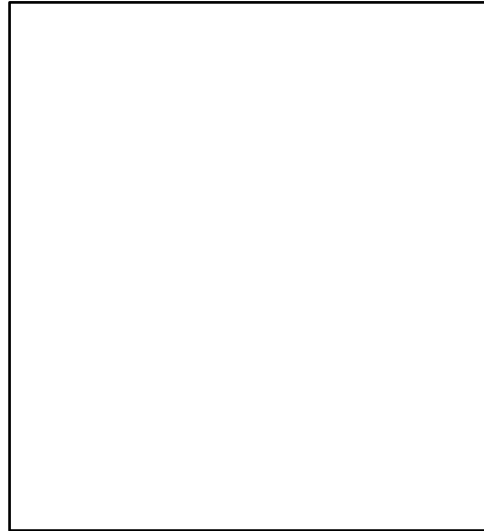


THE SPEECH-PRODUCING MECHANISM

Speech sounds are sound waves created in a moving stream of air. The air is expelled from the lungs, passes through the two vocal cords in the larynx (Adam's apple), and proceeds upward. This moving stream has two possible outlets. It can pass through the nasal cavity and emerge through the nose or it can pass through the oral cavity and come out through the mouth. Normally, the air stream can pass through one



Speech-producing mechanism

DIAGRAM 1

Meaning of abbreviations:

AR=alveolar ridge

NC=nasal cavity

OC=oral cavity

P=palate

TT=tongue tip

TF=tongue front

TM=tongue middle

TB=tongue back

V=velum

of the cavities because in speech one of them is ordinarily closed. In the production of oral sounds, the velum is raised to touch the back of the throat, cutting off the nasal cavity. The velum is a movable curtain of flesh which ends in a pear-shaped pendant, called the uvula. The velum is situated just behind the hard palate, which is a bony structure along the roof of the mouth. An example of the oral sound can be shown by the diagram 2 in the production of a-a-a-a-a-a.

In the production of the three nasal sounds of English, the velum rests in a hanging position as in diagram 1, and the oral cavity is blocked in either of the following three ways: with the lips (diagram 3), with the tongue tip (diagram 4), or with the tongue back (diagram 5).

With the oral cavity blocked off, the sound can emerge only through the nasal cavity. It is evident that every speech sound is either an oral or nasal sound. Sometimes,

a 'nasal twang' is heard. This is caused by the slight lowering of velum for sounds that are normally oral, thus permitting some of the air to go out through the nasal cavity. The three nasal sounds in English are symbolized by speech notations /m/, /n/, and /ŋ/. The /m/ is a bilabial nasal, made by the closing of lips. The /n/ is an alveolar nasal, made by stopping the flow of air with the tongue tip against the alveolar ridge. The /ŋ/ is a velar nasal, made by stopping the flow of air with the back of the tongue against the velum. In all three the air passes through the nasal cavity.

The sound is being originated by the vocal cords. They are two short bands of flesh and muscle stretching from front to rear, inside the larynx (Adam's apple). In breathing and during the production of some speech sounds, like f and s, these are held open, allowing free ingress and egress of air (diagram 6). But with many sounds they are pressed tightly together, and the air passing between them causes them to vibrate (diagram 7).

These vibrations are given resonance by the cavities of the mouth and nose and the result is the phenomenon called voicing. In the production of every speech sound, these vocal cords are either vibrating or not vibrating. If they are vibrating, the sound is called voiced. If they are not vibrating, the sound is called voiceless.



Diagram 2. Air passing through oral cavity.
tongue position for /a/.



Diagram 3. Air passing through nasal cavity.
lip position for /m/.

Exercise: (i) Nasal sounds: Pronounce and hold the final sounds of the following words: [rim, saw, bin, see, sing, tall, trim, pain, wrong]

As you hold the final sounds, stop the nose with your fingers, if this blocks the sound, the sound is a nasal.

(ii) Oral sounds: do the same with the above given words, if the sound continues even after blocking your nose then close your lips. If the sound cuts off, it is an oral sound.



Diagram 4. Air passing through nasal cavity.
tongue position for /n/.

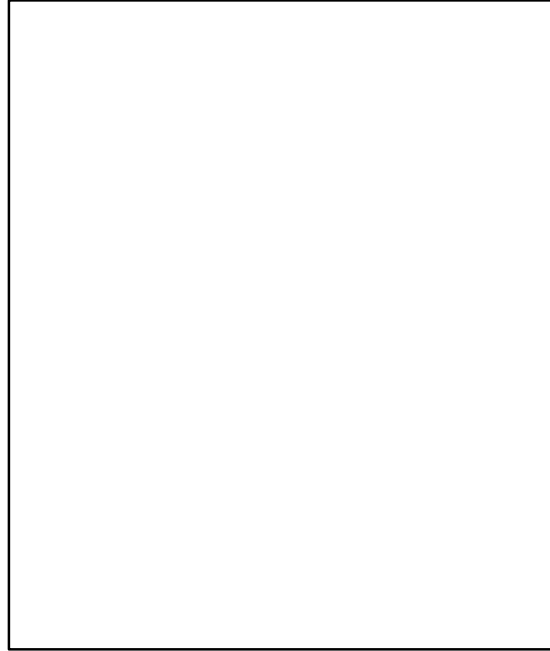


Diagram 5. Air passing through nasal cavity.
tongue position for /l/.

Diagram 6. position of vocal cords during exhalation.

Diagram 7. position of vocal cords when vibrating.

For voiced and voiceless words

hold your hand tightly over your ears and pronounce the last sound of the following words-less, hum, if, pin, sheath, among, mush, fin, song. Now try the same thing with the first sound of the following words: fine, vine, thin, then, seal, zeal, shock, late, rate. If you hear the hum of the vibration of your vocal cords, the words are voiced. If the hum is absent, then the sounds are voiceless.

THE PHONEME

"The phoneme is a speech sound that makes a difference in meaning." An example can be cited to make the definition clearer. The words dime and dine sound exactly alike except for the /m/ and the /n/, yet their meanings are different. Therefore, it must be the /m/ and /n/ which make the difference in meaning, and these two nasals are thereby established as English phonemes. Likewise, there is only one difference the sounds sin and sing: sin ends in the alveolar nasal /n/ and sing in the velar nasal /ŋ/. This contrast is evidence that /n/ and /ŋ/ are both phonemes. Pairs of words like those above which demonstrate a phonemic contrast are called minimal pairs. A phoneme may be pronounced in different ways, depending on its position in the utterance, and still remain the same phoneme. As an example, /l/ sound in the words 'lit' and 'well' is pronounced differently. The second /l/ seems to be pronounced with an 'uh' sound preceding it, whereas the first /l/ in lit is pronounced with the tongue tip touching the alveolar ridge and hence making it sound a little metallic. If someone pronounces 'well' with the /l/ of 'lit', the word will sound different, a little un-English, but the meaning will not be changed. The use of one or the other of these two /l/s never makes a difference in meaning; hence they are not two phonemes but merely variations of the /l/ phoneme. There is still another /l/ in words like 'play' and 'sled'. Here there may be a voiceless [l̥], whereas the [l]s of both well and lit were voiced. But whether the words well, lit, play, and certain other words are pronounced with a voiced or voiceless [l], the meaning remains unchanged; so this third [l̥] is another variant of the /l/ phoneme. Such variants of a phoneme are called allophones. Allophones are enclosed in brackets with the occasional addition of diacritical marks to indicate the exact pronunciation. Phonemes are enclosed in slants. Thus we may say that the /l/ phoneme has three allophones: [l̥] as in 'lit', [l] as in 'well', and [l] as in 'play'. In other words, a phoneme is not an individual sound but a small family of similar sounds.

THE ENGLISH PHONEMIC SYSTEM: VOWELS

It is difficult to define a vowel with precision, because the classification of English vowels is a complex and controversial matter. Four statements can be made about vowels to show their nature.

1. All vowels are oral sounds. In some dialects and in certain contexts, vowels may become partially nasal, but normally they are orals, not nasals.
2. All vowels are voiced.
3. Vowels are characterized by a free flow of air through the oral cavity.
4. The distinguishing features of the different vowels are determined largely by tongue position.

English may be said to have 12 vowels - five front, four back, and three central vowels.

FRONT VOWELS

In the pronunciation of the final sound of 'be', symbolized by /i/, the tongue front and middle is humped high in the mouth, leaving a narrow passage for the flow of



Front vowels

Diagram A

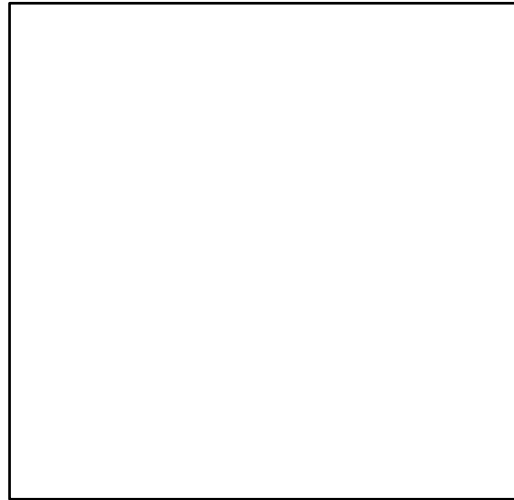
air between the hard palate and the surface of the tongue. The tongue position of /i/ is the top one on diagram A. But in the pronunciation of the first sound of 'add' symbolized by /a /, there happens a considerable drop in the tongue position. The jaw drops considerably and the tongue flattens. The tongue position of the vowel /a / is the bottom one on diagram A. Now, between these two extremes, /i/ and /a /, are three other vowels /I/, /e/, and /. The middle sound of the following words gives an account of the five front vowels in a descending order i.e. the top tongue position to the bottom one: beat /i/, bit /I/, bait /e/, bet / /, and bat /a /. These five vowels are called the front vowels, because they are formed in the front of the mouth by the position of the tongue front. For each front vowel the lips are spread, or unrounded. The tongue positions are shown on diagram A. English spelling cannot be used to represent accurately the speech sounds of English because of its inconsistencies. It is difficult, for example, to symbolize the vowel of 'bait' in English spelling. We cannot use 'ai' as in 'wait', 'eig' as in 'reign', 'ey' as in 'they', 'ay' as in 'say', 'a' as in 'late', 'ei' as in 'vein', 'au' as in 'gauge', 'ea' as in 'steak'. So to represent the sounds of words, we shall use a special alphabet in which one symbol always represents one and the same speech sound, and each sound is always represented by only one symbol. In this alphabet the five illustrative words in the preceding lines are written as follows:

beat = /bit/, bit = /bIt/, bait = /bet/, bet = /b t/, bat = /ba t/. The phonemic symbols and words written in these symbols are enclosed in slants, like /b t/.

BACK VOWELS

In the pronunciation of the final sound of 'too', symbolized by /u/, the lips are rounded and the back of the tongue is raised to a position near the velum, leaving a little

space for the air to flow. The tongue position of /u/ is the top one on diagram B. But in the pronunciation of the sound "aw", as in "Aw, come on," symbolized by /ɔ/, the tongue position is the bottom one on diagram B. Between these two extremes, there are two other vowels /U/ and /o/. There occurs a close rounding of the lips for /U/ and the open rounding for /ɔ/.



Back vowels

Diagram B

As the back of the tongue is lowered from the /u/ position, it reaches in turn the positions for the three other back vowels: /U/ as in pull, /o/ as in note, and /ɔ/ as in ought, law, and ball. And at each of these three positions the rounding of the lips is successively opened; the four back vowels, from top to bottom, are illustrated by this series:

fool = /fʊl/, full = /fʊl/, foal = /foʊl/, fall = /fɔl/.

These vowels are called the back vowels because they are formed in the back of the mouth by the position of the tongue back.

CENTRAL VOWELS

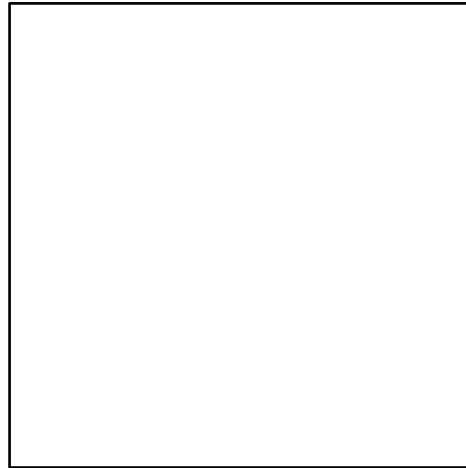
English has three central vowels. In the pronunciation of fur, sir, her, a central r-colored vowel is being uttered as the final sound. Majority of the Americans, who do not drop their 'r's, pronounce this. But there are other Americans who pronounce words like these with a /ɜ/ plus an r sound, as in hurry /hʌrɪ/, instead of with the single r-colored vowel. Thus a pair of symbols /r/ is used to represent both pronunciations - the single, central, r-colored vowel and also the schwa plus an r sound.

The tongue position of /ɜ/ is the top one on diagram C. The second central vowel may be illustrated by the first sound of up and upon. It is written /ʊ/, and the tongue position is the middle one on diagram C. It is heard as the pronunciation of the underlined vowels in the following words:

stressed: sun, done, flood,

Unstressed: sofa, alone, principal, specimen, science, kingdom, connect, difficult, suppose.

The /ə/ is a vowel of high frequency in English, especially in unstressed syllables, and is technically called "schwa"



Central vowels

Diagram C

The third central vowel is the sound we make when we open our mouth wide and say a-a-a-a-a. It is symbolized by /a/. In sounding this vowel, the mouth is widely opened and the tongue is nearly flat. The tongue position is the bottom one in diagram C. The central vowels from, top to bottom, are illustrated in the following series:

purr = /p r/, pup = /p p/, pot = /pat/

The twelve vowel phonemes of English can be seen in relation to one another on the vowel chart below.

This is a two-dimensional grid of tongue positions, the mouth being at the left and the throat at the right. This chart can give to the twelve vowels their descriptive names:

The names are these:

/i/ high front

/ɪ/ higher mid-central /U/ high back rounded

/I/ lower high front

// lower mid-central

/u/ lower high back rounded

/e/ higher mid front

/o/ mid back rounded

// lower mid front

/a/ low central

/a / low front

// low back rounded

It must be realized that the classification of vowels by tongue position is imprecise and generalized. Also, there are further classifications- tense, lax; close,

	front	central	back
high	i (beat)		u (moon)
	I (bit)		U (pull)
mid	e (say)	r (fur)	o (over)
	(set)	(up)	
low	a (sat)	a (father)	(fall)

CHART OF ENGLISH VOWEL PHONEMES

open; narrow, wide; long, short-that have been bypassed in this discussion for simplicity.

THE SYLLABLE

"A syllable is a sound or a short sequence of sounds which contains one peak of sonority." Certain sounds have a greater sonority or carrying power than others. For example, in soap /sop/, the /o/ has greater sonority than the /s/ or the /p/, even though all are spoken with equal force. If /sop/ is spoken at some distance from the listener, he may hear distinctly only the /o/. In potato /p teto/, the / /, /e/, and /o/ are more sonorous, more audible, than the /p/ and the /t/s. The sounds which have this greater sonority or carrying power are mostly the vowels.

Thus, in the utterance of connected discourse, there appears a peak-and-valley effect of sonority or audibility. The peaks of sonority are the vowels and the valleys of less distinctness are the consonants, as in 'create' /kri-et/. This peak of sonority is usually a vowel, and the vowel is said to be the center of the syllable. A segment of speech, then, contains as many syllables as there are peaks. Here are some examples of words with the peaks, or syllabic centers, underlined:

One syllable	be	/bi/
	string	/sti /
Two syllable	believe	/b liv/
	being	/biI /
	stirring	/st rI /
Three syllable	believing	/b livI /
Four syllable	unbelieving	/ nb livI /
Five syllable	unbelievably	/ nb livI li/

Apart from the vowels, which are considered to be the syllabic centers, there are four consonants - /m/, /n/, / /, and /l/- which also have considerable sonority and which can constitute syllables.

For example, the two-syllable expression stop'em can be uttered in two ways, the first is /stap m/. After the lips are closed to make the /p/, they are opened for the //, and closed again for the /m/. The second way is /stapm/. Here the lips are closed for /p/ and remain closed for the /m/. Since the lips close for /p/, no air can escape from the mouth. The closed lips shut off the mouth exit and the raised velum shuts off the nasal exit. For /m/, the velum is lowered and the air escapes with a slight explosion through the nasal cavity. This /m/ is a peak of sonority and constitutes a syllable. Hence, this /m/ is called a syllabic /m/.

The syllabic /n/ is formed similarly. This word can be pronounced two ways. For example, the tongue is dropped from the /t/ position on the alveolar ridge, uttering the //, and then replacing the tongue for the /n/ in /b t n/. for /b tn/ the tongue stays at the alveolar ridge. At the /t/ position the air is prevented from escaping by the tongue against the alveolar ridge and by the closed velum, which shuts off the nasal cavity. But if the tongue is in /t/ position and the velum is open, an /n/ is produced as the air escapes through the nasal cavity.

The syllabic /l/, which is less frequent, is heard in expression like "Jack and Jill" /Ja k Jil/.

Although it is easy to locate the peaks of sonority that indicate syllable centers, the vowels and syllabic consonants, it is sometimes impossible to find the boundary between syllable, that is, the point of minimum sonority. In the two-syllable 'hushing' /h s l /, for example, it is difficult to establish if the syllable boundary is after the //, before the /l/, or in the middle of /s/. It is like trying to establish in a valley the exact line separating the two hills on either side. For convenience, where the boundary is not audible, we can resort to an arbitrarily selected break.

DIPHTHONGS

A diphthong consists of two vowels which occur in the same syllable, the tongue moving smoothly from one position to the other without hiatus, as in 'sigh' /sai/, 'sow' (female pig) /sau/, and soy /s i/. The two vowels together represent the peak of sonority, though one always has greater prominence than the other. Many of the vowels are diphthongized in various subareas of English, and four of them are normally diphthongized in standard English: /i/, /e/, /u/, and /o/. There is, however, no phonemic difference between the pure vowels and the diphthongized vowels. According to the system used here, the diphthong phonemes are only three: /ai/ as in 'by', /au/ as in 'bough', and / i/ as in 'boy'.

	Bilabial	Labio-dental	Inter-dental	Alveolar	Alveo-palatal	Velar	Glottal
Stops V1	p			t		k	
Vd	b			d		g	
Fricatives V1		f			s		h
Vd		v			z		

Affricates	Vl			c	
	Vd			j	
Nasals	m		n		
Laterals			l		
Glides		r	y	w	

CHART OF ENGLISH CONSONANT PHONEMES

Vowels are characterized by a free flow of air. Consonants, on the other hand, are produced by stopping or obstructing this flow of air, except for the three nasals. The first six consonants presented in the chart /p b t d k g/ are produced by a stoppage of air.

STOPS: /p/, /b/: The consonant /p/ is produced by a sudden opening of the lips and an outward explosion of the air. This consonant is called a voiceless bilabial stop because -

1. the vocal cords do not vibrate.
2. two lips are used,
- and 3. a complete stop of the air flow is made.

If during the same process, there is a vibration of the vocal cords, the consonant /b/ is produced, which is a voiced bilabial stop.

/t/, /d/: instead of the lips, the air flow is stopped by holding the tongue against the alveolar ridge with the velum closed. A sudden removal of the tongue produces /t/ without the vocal cords vibrating. If the vocal cords are vibrating, then a /d/ is produced. So /t/ is a voiceless alveolar stop whereas /d/ is a voiced alveolar stop.

/k/, /g/: this pair of stops is produced by raising the tongue back against the velum, which is also raised to cut off the nasal cavity. When the tongue back is released, the outrushing air results in a /k/, a voiceless velar stop, or a /g/, a voiced velar stop, depending on whether or not the vocal cords are vibrating.

FRICATIVES: English contains nine consonants which are produced by an obstruction of the air stream causing audible friction. These nine fricatives are:

/f v s z θ ð h/

The first pair, /f/ and /v/, are heard in 'fail' and 'vale'. They are produced when the outgoing air is obstructed by the lower lip touching the upper teeth. The /f/ is called a voiceless labiodental fricative, and /v/ a voiced labiodental fricative. They differ only in the fact that /v/ is voiced.

The second pair, /θ/ and /ð/, are heard in 'ether' and 'either'. They are made with the tongue obstructing the air stream between the upper and lower teeth, or at the bottom of the upper teeth. The /θ/ is a voiceless interdental fricative, and /ð/ a voiced interdental fricative.

The third pair is /s/ and /z/, as in 'face' and 'faze'. These are pronounced by the tongue permitting a small stream of air to hiss over its surface at the alveolar ridge. The /s/ is a voiceless alveolar fricative, and /z/ a voiced alveolar fricative.

The fourth pair of fricatives are /ʃ/, the third consonant in 'dilution', and /ʒ/, the third consonant in 'delusion'. These are made by the friction of moving air between the

tongue front and the palatal region just behind the alveolar ridge. The /s/ is a voiceless alveopalatal fricative, and /z/ a voiced alveopalatal fricative.

The last fricative is /h/, as in 'hat' contrasted with 'at'. This is produced by the breath rushing through the open vocal cords. The fricative's tongue and lip position is that of the following vowel. The position of the tongue and the lips are different in the following sounds - ha, he, and who. It is called the voiceless glottal fricative, the glottis being the space between the vocal cords.

AFFRICATES: English has two affricates - the voiceless /tʃ/ as in 'chill' and the voiced /dʒ/ as in /jill/. The /tʃ/ begins with the voiceless stop /t/, which is exploded as a voiceless fricative /ʃ/. It is a voiceless alveopalatal affricate. The /dʒ/ consists of a voiced stop /d/, which is exploded as a voiced fricative /ʒ/. It is called the voiced alveopalatal affricate. /tʃ/ is sometimes written as /ts/ and /dʒ/ as /dz/.

NASALS: /m/, /n/, and /ŋ/, the /m/ is a bilabial nasal, made by closing the two lips. The /n/ is an alveolar nasal, made by stopping the flow of air with the tongue tip against the alveolar ridge. The /ŋ/ is a velar nasal, made by stopping the flow of air with the back of the tongue against the velum. In all three the air moves through the nasal cavity.

LATERAL: The lateral /l/ as in 'louse', is made by placing the tongue tip on the alveolar ridge and vibrating the vocal cords as the air passes out on one or both sides of the tongue.

GLIDES: The three glides - /y/, /r/, and /w/- are signaled by a moving, not a stationary, tongue position. They are all voiced.

With /y/, as in 'yoke' contrasted with 'oak', the tongue begins in the /i/ region and moves toward the position for the following vowel. It is called the high front glide. In the case of /r/, as in 'rate' contrasted with 'ate', the tongue begins in the position of the r-colored vowel of 'purr', /r/, and moves toward the following vowel. It is known as the retroflex alveolar glide, though this name is not descriptive of some /r/s.

The third glide is /w/, as in 'will' versus 'ill'. Here the tongue takes a /u/ position and then moves into the following vowel. It is called the high back glide.

So, altogether, English contains 12 vowels, 3 diphthongs, and 25 consonants.

SOUND SYMBOLS

CONSONANTS	SYMBOLS	EXAMPLE
STOPS	p	pare, spare, stop
	b	bare
	t	tare
	d	dare
	k	care
	g	get
FRICATIVES	f	fish
	v	vase
		think
		then

	s	sink
	z	zink
	s,	sure
	z,	pleasure
	h	how
	c	huge(with friction between-tongue front and palate)
AFFRICATES	, t , ts	church
	j, d , dz	judge
NASALS	m	my
	n	nose sing
LATERAL	l	like
GLIDES	r	rye
	y, j	yes
	w	wait
	hw	sweet(voiceless w), why

SOUND SYMBOLS

VOWELS	SYMBOLS	EXAMPLE
FRONT	i	meet
	I	sit
	e	rate
	a	pet pat
CENTRAL	r, r ,	fur
	,	sofa, up
	a,	father
BACK	u	fool
	U	full
	o	note

		saw
		not(British slightly rounded)
DIPHTHONGS	aɪ, ai, I, i	mine
	I, i, oɪ, oi	boy
	aʊ, au, U, u	how
	iu	few
	ju	use

PHONETIC CHANGE ASSIMILATION

When we speak, we do not utter a series of individual units of sound. Rather, we speak in a continuous flow of sounds; vowels and consonants are constantly jostling each other, often blurring or wearing away the edges of adjoining sounds especially. In other words, under certain conditions phonetic change takes place. One type of phonetic change is assimilation.

Changes in the language are usually imperceptible till afterwards, and often seemingly capricious. Analysis of the historical changes shows, however, that the patterns of development are usually clear in retrospect, and that definite causes can be assigned to some of them. The following type of historical change will throw light on the changeable nature of present-day speech.

The usual pronunciation of 'income' is ['ɪn,k m], with primary stress on the first syllable, secondary stress on the second syllable, and a distinct syllabic division between [n] and [k]. When we use the word as an adjective, however, in the phrase 'income tax', the pronunciation may be ['ɪn,k m 'tɑ ks], but often it changes to ['ɪ k m, tɑ ks]. The reduced vowel represents reduced stress. The change from [n] to

[] illustrates what is known as 'assimilation'. So "assimilation is a type of phonetic change which occurs frequently enough to warrant detailed examination."

Assimilation may also be defined as "the process whereby one sound is changed to a second under the influence of a third." For example, in 'income tax', the alveolar [n] changes to the velar [ŋ] under the influence of the velar [k]. According to Bloomfield, in assimilation "the position of the speech agents for the production of one sound is altered to a position more like that of a neighboring sound."

All assimilations start in a manner similar to that of 'income tax'. The change may take place as soon as the two original sounds come close together. Many assimilations start as accidental mispronunciations of an accepted sequence of sounds. A substandard pronunciation of 'length' illustrates this accidental type of assimilation. The shift from the velar [ŋ] to the linguadental [ɲ] is apparently too great for some speakers' muscular control. Consequently the tongue anticipates the dental position, [ŋ] changes to [n] in anticipation of the following [l], and the new pronunciation [l n], rhyming with 'tenth' [t n] is heard. Most people pronounce 'length' as [l ŋ] or [l k], the added [k] serving as an insulation against assimilation.

Another example of assimilation is the derivatives of the Latin preposition 'cum' which occur in English with all three nasal consonants. The original [m] of 'cum' survives in such words as 'combine' [k m'bain], compare [k m'p r] and comfort [k mf t]; but it has become [n] in such words like 'contact' ['kan,ta kt], 'condemn' [k n'd m], and 'constant' ['kanst nt], and has become [ŋ] in such words as 'congress' ['ka gr s] and 'conquer' ['ka k ŋ]. A glance at the consonant which follows the nasal shows that in every word the nasal has approximated the position of the following consonant; [m] has assimilated to [n] before alveolar consonants, and to [ŋ] before velars.

Considering the inherent nature of the assimilative process, it is apparent that the most common type of assimilation is the one in which the preceding sound has been influenced in anticipation of the sound that follows. When the preceding sound influences the sound, the assimilation is classified as 'progressive'. When second sound influences the preceding sound as in 'length' and 'congress', we call it 'regressive assimilation'.

Finally, there is a third assimilative classification known as 'reciprocal', in which the two sounds influence each other and combine to produce a single sound which is a compromise between the two. The word 'sure', for instance, was formerly pronounced as [sjur]; but the sequence [sj] required a more delicate adjustment and consequently, the tongue slipped further back for [s] and further forward for [j] or [sc], until the two sounds came together at a position for [ʃ] and gave us the present pronunciation [ʃur]. Some other examples of reciprocal assimilation are ['vI n], from ['vIzj n], ['I u] from ['Isju], and [Men] from [hwen].

