abrupt release].

Although nasal consonants are [+abrupt release], this feature is considered a redundant feature for them; therefore, it is not part of the inventory of features that are used to identify nasal consonants. In fact, the feature [+nasal] results in a [+abrupt release] phoneme. This is, like it was mentioned earlier, usually captured in a redundancy rule. Remember from preceding sections and chapters that a redundancy rule is a schematic representation (or a specific notation) that indicates a specific relationship between two entities. Redundancy rules usually go by the general format of

$$X \rightarrow Y$$

where X represents one of the entities and Y the other. In phonetics, for example, X stands for one phoneme or one phonetic feature, and Y for another phoneme or phonetic feature. The symbol (is read as "rewrites as." Take the following example:

$$[+\text{nasal}] \rightarrow [+\text{abrupt release}]$$

This rule means that if a phoneme contains the feature [+nasal], it "automatically" contains the feature [+abrupt release]. Therefore, the feature [+abrupt release] does not need to be specifically mentioned in the bundle of features used to identify a [+nasal] phoneme. In fact, redundancy rules tell us how we can predict phoneme qualities and the basis of the qualities we have already identified.

Also notice that the features [abrupt release] and [continuant] are the opposite of each other. That is, when a phoneme is [+abrupt release], it will certainly be [-continuant]. This relationship, too, can be represented in the form of the following redundancy rule;

$$[+\text{abrupt release}] \rightarrow [-\text{continuant}]$$

4.5. CONTINUANT

Whenever the pulmonic egressive airstream is constricted in the oral cavity, it is possible to lengthen the duration of the articulation of phonemes that are produced in that way. In fact, the phonetic feature [continuant] is used to describe the degree of obstruction involved in the articulation of a sound. If the obstruction is partial, with enough space for the air to pass out through the mouth, the phoneme involved is [+continuant]. Otherwise, the phoneme will be [-continuant]. In English, all vowels, glides, liquids, and fricative consonants are [+continuant]. Only affricates, stops, and nasal are [-continuant]. The feature [+continuant] is a redundant feature for liquids and vowels. This is because the feature [+vocalic] means that the phoneme is automatically [+continuant].

Therefore, [+continuant] is not included in the bundle of features used to identify [+vocalic] phonemes. The relationship between [+vocalic] and [+continuant] phonemes can be shown in the following redundancy rule:

$$[+\text{vocalic}] \rightarrow [+\text{continuant}]$$

4.6. ASPIRATION

In the articulation of word-initial voiceless stop consonants, in English, there is a build up of air pressure in the oral cavity prior to the actual release of airstream at the place of articulation. Therefore, when these phonemes are produced, an increased puff of air accompanies the phonemes. This quality is called aspiration. If you light a match, hold it in front of your mouth, and say a word like Peter, you may be able to feel the aspiration of the first sound.

In broad transcription (see chapter 2), aspiration is not identified; in narrow transcription, however, a superscript h is used to identify the aspiration of the phoneme:

toe	[tʰəʊ]	kill	[kʰɪɫ]
Peter	[pʰi:tə]	chat	[ʧʰæt]

As you see, in these examples the feature [+aspiration] is represented with a small superscribed h. The following phonological rule identifies aspirated phonemes of English:

$$\begin{bmatrix} -\text{voice} \\ -\text{continuant} \end{bmatrix} \rightarrow [+\text{aspiration}] \;/\; \# \begin{bmatrix} \_\_\_ \end{bmatrix}$$

The rule tells us that when [-voice, -continuant] phonemes occur at the beginning of a word (i.e., occupy the word-initial position), they are always [+aspiration]. This rule is not a redundancy rule because it identifies a condition for the change to take place. That is, the "word-initial position" condition must first be met if the [-voice, -continuant] phoneme wants to become [+aspiration].

Therefore, unlike redundancy rules, phonological rules need specific conditions. They are if…then rules. The schematic representation of phonological rules (also called phonological processes) is:

$$X \rightarrow Y \; / \; Z$$

The rule tells us that X will rewrite as Y if the conditions identified by Z are met (that is, X rewrites as Y in the context of Z). In the above example, the phoneme which is [-voice, -continuant] rewrites as [+aspiration] in the context of word-initial position. Phonological and redundancy rules will be treated with greater detail in the following chapters.

When phonemes containing [-voice] and [-continuant] features occur at the end of a word, they are sometimes [+aspiration]. When they occur after /s/, these phonemes are always [-aspiration]. The absence of aspiration can easily be felt by comparing the pronunciation of pairs of words like pie/spy, toe/stow, kill/skill, and so on. Also notice that there is never any puff of air accompanying the articulation of English phonemes which share such phonetic features as [+voice, -continuant], or [-voice, +continuant], or [+voice, +continuant]. Such English phonemes are [-aspiration]. Thus, fricatives, nasals, liquids, glides, vowels, voiced affricates, and voiced stop consonants are always [-aspiration].

4.7. ANTERIOR

The feature [anterior] is related to the place or position of articulation of phonemes.

Any phoneme which is articulated at a position closer to the frontal parts of the mouth is [+anterior]; any phoneme articulated at a position closer to the dorsal (or the posterior) parts of the mouth is considered to be [-anterior].

As figure 4.2 shows, all bilabial, labio-dental, dental, alveolar, and palatal sounds are [+anterior]. In contrast, all palato-velar, velar, velo-uvular, uvular, pharyngeal, and glottal sounds are [-anterior].

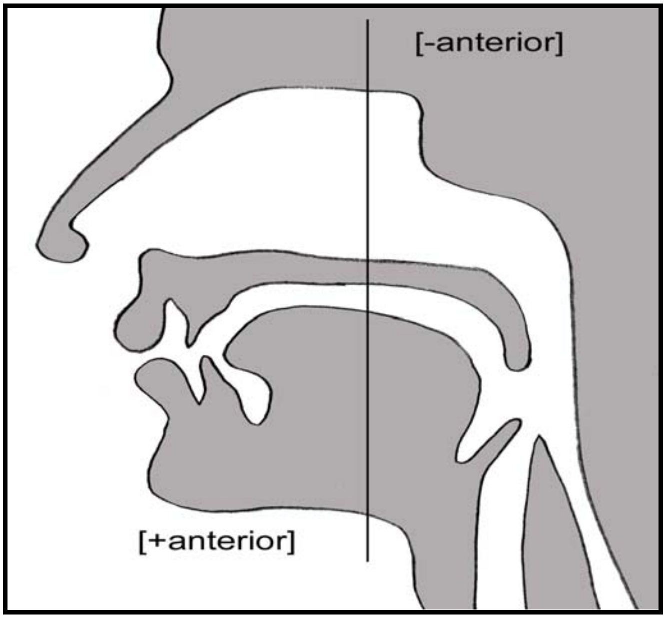


Figure 4.2: Anterior and posterior positions of articulation

In English, liquids are [+anterior] while glides (also less technically called semi-vowels) are [-anterior].

$$\begin{bmatrix} \text{+vocalic} \\ \text{+consonantal} \end{bmatrix} \rightarrow \text{[+anterior]}$$

$$\begin{bmatrix} \text{-consonantal} \\ \text{-vocalic} \end{bmatrix} \rightarrow \text{[-anterior]}$$

In addition the phonemes /ʃ/, /ʒ/, /ʧ/, /ʤ/, and /ŋ/ are [-anterior]. Other English phonemes are [+anterior]. Some phoneticians may use the feature [posterior] instead of the feature [anterior]. In other words, they use [-posterior] instead of [+anterior], and [+posterior] instead of [-anterior].

$$[\pm\text{anterior}] \rightarrow [\mp\text{posterior}]$$

These rules can be collapsed together to form a single and more general rule. If we use the mathematical symbol α to represent ±, then we can rewrite the above two rules as:

$$[\alpha \text{ anterior}] \rightarrow [-\alpha \text{ posterior}]$$

In most cases, however, phoneticians normally prefer not to use the feature [posterior]. The students' attention is drawn to this point so that confusions will not arise if the feature [posterior] is in any writing.

4.8. CORONAL

Like the feature [anterior], the feature [coronal], too, is related to the place or position of articulation of phonemes. The feature [coronal], however, shows whether the tip and the blade of tongue are involved in the articulation of phonemes or not. Any phoneme which is articulated by the involvement of the tip (i.e., apex) or the blade of tongue is considered to be [+coronal]; any phoneme which is not articulated through the involvement of the tip or the blade of tongue is considered to be [-coronal].

It can easily be felt that usually three parts of the human tongue are involved in the production of phonemes: (a) the very tip or apex, (b) the blade, and (c) the back or dorsum:

All interdental, alveolar, and palatal consonants are [+coronal]. Vowels and liquids are [-coronal]:

$$\begin{bmatrix} +\text{vocalic} \\ \pm\text{consonantal} \end{bmatrix} \rightarrow [-\text{coronal}]$$

The coronal phonemes of English are /θ/, /ð/, /t/, /d/, /n/, /s/, /z/, /ʒ/, /ʃ/, /ʧ/, /ʤ/, /l/, /r/, and /j/. In formal notations, the features [+anterior] and [coronal] usually occur together; in other words, whenever a phoneme is [+anterior], it is usually necessary to identify whether it is [+coronal] or [-coronal].

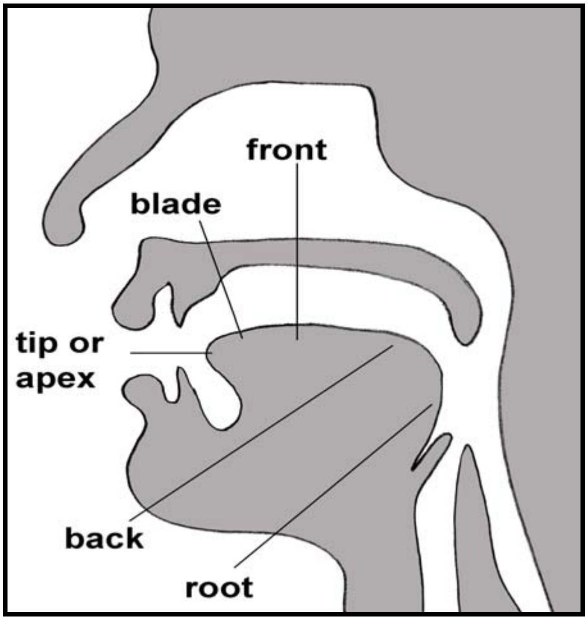


Figure 4.3: Different parts of human tongue

4.9. STRIDENT

This feature is in fact related to the manner of articulation of phonemes. As you will certainly remember, consonants are produced when a certain type of obstruction is introduced to the passage of air in the mouth. Some consonants are explosive in the sense that the blocked air is suddenly released; some others are fricatives or affricates because the air moves out of the mouth through a narrow opening. In this latter case, sometimes there is a turbulence in the air (like in the case of /s/, /z/, /ʃ/, /ʒ/, /ʧ/, /ʤ/, /f/, and /v/), and sometimes there is no turbulence (like in the case of phonemes other than these). The feature [strident] simply says whether this turbulence exists or not. English liquids, glides, and vowels are [-strident].

$$[+\text{vocalic}] \rightarrow [-\text{strident}]$$

$$\begin{bmatrix} -\text{vocalic} \\ -\text{consonantal} \end{bmatrix} \rightarrow [-\text{strident}]$$

Stridency is a feature of fricative (sometimes called spirant) and affricate consonants. Therefore, plosive consonants (i.e., those with the [+abrupt release] feature) are predictably [-strident]:

$$[+\text{abrupt release}] \rightarrow [-\text{strident}]$$

[+Strident] sounds are categorized into two subtypes based on their place of articulation. /f/ and /v/ are [-sibilant]. However, /s/, /z/, /ʃ/, /ʒ/, /ʧ/, and /ʤ/ are [+sibilant]. In fact, the term sibilant is used for hushers and hissers. During the articulation of [+sibilant] sounds, a groove is formed on the tongue so that the outgoing air takes on a hissing overtone. This hissing sound is not felt during the articulation of /f/ and /v/ (which are [+strident] labiodentals).

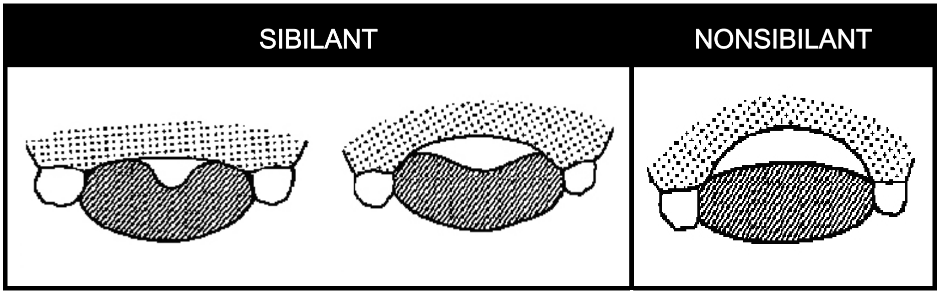


Figure 4.4: Sibilant versus non-sibilant phonemes

4.10. LATERAL

To produce a lateral sound, air is obstructed by the tongue at a point along the centre of the mouth but the sides of the tongue are left low so that air can escape over its sides. In fact, the tongue is strongly flexed and the air is forced through a narrow oval cavity, producing a hushing sound. /l/ is the clearest example of a [+lateral] sound in English. Both the clear [l] and the word-final dark [ɫ] allophones (i.e., the variants of the same phoneme) of /l/ are [+lateral] sounds. All other sounds of the English language are [-lateral]. Since /l/ is a liquid, we can safely claim that vowels, consonants, and glides are all [-lateral]:

$$[-\text{vocalic}] \rightarrow [-\text{lateral}]$$

$$\begin{bmatrix} +\text{vocalic} \\ -\text{consonantal} \end{bmatrix} \rightarrow [-\text{lateral}]$$

When an alveolar plosive is followed by the lateral /l/, then what happens is that we simply lower the sides of the tongue to release the compressed air, rather than raising and lowering the blade of the tongue. If you say the word 'bottle' to yourself you can feel the sides of the tongue lower to let out the air. Also notice that, since /l/ and /r/ are the only liquid sounds (i.e., [+vocalic] and [+consonantal] sounds) of English, the feature [lateral] is the most important distinctive feature that serves to distinguish between them.

4.11. SONORANT

The nonnasal stops, the fricatives, and the affricates form a major class of sounds technically termed **obstruents**. Nasals, on the other hand, form a class of sounds called sonorants. When the passage of air is blocked or constricted in the mouth, and the air passes out through the nose, the sound is [+sonorant]. When, by way of contrast, the air cannot escape through the nose, the sound is [-sonorant]. Fricatives are continuant obstruents because (a) the air is constricted in the mouth, and (b) the air cannot escape through the nose. Nonnasal stops and affricates are noncontinuant obstruents because (a) the air is completely blocked in the mouth, and (b) it cannot escape through the nose. Obstruents are, therefore, [-sonorant]. Notice that the feature [+sonorant] is a redundant feature for [+nasal] sounds:

$$[+\text{nasal}] \rightarrow [+\text{sonorant}]$$

Also notice that [+sonorant] phonemes are redundantly [-strident].

$$[+\text{sonorant}] \rightarrow [-\text{strident}]$$

4.12. SYLLABIC

Vowels usually constitute the "main core" or the **nucleus** of syllables; every vowel is at the center of a single syllable. Vowels can even occur as syllables. For instance, the word ago /ə'gəʊ/ is composed of two syllables: /ə/ and /gəʊ/. The first syllable is a single vowel. Even in the second syllable, the vowel (or the diphthong) takes the prominence (or stress) on itself. Therefore, vowels are said to be [+syllabic].

It was pointed out earlier that every language of the world contains the two basic classes of speech sounds often referred to by the cover terms **consonants** and **vowels.** This classification implies that consonants and vowels differ. In the production of consonants the flow of air is obstructed as it travels through the mouth. Vowels are produced with no obstruction in the mouth whatsoever. Oral and nasal stops, fricatives, affricates, liquids, and glides all have some degree of obstruction and are therefore consonants.

Consonants do not correspond exactly to the sounds specified as [+consonantal], because glides are [-consonantal], forming a subclass with vowels (which is why they often occur as part of a vocalic diphthong, and are also called semi-vowels). However, unlike vowels, glides are produced with some small oral obstruction and therefore do not constitute syllable peaks (i.e., they can neither occur as syllables nor do they govern stress), as do vowels. Vowels, like glides, are [-consonantal] and [+sonorant]. They differ from glides because they constitute syllable peaks; so vowels are [+syllabic], whereas glides are [-syllabic].

Liquids and nasals can also be syllabic—function as a syllable. Liquids and nasals are not always syllabic; rather, they *can* be syllabic. That is:

(a) They may constitute separate syllables, as shown by the words Rachel [reɪʧl̩], faker [feɪkr̩], rhythm [rɪðm̩], and button [bʌtn̩]. These words could also be pronounced as /ˈreɪʧəl/, /ˈfeɪkər/, /ˈrɪðəm/, and /ˈbʌtən/.

(b) They may be nonsyllabic, as in the words lead /li:d/, read /ri:d/, deal /di:l/, name /neɪm/, or mean /mi:n/.

Similarly, the vowel sound in words like bird and verb are sometimes written as a syllabic /r/ (i.e., [br̩d] and [vr̩b]); most dictionaries, however, prefer to transcribe syllabic vowels as /ə/ (e.g., [bərd], [vərb]). This is the only instance where a schwa represents a stressed vowel. The diacritic symbol [ˌ] usually appears under the phoneme to represent its [+syllabic] feature.

SYMBOL	MEANING	EXAMPLE	
[l̩]	syllabic l	Rachel	[reɪʧl̩]
[r̩]	syllabic r	faker	[feɪkr̩]
[m̩]	syllabic m	rhythm	[rɪðm̩]
[n̩]	syllabic n	button	[bʌtn̩]

Table 4.4 shows how four different classes of speech sounds may be distinguished by the features [± consonantal], [± syllabic], and [±sonorant].

	obstruents	vowels	glides	nasals and liquids
consonantal	+	-	-	+
syllabic	-	+	-	±
sonorant	-	+	+	+

Table 4.4: Specifications of major classes of sounds

By the system shown in Table 4.3, obstruents and vowels are distinct classes; they do not share any feature. Glides are like consonants in the sense that they are in the class of [-syllabic] sounds, but they are like vowels in the sense that they are [-consonantal] and [+sonorant]. Similarly, liquids and nasals are in the [+consonantal] class with obstruents, but share the feature [+sonorant] with vowels (and sometimes are [+syllabic]). Nonsyllabic liquids, nasals, and glides (or semivowels) are in the class of [+sonorant, -syllabic] sounds.

The syllabicity of liquids and nasal consonants may also be shown by describing the words in which they function as syllables as having short vowels before the liquids. So, whenever a short vowel precedes a liquid or a nasal consonant, the nasal consonant or the liquid becomes [+syllabic].

$$[+\text{nasal}] \rightarrow [+\text{syllabic}] \quad / \quad \begin{bmatrix} +\text{vocalic} \\ -\text{consonantal} \\ -\text{length} \end{bmatrix} \begin{bmatrix} \_\_ \end{bmatrix}$$

$$\begin{bmatrix} +\text{consonantal} \\ +\text{vocalic} \end{bmatrix} \rightarrow [+\text{syllabic}] \quad / \quad \begin{bmatrix} +\text{vocalic} \\ -\text{consonantal} \\ -\text{length} \end{bmatrix} \begin{bmatrix} \_\_ \end{bmatrix}$$

In many English dictionaries, short vowels that occur before syllabic phonemes are placed within parentheses to indicate that they are optional. Take the following examples:

medal /med(ə)l/ button /bʌt(ə)n/

In these cases, the schwa can be deleted due to the syllabicity of the phonemes that follow it. As it was mentioned earlier, this is the only instance where a schwa represents a stressed vowel.

EXERCISE

1. The following sets of minimal pairs show that English /p/ and /b/ contrast in initial, medial, and final positions.

Initial	Medial	Final
pit/bit	rapid/rabid	cap/cab

Find similar sets of minimal pairs for each pair of consonants given:

A	/k/-/g/	D	/b/-/v/	G	/s/-/ʃ/
B	/m/-/n/	E	/b/-/m/	H	/ʧ/-/ʤ/
C	/l/-/r/	F	/p/-/f/	I	/s/-/z/

2. For each group of vowels listed, state the phonetic feature(s) they all share.

	VOWEL GROUP	SHARED FEATURE(S)
A	/æ/ /ɑː/	
B	/ʌ/ /uː/ /ʊ/ /ɔː/ /ɒ/ /ɑː/	
C	/uː/ /ʊ/ /ɔː/ /ɒ/	
D	/iː/ /e/ /uː/ /ɔː/	
E	/iː/ /ɪ/ /uː/ /ʊ/	

3. For each group of sounds listed, state the phonetic feature(s) they all share.

	SOUND GROUP	SHARED FEATURE(S)
A	/m/ /n/ /ŋ/	
B	/ʧ/ /s/ /p/ /k/ /f/ /ʃ/ /t/ /θ/ /h/	
C	/f/ /v/ /θ/ /s/ /r/ /w/	
D	/h/ /w/ /j/	
E	/ʤ/ /z/ /b/ /g/ /v/ /ʒ/ /d/ /ð/	
F	/l/ /r/	

4. What phonetic property or feature distinguishes the following pairs of sounds?

	SOUND PAIR	FEATURE
A	/m/ /n/	
B	/j/ /w/	
C	/p/ /b/	
D	/j/ /h/	
E	/g/ /d/	

	SOUND PAIR	FEATURE
F	/ʧ/ /ʤ/	
G	/ɔ/ /ɒ/	
H	/θ/ /ð/	
I	/m/ /ŋ/	
J	/ŋ/ /n/	

5. List the distinctive features of each of the following sounds in the square brackets provided.

[p^h]	[ĩ]	[m̩]
[ũ]	[r̩]	[n̩]
[l̩]	[k^h]	[ʔ]

6. In systematic phonetics, a word may be described as a linear array of bundles of phonetic features. Use the same strategy to describe the words *pen*, *cat*, and *leak*.

pen

[ ] [ ] [ ]

cat

[ ] [ ] [ ]

leak

[ ] [ ] [ ]

CHAPTER FIVE
SUPRASEGMENTALS

1. INTRODUCTION

The term **suprasegmental** refers to those properties of an utterance which are not properties of any single segment. In chapter four, you were familiarized with the binary features that sought to distinguish between the different sound segments of English. Since these features were considered to be the qualities of sound segments, they are sometimes called segmentals. It is also vital to note that speech is not a sketchy haphazard use of random sounds. Rather, it is a more or less continuous, unified, and organized use of sounds that form syllables which, in turn, form words, phrases, sentences, and so on. Therefore, it is readily perceivable that features of speech are not confined to the binary features of sound segments. In fact, there are certain other features that belong not to any single sound segment, but to groups of them. These other features are called suprasegmental features, or suprasegmentals.

2. SUPRASEGMENTALS

Phoneticians talk of five major types of suprasegmental features that exist in almost all languages of the world. They include (a) stress, (b) tone, (c) intonation, (d) length, and (e) syllable. Notice that while a language like Chinese makes full use of tone, a language like English does not employ this suprasegmental feature. Therefore, your attention is drawn to the fact that the languages of the world make their own choices of suprasegmentals.

2.1. STRESS

The word stress means "loudness." Stress is a term that we apply to words in isolation which have more than one syllable. It refers to the property that certain syllables carry which makes them stand out from the rest of the word. It seems certain that stressed syllables are made with more effort than unstressed ones; the muscles in the lungs seem to expel air more aggressively than for unstressed syllables. Before we identify what stress means, try this short exercise: Say this sentence aloud and count how many seconds it takes.

The beautiful Mountain appeared transfixed in the distance.

How much time did it require? Probably about 5 seconds. Now, try saying this sentence aloud.

He can come on Sundays as long as he doesn't have to do any homework in the evening.

How much time did this one require? Probably about 5 seconds. The first sentence is much shorter than the second sentence. Compare them:

The beautiful Mountain appeared transfixed in the distance

He can come on Sundays as long as he doesn't have to do any homework in the evening

This simple example makes a very important point about how native speakers speak and use English. Namely, English is considered a stressed language while many other languages are considered syllabic. What does that mean? It means that, in English, speakers give stress to certain words (or more precisely syllables) while other words are quickly spoken. In other languages, such as French or Italian, each syllable receives equal importance.

Many speakers of syllabic languages do not understand why English people quickly speak, or swallow, a number of words in a sentence. In syllabic languages each syllable has equal importance, and therefore equal time is needed. English however, spends more time on specific stressed words while quickly gliding over the other, less important, words. For example, native speakers quickly glide over the modal verb "can" and it is hardly pronounced when it is used in the positive form.

They can come on Friday. (stressed words are underlined)

On the other hand, when they use the negative form "can't" they tend to stress the fact that it is the negative form by also stressing "can't".

They can't come on Friday.

As you can see from the above example the sentence, "They can't come on Friday" is longer than "They can come on Friday" because both the modal "can't" and the verb "come" are stressed. So, you see that stress adds to the length of the syllable that carries it. We mentioned earlier that only some syllables in a word carry stress. Any good dictionary provides you with information regarding the stress-carrying syllables of a given word. Take the following examples:

confine	/kən'faɪn/
album	/ˈælbəm/
commemoration	/kəˌmemə'reɪʃən/

In these examples, primary stress is marked with a raised vertical line, and secondary (or medium) stress is marked with lowered vertical line. In British dictionaries, both marks come at the beginning of the stressed syllable where as in American dictionaries they come at the end of stressed syllables. They apply to the entire syllable, not to any single segment within the syllable.

2.1.1. THE PLACE OF STRESS

Notice that English has four levels of stress placement: (a) **weak** or quiet, (b) **tertiary** or regular, (c) **secondary** or louder, and (d) **primary** or loudest. In the last example above, only two levels of stress have been identified: primary and secondary. As you have certainly noticed, only one syllable has remained intact (or untouched) between the two stressed syllables [kəˌ**me**mə'**reɪ**ʃən]. In fact, this observation has an implication: in English, after you identify the primary stress, you can identify the other levels of stress but you must be careful to leave one syllable intact between the levels. However, usually only two levels of stress are identified because no word in English gets long enough to require more that two levels of stress. We will return to this point in the following sections.

2.1.2. STRESS VERSUS PROMINENCE

When a syllable receives stress, it becomes prominent. It would have been logically possible for every syllable to have exactly the same loudness, pitch, and so on. In fact, some early attempts at speech synthesizers sounded like this. However, human languages have ways to make some syllables more prominent than others. A syllable might be more prominent by differing from the surrounding syllables in terms of (a) **loudness**, (b) **pitch**, or (c) **length**.

Pitch is an auditory sensation that places sounds on a scale from high to low. Every syllable has pitch, however, any syllable that is articulated with a noticeably different pitch will be deemed to carry stress. This can go either way: if all the syllables are said in a low pitch except one, then that higher pitch syllable will be deemed to carry the stress of the word. Pitch also plays a central role in intonation. Length, too, seems to play a role in stress. Generally, if one syllable has a longer length than the others in the word then it is deemed to be the one carrying stress. Length is one of the more important determiners of stress. In addition to length and pitch, loudness can result in prominence. It seems obvious that if one syllable is articulated louder than the others then it will have achieved some prominence from the other syllables. This prominence would then make that syllable the stressed syllable. However, it is very difficult to make a sound louder without affecting the length, pitch or quality of that syllable. If you could only change the loudness of a sound then the perceptual change is not as great as you would expect.

Prominence is relative to the surrounding syllables, not absolute. A stressed syllable that is nearly whispered will be quieter than an unstressed syllable that is shouted. In English, the three ways to make a syllable more prominent are to make it: (a) louder, (b) longer, or (c) higher pitched. The last one (i.e., change of pitch) is, in fact, the most usual way of achieving prominence in English. However, English typically uses all three kinds of prominence simultaneously. Other languages might use only one or two of them. The cues can also be used differently in other languages. In Swedish, stressed syllables are usually lower in pitch (which is, in effect, one of the most noticeable features of a Swedish accent).

Even in English, stress does not always mean higher pitch. In one of the intonation contours used to convey surprised disbelief, the most strongly stressed syllable of the utterance has the lowest pitch. Take the following example:

You're taking phonetics!

One of the most interesting points about stress in English is that vowels in unstressed syllables are systematically reduced. English speakers will not try to control the position of the tongue body during the vowel of an unstressed syllable. Instead, the tongue body will reach whatever point is convenient in getting from the preceding consonant to the following consonant. The average position reached is mid-central schwa /ə/. Failing to reduce unstressed vowels is one of the major contributors to an accent in non-native speakers of English. Reducing vowels inappropriately is one of the major contributors to an English accent in other languages. In general, the differences between stressed and unstressed syllables are more extreme in English than in most languages.

2.1.3. THE IMPORTANCE OF STRESS

There are many ways in which stress can be considered as an important quality of any language. However, all these ways can be captured under a cover term: **contrast**. The job of stress is to create contrast. In other words, stress is the safest way to reduce fuzziness in speech.

2.1.3.1. BOUNDARY MARKING

In normal speech, words and phrases simply do not have little pauses between them. In fact, a spoken sentence will definitely appear like a very long word to a person who is not familiar with the language. In such a situation, stress or prominence can help indicate where the boundaries are. This will make life easier for the listener.

In this connection, it is interesting to notice that a language like French usually gives prominence to the syllable at the end of a word or phrase. Persian, too, appears to follow a more or less similar pattern except for its Yazdi accent. Many other languages, however, give prominence to the initial syllables of words (e.g., Icelandic, Hungarian). There seems to be a bias for English listeners to interpret a stressed syllable as the beginning of a new word. As such, stress is a good help for listeners.

2.1.3.2. ADDITIONAL CONTRASTS

In many languages, changing which syllable is stressed can change the meaning of a word. Take the following examples:

	verb	noun
convert	/kənˈvɜːt/	/ˈkɒnvət/
contrast	/kənˈtræst/	/ˈkɒntræst/

Here again, the job of stress is to help the listener determine what the speaker means. Therefore, stress is important in the process of speech perception.

Listening is not the only aspect of your language learning which is affected by stress. Stress is also important in the articulation of speech. Let us take an example to demonstrate how stress affects speech:

> The beautiful Mountain appeared transfixed in the distance. (14 syllables)
>
> He can come on Sundays as long as he doesn't have to do any homework in the evening. (22 syllables)

Even though the second sentence is approximately 30% longer than the first, the sentences take the same time to speak. This is because there are 5 **stressed** words in **each** sentence. From this example, you can see that you do not need to worry about pronouncing every word clearly to be understood (native speakers certainly do not). You should however, concentrate on pronouncing the **stressed** words clearly.

Now, you can do some listening comprehension or speak to your native English speaking friends and listen to how they concentrate on the stressed words rather than giving importance to each syllable. You will soon find that you can understand and communicate more because you begin to listen for (and use in speaking) stressed words. All those words that you thought you did not understand are really not crucial for understanding the sense or making yourself understood. You can safely conclude that **stressed** words are the key to excellent pronunciation and understanding of English. Stress in English will help you to improve your understanding and speaking skills.

2.1.4. RULES OF STRESS

Many EFL (i.e., English as a foreign language) students find it very difficult to memorize the pronunciation of each and every English word. In fact, there is the misconception among EFL (English as a foreign language) students that they have no choice other than memorizing the pronunciations of words. This is far from being correct. Knowing a few rules may make it much easier for you to identify the exact pronunciation of a lot of English words.

2.1.4.1. CONTENT AND FUNCTION WORDS

Before we go on with the discussion of these rules, we should distinguish between two major groups of words: content words and function words. Content words are distinguished from function words by the fact that they carry the encyclopedic information contained in the sentence. In other words, content words have their own independent meanings and can be used independently. They can refer to things, states, qualities, or actions in the outside world. They have their own referents. Function words, by way of contrast, do not refer. They do not have independent encyclopedic meanings and can only be used in connection to content words. They usually signal grammatical relationships among content words in the sentence.

Content words include:

- (most) principle verbs (e.g. visit, construct)
- Nouns (e.g. kitchen, Peter)
- Adjectives (e.g. beautiful, interesting)
- Adverbs (e.g. often, carefully)

Function words include:

- Determiners (e.g. the, a, some, a few)
- Auxiliary verbs (e.g. don't, am, can, were)
- Prepositions (e.g. before, next to, opposite)
- Conjunctions (e.g. but, while, as)
- Pronouns (e.g. they, she, us)

2.1.4.2. WORD STRESS PATTERNS

English is not a language that follows precise rules for the placement of stress. This, of course, does not mean that it is not possible to find some rules for the placement of stress in English. The unfortunate point is that English has a very complex set of procedures that determine stress. It should be noted that nearly all English speakers agree on where stress should be placed in individual words. This means that the English system of stress does have some method to its chaos. One main consideration is that the presence of stress results in the prominence of one syllable over others in a word. This may lead the reader to the conclusion that mono-syllabic words (i.e., words with one syllable or only one vowel) may not need stress inside the sentence since there is no other syllable in them to require any contrast whatsoever.

2.1.4.2.1. MONOSYLLABIC WORDS

Obviously, these words do not present any problems for you because, when pronounced in isolation, they receive the primary stress on their only syllable. Notice that we do not show the stress of monosyllabic words. When part of a sentence, monosyllabic content words are pronounced with full forms of vowels whereas monosyllabic function words are normally pronounced with the weak forms of vowels (i.e., with /ə/ or /ɪ/). The vowels of monosyllabic function words are pronounced fully in two contexts: (a) when the word appears in a sentence-final position, and (b) when the word is used emphatically (or for purposes of contrast). You can feel easy to reduce all unstressed vowels (whether part of a content word or a function word) into /ə/. This is, after all, what many dialects (local forms) of English do. Take the following examples:

SENTENCE	PRONUNCIATION
Jack is from London	/dʒæk əz frəm 'lɒndən/
Is Jack from New York?	/əz dʒæk frəm nju:jɔ:k/
No, he is from London.	/nəʊ hɪ əz frəm 'lɒndən/
Where is Jack from?	/wərəz dʒæk frɒm/

Notice that in all of these sentences the word 'from' is pronounced as /frəm/ except for the last sentence where it is pronounced as /frɒm/. This is because of its sentence-final position. Also notice that the vowel of the second syllable in the word 'London' has been reduced into the weak form /ə/ due to the fact that this syllable is not stressed.

2.1.4.2.2. BI-SYLLABIC WORDS

In two syllable words, the choice of stress placement is still somewhat simple: either the first or the second syllable will be stressed. Bi-syllabic function words usually receive stress on their last syllable; there are some exceptions, though, which must be memorized. Take the following examples:

between	/bət'wi:n/	before	/bə'fɔ:/
below	/bə'ləʊ/	above	/ə'bɒv/
upon	/ə'pɒn/		

Bi-syllabic content words are stressed in a different way. The following hints are helpful for the placement of stress in bi-syllabic content words:

Bi-syllabic verbs are usually stressed according to their second syllable:

1) If the second syllable of the bi-syllabic verb contains a long vowel or diphthong, then the second syllable is stressed. Take the following examples:

increase	/ɪŋ'kri:s/	encroach	/əŋ'krəʊʧ/

2) If the bi-syllabic verb ends with more than one consonant, then the second syllable is stressed. Take the following examples:

collapse	/kə'læps/	condense	/kən'dens/

3) If the final syllable contains a short vowel and one or no final consonant, then usually, the first syllable will be stressed. Take these examples:

open	/ˈəʊpən/	envy	/ˈenvɪ/

Two syllable adjectives are stressed in the same manner. That is, the three rules mentioned for verbs hitherto, also apply to bi-syllabic adjectives. Take the following examples:

discreet	/dɪsˈkri:t/	disguised	/disˈgaɪzd/
fishy	/ˈfɪʃɪ/	handsome	/ˈhændsəm/

For adverbs, always notice that the ending –ly does not affect stress. Therefore, it is easy to put the –ly aside and identify the stress for the remaining adjective. The ending '-ly' is usually pronounced as /lɪ/. Take the following examples:

handsomely	/ˈhændsəmlɪ/	manly	/ˈmænlɪ/
lazily	/ˈleɪzɪlɪ/	discreetly	/dɪsˈkri:tlɪ/

Bi-syllabic nouns generally follow a different stress placement pattern.

1) If the second syllable contains a short vowel, then the stress usually comes on the first syllable. Take these examples:

labrum	/ˈleɪbrəm/	chimney	/ˈʧɪmnɪ/

2) If the bi-syllabic noun does not go by rule 1, its stress will have to be placed on the second syllable. Nouns that fall in this category are very rare. Take this example:

increase /ɪŋˈkri:s/

2.1.4.2.3. MULTI-SYLLABIC WORDS

Deciding which syllable to give stress becomes very complicated when words include more than two syllables. With poly-syllabic or multi-syllabic words (sometimes called heavy words), the rules of stress placement start to become ever more arbitrary—with more exceptions than can easily be explained away. However, this does not mean that there is no rule for this purpose.

One general hint to remember is that weak syllables never carry stress. Therefore, any unstressed syllable will reduce its vowel into a weaker one—like /ə/ or /ɪ/. Another hint is that multi-syllabic function words happen, for the most part, to carry stress on their final or ultimate syllable. Take the following examples:

nevertheless	/ˌnev(ə)rðə'les/	nonetheless	/ˌnʌnðə'les/
inasmuch (as)	/ˌɪnəz'mʌʧ/ (/əz/)	although	/ɔːl'ðəʊ/

Despite the large amount of research done on the subject of stress, it is still one area of which phoneticians have little understanding. They have, however, developed a few rules that can be safely used to identify the stress patterns of the majority, though not all, of the English poly-syllabic content words. Very often the exceptions are words borrowed from other languages (which are normally stressed according to the stress patterns of the languages from which they were originally borrowed.

These rules, and a few examples for each, are provided here. Notice that these rules do not identify the stress patterns of all content words. They are not exhaustive and all-inclusive.

1) Prefixes and the adverbial suffix -ly (used to make adverbs) usually do not change the pattern of stress. Take the following examples:

capitulate	/kə'pɪʧəleɪt/	recapitulate	/ˌriːkə'pɪʧəleɪt/
iterate	/'ɪtəreɪt/	reiterate	/riː'ɪtəreɪt/
sensitive	/'sensətɪv/	oversensitive	/ˌəʊvə'sensətɪv/
moderate	/'mɒdərət/	moderately	/'mɒdərətlɪ/
tangible	/'tænʤɪbl/	tangibly	/'tænʤɪblɪ/

2) Verbs that end in –ate or –ize receive stress on their antepenultimate (i.e., last but two, or the third from right) syllables. Notice that the endings are pronounced as /eɪt/ and /aɪz/ respectively. Take the following examples:

confiscate	/ˈkɒnfɪskeɪt/	demonstrate	/ˈdemənstreɪt/
evaporate	/ɪˈvæpəreɪt/	eradicate	/ɪˈrædɪkeɪt/
vaporize	/ˈveɪpəraɪz/	recognize	/ˈrekəgnaɪz/
romanticize	/rəˈmæntɪsaɪz/	sexualize	/ˈsekʃʊəlaɪz/

3) Verbs that end in –ify usually take stress on the syllable prior to the –ify ending. Take the following examples:

ramify	/ˈræməfaɪ/	beautify	/ˈbju:təfaɪ/
testify	/ˈtestəfaɪ/	bokonovskify	/ˌbɒkəˈnɒfskəfaɪ/

4) Other verbs usually take stress on their last syllables, unless when the last syllable contains a short vowel. Take the following examples:

intervene /ˌɪntəˈvi:n/ intercede /ˌɪntəˈsi:d/

5) Adjectives that end in –ate receive stress on their antepenultimate syllable (like verbs) but the –ate ending is pronounced as /ət/. Take the following examples:

moderate /ˈmɒdərət/ elaborate /ɪˈlæbərət/

6) Adjectives that end in –ese usually receive stress on the syllable containing –ese. Take the following examples:

Japanese /ˌʤæpəˈni:z/ Javanese /ˌʤævəˈni:z/

7) Adjectives which end in –ious, -uous, -eous, -ieous, -ic, -ical, -ian, -ible, –ial, or -ive (except for those ending in -tive) usually receive stress on the syllable prior to these endings. There are a few exceptions in connection to the ending –ic (e.g., Arabic, lunatic, and rhetoric). Take the following examples:

impressive	/ɪm'presɪv/	comprehensible	/ˌkɒmprə'hensɪbəl/
public	/'pʌblɪk/	spontaneous	/spən'teɪnɪəs/
biblical	/'bɪblɪkəl/	presumptuous	/prɪ'zʌmptʃʊəs/
grammarian	/grə'meərɪən/	dubious	/'du:bɪəs/
secretarial	/ˌsekrə'teərɪəl/		

8) Adjectives that end in –able, -al, and –ous usually take stress on their antepenultimate syllables (i.e., the third from right). Take these examples:

corporal	/'kɔ:pərəl/	admirable	/'ædmərəbl/
scrupulous	/'skru:pjələs/		

9) Nouns that end in –ity, -ety, -al, -ion, -ence, -ance, -acy, and –ian usually take stress on the syllable prior to these endings. Take the following examples:

piety	/'paɪətɪ/	importance	/ɪm'pɔ:təns/
ability	/ə'bɪlɪtɪ/	correspondence	/ˌkɒrəs'pɒndəns/
proposal	/prə'pəʊzəl/	recognition	/ˌrekəg'nɪʃən/
historian	/hɪs'tɔ:rɪən/	democracy	/dɪ'mɒkrəsɪ/

10) Other heavy nouns usually receive stress on their antepenultimate syllables. Take these examples:

photography	/fə'tɒgrəfɪ/	democrat	/'deməkræt/

sovereignty	/'sɒvrəntɪ/	decubitus	/dɪ'kju:bɪtəs/
parameter	/pə'ræmɪtə/	diplomat	/'dɪpləmæt/

11) The endings –ist, and –ism do not change stress. Take the following examples:

organ	/'ɔ:gən/	organism	/'ɔ:gənɪzəm/
physics	/'fɪzɪks/	physicist	/'fɪzɪsɪst/

12) Nouns ending in –ee usually receive stress on the ending itself; there are some exceptions though. Often the exceptions are those nouns that include double consonants prior to the –ee ending (e.g., committee, coffee, etc.). Take the following examples:

referee	/ˌrefə'ri:/	absentee	/ˌæbsən'ti:/
flunkee	/flən'ki:/	devotee	/ˌdevə'ti:/

13) –ly does not change stress. Therefore, for adverbs ending in –ly, the easiest way is to ignore the –ly ending and to identify the stressed syllable of the adjective. –ly in adverbs is usually pronounced as /lɪ/.

2.2. INTONANTION

There are other suprasegmental features, besides stress, that govern appropriate English pronunciation. You will certainly remember that these other features include intonation, tone, and length. Due to its importance in native-like pronunciation, the suprasegmental feature 'intonation' will be explained here. 'Intonation' means when, why and how a speaker chooses to raise or lower or sustain the pitch of her or his voice at particular points while speaking. These choices are as much a part of the grammatical system of the language as, say, using the correct auxiliary verb, or remembering when and where to place the person markers (I, you, he/she ...) which tell us who is carrying out an action; who or what is acted upon.

2.2.1. WHAT INTONATION DOES

Intonation in American English is the tools for achieving at least five important aims: (a) expressing new information, (b) showing contrast, (c) **expressing meaning**, (d) **showing pronunciation**, and (e) **showing mood or personality**. Language conveys very specific information, such as how to get somewhere or what someone is doing. It can be also used beyond the exact meaning of the words to indicate how the speaker feels about what he is saying, or how he personally feels at that moment. Generally speaking, if English is not your first language, this is where you start running into difficulty. Even if you pronounce each word clearly, if your intonation patterns are non-standard, your meaning will probably not be clear. Also, in terms of comprehension, you will lose a great deal of information if you are listening for the actual words used.

Each language deals with expressing these emotional ranges and contextual importances in different ways. Some languages, such as French and other Romance languages, stress the end of a sentence, and then use word order to indicate an important change. Other languages, such as Chinese, have a pitch change that indicates different vocabulary words, and then superimpose further pitch change to change meaning or emotion. Because English has a fairly strictly fixed word order, it is not an option to rearrange the words when we want to make a point about something. Intonation in American English is the rise and fall of pitch in order to convey a range of meanings, emotions or situations, within the confines of standard grammar and fixed word order.

2.2.1.1. EXPRESSING NEW INFORMATION

This is the starting point of standard intonation. When we say that we need to stress the new information, it is logical to think, "Hmmm, this is the first time I am saying this sentence, so it is all new information. I had better stress every word." Well, not quite. In Standard English, we consider that the nouns carry the weight of a sentence, when all else is equal. Although the verb carries important information, it does not receive the primary stress of a first-time noun.

Dogs eat **bones.**

After the information has been introduced, or is being repeated through the use of pronouns, the intonation shifts over to the verb. Notice how the intonation changes when a sentence changes from nouns to pronouns:

Dogs eat **bones.** They **eat** them.

2.2.1.2. CONTRAST

Once the intonation of new information is established, you will soon notice that there is a pattern that breaks that flow. When you want to emphasize one thing over another, you reflect this contrast with pitch change. Notice how the intonation indicates contrast:

Bob studies **English.**
Bob **studies** English, but he doesn't **use** it.

If a person consistently stresses "contrast words" as opposed to "new information words", he can end up sounding permanently argumentative or bad-tempered:

I said it is good.
He **doesn't** like it. Where **are** you going?

Additionally, mixed messages occur when modals or verbs of perception are stressed—you end up with the opposite meaning!

People **should** exercise more, but . . .
They **would** help us, if . . .
It **looks** like chance, but at that price, it's a knock-off.
He **seems** like a nice guy, but once you get to know him. . .

2.2.1.3. MEANING

A good exercise to demonstrate the variety of meaning through intonation changes is to take a single sentence, try stressing each word in turn, and see the totally different meanings that come out.

1. **I** didn't say he stole the money.
2. **I didn't** say he stole the money.
3. I didn't **say** he stole the money.
4. I didn't say **he** stole the money.
5. I didn't say he **stole** the money.
6. I didn't say he stole **the** money.
7. I didn't say he stole the **money.**

Once you are clear on the intonation changes in the seven sentences, you can add context words to clarify the meaning:

1. **I** didn't say he stole the money, someone **else** said it.
2. **I didn't** say he stole the money, **that's** not true at **all**.
3. I didn't **say** he stole the money, I only **suggested** the **possibility**.
4. I didn't say **he** stole the money, I think someone **else** took it.
5. I didn't say he **stole** the money, maybe he just **borrowed** it.
6. I didn't say he stole **the** money, but rather some **other** money.
7. I didn't say he stole the **money**, he may have taken some **jewelry**.

2.2.1.4. PRONUNCIATION

In any language, there are areas of overlap, where one category has a great deal in common with a different category. In this case, intonation and pronunciation have two areas of overlap. (a) The first is the pronunciation of the letter T. When a T is at the beginning of a word (such as table, ten, take), it is a clear sharp /t/ sound. It is also clear in combination with certain other letters, (contract, contain, etc.) When T is in the middle of a word, between two vowels, or in an unstressed position, it turns into a softer /d/ sound.

Betty bought a bit of better butter. /bedɪ bɔ:d ə bɪd ə bedə bʌdə/

It is this intonation/pronunciation shift that accounts for the difference between photography /fə'tɒgrəfɪ/ and photograph /'fəʊdəgræf/. (b) The second overlap has to do with the syllable (within a word) that receives the prominence of stress and pitch. In fact, the same syllable that receives the stress, will also govern pitch in the sentence if it occurs in the last content word, or if it is the focus of contrast in the sentence.

2.2.1.5. MOOD AND PERSONALITY

This is an extremely important aspect of intonation, as it goes beyond what you are trying to say--it dictates how your listener will relate to you as an individual. In fact, intonation determines if you will be considered charming or rude, confident or nervous, or informed or unfamiliar.

An extremely important part of intonation is inside a one-syllable word. You may wonder if intonation exists in a one-syllable word at all. However, we often put little sounds in a one-syllable word that are not in the written language, but that convey a great deal of information in terms of who we are. These extra sounds are explained in the following sections under the heading of liaisons.

When we contrast two similar words, one ending with a **voiced** consonant (d, z, g, v, b) and the other with an **unvoiced** consonant (t, s, k, f, p), you will hear the difference in the preceding **vowel**, specifically in the length or duration of that vowel.

Simply put, words that end in a voiced consonant have a doubled vowel sound. For example, if you say **bit**, it is a quick, sharp sound—a single musical note. If you say **bid**, however, the word is stretched out, it has two musical notes, the first one higher than the second, **bi-id**.

	single	double
tense	beat	bead
lax	bit	bid

2.2.2. TYPES OF INTONATION

Before we embark on any discussion of the types of intonation, it is vital to note that the last content word in the sentence is usually the one that carries sentence stress. It is, therefore, the one that affects the intonation of the sentence. In fact, in its most technical sense, intonation refers to (a) categories of pitch 'peaks' and 'valleys' as well as their combinations at each sentence stress position (i.e., the last content word of the sentence), (b) types of pitch category concatenation, and (c) pitch of sentence fractions occurring before the first sentence stress.

All vowels with 'primary' or 'secondary' sentence stress, receive intonation features, which may be either 'valleys' (i.e., falls) or 'peaks' (i.e., rises). In the case of peaks, they may contain a unidirectional fall, or rise, again at the end resulting in either a rise-fall or fall-rise intonation pattern. Peaks are characterized by a quick rise confined to the vicinity of a sentence-stressed syllable. Peaks are of three types: early-peaks, medial peaks, and late peaks. When the rise precedes the onset of syllable stress, and is narrow in temporal extension, an 'early peak' will form. When the rise extends into the first half of the stressed syllable, it results in a medial peak. In the late peak, the rise starts after the stressed vowel onset and continues into the second half of the stressed syllable or beyond. Valleys, on the other hand, have a continuous rise, starting before the stressed syllable (early) or inside it (non-early) and extending as far as the beginning of the following sentence-stressed syllable. If there are several unstressed syllables between two sentence stresses a valley is thus realized as more gradual compared with the much quicker rise for a late peak. The less distance there is between stressed syllables the more difficult it becomes to distinguish between a "valley+peak" and a late "peak+peak" sequence.

Since it receives the most important stress of the sentence, the last content word of the sentence is considered to be the cradle of intonation. Therefore, the ups and downs of pitch in connection to this content word are often used to identify the intonation of a sentence. In general, English uses six different forms of intonation patterns: (a) falling, (b) rising, (c) falling-rising, (d) rising-falling, (e) take-off, and (f) level.

2.2.2.1. FALLING INTONATION

Falling intonation is characterized by a clear fall of the pitch after the stressed syllable of the last content word towards the end of the sentence. In other words, the pitch will have to fall after the stressed syllable of the last content word. Falling intonation is used in many situations. Statements, WH-questions, and confirmatory tag questions (i.e., tag questions in which you expect the listener to confirm what you say) all use falling intonation.

Statements are those sentences that provide information for the listener. They may be either positive or negative. They may sometimes be imperative. All of these forms of statements require falling intonation. For statements, you should stress the nouns and let the tone fall at the end of the sentence (e.g., **Dogs** eat **bones**). Take the following examples:

> Open the door, please.
> Ali will go to Tehran tomorrow.
> We've just come from the airport.
> Her father makes sure she doesn't sell favor to friends.

WH-questions are those questions that ask for information by having the question word (e.g., where, when, how, what, which, why, who, etc.) placed at the front of the sentence (e.g., Where did she go?). You have certainly noticed that the pitch falls for the questions that begin with WH-words like where, when, etc. Take the following examples:

Hey! Where are you going?
What should we do in Senggigi?
Which temples should I visit?
Who would eat it?
How does her father ensure she doesn't sell favor to friends?
Why not visit the temples at Lingsar?
Where will we buy mangoes next time?

Tag questions are very small questions (usually in the form of auxiliary verbs followed by subject) that come at the end of statements (often for politeness purposes). They can receive either falling or rising intonation depending on the predictions the speaker can make. When the speaker believes that the listener will confirm what he says, he will use falling intonation; when, on the other hand, the speaker is not sure whether the listener will confirm his position, he may prefer to use rising intonation. Rising intonation is often more polite. Take the following examples:

Jack will come over to lunch, won't he?
You come from London, don't you?
Let's go to the movies, shall we?

2.2.2.2. RISING INTONATION

Rising intonation is characterized by a clear rise of the pitch after the stressed syllable of the last content word of the sentence onwards. In other words, the pitch will have to rise after the stressed syllable of the last content word. Rising intonation is used in yes-no questions, repeated questions, question-statements, and tag questions. English clearly differentiates between two types of questions: WH-questions and yes-no questions. Yes-no questions are simply those that require a yes or a no answer (with perhaps a little further elaboration, depending on the question.) In English, yes-no questions are often differentiated from question-word questions (also called information questions or WH-questions) in terms of intonation. Yes-no questions require rising intonation. Take the following examples:

Have you just come from the airport?
Is there a local bus terminal nearby?
Is it possible to get to Flores from Senggigi?
Do you really want to see a monitor lizard?

Notice that pitch on a regular question goes up (compared with a statement), but drops back down at the end.

Do **dogs** eat **bones**?

A repeated, rhetorical or emotional question goes up, and then up again at the end.

Do **dogs** eat **bones**?!

Rising intonation is also used in tag questions in which the speaker is not sure if the listener will confirm what he says. Take the following examples:

Jack will come over to lunch, won't he?
You come from London, don't you?
Let's go to the movies, shall we?

Sometimes speakers refrain from using direct questions and, instead, modulate the intonation of statements to change them into questions. In these cases, too, rising intonation is required. Take the following examples:

Ali will go to **Tehran** tomorrow?

Her father makes sure she doesn't sell favor to friends?

Questions of this type are more of a rhetorical or emotional nature than pure yes-no questions. They are often used to express the speakers' disbelief in the ideas expressed by the statement.

2.2.2.3. RISING-FALLING INTONATION

Rising-falling intonation is characterized by a combination of a rise followed by a fall. Rising-falling intonation is characterized by a combination of a rise followed by a fall. In this case, the pitch will go up on one part of the sentence and then fall down on another part. Rising-falling intonation is often used in two-part statements, intro-phrases, and listings.

In two-part sentences, the first half of the sentence usually sets up the second half. That is, the first half requires rising intonation, and the second half will take on falling intonation. The cradle of intonation in both parts is the last content word of each part. Take the following example:

Dogs eat **bones**, but **cats** eat **fish**.

Intro-phrases, too, require rising-falling intonation patterns. When you want to preface your statement, use a rising tone. In the following example, the first part requires rising intonation because it is an intro phrase. The second part will then be articulated with falling intonation.

As we all know, **dogs** eat **bones**.

Rising-falling intonation is also used in listings. With more than one item in a list, all but the last one have a rising tone. The last item in the list will, however, require falling intonation.

Dogs eat **bones**, **kibbles**, and **meat**.

If there are more than two sentences joined by conjunctions, the same pattern will be used. That is, all but the last one have a rising tone. The last part will, however, require falling intonation. Take the following example:

> **Dogs** eat **bones**, **cats** eat **cheese**, and **mice** eat **wheat**.

In this example, the parts prior to the conjunction "and" require rising intonation, and the part following it requires falling intonation.

2.2.2.4. FALLING-RISING INTONATION

Falling-rising intonation is characterized by a combination of a fall followed by a rise. This intonation pattern is the reverse of the rising-falling pattern. Falling-rising intonation is often used for soothing and politeness. When small children feel uneasy, for example, parents use falling-rising intonation to soothe them. Soothing refers to statements that make pain or discomfort less severe. Take the following examples:

> Don't cry.
> I'll take you to the park.

For purposes of politeness, too, falling-rising intonation is a good tool. Take the following example:

> Thank you

If the level of pitch falls on "thank" and then rises over "you", the expression will be more polite than when it is pronounced with falling intonation.

Falling-rising intonation pattern is sometimes used for pleading. Pleading refers to statements in which the speaker makes an earnest or urgent entreaty, often in emotional terms. Take this example:

> I plead with you to stay.

You will have certainly noticed that falling-rising intonation is a good tool for the expression of emotions. Politeness, soothing, and pleading are all emotionally charged.

2.2.2.5. TAKE-OFF INTONATION

Take-off intonation is characterized by a take-off pattern after the most important contrastive stress of the sentence. In other words, the speaker begins with a regular or level tone and then gradually raises the pitch (similar to the way airplanes run along the runway before take-off). The point at which the pitch comes to a full rising pattern depends on the speaker's choice. This is normally the point the speaker finds most important. Take-off intonation is often used in statements with negatively-charged emotions. For example, take-off intonation can freely be used for grumbling. Take the following example:

> You shouldn't have given him all that **money you silly boy.**

Grumbling is a kind of emotional expression, one that is normally negative. Other forms of negative emotional expressions, too, require take-off intonation. Cursing, blasphemy, etc. often take on this intonation pattern.

2.2.2.6. LEVEL INTONATION

Level intonation is characterized by no or very little pitch modulation throughout a sentence. In rising or falling intonation patterns (or their combinations), the pitch would move from one level to another (i.e., a higher or lower level). In level intonation, however, the modulations of pitch are not remarkable enough to move across different pitch levels. That is, the pitch will stay within the same level throughout the statement. Level intonation has a very strict usage. It is used in places where frozen style is best. One such situation is in the church. The priest usually uses this intonation pattern when he reads the holy book to the church-goers. Sometimes, in courts of law, too, this intonation pattern is used. Learners of English do not need to worry about this intonation pattern since they will almost never need it.

2.2.3. PHRASING VERSUS INTONATION

In addition to the intonation of a statement, there is another aspect of speech that indicates meaning—phrasing. Have you ever caught just a snippet of a conversation in your own language, and somehow managed to piece together what came before or after the part you heard? This has to do with your "grammar of expectancy" or your natural understanding of phrasing. In a sentence, phrasing tells you where the speaker is at the moment, where he is going, and if he is finished or not. Notice that the intonation stays on the nouns.

2.3. TONE

So far this chapter has been mainly concerned with the major suprasegmental aspects of phonology (i.e., stress and intonation). There are, however, other linguistically important suprasegmentals like tone, length, and syllable.

The way pitch is used linguistically differs from language to language. In English, it does not much matter whether you say 'cat' with a high pitch or a low pitch. It will still mean "cat." In other words, tone does not change within a single syllable. Therefore, English is a non-tonal language. However, this is not the case in many languages. For example, if you say 'ba' with a high pitch in Nupe (a language spoken in Nigeria), it will mean "to be sour;" if you say 'ba' with a low pitch, it will mean "to count." Languages that use the pitch of individual syllables to contrast meanings are called tone languages (or tonal languages). In other words, tone refers to the changes of pitch within a single syllable that result in changes in meaning. Thai is a language that has contrasting pitches, or tones. The same string of "segmental" sounds represented by [na:] will mean different things when spoken with a low pitch, a mid pitch, a high pitch, a falling pitch from high to low, or a rising pitch from low to high:

SYLLABLE	TYPE OF TONE	MEANING
na	low tone	a nickname
na	mid tone	rice paddy
na	high tone	young maternal uncle or aunt
na	falling tone	face
na	rising tone	thick

In some tone languages the pitch of each tone is "level"; in others, the direction of the pitch (whether it glides from high to low, or from low to high) is important. Tones that "glide" are called contour tones; tones that do not are called level or register tones. In Thai, there are three level or register tones (low, mid, and high) and two contour tones (rising and falling). Since English is not a tonal language, this brief explanation of tone seems to suffice here.

2.4. LENGTH

Length, too, seems to play a role in stress. Generally, if one syllable has a longer length than the others in the word, then it is deemed to be the one carrying stress. Length is one of the more important determiners of stress.

2.5. SYLLABLE

The notion of syllable has already been discussed in chapter two. It is, therefore, unnecessary to explain it here again. However, your attention is drawn to one important point: not all random combinations of phonemes can result in syllables in a language. Any language imposes some limitation on the types of phonemes that go together to form syllables. These restrictions are called sequential constraints because they constrain the sequences of sounds that go together in a syllable. You will remember, for example, that Persian does not allow initial consonant clusters in words. The same is true about Persian syllables: Persian does not allow initial consonant clusters in syllables either.

EXERCISE

1. Use the rules of stress discussed above to transcribe the following words. If the rules do not apply, write "impossible" in the corresponding cell.

WORD	TRANSCRIPTION
syllabic	
photography	
presumptuous	
sporadic	
parameter	
democracy	
chemist	
political	
referee	
survival	

WORD	TRANSCRIPTION
eradicate	
although	
rapid	
devotee	
redemption	
elaborate	
corporal	
historian	
comedy	
diplomat	

2. Say the following sentences with their correct intonation.

How are you today?
Are Dan and Megan studying?
They are in London.
He is from Peru.
Where does your brother work?
Your cough sounds terrible. Take some cough drops.
John will win the price, will he?
What city was your English school in?

3. Underline the word in each of the following sentences that normally receives the sentence stress.

    How do you spend your day, Helen?
    They are perfect for you.
    These jackets are nice.
    What kind of music do you like?
    Tell me about your brother and sister, Sue.
    Wow! and what does your brother do?
    He has an exhibition there.
    What do you do after class?
    What city was your English school in?
    Do you want to play sometime?
    What a wonderful thing democracy is.

4. The utterance 'I thought it would rain' can have two implications based on the kind of intonation with which it is said:
    (a) 'I thought it would rain' implying that it rained.
    (b) 'I thought it would rain' implying that 'I am surprised why it did not rain'.

    Try saying the sentence with the appropriate intonation for each of these implications.

5. What is the difference between phrasing and intonation?

6. How is tone different from intonation?

7. What are the different functions of intonation in English?

8. Choose some utterances to illustrate the functions of intonation in English.

CHAPTER SIX
PHONEMICS

1. INTRODUCTION

Phonetics describes the observable and concrete aspects of language (i.e., speech or sounds). There are other abstract and mental aspects of language that are not the subject of phonetics. We plan speech and then speak these plans. When we speak our planned speech, some changes take place due to biological and physiological constraints.

Phonetics can only describe the end result of our plans—speech. It does not describe the changes that occur to our plans; neither does it tell us what goes on in the human brain during the planning of speech. For instance, we all know that /ʃænbe/ is the correct form of the word meaning Saturday in Persian, yet we all pronounce it, unconsciously of course, as /ʃæmbe/. This shows that what we actually speak differs from what we plan to speak. In other words, speech planning and speech execution are somewhat different stories. That branch of phonology that is concerned with the explanation of these differences is called **phonemics**. As such, phonetics and phonemics are the two sides of a coin called phonology. Like phonetics, phonemics, too, is of two types: traditional and modern (or systematic). In traditional phonemics, abbreviations like the ones used in broad and narrow transcriptions were used. In modern phonemics, phonetic features are used.

Phonemics is the study of the **system** that governs the sounds of a language. Phonemics can also interact with morphology to bring about what is known as **morphophonemics**. We all know, for example, that the plural ending "s" is used in English to pluralize nouns. We also know that this ending can be pronounced in three different ways: /z/ in dogs, /s/ in books, and /əz/ or /ɪz/ in busses (usually called the **allomorphs** of the plural morpheme). This means that the same morpheme (i.e., the plural morpheme) has three surface realizations whereas it has only one mental or abstract realization. That is, in our mental dictionary (or our lexicon), there is only one entry for the plural morpheme; however, in actual speech, we pronounce this single morpheme in three different ways. This is because the surrounding sounds affect the surface realization of this morpheme in speech. Morphophonemics describes the mental processes that bring about such changes in the pronunciation of morphemes.

Phonemics accounts for at least two types of knowledge the speakers of a language possess: (a) the knowledge of sequential/phonological constraints, and (b) the knowledge of phonological processes. Phonological/sequential constraints refer to those aspects of human linguistic knowledge that explain why certain sound combinations are possible while certain other sound combinations impossible. Phonological processes explain the rules that account for the differences that exist between planned speech (i.e., phonemic representation of speech) and executed speech (i.e., phonetic representation of speech).

2. SEQUENTIAL/PHONOLOGICAL CONSTRAINTS

All native speakers of a language know, unconsciously of course, that the phonemes of their language cannot be strung together in any random order to form words. In fact, the phonological system of their language determines which phonemes can begin a word, end a word, and follow each other. Very often, sequential constraints are biologically-driven. That is, there is a biological explanation for them. Since breathing is the primary function of human vocal organ, speech is considered to be its subsidiary function. Therefore, any form of speech that can interfere with breathing is impossible; hence, sequential constraints.

Native speakers of English, for example, tacitly know that English words can begin with such sequences as **sp**, **sk**, **dr**, **tr**, and **str** but not with such sequences as **bp**, **pb**, **lbk**, etc. This unconscious knowledge tells them that no English word will ever begin with a sequence that is phonologically constrained in their language. New words enter the language but they all conform to the sequential constraints of the language. **Wog** is phonologically possible and may be added to the English language in future as a new word but ***wgo** is phonologically impossible and will never enter the English language as a new word. In fact, it can safely be argued that wog is a potential word (a word to which no meaning has been assigned yet). Such potential words are sometimes called accidental gaps.

3. PHONOLOGICAL PROCESSES

In connected speech, the usual aim is for ease of communication rather than complete accuracy. Therefore, speakers unconsciously draw on a number of phonological processes to aid that ease of communication. There are a good number of such processes. However, four of them are more important than the rest: assimilation, elision, neutralization, and linking/joining. Sometimes speakers make adjacent/close sounds more like each other (assimilation), sometimes they leave some sounds out altogether (elision), sometimes, they neutralize the differences between members of a natural class of sounds (neutralization), and other times they insert an unrelated sound to ease the transition from one sound to another (dissimilation or joining/linking).

3.1. ASSIMILATION

Assimilation is when one sound is influenced and changed by a neighboring sound. Many phoneticians use the term assimilation to refer to cases in which one sound affects the sound that appears on its left side. These scholars do not refer to right-ward sound change as assimilation.

Generally speaking, there are three types of assimilation that can be identified: assimilation of place, assimilation of manner, and/or assimilation of voice. **Assimilation of place** can be demonstrated by *ratbag* (that is, ['ræpbæg]) or *oatmeal* (that is, ['əʊpmi:ɫ]) where the /t/ sound is replaced by a /p/ sound. This is because the alveolar plosive /t/ is simplified into the /p/ sound which is closer to the bilabial plosive /b/. **Assimilation of manner** refers to when two different manners of articulation influence each other to form a different type of manner of articulation. Examples of this are the pronunciation of [Indian] as ['ɪnʤən] or the pronunciation of ['səʊlʤə(r)]. This is because the plosive /d/ combines with the approximant /j/ to form the affricate /ʤ/. **Assimilation of voice** is illustrated by the pronunciation of "have to" with an /f/ sound rather than the /v/ sound (i.e., ['hæftə]) as the voiced fricative is followed by a voiceless consonant.

The reason for assimilation is because (a) the tongue cannot always move quickly enough to get from one position to another in order to articulate the next sound, or because (b) the mouth is to busy anticipating the following sound. In either case, it approximates the sound before moving on to the next segment of sound. This approximation is technically referred to as **assimilation**. The following rules explain the most common types of assimilation that take place in English.

1. When a /d/ sound occurs before a /g/ sound, it is usually converted into the /g/ sound: (e.g., sad girl [sæggɜ:l])

$$\begin{bmatrix} +\text{consonantal} \\ -\text{vocalic} \\ +\text{voice} \\ +\text{abrupt release} \\ -\text{nasal} \\ +\text{coronal} \end{bmatrix} \rightarrow \begin{bmatrix} +\text{consonantal} \\ -\text{vocalic} \\ +\text{voice} \\ +\text{abrupt release} \\ -\text{nasal} \\ -\text{anterior} \end{bmatrix} \Big/ \begin{bmatrix} \underline{\quad} \end{bmatrix} \begin{bmatrix} +\text{consonantal} \\ -\text{vocalic} \\ +\text{voice} \\ +\text{abrupt} \\ \text{release} \\ -\text{nasal} \\ -\text{anterior} \end{bmatrix}$$

2. When /t/ is followed by /m/, it usually changes into a /p/ sound. Take the following example: not me [nɒpmi:].

$$\begin{bmatrix}\text{+consonantal}\\ \text{-vocalic}\\ \text{-voice}\\ \text{+abrupt release}\\ \text{+coronal}\end{bmatrix} \rightarrow \begin{bmatrix}\text{+consonantal}\\ \text{-vocalic}\\ \text{-voice}\\ \text{+abrupt release}\\ \text{+anterior}\\ \text{-coronal}\end{bmatrix} \;/\; \begin{bmatrix}\;\underline{\quad}\;\end{bmatrix} \begin{bmatrix}\text{+consonantal}\\ \text{-vocalic}\\ \text{+nasal}\\ \text{+anterior}\\ \text{-coronal}\end{bmatrix}$$

3. When /t/ is followed by /p/, it usually changes into a /p/ sound. Take the following example: that person ['ðæppɜ:sən].

$$\begin{bmatrix}\text{+consonantal}\\ \text{-vocalic}\\ \text{-voice}\\ \text{+abrupt release}\\ \text{+coronal}\end{bmatrix} \rightarrow \begin{bmatrix}\text{+consonantal}\\ \text{-vocalic}\\ \text{-voice}\\ \text{+abrupt release}\\ \text{+anterior}\\ \text{-coronal}\end{bmatrix} \;/\; \begin{bmatrix}\;\underline{\quad}\;\end{bmatrix} \begin{bmatrix}\text{+consonantal}\\ \text{-vocalic}\\ \text{-voice}\\ \text{+abrupt release}\\ \text{+anterior}\\ \text{-coronal}\end{bmatrix}$$

4. When /t/ is followed by /b/, it usually changes into a /p/ sound. Take the following two examples: that boy ['ðæpbɔɪ]; light blue ['laɪpblu:].

$$\begin{bmatrix}\text{+consonantal}\\ \text{-vocalic}\\ \text{-voice}\\ \text{+abrupt release}\\ \text{+coronal}\end{bmatrix} \rightarrow \begin{bmatrix}\text{+consonantal}\\ \text{-vocalic}\\ \text{-voice}\\ \text{+abrupt release}\\ \text{+anterior}\\ \text{-coronal}\end{bmatrix} \;/\; \begin{bmatrix}\;\underline{\quad}\;\end{bmatrix} \begin{bmatrix}\text{+consonantal}\\ \text{-vocalic}\\ \text{+voice}\\ \text{+abrupt release}\\ \text{+anterior}\\ \text{-coronal}\end{bmatrix}$$

5. When /d/ is followed by /b/, it usually changes into a /b/ sound (e.g., good boy ['gʊbbɔɪ]).

$$\left[\begin{array}{l} +\text{consonantal} \\ -\text{vocalic} \\ +\text{voice} \\ +\text{abrupt release} \\ -\text{nasal} \\ +\text{coronal} \end{array}\right] \rightarrow \left[\begin{array}{l} +\text{consonantal} \\ -\text{vocalic} \\ +\text{voice} \\ +\text{abrupt release} \\ +\text{anterior} \\ -\text{coronal} \end{array}\right] / \left[\begin{array}{c} \_\_\_ \end{array}\right] \left[\begin{array}{l} +\text{consonantal} \\ -\text{vocalic} \\ +\text{voice} \\ +\text{abrupt release} \\ +\text{anterior} \\ -\text{coronal} \end{array}\right]$$

6. When /d/ is followed by /p/, it usually changes into a /b/ sound (e.g., good practice ['gʊbpræktɪs]).

$$\left[\begin{array}{l} +\text{consonantal} \\ -\text{vocalic} \\ +\text{voice} \\ +\text{abrupt release} \\ -\text{nasal} \\ +\text{coronal} \end{array}\right] \rightarrow \left[\begin{array}{l} +\text{consonantal} \\ -\text{vocalic} \\ +\text{voice} \\ +\text{abrupt release} \\ +\text{anterior} \\ -\text{coronal} \end{array}\right] / \left[\begin{array}{c} \_\_\_ \end{array}\right] \left[\begin{array}{l} +\text{consonantal} \\ -\text{vocalic} \\ -\text{voice} \\ +\text{abrupt release} \\ +\text{anterior} \\ -\text{coronal} \end{array}\right]$$

3.2. ELISION

Elision refers to when a sound or syllable is lost or omitted. It particularly affects: (a) consonant clusters, (b) weakly stressed syllables that are not especially missed and (c) words that end in an alveolar consonant and that are immediately followed by a word beginning with a consonant. The sounds that are elided are those sounds that are so weakly articulated that they no longer have any auditory significance. Contracted forms of words are caused by elision.

Some elided syllables are represented in standard punctuation, for example, /I'm/ should be /I am/. In standard speech, the missing vowel is understood and so meaning does not suffer from this contraction. Elision is one of the reasons for the great mismatches found in English between a word's spelling and its pronunciation. [Wednesday], for example, was originally a contraction of [Odin's day] while today, Odin is barely discernable as the /d/ is no longer pronounced. Elision is most commonly used in, but is not exclusive to, connected speech. The faster the speech, the more likely that sounds and syllables will be elided. Notice that when a vowel is elided, it is usually a weak vowel, typically the schwa. When a consonant is elided, it is usually because it is in an environment with other consonants.

One cause of elision is the loss of a weak vowel after the voiceless plosives /p/, /t/ and /k/. The word **permit** is often pronounced as [p'mit] which is an example of the schwa sound /ə/ being elided after /p/.

$$\begin{bmatrix} \text{-consonantal} \\ \text{+vocalic} \\ \text{-low} \\ \text{-high} \\ \text{-back} \\ \text{-tense} \end{bmatrix} \rightarrow \varnothing \;/\; \begin{bmatrix} \text{+consonantal} \\ \text{-vocalic} \\ \text{+abrupt release} \\ \text{-voice} \end{bmatrix} \begin{bmatrix} \_\_\_\_ \end{bmatrix}$$

Another cause of elision is when a weak vowel is elided before the syllabic consonants /l/, /m/, /n/ and sometimes /r/. The pronunciation of the word *seven* ['sevn] demonstrates the loss of the schwa /ə/ before the /n/ sound.

$$\begin{bmatrix} \text{-consonantal} \\ \text{+vocalic} \\ \text{-low} \\ \text{-high} \\ \text{-back} \\ \text{-tense} \end{bmatrix} \rightarrow \varnothing \;/\; \begin{bmatrix} \_\_\_\_ \end{bmatrix} \begin{bmatrix} \text{+consonantal} \\ \text{-vocalic} \\ \text{+syllabic} \end{bmatrix}$$

Complex clusters are often elided in order to simplify the saying of the sound. The word clothes /kləʊðz/ is often elided to the much simpler pronunciation /kləʊz/.

$$\begin{bmatrix} \text{+consonantal} \\ \text{-vocalic} \\ \text{+voice} \\ \text{+continuant} \\ \text{-strident} \\ \text{+coronal} \end{bmatrix} \rightarrow \varnothing \;/\; \begin{bmatrix} \\ \\ \\ \_\_\_\_ \\ \\ \\ \end{bmatrix} \begin{bmatrix} \text{+consonantal} \\ \text{-vocalic} \\ \text{+voice} \\ \text{+strident} \\ \text{+anterior} \\ \text{+coronal} \end{bmatrix}$$

/v/ is often elided when it comes before a consonant. The Russian proper name Pavlov is, therefore, sometimes elided to ['pæləf] which is a much simpler pronunciation.

$$\begin{bmatrix} \text{+consonantal} \\ \text{-vocalic} \\ \text{+voice} \\ \text{+strident} \\ \text{-coronal} \end{bmatrix} \rightarrow \varnothing \;/\; \begin{bmatrix} \\ \\ \_\_\_\_ \\ \\ \\ \end{bmatrix} \begin{bmatrix} \\ \text{+consonantal} \\ \text{-vocalic} \\ \\ \\ \end{bmatrix}$$

Some elisions are just by convention or to speed up or simplify the way we speak. For example in the phrase **horse shoe**, the /s/ sound is usually elided to make the pronunciation of the phrase easier; hence, ['hɔ:ʃu:]. In other words, the /ʃ/ sound affects any /s/ or /z/ sound that comes before it. Take the following examples:

this shoe [ðɪ'ʃu:]

for those shoes [fəðəʊ'ʃu:z]

$$\begin{bmatrix} +\text{consonantal} \\ -\text{vocalic} \\ -\text{voice} \\ +\text{strident} \\ +\text{anterior} \\ +\text{coronal} \end{bmatrix} \rightarrow \varnothing \;/\; \begin{bmatrix} \_\_\_\_ \end{bmatrix} \begin{bmatrix} +\text{consonantal} \\ -\text{vocalic} \\ -\text{voice} \\ +\text{strident} \\ +\text{continuant} \\ -\text{anterior} \\ +\text{coronal} \end{bmatrix}$$

$$\begin{bmatrix} +\text{consonantal} \\ -\text{vocalic} \\ +\text{voice} \\ +\text{strident} \\ +\text{anterior} \\ +\text{coronal} \end{bmatrix} \rightarrow \varnothing \;/\; \begin{bmatrix} \_\_\_\_ \end{bmatrix} \begin{bmatrix} +\text{consonantal} \\ -\text{vocalic} \\ -\text{voice} \\ +\text{strident} \\ +\text{continuant} \\ -\text{anterior} \\ +\text{coronal} \end{bmatrix}$$

3.3. LIAISONS (JOINING)

If you have ever listened to people speaking a foreign language that was unknown to you, you may have noticed that it was impossible to pick out individual words from the string of sounds that you heard. This is because, in real connected speech, words are linked to one another. This is an especially important phenomenon in RP where the phoneme /r/ does not occur in syllable-final position unless the word with a final 'r' is followed by a word beginning with a vowel. Furthermore, when two vowels meet over a word boundary, an extra letter is frequently added in order to help the transition. While /r/ is the most common linking sound, /w/ and /j/ can also be activated between two vowels. This kind of liaison in which a sound segment is inserted between two other sound segments to differentiate between them is often called **dissimilation**.

Dissimilation is the opposite of assimilation. One example of linking is to re-activate an /r/ sound:

e.g., the teacher of the school /ðə’ti:ʧərəvðə’sku:l/.

Another example of linking is to insert an /r/ between two vowels:

e.g., the idea of /ðə’aɪdɪərəv/.

Linking is not limited to the /r/ sound it can also be used by /j/ and /w/. for example, in English the [+tense, -low] vowels [i:, e, u:, ɔ:] are predictably followed by a glide:

$$\varnothing \rightarrow \begin{bmatrix} -\text{vocalic} \\ -\text{consonantal} \\ +\text{high} \\ \propto\text{back} \\ \beta\text{round} \end{bmatrix} / \begin{bmatrix} +\text{vocalic} \\ -\text{consonantal} \\ +\text{tense} \\ -\text{low} \\ \propto\text{back} \\ \beta\text{round} \end{bmatrix} \begin{bmatrix} \quad \underline{\qquad} \quad \end{bmatrix}$$

In American English, words are not pronounced one by one. Usually, the end of one word attaches to the beginning of the next word. This is also true for initials, numbers, and spelling. Part of the glue that connects sentences is an underlying hum or drone that only breaks when you come to a period, and sometimes not even then. You have this underlying hum in your own language and it helps a great deal toward making you sound like a native speaker. Once you have a strong intonation, you need to connect all sounds together so that each sentence sounds like one long word.

The dime easier
They tell me the dime easier.
They tell me the dime easier to understand.
They tell me that I'm easier to understand.

The last two sentences above should be pronounced exactly the same, no matter how they are written. It is the **sound** that is important, not the spelling. This stringing together of sounds is usually called **liaison**. There are four main points where liaisons happen in English: (a) when consonants are followed by vowels; (a) when consonants are followed by other consonants; (c) when vowels are followed by other vowels; and (d) when /t/, /d/, /s/, and /z/ are followed by /j/.

3.3.1. CONSONANT-VOWEL SEQUENCES

Words are connected when a words ends in a **consonant** sound and the next word starts with a **vowel** sound, including the semivowels /w/, and /j/ and the liquid /r/.

Spelling	Pronunciation
My name is Ann.	/maɪ neɪ mɪ zæn/
American accent	/ə'mer k 'næksənt/

You also use liaisons in spelling and numbers.

Spelling	Pronunciation
LA	/eh leɪ/
909-5068	/naɪ nəʊ naɪn, faɪ vəʊ sɪk seɪt/

3.3.2. CONSONANT-CONSONANT SEQUENCES

Words are connected when a word ends in a consonant sound and the next word starts with a consonant that is in a similar position of articulation. For example, if a word ends with an alveolar or palatal sound and the next word starts with another alveolar or palatal sound, these words are going to naturally join together. This is the same for almost all places of articulation.

Spelling	Pronunciation
I just didn't get the chance	/aɪ ʤʌsdɪdən geðə ʧæns/
I've been late twice.	/aɪvbɪn leɪtwaɪs/

3.3.3. VOWEL-VOWEL SEQUENCES

When a word ending in a vowel sound is next to one beginning with a vowel sound, they are connected with a glide (or sometimes the liquid /r/) between the two vowels.

Spelling	Pronunciation
Go away.	/gəʊ(w)əweɪ/
I also need the other one.	/aɪ(j)a:lsəʊ ni:d ðə(j)ʌðə wʌn/
The idea of space travel is new.	/ðəaɪdɪərəv speɪs trævəlɪz nu:/

A glide is either a slight /j/ sound or a slight /w/ sound. How do you know which one to use? This will take care of itself: the position your lips are in will dictate either /j/ or /w/. For example, if a word ends in /əʊ/, your lips are going to be in the forward position, so a /w/ quite naturally leads into the next vowel sound: /gəʊ(w)əweɪ/. After a long /ɜ:/ sound, your lips will be pulled back far enough to create a /j/ glide or liaison: /aɪ(j)a:lsəʊ ni:d ðə(j)ʌðə wʌn/. Do not force this sound too much, though. It is not a strong pushing sound. When one word ends with /ə/ and the adjacent one begins with /ə/, the use of /r/ is almost unassailable.

3.3.4. /T, D, S, Z/+/Y/ SEQUENCES

When the /t, d, s, z/ sounds are followed by a word that starts with /j/, both sounds are connected. These sounds connect not only with written /j/ sounds, but they do so as well with the initial unwritten /j/ sound of syllables and words. They form a combination that changes the pronunciation:

a) Use /ʧ/ instead of a /t/ sound followed by a /j/ sound.

Spelling	Pronunciation
What's your name?	/wʌʧərneɪm/
Can't you do it?	/kænʧə du:ɪt/
Don't you like it?	/dəʊnʧə laɪkɪt/
actually	/æʧəlɪ/

b) Use /ʤ/ instead of a /d/ sound followed by a /j/ sound.

Spelling	Pronunciation
What did you do?	/wʌ(dɪ)ʤə du:/
Would you help me?	/wʊʤʊ help mi:/
Did you like it?	/dɪʤə laɪkɪt/
graduation	/græʤʊ(w)eɪʃən/

c) Use /ʃ/ instead of a /s/ sound followed by a /j/ sound.

Spelling	Pronunciation
insurance	/ɪnʃʊrəns/
sugar	/'ʃu:gə/

d) Use /ʒ/ instead of a /z/ sound followed by a /j/ sound.

Spelling	Pronunciation
How's your family?	/hɑʊʒə fæmɪlɪ/
Who's your friend?	/hu:ʒə frend/
casual	/kæʒʊwl/
usual	/ju:ʒʊwl/

3.4. NEUTRALIZATION

With some pairs of sounds, it is not always clear which of the two sounds is being articulated. It is as if it lay between the two distinct sounds. This intermediate sound is known as **neutralization**. This process can be found in both vowels and some consonants. The most common examples of neutralization are the vowels /u/ and /i/ and the consonants /b/, /d/ and /g/ with their voiceless equivalents /p/, /t/ and /k/. Also, it is sometimes difficult to hear the contrast in the sounds /s/ and /z/.

When the article **'the'** is followed by a vowel, its final vowel is deemed to be /i/. The schwa is inappropriate, the long /i:/ is too long and the short one /ɪ/ is too short and so the in-between /i/ is used:

e.g., the orange [ði'ɒrɪnʤ]

In addition, after some consonants /i/ is used:

e.g., happy [hæpi], easy ['i:zi]

Along the same lines, /i/ is used in an unstressed prefix that is followed by a vowel:

e.g., react [ri'ækt]

It is also used in short unstressed words followed by a word beginning with a vowel:

e.g., he asked [hi'æskt]

The same is true about the vowel /u/ which is an in-between vowel differing from both /ʊ/ and /u:/. The vowel /u/ is used when it is in an unstressed syllable followed by a vowel, or the vowel like consonants (i.e., glides or semi-vowels). Take the following examples:

e.g. to ask [tu'æsk] who would [hu'wʊd]

When the plosives [p, t, k, b, d, g] follow /s/ in syllable-initial position, they become neutralized. An in-between sound is produced which is really difficult to place as one or the other. In fact, it is only the convention of regular spelling that determines how we transcribe them. Take the following examples:

The word **spill** could be transcribed as **[sbɪɫ]**
The word **still** could be transcribed as **[sdɪɫ]**
The word **skill** could be transcribed as **[sgɪɫ]**

3.5. VOICE ONSET TIME (VOT)

The term **Voice Onset Time (VOT)** refers to the timing of the beginning of vocal cord vibration in CV sequences (i.e., CV syllables) relative to the timing of the consonant release. The theory proposes that this timing is critical for accurate perception of the voiced/voiceless phonological contrast between consonants. This holds for stop consonants and to some extent for fricative consonants (that is, when the consonant is a plosive and often when the consonant is a fricative).

In phonology, consonants are contrasted and differentiated on the [voice] feature. Thus /b, d, g/ are marked [+voice] and /p, t, k/ are marked [–voice] in phonology. The abstract nature of phonology implies firm boundaries to the segment, and a straightforward conversion from abstract to concrete (phonemic representation to phonetic representation) as phonetics realizes phonology. In fact, it can be shown that in English and many other languages, vocal cord vibration does not occur throughout a [+voice] stop, and that it does not begin simultaneously with the beginning of the vowel following a [–voice] stop.

VOT may be **negative, zero** or **positive**. Zero VOT indicates that vocal cord vibration has begun simultaneously with the release of the plosive consonant; negative VOT indicates vibration beginning earlier than the release of the plosive consonant; and, positive VOT indicates vibration beginning after the release. Different languages have different methods of phonetic realization of this phonological feature. Notice that zero VOT in French cues the perception of a [–voice] stop, whereas in English the same auditory cue indicates a [+voice] stop.

Experiments with VOT have shown that abstract phonological specifications like [+voice] or [–voice] do not always have a direct one-to-one phonetic realization. **Phonetic features** (in this case that of vocal cord vibration) do not necessarily correspond directly with phonological features (in this case [voice]), nor do phonetic features necessarily reflect the abstract segmental divisions of phonology, but tend to 'blur' very often across where a segmental boundary might be. Notice also that different languages often realize the same abstract phonological specifications (i.e., phonemic representations) differently in their phonetics (i.e., phonetic representations). Here the phonological feature [+voice] is realized in English by a zero VOT, but in French by a comparatively long negative VOT, whereas the [–voice] feature was realized in English by a long positive VOT, but in French by a zero or very short positive VOT.

3.6. ASPIRATION

Closely related to the notion of voice onset time (VOT) is the notion of aspiration. The term aspiration means "pronunciation accompanied by breathing out." In the studies of voice onset time (VOT) in traditional phonetics, the focus of attention was on the sound or the airflow which filled the voice onset time (usually in the form of a puff of air). This quality was usually referred to as **aspiration**. When voiceless stop consonants take the word-initial position, they are often articulated with this added puff of air. Aspiration can be represented with the formal notations of systematic phonemics as the following phonological rule:

$$\begin{bmatrix} \text{+consonantal} \\ \text{-vocalic} \\ \text{-voice} \\ \text{+abrupt release} \end{bmatrix} \rightarrow \begin{bmatrix} \text{+aspiration} \end{bmatrix} \; / \quad \# \quad \begin{bmatrix} \_\_ \end{bmatrix}$$

The symbol # marks word boundaries. If placed to the left (like in the above rule) it marks the initial position; when placed to the right, it marks the final position. As such, the above rule reads as: Voiceless stop consonants become aspirated in word-initial contexts. Compare the /p/ sound in words like **pit** and **spit**.

3.7. NASALIZATION

Nasalization (sometimes called nasal assimilation) occurs when a vowel is followed by a nasal consonant like in the words moon, room, etc. In these cases, the vowel usually takes on a nasal quality as a result of the speakers' anticipation of the following nasal sound. Nasalization can be expressed in the following phonological rule:

$$\begin{bmatrix}-\text{consonantal}\\ +\text{vocalic}\end{bmatrix} \rightarrow \begin{bmatrix}+\text{nasal}\end{bmatrix} \;/\; \begin{bmatrix}\_\_\end{bmatrix}\begin{bmatrix}+\text{nasal}\end{bmatrix}$$

A special kind of nasalization occurs when one nasal consonant changes into another nasal consonant. This can be observed in such words as impolite, imbalance, etc. The prefix [ɪn] can change into [ɪm] when it is followed by [+anterior, -coronal] consonants. The prefix [ɪn] can also change into [ɪŋ] when it is followed by [-anterior, -coronal] consonants as in words like incomplete and inglorious. In fact, the nasal [n] assumes the same value [±] for [anterior] and [coronal] as the following consonants:

$$\begin{bmatrix}+\text{anterior}\\ +\text{coronal}\end{bmatrix} \rightarrow \begin{bmatrix}\propto\text{anterior}\\ \beta\text{coronal}\end{bmatrix} \;/\; \begin{bmatrix}\_\_\_\_\\ +\text{nasal}\end{bmatrix}\begin{bmatrix}\propto\text{anterior}\\ \beta\text{coronal}\end{bmatrix}$$

Since this kind of nasalization affects the structure of the morpheme **in,** it is often the subject of morphophonemics rather than being the subject of phonemics.

3.8. NON-CONTIGUOUS ASSIMILATION

Assimilation does not always involve sounds that are adjacent. Sometimes, a sound can affect another sound which is not adjacent or close to it. In cases where assimilation involves non-adjacent sounds, the term non-contiguous assimilation may be used. Two such assimilation processes are **vowel harmony** and **umlaut**. In umlauting, the vowel of a root assimilates to the vowel of a suffix:

$$\begin{bmatrix} \text{-consonantal} \\ \text{+vocalic} \\ \text{+back} \end{bmatrix} \rightarrow \begin{bmatrix} \text{-back} \end{bmatrix} / \begin{bmatrix} \_\_ \end{bmatrix} \ldots\nparallel\ldots \begin{bmatrix} \text{-consonantal} \\ \text{+vocalic} \\ \text{-back} \end{bmatrix}$$

In vowel harmony the reverse occurs: the vowel of a suffix assimilates to the vowel of a root:

$$\begin{bmatrix} \text{-consonantal} \\ \text{+vocalic} \end{bmatrix} \rightarrow \begin{bmatrix} \alpha\text{round} \\ \beta\text{back} \end{bmatrix} / \begin{bmatrix} \alpha\text{round} \\ \beta\text{back} \end{bmatrix} \ldots\nparallel\ldots \begin{bmatrix} \_\_\_ \end{bmatrix}$$

The symbol $\ldots\nparallel\ldots$ indicates that the sounds are not adjacent or contiguous (i.e., some other sounds come in between).

Umlauting is not a feature of English, but many Germanic languages, Hungarian, Finnish, Turkish, and some other languages make use of this phonological process. Since umlauting and vowel harmony affect the structure of morphemes, they are often the subject of morphophonemics.

EXERCISE

1. The following narrow transcriptions show words and non-words in English. Put an asterisk (*) to the left of the wrong transcriptions to identify them as non-words.

Example: *[bæŋ] [bæ̃ŋ]

[pi:tə]	[k^{h}ɪɫ]	[bĩ]	[ɫu:s]	[ɪŋk]
[mu:n]	[ræmp^{h}]	[leɪ]	[fət^{h}ɒgrəfɪ]	[ʧæp]
[lʊk]	[metəl̩]	[bĩ:n]	[nũ:n]	[θɪŋg]

2. The following are statements of some predictable aspects of the phonetic features of English. Convert each statement into a formal phonological or redundancy rule. Use the fewest number of features necessary to identify the sounds to which the rules apply. Write your answers in the boxes provided. The first one is done as an example for you.

Example: Nasalize vowels before nasal sounds.

$$\begin{bmatrix}\text{-consonantal}\\ \text{+vocalic}\end{bmatrix} \rightarrow \begin{bmatrix}\text{+nasal}\end{bmatrix} \;/\; \begin{bmatrix}\_\_\_\end{bmatrix}\begin{bmatrix}\text{+nasal}\end{bmatrix}$$

A: Liquids are lax and voice.

B: Of the nasal consonants, only /m/ and /n/ can occur in word-initial position.

C: Delete a consonant before a word beginning with any consonant that is not a glide.

3. Rewrite the following formal phonological and redundancy rules as statements that explain those phonological processes.

$$\text{Ø} \rightarrow \begin{bmatrix} \text{-vocalic} \\ \text{-consonantal} \\ \text{+high} \\ \propto\text{back} \\ \beta\text{round} \end{bmatrix} \Big/ \begin{bmatrix} \text{+vocalic} \\ \text{-consonantal} \\ \text{+tense} \\ \text{-low} \\ \propto\text{back} \\ \beta\text{round} \end{bmatrix} \begin{bmatrix} \underline{\qquad} \end{bmatrix}$$

YOUR ANSWER:

$$\begin{bmatrix} \text{+consonantal} \\ \text{-vocalic} \\ \text{+voice} \\ \text{+abrupt release} \\ \text{-nasal} \\ \text{+coronal} \end{bmatrix} \rightarrow \begin{bmatrix} \text{+consonantal} \\ \text{-vocalic} \\ \text{+voice} \\ \text{+abrupt release} \\ \text{+anterior} \\ \text{-coronal} \end{bmatrix} \Big/ \begin{bmatrix} \underline{\quad} \end{bmatrix} \begin{bmatrix} \text{+consonantal} \\ \text{-vocalic} \\ \text{-voice} \\ \text{+abrupt release} \\ \text{+anterior} \\ \text{-coronal} \end{bmatrix}$$

YOUR ANSWER:

$$\begin{bmatrix} \text{-consonantal} \\ \text{+vocalic} \end{bmatrix} \rightarrow \begin{bmatrix} \propto\text{round} \\ \beta\text{back} \end{bmatrix} \Big/ \begin{bmatrix} \propto\text{round} \\ \beta\text{back} \end{bmatrix} \ldots \not\# \ldots \begin{bmatrix} \underline{\qquad} \end{bmatrix}$$

YOUR ANSWER:

CHAPTER SEVEN
ACOUSTIC PHONETICS[1]

1. INTRODUCTION

Acoustic phonetics is a technical area of linguistics. Phoneticians depict and analyze sound waves using machines and computer programs. Acoustic phonetics can be defined as the study of sound waves made by the human vocal organs for communication. Speech sounds, like sounds in general, are transmitted through the air as small, rapid variations in air pressure that spread in longitudinal waves from the speaker's mouth and can be heard, recorded, visualized, and measured.

Differences between individual speech sounds are directly reflected as differences in either one or several or all such parameters as **duration**, **pitch**, **loudness** and **quality** of the belonging speech waves. Basically, acoustic phonetics deals with the study and description of three topics: (a) the acoustical properties of individual phonemes (i.e., speech sounds), (b) prosody, and (c) voice quality. It forms not only the immediate link between speech production (i.e., Articulatory Phonetics) and speech perception (i.e., Auditory Phonetics), but is also important for applications in the fields of signal processing and speech technology. As such, acoustic phonetics goes hand in hand with an almost new branch of phonetics called experimental phonetics. Experimental phonetics also lends support to articulatory phonetics discussed in the following chapters.

In some ways experimental phonetics, unlike experimental work in other areas of linguistics, is like experimental work in the physical sciences. For example, we can (a) examine in detail the acoustic waveform of the sounds of speech

---

[1] Most of the discussions of this chapter were adopted from *Encyclopedia Encarta 2004* © 1993-2003 Microsoft Corporation.

(i.e., acoustic phonetics) or (b) can examine the behavior of the musculature during articulation (i.e., articulatory phonetics). The data collected in such experiments reflect what actually happens when a speaker speaks, and are thus called **real world data**.

As you have already noticed, phonetics is interested in discovering what the speaker actually does when he speaks, not what he feels or thinks about speech. Of course, people do have feelings about speech and it is important for linguistics to understand the mental, as well as the physical aspects of speaking. It has already been mentioned in chapter one that the study of the **underlying mental processes** involved in speaking is treated in linguistics for the most part under the heading "phonology." Phonetics studies what people actually do when they speak. Of course, the relatively new area of **Cognitive Phonetics** tries to characterize some of the mental aspects of speaking which are not truly phonological. Cognitive phonetics has made the boundary between phonology and phonetics somewhat fuzzy.

2. WAVE MOTION

Wave Motion, in physics refers to a mechanism by which energy is conveyed from one place to another in mechanically propagated waves without the transference of matter. At any point along the path of transmission a periodic (sine) displacement, or oscillation, occurs about a neutral position. The oscillation may be of air molecules, as in the case of sound traveling through the atmosphere; of water molecules, as in waves occurring on the surface of the ocean; or of portions of a rope or a wire spring. In each of these cases the particles of matter oscillate about their own equilibrium position and only the energy moves continuously in one direction.

Such waves are called mechanical because the energy is transmitted through a material medium, without a mass movement of the medium itself. The only form of wave motion that requires no material medium for transmission is the electromagnetic wave; in this case the displacement is of electric and magnetic fields of force in space.

3. TYPES OF WAVES

Waves are divided into types according to the direction of the displacements in relation to the direction of the motion of the wave itself. If the vibration is parallel to the direction of motion, the wave is known as a longitudinal wave. The **longitudinal** wave (or P wave) is always mechanical because it results from successive **compressions** (state of maximum density and pressure) and **rarefactions** (state of minimum density and pressure) of the medium. Sound waves typify this form of wave motion. The successive compressions and rarefactions (i.e., oscillation) of the particles of matter (or the medium) help the longitudinal wave to travel. When a person speaks, the molecules of air (or the medium) oscillate to produce a longitudinal wave that can reach listeners' ears (see figure 7.1).

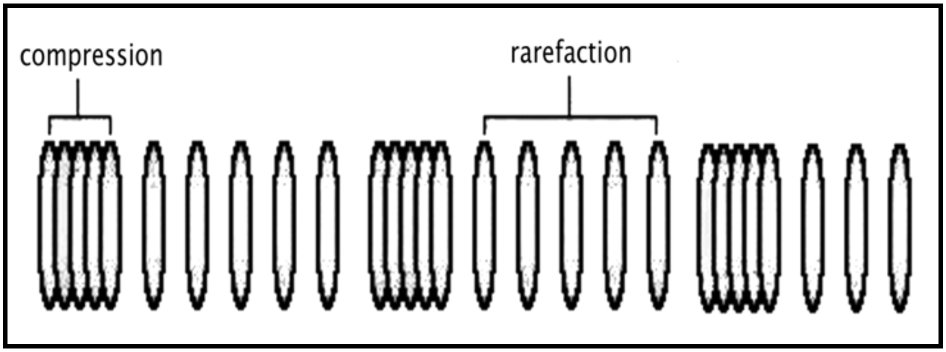


Figure 7.1: Longitudinal wave (Reprinted from *Encyclopedia Encarta 2004* © 1993-2003 Microsoft Corporation).

Another type of wave is the **transverse** (or S wave) wave, in which the vibrations are at right angles (90 degrees) to the direction of motion. A transverse wave may be mechanical, such as the wave projected in a taut string that is subjected to a transverse vibration; or it may be electromagnetic, such as light, X ray, or radio waves. Some mechanical wave motions, such as waves on the surface of a liquid, are combinations of both longitudinal and transverse motions, resulting in the circular motion of liquid particles (see figure 7.2).

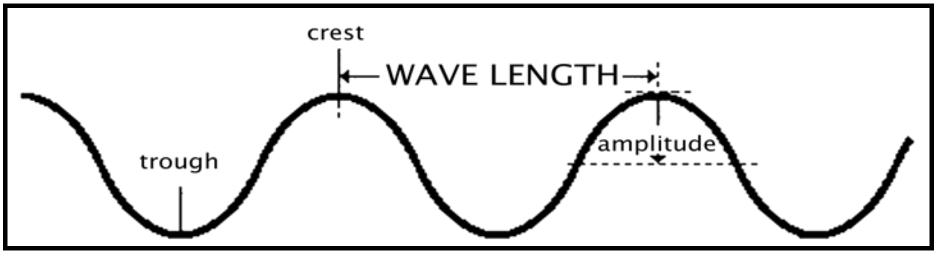


Figure 7.2: Transverse wave (Reprinted from *Encyclopedia Encarta 2004* © 1993-2003 Microsoft Corporation).

For a transverse wave, the wavelength is the distance between two successive crests or troughs. For longitudinal waves, it is the distance from compression to compression or rarefaction to rarefaction. The frequency of the wave is the number of vibrations per second. The velocity of the wave, which is the speed at which it advances, is equal to the wavelength times the frequency. The maximum displacement involved in the vibration is called the amplitude of the wave. The sequence of a trough and a crest is called a **cycle**. Frequency is the technical term used to refer to the number of cycles that occur in a given period of time (usually one second). Hz (hertz) is the measurement unit that is used to measure wave frequencies. One Hz is equal to one cycle per second.

4. TYPES OF WAVE MOTION

Waves, such as water or sound waves, are periodic (sine or regular) disturbances of the medium through which they travel. For longitudinal waves, the medium is displaced in the direction of travel. For example, air is compressed and expanded (figure 7.1) in the same direction that a sound wave travels. For transverse waves, the medium is displaced perpendicular to the direction of travel. Ripples on the surface of a pond are an example of a transverse wave: the water is displaced vertically (figure 7.2), while the wave itself travels horizontally. Earthquakes generate both P (compression, or longitudinal) and S (shear, or transverse) waves, which travel at different speeds and follow different paths.

5. SOUND

Sound is a physical phenomenon that stimulates the sense of hearing. In humans, hearing takes place whenever vibrations of frequencies from 15 hertz to about 20,000 hertz reach the inner ear. The hertz (Hz) is a unit of frequency equaling one vibration or cycle per second. Such vibrations reach the inner ear when they are transmitted through air. The speed of sound varies, but at sea level it travels through cool, dry air at about 1,190 km/h.

The term **sound** is sometimes restricted to such airborne vibrational waves. Modern physicists, however, usually extend the term to include similar vibrations in other gaseous, liquid, or solid media. Physicists also include vibrations of any frequency in any media, not just those that would be audible to humans. Sounds of frequencies above the range of normal human hearing, higher than about 20,000 Hz, are called ultrasonic. Sounds of frequencies below the range of normal human hearing, lower than 15 Hz, are called infrasonic.

In general waves can be propagated, or transmitted, transversely or longitudinally. In both cases, only the energy of wave motion is propagated through the medium; no portion of the medium itself actually moves very far. In transverse waves, the material through which the wave is transmitted vibrates perpendicular to the wave's forward movement. As a simple example, a rope may be tied securely to a post at one end, and the other end pulled almost taut and then shaken once. A wave will travel down the rope to the post, and at that point it will be reflected and returned to the hand. No part of the rope actually moves longitudinally toward the post, but each successive portion of the rope moves transversely. This type of wave motion is called a transverse wave (see figure 7.3).

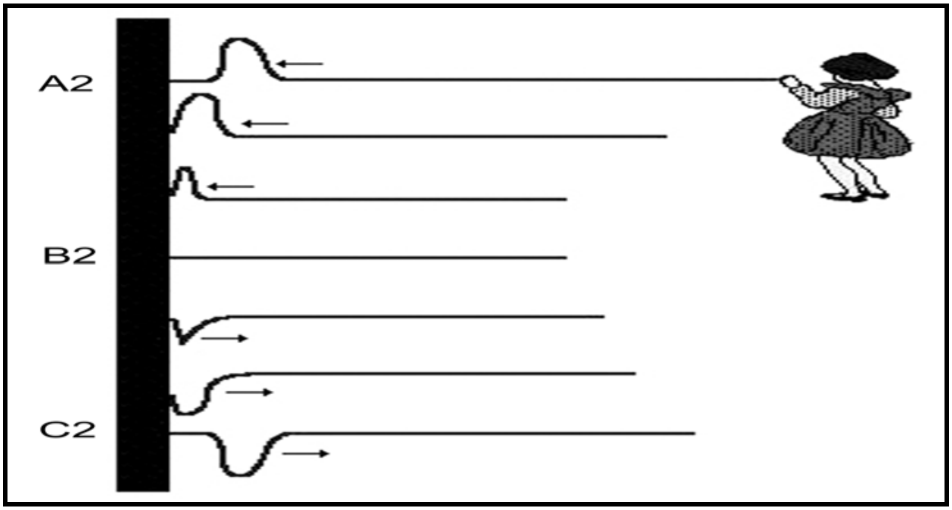


Figure 7.3: Movement of a transverse wave (Reprinted from *Encyclopedia Encarta 2004* © 1993-2003 Microsoft Corporation).

Similarly, if a rock is thrown into a pool of water, a series of transverse waves moves out from the point of impact. A cork floating near the point of impact will bob up and down, that is, move transversely with respect to the direction of wave motion, but will show little if any outward or longitudinal motion.

A sound wave, on the other hand, is a longitudinal wave. As the energy of wave motion is propagated outward from the center of disturbance, the individual air molecules that carry the sound move back and forth, parallel to the direction of wave motion. Thus, a sound wave is a series of alternate increases and decreases of air pressure. Each individual molecule passes the energy on to neighboring molecules, but after the sound wave has passed, each molecule remains in about the same location.

6. PHYSICAL CHARACTERISTICS

Any simple sound, such as a musical note, may be completely described by specifying three perceptual characteristics: pitch, loudness (or intensity), and quality (or timbre). These characteristics correspond exactly to three physical characteristics: frequency, amplitude, and harmonic constitution, or waveform, respectively. Noise is a complex sound, a mixture of many different frequencies or notes not harmonically related.

6.1. FREQUENCY

Sounds can be produced at a desired frequency by different methods. Sirens emit sound by means of an air blast interrupted by a toothed wheel with 44 teeth. The wheel rotates at 10 revolutions per second to produce 440 interruptions in the air stream every second. Similarly, hitting the A above middle C on a piano causes a string to vibrate at 440 Hz (i.e., fundamental frequency). The sound of the speaker and that of the piano string at the same frequency are different in quality, but correspond closely in pitch. The next higher A on the piano, the note one octave above, has a frequency of 880 Hz, exactly twice as high. Similarly, the notes one and two octaves below have frequencies of 220 and 110 Hz, respectively. Thus, by definition, an octave is the interval between any two notes whose frequencies are in a two-to-one ratio (see figure 7.4).

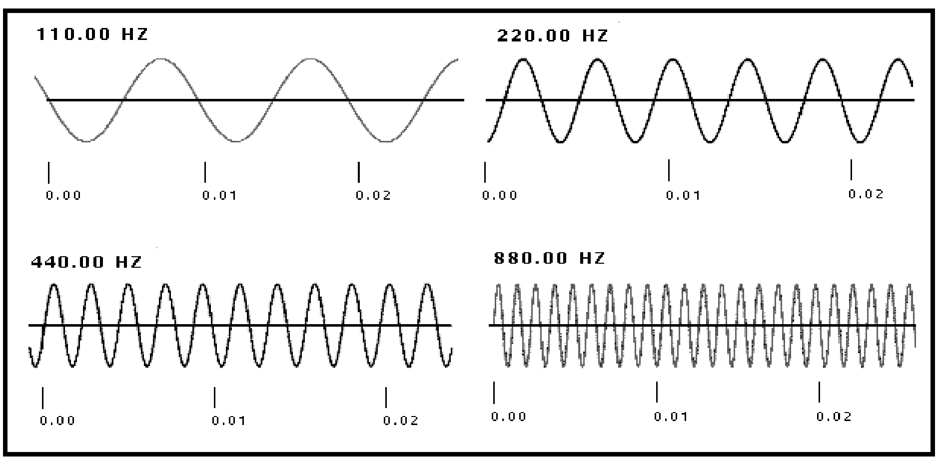


Figure 7.4: Sound waves with different octaves (Reprinted from *Encyclopedia Encarta 2004* © 1993-2003 Microsoft Corporation).

We perceive frequency as “higher” or “lower” sounds. In the figure above, the frequency of each higher wave is double that of the one below, producing the same note at different frequencies, from 110.00 Hz to 880.00 Hz. Waves propagate at both higher and lower frequencies, but humans are unable to hear them outside of a relatively narrow range (between 15 Hz and 20,000 Hz).

Frequencies of oscillating objects can cover a wide range of values. The tremors of earthquakes may have a frequency of less than one, while the rapid electromagnetic oscillations of gamma rays may have frequencies of 10^{20} or more. In almost all forms of mechanical vibration, a relationship exists between frequency and the physical dimensions of the vibrating object. Thus, for example, the time required by a pendulum to make one complete swing is partly determined by the length of the pendulum; the frequency or speed of vibration of a string of a musical instrument is partly determined by the length of the string. In each instance, the shorter the object, the higher the frequency of vibration.

In wave motion of all kinds, the frequency of the wave is usually given in terms of the number of wave crests that pass a given point in a second. The velocity of the wave and its frequency and wavelength are interrelated. The wavelength (the distance between successive wave crests) is inversely proportional to frequency and directly proportional to velocity.

Frequency is expressed in hertz (Hz); a frequency of one Hz means that there is one cycle or oscillation per second. The unit is named in honor of the German physicist Heinrich Rudolf Hertz, who first demonstrated the nature of electromagnetic wave propagation. Kilohertz (kHz), or thousands of cycles per second, megahertz (MHz), or millions of cycles per second, and gigahertz (GHz), or billions of cycles per second, are employed in describing certain high-frequency phenomena, such as radio waves. Radio waves and other types of electromagnetic radiation may be characterized either by their wavelengths, or by their frequencies. Electromagnetic waves of extremely high frequencies, such as light and X rays, are usually described in terms of their wavelength measure, which is often expressed in angstrom units (symbolized as Å meaning hundred-millionths of a cm). An electromagnetic wave that has a wavelength of one Å has a frequency of about three billion GHz.

6.2. AMPLITUDE

The amplitude of a sound wave is the degree of motion of air molecules within the wave, which corresponds to the changes in air pressure that accompany the wave. The greater the amplitude of the wave, the harder the molecules strike the eardrum and the louder the sound that is perceived. The amplitude of a sound wave can be expressed in terms of absolute units by measuring the actual distance of displacement of the air molecules, the changes in pressure as the wave passes, or the energy contained in the wave.

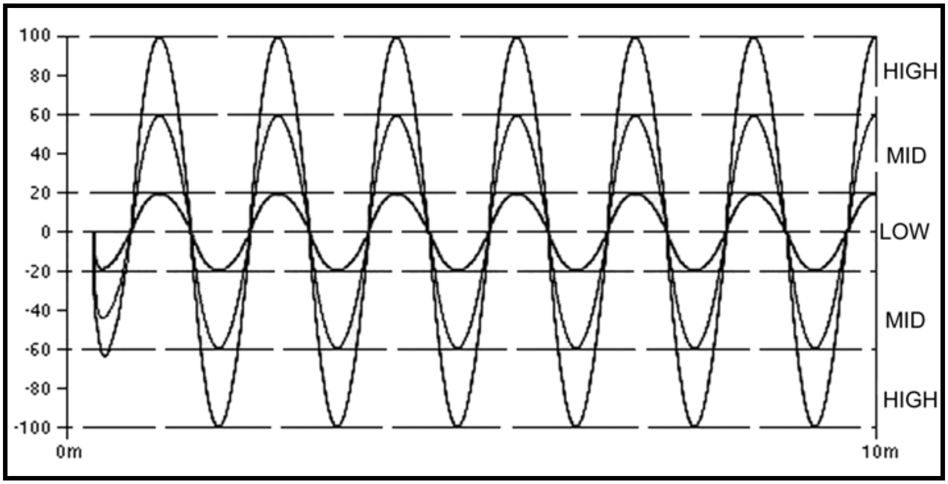


Figure 7.5: Sound waves with different amplitudes but same frequency (Reprinted from *Encyclopedia Encarta 2004* © 1993-2003 Microsoft Corporation).

Ordinary speech, for example, produces sound energy at the rate of about one hundred-thousandth of a watt. All of these measurements are extremely difficult to make, however, and the intensity of sounds is generally expressed by comparing them to a standard sound, measured in decibels. Amplitude is the characteristic of sound waves that we perceive as volume. The maximum distance a wave travels from the normal, or zero, position is the amplitude; this distance corresponds to the degree of motion in the air molecules of a wave. As the degree of motion in the molecules is increased, they strike the ear drum with progressively greater force. This causes the ear to perceive a louder sound. A comparison of samples at low, medium, and high amplitudes demonstrates the change in sound caused by altering amplitude. The three waves in figure 7.5 have the same frequency, and so should sound the same except for a perceptible volume difference.

6.3. INTENSITY

The distance at which a sound can be heard depends on its intensity. Intensity is the average rate of flow of energy per unit area perpendicular to the direction of propagation, similar to the rate at which a river flows through a gate in a dam. In the actual propagation of sound through the atmosphere, changes in the physical properties of the air, such as temperature, pressure, and humidity, produce damping and scattering of the directed sound waves. You already remember from chapter two that loudness or intensity is measured in decibels. The softest audible sound to humans is 0 decibels, while painful sounds are those that rise above 120 decibels. 0 dB marks the threshold of hearing in human beings. 120 dB marks the threshold of pain in human beings.

6.4. QUALITY

If A above middle C is played on a violin, a piano, and a tuning fork, all at the same volume, the tones are identical in frequency and amplitude, but different in quality. Of these three sources, the simplest tone is produced by the tuning fork; the sound in this case consists almost entirely of vibrations having frequencies of 440 Hz. Because of the acoustical properties of the ear and the resonance properties of the ear's vibrating membrane (ear drum or tympanic membrane), however, it is doubtful that a pure tone reaches the inner hearing mechanism in an unmodified form. The principal component of the note produced by the piano or violin also has a frequency of 440 Hz, but these notes also contain components with frequencies that are exact multiples of 440, called **overtones**, at 880, 1320, and 1760 Hz, for example. The exact intensity of these other components, which are called **harmonics** (i.e., formant frequencies or **formants**), determines the quality, or timbre, of the note. The technical term "harmonics" refers to a series of subsidiary vibrations that accompany a primary, or fundamental (i.e., wave-motion) vibration.

Harmonics are most notably observed in musical instruments. They result when the vibrating body, for example, a stretched string or an enclosed air column, vibrates simultaneously as a whole and in equal parts (halves, thirds, fourths, and so on), producing wave frequencies that are in simple ratios with the fundamental frequency (2:1, 3:1, 4:1, and so on). In musical sound the full-length vibration produces the fundamental tone (or **first harmonic** or first partial), which is usually perceived as the **basic pitch** of the musical sound. The **subsidiary vibrations** produce **faint overtones** (second and higher harmonics or partials). As the series progresses, the vibrating segments become smaller, the frequencies higher, and the musical pitches closer together. The harmonic series for the tone **C** is given in the accompanying notation; black notes show pitches that do not correspond exactly with the Western tuning system. Harmonics contribute to the ear's perception of the quality, or timbre, of a sound: On a flute, certain harmonics of the series are most prominent; on a clarinet, others.

In "overblowing" a wind instrument, the player isolates and makes predominant one of the higher harmonics, thus extending the range of the instrument upward. Unvalved brass instruments, such as the bugle, produce only the tones of the harmonic series; the valves of the trumpet and the slide of a trombone add extra tubing, giving the player a new fundamental wavelength with a new harmonic series; the instrument can thus produce more tones. The harmonics of string players are flutelike tones produced when they cause the string to vibrate solely in halves (or thirds, and so on). Harmonics are a specialized class of overtones; vibrating membranes and solids, such as bells, produce complex partials that lie outside the harmonic series.

7. SPEED OF SOUND

The frequency of a sound wave is a measure of the number of waves passing a given point in one second. The distance between two successive crests of the wave is called the wavelength. The product of the wavelength and the frequency equals the speed of the wave. The speed is the same for sounds of all frequencies and wavelengths (assuming the sound is propagated through the same medium at the same temperature). The speed of sound in dry, sea level air at a temperature of 0°C (32°F) is 332 m/sec (1,088 ft/sec). The speed of sound in air varies under different conditions. If the temperature is increased, for example, the speed of sound increases; thus, at 20°C (68°F), the speed of sound is 344 m/sec (1,129 ft/sec). The speed of sound is different in other gases of greater or lesser density than air. The molecules of some gases, such as carbon dioxide, are heavier and move less readily than molecules of air. Sound progresses through such gases more slowly. Conversely, sound travels through helium and hydrogen faster than through air because atoms of helium and hydrogen are lighter than molecules of air.

Sound generally moves much faster in liquids and solids than in gases. In both liquids and solids, density has the same effect as in gases. The speed of sound also varies directly as the square root of the elasticity of the medium. Elasticity is the ability of a substance to regain its original shape and size after being deformed. The speed of sound in water, for example, is slightly less than 1,525 m/sec (5,000 ft/sec) at ordinary temperatures—almost five times as fast as in air. The speed of sound in copper, which is more elastic than water, is about 3,350 m/sec (about 11,000 ft/sec) at ordinary temperatures. In steel which is even more elastic sound moves at a speed of about 4,880 m/sec (about 16,000 ft/sec). Sound is propagated very efficiently in steel.

8. REFRACTION, REFLECTION, AND INTERFERENCE

Sound moves forward in a straight line when traveling through a medium having uniform density. Like light, however, sound is subject to refraction, which bends sound waves from their original path. In polar regions, where air close to the ground is colder than air that is somewhat higher, a rising sound wave entering the warmer less dense region, in which sound moves with greater speed, is bent downward by refraction. The excellent reception of sound downwind and the poor reception upwind are also due to refraction. The velocity of wind is generally greater at an altitude of many meters than near the ground; a rising sound wave moving downwind is bent back toward the ground, whereas a similar sound wave moving upwind is bent upward over the head of the listener.

Sound is also governed by reflection, obeying the fundamental law that the angle of incidence equals the angle of reflection. An echo is the result of reflection of sound. Sonar depends on the reflection of sounds propagated in water. A megaphone is a funnel-like tube that forms a beam of sound waves by reflecting some of the diverging rays from the sides of the tube. A similar tube can gather sound waves if the large end is pointed at the source of the sound; an ear trumpet is such a device.

Sound is also subject to diffraction and interference. If sound from a single source reaches a listener by two different paths—one direct and the other reflected—the two sounds may reinforce one another; but if they are out of phase they may interfere, so that the resultant sound is actually less intense than the direct sound without reflection. Interference paths are different for sounds of different frequencies, so that interference produces distortion in complex sounds. Two sounds of different frequencies may combine to produce a third sound, the frequency of which is equal to the sum or difference of the original two frequencies.

9. TYPES OF ORDINARY SOUND

In **speech**, **music**, and **noise**, pure tones are seldom heard. A musical note contains, in addition to a fundamental frequency, higher tones that are harmonics of the fundamental frequency. Speech contains a complex mixture of sounds, some (but not all) of which are in harmonic relation to one another. Noise consists of a mixture of many different frequencies within a certain range; it is thus comparable to white light, which consists of a mixture of light of all different colors. Different noises are distinguished by different distributions of energy in the various frequency ranges.

When a musical tone containing some harmonics of a fundamental tone, but missing other harmonics or the fundamental itself, is transmitted to the ear, the ear forms various beats in the form of sum and difference frequencies, thus producing the missing harmonics or the fundamental not present in the original sound. These notes are also harmonics of the original fundamental note. This incorrect response of the ear may be valuable. Sound-reproducing equipment without a large speaker, for example, cannot generally produce sounds of pitch lower than two octaves below middle C; nonetheless, a human ear listening to such equipment can resupply the fundamental note by resolving beat frequencies from its harmonics. Another imperfection of the ear in the presence of ordinary sounds is the inability to hear high-frequency notes when low-frequency sound of considerable intensity is present. This phenomenon is called masking.

In general, reproductions of speech and musical themes are recognizable even if only a portion of the frequencies contained in the originals are copied. Frequencies between 250 and 3,000 Hz, the frequency range of ordinary telephones, are normally sufficient. A few speech sounds, such as /ð/ require frequencies as high as 6,000 Hz. For high quality reproduction, however, the range of about 100 to 10,000 Hz is necessary. Sounds produced by a few musical instruments can be accurately reproduced only by adding even lower frequencies, and a few noises can be reproduced only at somewhat higher frequencies.

10. SPEECH SIGNAL PROCESSING

Closely related to acoustic phonetics is the notion of speech signal processing. Analog and digital processing of speech signals belongs to the realm of the engineering sciences and aims at the extraction of information from the acoustical speech wave. In addition to developing (and continuously refining!) methods of direct measurement and observation, researchers in the field of speech signal processing also deal with the important tasks of signal representation and signal transformation, both of which are indispensable prerogatives for a variety of applications such as speech storage and transmission, speech synthesis, speech recognition, speaker identification and verification, speech signal enhancement and the development of aids to the handicapped.

EXERCISE

1. Ensure that you understand the following terms:

waveform	amplitude	sound quality
cognitive phonetics	basic pitch	transverse wave
longitudinal wave	subsidiary vibrations	interference
rarefaction	formant	noise
cycle	intensity	harmonic
faint overtone	frequency	

2. How is frequency measured?

3. Define the following terms:

crest
wave length
trough
duration
pitch
loudness

4. The following terms are related to acoustic phonetics, but they do not appear in this chapter. Search the library or the Internet for books or papers on acoustic phonetics, and try to find brief definitions for these terms:

periodic vibration
sinusoidal vibration
white noise
damped vibration
resonant frequency
spectral envelop
locus
mel
discrete Fourier transform
phantom fundamental
digitization
quantization
sampling rate

CHAPTER EIGHT
AUDITORY PHONETICS[1]

1. INTRODUCTION

In chapter one, phonetics was defined to be a branch of linguistics which is concerned with the production, physical nature, and perception of speech sounds. We also divided phonetics into three main branches: auditory, acoustic, and articulatory. Auditory phonetics is the field involved in determining how speech sounds are perceived by the human ear.

Ear is the organ of hearing and balance. Only vertebrates, or animals with backbones, have ears. Invertebrate animals, such as jellyfish and insects, lack ears, but have other structures or organs that serve similar functions. The most complex and highly developed ears are those of mammals, or animals that have breasts.

2. STRUCTURE OF THE HUMAN EAR

Like the ears of other mammals, the human ear consists of three sections: the **outer**, **middle**, and **inner** ear. The outer and middle ears function only for hearing, while the inner ear also serves the functions of balance and orientation.

The outer ear includes the auricle (pinna), the visible part of the ear that is attached to the side of the head, and the waxy, dirt-trapping auditory canal. The tympanic membrane (eardrum) separates the external ear from the middle ear which is an air-filled cavity. Bridging this cavity are three small bones—the

---

[1] Most of the discussions of this chapter were reprinted from *Encyclopedia Encarta 2004* © 1993-2003 Microsoft Corporation.

malleus (hammer), the **incus** (anvil), and the **stapes** (stirrup). The cochlea and semicircular canals make up the inner ear.

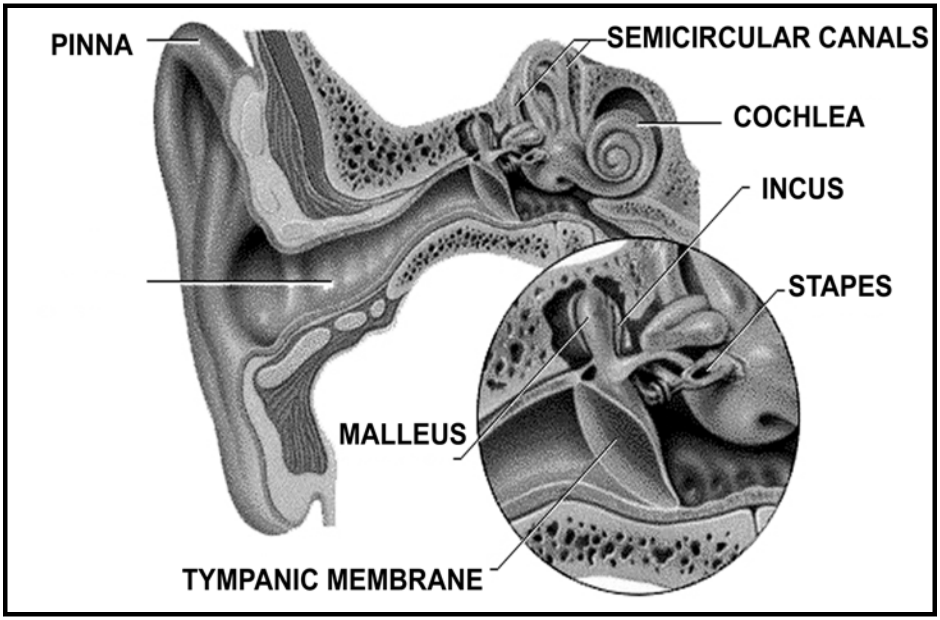


Figure 8.1: Structure of the human ear (Reprinted from *Encyclopedia Encarta 2004* © 1993-2003 Microsoft Corporation).

2.1. OUTER EAR

The outer ear is made up of the auricle, or pinna, and the outer auditory canal. The auricle is the curved part of the ear attached to the side of the head by small ligaments and muscles. It consists largely of elastic cartilage, and its shape helps collect sound waves from the air. The earlobe, or lobule, which hangs from the lower part of the auricle, contains mostly fatty tissue.

The outer auditory canal, which measures about 3 cm (about 1.25 in) in length, is a tubular passageway lined with delicate hairs and small glands that produce a wax-like secretion called **cerumen**. The canal leads from the auricle to a thin taut membrane called the eardrum or tympanic membrane, which is nearly round in shape and about 10 mm (0.4 in) wide. It is the vibration of the eardrum that sends sound waves deeper into the ear, where they can be processed by complex organs and prepared for transmission to the brain. The cerumen in the outer auditory canal traps and retains dust and dirt that might otherwise end up on the eardrum, impairing its ability to vibrate.

The inner two-thirds of the outer auditory canal is housed by the temporal bone, which also surrounds the middle and inner ear. The temporal bone—the hardest in the body—protects these fragile areas of the ear.

2.2. MIDDLE EAR

The eardrum separates the outer ear from the middle ear. A narrow passageway called the **eustachian tube** connects the middle ear to the throat and the back of the nose. The eustachian tube helps keep the eardrum intact by equalizing the pressure between the middle and outer ear. For example, if a person travels from sea level to a mountaintop, where air pressure is lower, the eardrums may cause pain because the air pressure in the middle ear becomes greater than the air pressure in the outer ear. When the person yawns or swallows, the eustachian tube opens, and some of the air in the middle ear passes into the throat, adjusting the pressure in the middle ear to match the pressure in the outer ear. This equalizing of pressure on both sides of the eardrum prevents it from rupturing.

The middle ear is a narrow, air-filled chamber that extends vertically for about 15 mm (about 0.6 in) and for nearly the same distance horizontally. Inside this chamber is a linked chain of three **ossicles**, or very small bones. Both the Latin and common names of these bones are derived from their shapes. They are called the malleus, or hammer; the incus, or anvil; and the stapes, or stirrup, which is the tiniest bone in the body, being smaller than a grain of rice.

The hammer is partly embedded in the eardrum, and the stirrup fits into the oval window, a membrane that fronts the inner ear. Vibrations of the eardrum move the hammer. The motion of the hammer moves the anvil, which in turn moves the stirrup. As sound vibrations pass from the relatively large area of the eardrum through the chain of bones, which have a smaller area, their force is concentrated. This concentration amplifies, or increases, the sound just before it passes through the oval window and into the inner ear. When loud noises produce violent vibrations, two small muscles, called the tensor tympani and the stapedius, contract and limit the movement of the ossicles, thus protecting the middle and inner ear from damage.

2.3. INNER EAR

The chain of bones in the middle ear leads into the convoluted structures of the inner ear, or labyrinth, which contains organs of both **hearing** and **balance**. The three main structures of the inner ear are the **cochlea**, the **vestibule**, and the three **semicircular canals**.

The cochlea is a coiled tube that bears a close resemblance to the shell of a snail, which is what the word means in Greek. Along its length the cochlea is divided into three fluid-filled canals: the vestibular canal, the cochlear canal, and the tympanic canal. The partition between the cochlear canal and the tympanic canal is called the **basilar membrane**. Embedded in the basilar membrane is the spiral-shaped **organ of Corti**. The sensory cells in the organ of Corti have thousands of hairlike projections that receive sound vibrations from the middle ear and send them on to the brain via the **auditory nerve**. In the brain they are recognized and interpreted as specific sounds.

3. HEARING

Sound is a series of vibrations moving as waves through air or other gases, liquids, or solids. A ringing bell, for example, sets off vibrations in the air. Detection of these vibrations, or sound waves, is called hearing. The detection of vibrations passing through the ground or water is also called hearing. Some animals can detect only vibrations passing through the ground, and others can hear only vibrations passing through water.

Humans, however, can hear vibrations passing through gases, solids, and liquids. Sometimes sound waves are transmitted to the inner ear by a method of hearing called bone conduction. For example, people hear their own voice partly by bone conduction. The voice causes the bones of the skull to vibrate, and these vibrations directly stimulate the sound-sensitive cells of the inner ear. Only a relatively small part of a normal person's hearing depends on bone conduction. If you crunch a biscuit in your mouth, for instance, with your mouth tightly closed, you can hear the noise made by your chewing act through bone conduction.

Humans hear primarily by detecting airborne (i.e., carried by air molecules) sound waves, which are collected by the auricles. The auricles also help locate the direction of sound. Although some people have auricular muscles so well-developed that they can wiggle their ears, human auricles, when compared to those of other mammals, have little importance. Many mammals, especially those with large ears, such as rabbits, can move their auricles in many directions so that sound can be picked up more easily.

After being collected by the auricles, sound waves pass through the outer auditory canal to the eardrum, causing it to vibrate. The vibrations of the eardrum are then transmitted through the ossicles, the chain of bones in the middle ear. As the vibrations pass from the relatively large area of the eardrum through the chain of bones, which have a smaller area, their force is concentrated. This concentration amplifies, or increases, the sound.

When the sound vibrations reach the stirrup, the stirrup pushes in and out of the oval window. This movement sets the fluids in the vestibular and tympanic canals in motion. To relieve the pressure of the moving fluid, the membrane of the oval window bulges out and in. The alternating changes of pressure in the fluid of the canals cause the basilar membrane to move. The organ of Corti, which is part of the basilar membrane, also moves, bending its hairlike projections. The bent projections stimulate the sensory cells to transmit impulses along the auditory nerve to the brain.

Human ears are capable of perceiving an extraordinarily wide range of changes in **loudness**, the tiniest audible sound being about one trillion times less intense than a sound loud enough to cause the ear pain. The loudness or intensity of a noise is measured in a unit called the decibel. The softest audible sound to humans is 0 decibels, while painful sounds are those that rise above 140 decibels. The decibel is a unit of measure (abbreviated dB) originally used to compare sound intensities and subsequently electrical or electronic power outputs. Today, it is also used to compare voltages. An increase of 10 dB is equivalent to a 10-fold increase in intensity or power, and a 20-fold increase in voltage. A whisper has an intensity of 20 dB. A jet aircraft taking off nearby (140 dB) is the threshold of pain.

DECIBELS		TYPICAL SOUND
0	dB	threshold of hearing
10	dB	rustle of leaves in gentle breeze
10	dB	quiet whisper
20	dB	average whisper
20-50	dB	quiet conversation
40-45	dB	hotel; theater (between performances)
50-65	dB	loud conversation
65-70	dB	traffic on busy street
65-90	dB	train
75-80	dB	factory (light/medium work)
90	dB	heavy traffic
90-100	dB	thunder
110-140	dB	jet aircraft at takeoff
130	dB	threshold of pain
140-190	dB	space rocket at takeoff

Besides loudness, the human ear can detect a sound's **pitch**, which is related to a sound's vibration frequency, or the number of sound waves passing into the ear in a given period. The greater the frequency, the higher the pitch. The maximum range of human hearing includes sound frequencies from about 15 to about 18,000 waves, or cycles, per second (i.e., 15 Hz to 18000 Hz). Because the human ear cannot hear very low frequencies, the sound of one's own heartbeat is inaudible. At the other end of the scale, a highly pitched whistle producing 30,000 cycles per second is not audible to the human ear, but a dog can hear it.

The third characteristic of sound detected by the human ear is **tone**. The ability to recognize tone enables humans to distinguish a violin from a clarinet when both instruments are playing the same note. The least noticeable change in tone that can be picked up by the ear varies with pitch and loudness.

Another sonic phenomenon, known as **masking**, occurs because lower-pitched sounds tend to deafen the ear to higher-pitched sounds. To overcome the effects of masking in noisy places, people are forced to raise their voices.

The ear, oddly enough, not only receives sounds but can generate its own noises, called spontaneous otoacoustic emissions. About 30 percent of all people experience them. Some hear their ears' sounds, typically a faint buzzing, if they are in a completely quiet room, but most people do not notice them at all. In particularly bad cases, people have walked into doctors' offices complaining of a ringing in their ears, and the doctors could hear the noises while standing next to the patients. Such sounds, while certainly annoying, have never been thought to be more than that. But psychologists at the State University of New York at Buffalo have found that otoacoustic emissions may actually produce an unusual type of hearing loss, one that occurs without physical damage to any parts of the ear.

The researchers' evidence comes not from humans but from chinchillas. These small rodents are prized by some people for their long, silky fur, but to a sensory psychologist they have a different appeal: they have a hearing range similar to that of humans. Psychologist Salvi and his colleagues were experimenting with a chinchilla when, with a small microphone, they discovered a sound coming from the animal's ear. The rodent's ears were producing otoacoustic emissions at a frequency of 4,000 hertz (vibrations per second).

To see what effect these sounds might have on the chinchilla's hearing, the researchers placed microelectrodes in the animal's auditory nerve (an experiment they couldn't have done with humans) to test it for sensitivity to specific frequencies. They found that the chinchilla was insensitive to external sounds at frequencies around 4,000 hertz. The constant buzz from its own ears was apparently causing its brain to ignore that frequency.

Otoacoustic emissions are thought to be produced by delicate microscopic hair cells in the outer ear. Why they are produced by chinchillas and people, and why they are produced by some people and not by others, remains largely a mystery. But Salvi now thinks the sounds create a sort of busy signal in the auditory nerve. “The ear is the cause of its own hearing loss,” he says. “The hearing loss is the result of the masking effect only, and there is no cellular or long-term damage to the ear.” Salvi next hopes to find out whether humans experience similar types of hearing loss.

4. SPEECH PERCEPTION

Fromkin and Rodman (1988) provide the simplest explanation for speech perception. They argue that when you hear a car backfire, you may wonder whether the sound represented a gunshot or a backfiring. Your perception of the particular acoustic signal and your knowledge of what creates different sounds result in your assigning some "meaning" to the sounds you heard. Similarly, when you hear the sounds represented by the phonetic transcription [ʤæk ɪz ən 'i:dɪət], you assign the meaning "Jack is an idiot" to the sound signal. The acoustic signal, however, does not reach our ears in phonemic segment-sized chunks; it is a semicontinuous signal. In order for us to process it as speech, it must be segmented into phonemes, words, phrases, and sentences.

Speech perception is a process by which we segment the continuous signal and, in so doing, may "mischunk" or misperceive the speaker's intended utterance. The difficulties inherent in speech perception are compounded by the fact that the speech signal for the "same" utterance varies greatly from speaker to speaker and from one time to the next by the same speaker. Nevertheless the brain is able to analyze these different signals, conclude that they are the same linguistically, segment the utterance into a phonetic/phonological string of words, "look up" the meaning of these words in the mental dictionary, technically called the lexicon, analyze the linear string of words into a hierarchical syntactic structure, and, most of the time, end up with the intended meaning.

All this work is done so quickly we are unaware that it is going on at all. Despite the variation between speakers and occurrences, there must be certain invariant features of speech sounds that permit us to perceive a /d/ or an /ə/ produced by one speaker as identical phonologically with a /d/ and /ə/ produced by another. The relations between the formants (i.e., a frequency range where vowel sounds are at their most distinctive and characteristic pitch) of the vowels of one speaker are similar to those of another speaker of the same language, even though the absolute frequencies may differ.

When a stop consonant is produced, the signal is interrupted slightly, and the frequency of the "explosion" that occurs at the release of the articulators in producing stop consonants differs from one consonant to another. The transitions between consonants and vowels provide important information as to the identity of the consonants. After voiceless consonants, the onset of vowel formants starts at higher frequencies than after voiced consonants. Different places of articulation influence the starting frequencies of formant onsets.

There are many such acoustic cues that, together with our knowledge of the language we are listening to, permit us to perform a "phonetic analysis" on the incoming acoustic signal. Confusions may also be disambiguated by visual, lexical, syntactic, and semantic cues. Speech communication often occurs in a noisy environment, but we can still pick out of the sound signal those aspects that pertain to speech. We are thus able to ignore large parts of the acoustic signal in the process of speech perception, which has led to the view that the human auditory system—perhaps in the course of evolution—has developed a special ability to detect and process speech cues.

To understand an utterance we must, in some fashion, retrieve information about the words in that utterance, discover the structural relationship and semantic properties of those words, and interpret these in the light of the various pragmatic and discourse constraints operating at the time. Further, all of this takes place at a remarkably rapid pace. Analyzing the speech signal in speech perception is a necessary but not sufficient step in understanding a sentence or utterance. Suppose you heard someone say:

A sniggle blick is procking a slar

and were able to perceive the sounds as

/ə snɪgəl blɪk ɪz prɒkɪn ə slɜː/

You would still be unable to assign a meaning to the sounds, because the meaning of a sentence depends on the meanings of its words, and the only English lexical items in this string are the morphemes *a, is,* and *-ing.* The sentence lacks any English content words. You can only know that the sentence has no meaning by attempting a lexical lookup of the phonological strings you construct; finding no entries for *sniggle, blick, prock,* or *slar* in your mental dictionary tells you that the sentence is composed of nonsense strings. If instead you heard someone say

The cat chased the rat

through a lexical lookup process you would conclude that an event concerning a cat, a rat, and the activity of chasing had occurred. Who chased whom is determined by syntactic processing. That is, processing speech to get at the meaning of what is said requires syntactic analysis as well as knowledge of lexical semantics. Stress and intonation provide some cues to syntactic structure. We know, for example, that the different meanings of the sentences

He lives in the white house

He lives in the White House

can be signaled by differences in their stress patterns. Relative loudness, pitch, and duration of syllables thus provide important information in speech perception.

5. SENSATIONS OF TONE

The ears of an average young person are sensitive to all sounds from about 15 Hz to 20,000 Hz. The hearing of older persons is less acute, particularly to the higher frequencies. The degree to which a sensitive ear can distinguish between two pure notes of slightly different loudness or slightly different frequency varies in different ranges of loudness and frequency. Differences of about 20 percent in loudness (about one decibel, dB), and 0.33 percent in frequency (about 1/20 of a note) can be distinguished in sounds of moderate intensity at the frequencies to which the ear is most sensitive (about 1,000 to 2,000 Hz). In this same range, the difference between the softest sound that can be heard and the loudest sound that can be distinguished as sound (louder sounds are “felt,” or perceived, as painful stimuli) is about 120 dB (about one trillion times as loud).

All of these sensitivity tests refer to pure tones, such as those produced by a tuning fork. Even for such pure tones the ear is imperfect. Notes of identical frequency but greatly differing intensity may seem to differ slightly in pitch. More important is the difference in apparent relative intensities with different frequencies. At high intensities the ear is approximately equally sensitive to most frequencies, but at low intensities the ear is much more sensitive to the middle high frequencies than to the lowest and highest. Thus, sound-reproducing equipment that is functioning perfectly will seem to fail to reproduce the lowest and highest notes if the volume is decreased.

EXERCISE

1. Ensure that you understand the following terms:

auditory phonetics	ossicles	organ of Corti
stapes	hearing	auditory nerve
incus	cochlea	bone conduction
malleus	vestibule	loudness
cerumen	semicircular canals	masking
eustachian tube	basilar membrane	

2. What is a decibel and how does it relate to sound pressure level and acoustic intensity?

3. Review the following terms:

hearing
speech perception
sensation of tone
tone
loudness
balance

BIBLIOGRAPHY

Bird, S. & Stegen, O. (1993). *Principles of F0 measurement*. Poster paper, Fourth Conference on Laboratory Phonology. Oxford.

Bird, S., & Stegen, O. (1995). *The Bamileke Dschang associative construction: instrumental findings* (EUCCS/RP-66). Centre for Cognitive Science, University of Edinburgh.

Clements, N. G. (1981). Downdrift in a tone language with four tone levels: a report on some preliminary measurements. Unpublished ms. (Published in *York Papers in Linguistics*, Vol. 15 1991)

Clements, G. N. (1990). The status of register in intonation: comments on the paper by Ladd and by Inkelas and Leben. In J. Kingston & M. Beckman (Ed.), *Papers in Laboratory Phonology I: Between the grammar and physics of speech* (pp. 58-71). Cambridge: Cambridge University Press.

Connell, B. A. & Ladd, D. R. (1990). Aspects of Pitch Realization in Yoruba. *Phonology*, *7* 1, 1-29.

Falk, J. S. (1978). *Linguistics and language: A survey of basic concepts and implications*. New York: John Wiley & Sons.

Fromkin, V., Rodman, R., & Hyams, N. (2003). *An introduction to language*. United States: Thomson.

Gandour, J., Potisuk, S. & Dechongkit, S. (1994). Tonal coarticulation in Thai. *Journal of Phonetics*, *22* 4, 477-492.

Gussenhoven, C. & Rietveld, A. (1988). Fundamental frequency in Dutch: Testing three hypotheses. *Journal of Phonetics*, *16* 355-369.

Hombert, J.-M. (1974). Universals of Downdrift: their phonetic basis and significance for a theory of tone. *Studies in African Linguistics*, *Suppl. 5* 169-183.

Hombert, J.-M. (1977). Consonant types, vowel height, and tone in Yoruba. *Studies in African Linguistics*, *8* 2, 173-190.

Hombert, J.-M. (1978). Consonant types, vowel quality, and tone. In F. V. (Ed.), *Tone: a linguistic survey* (pp. 77-111). New York: Academic Press.

Katamba, F. (1988). *An introduction to phonology*. London: Longman Group UK Limited.

Ladd, D. R. (1984). Declination: a review and some issues. *Phonology Yearbook*, *1* 53-74.

Ladd, D. R. (1990). Metrical representation of pitch register. In J. Kingston & M. Beckman (Ed.), *Papers in Laboratory Phonology I: Between the grammar and physics of speech* Cambridge: Cambridge University Press.

Ladd, D. R. (1992). An introduction to intonational phonology. In G. J. Docherty & D. R. Ladd (Ed.), *Papers in Laboratory Phonology II: Gesture, segment, prosody.* Cambridge: Cambridge University Press.

Ladd, D. R. & Johnson, C. (1987). 'Metrical' factors in the scaling of sentence-initial accent peaks. *Phonetica*, *44* 238-245

Ladefoged, P. (1982). *A course in phonetics.* New York: Harcourt, Brace Jovanovich.

Laniran, Y. (1992a). *Intonation in a tone language: the phonetic implentation of tone in Yoruba*. Ph.D., Cornell,

Laniran, Y. (1992b). Phonetic aspects of tone realization in Igbo. *Progress Reports from Oxford Phonetics*, *5* 35-51.

Laver, J. (1994). *Principles of Phonetics*. Cambridge: Cambridge University Press.

Leben, W. R. (1996). *Prosodic domains and the theory of tone*. Paper presented to the 27th ACAL, Gainesville, FL.

Liberman, M. & Pierrehumbert, J. (1984). Intonational invariance under changes in pitch range and length. In M. Aronoff & R. Oehrle (Ed.), *Language and Sound Structure* (pp. 157-233). Cambridge MA: MIT.

Liberman, M., Schultz, J. M., Hong, S. & Okeke, V. (1993). The phonetic interpretation of tone in Igbo. *Phonetica*, *50* 147-160.

Lieberman, P., & Blumstein, S. E. (1988). *Speech physiology, speech perception, and acoustic phonetics*. Cambridge: Cambridge University Press.

Lindau, M. (1986). Testing a model of intonation in a tone language. *Journal of the Acoustical Society of America*, *80* 757-764

Microsoft Corporation. (2004). Microsoft Encarta reference library [Computer Software]. Redmond, WA.: Microsoft Corporation.

Ohala, J. J. (1978). The production of tone. In V. Fromkin (Ed.), *Tone: a linguistic survey* (pp. 5-39). New York: Academic Press.

Perrin, M. J. (1974). Mambila. In J. Bendor-Samuel (Ed.), *Ten Nigerian Tone Systems* (pp. 93-108). Jos: Institute of Linguistics.

Perrin, M. (n.d.). *The tone system in Mambila: some further comments*.

Pierrehumbert, J. & Beckman, M. (1988). *Japanese Tone Structure*. Cambridge MA: MIT Press.

Snider, K. & van der Hulst, H. (1993). Issues in the representation of tonal register. In H. van der Hulst & K. Snider (Ed.), *The Phonlogy of Tone* (pp. 1-27). Berlin: Mouton de Gruyter.

Stewart, J. M. (1983). Downstep and Floating Low Tones in Adioukrou. *Journal of African Languages and Linguistics*, *5* 57-78.

Welmers, W. (1973). *African Language Structures*. Berkeley: University of California Press.

Zee, E. (1978). The interaction of tone and vowel quality. *UCLA Working Papers in Phonetics*, *41* 53-67.