This chapter provides an overview of the sound system of English. It begins with a discussion of the smallest unit of sound, the **phoneme**, and continues with a description of the **phonetic alphabet** and how it differs from the English alphabet. The phonetic symbols for English consonants and vowels are then presented and classified according to three criteria: **voicing** (whether the vocal cords vibrate or not), **place of articulation** (where in the mouth the sound is produced), and **manner of articulation** (how the airstream flows in the mouth during the articulation).
Introduction

The study of speech sounds can involve either segments or suprasegmentals. Analyses of speech segments are focused on the individual sounds in a given word. For instance, the word hat has three segments: two consonants beginning and ending the word and a single vowel between the two consonants. To describe these sounds, linguists use a set of symbols from the phonetic alphabet, an alphabet in which each symbol corresponds to one (and only one) sound. Thus, the word hat would be transcribed as /hæt/. A phonetic alphabet is necessary because in the English alphabet, for instance, a single symbol can represent more than one sound: the pronunciation of orthographic a in hat is different from its pronunciation in talk.

The study of suprasegmentals moves the analysis beyond individual speech sounds to syllables within a given word or to intonational patterns across words, phrases, and clauses. In a word such as recording, for instance, one can say that the primary stress is on the second syllable: re′cor.ding. The sentence When we arrived at the party, everyone was having fun can be analyzed into two tone units. In each of the two tone units, the pitch will rise, peak on one syllable, and then fall. The syllable with the highest pitch will receive the greatest stress (indicated by the capital letters below) of any syllable in the tone unit:

when we arrived at the PARty
everyone was having FUN

Of course, the stress could be placed elsewhere in each of the tone units if some kind of emphasis is desired. For instance, in the second unit, the first syllable of everyone could receive the primary stress if the speaker wished to emphasize that all people at the party were having fun:

Everyone was having fun

But the point when studying suprasegmentals is that sound can be examined beyond individual speech segments.

This chapter explores in detail how speech segments and suprasegmentals are studied. It opens with a discussion of segments – how they are identified, transcribed, and classified – and concludes with an overview of how stress is placed on syllables in English words and how pitch and stress are assigned in tone units.

Speech segments

Speech segments can be either phonemes or allophones. Phonemes are distinctive speech sounds; that is, they create meaningful differences in words. One way to determine whether a speech sound is distinctive is to examine minimal pairs: words that differ by only a single phoneme in the same position in a word. For instance, the words bat and cat differ by only
one sound: the second and third segments are the same vowel and consonant – /æ/ and /t/, respectively – but the two initial sounds are different: *bat* begins with /b/ and *cat* with /k/. That *bat* and *cat* are different words provides evidence that the sounds /b/ and /k/ in English are phonemes. Indeed, considering other minimal pairs with these sounds points to their status as phonemes:

- **tack/tab**
- **cake/bake**
- **kind/bind**

Phonemes are abstract representations of speech segments. Consequently, the words *pot* and *spot* both contain the phoneme /p/. However, if the actual pronunciation of these words is considered, it turns out that the phoneme /p/ is pronounced differently in the two words. When /p/ occurs at the start of a syllable, as in *pot*, it is **aspirated**: a puff of air accompanies the pronunciation of this sound. In contrast, when /p/ occurs in the middle of a syllable, as in *spot*, or at the end of a syllable, as in *top*, it is unaspirated. It is possible to actually feel the presence or absence of air by placing your hand in front of your mouth while pronouncing each of these three words. But while aspirated and unaspirated /p/ are different sounds, they are not phonemes (at least in English) because they are not distinctive. It is not possible to create minimal pairs with these two sounds: no way to create two separate words in English that differ only by aspirated and unaspirated /p/. These two sounds are therefore considered allophones: predictable variations in pronunciation of a phoneme. The phoneme /p/ is aspirated initially in a syllable and unaspirated elsewhere. A later section will consider in greater detail other types of allophonic variation in English.

Languages vary in terms of the inventory of phonemes that they contain. While aspirated and unaspirated /p/ are not distinctive in English, in Hindi they are. English has the phoneme /ð/ at the beginning of a word such as *the*. German, a language that is very closely related to English, lacks this phoneme, using /d/ to begin words for the definite article: *die*, *der*, and *das*. English distinguishes /s/ and /l/ in words such as *right* and *light*; many Asian languages, such as Japanese, do not. There is tremendous variation in the number of phonemes across languages, with the “range in size from around a dozen phonemes to nearer a hundred depending upon the language” (*Handbook of the International Phonetics Association*, p. 27).

**The phonetic alphabet**

To study phonemes, it is important to use a system of symbols that represent one and only one sound. To see why such a system is necessary, it is useful to compare the English alphabet with a phonetic alphabet. In alphabetic writing systems, there is a (loose) association between letters of the alphabet, or **graphemes**, and sounds. In English, the graphemes c-a-t in the word *cat* correspond to the three phonemes in this word: /k/, /æ/,
and /t/. But as is the case in most alphabetic writing systems, there is not a one-to-one correspondence between sound and grapheme – graphemes can have more than one pronunciation. This is why the grapheme \(a\) in cat can be associated with many different sounds: /æ/ in broadway, /i/ in substance, and /ə/ in addiction.

There are a number of reasons why over time English graphemes have deviated from English pronunciation. First of all, English uses the Roman alphabet, an alphabet originally designed to spell Latin. Because Latin and English have different sounds, the Roman alphabet had to be adapted to spell certain English sounds. For instance, Latin lacks the phonemes /ð/ and /θ/ found at the beginning of English words such as the and thin, respectively. To spell these sounds in English, the digraph th had to be created. Writing is also more conservative than speech. As a result, changes in pronunciation are not reflected in spelling. The word knight has six graphemes but only three phonemes: /n/, /ai/, and /t/. However, during Chaucer’s time, the word had a pronunciation that more closely reflected its spelling: /kænɪxt/. The spelling of knight therefore reflects the history of the word: additional phonemes that are no longer pronounced, and one phoneme, /ʃ/, spelled with the digraph gh, that no longer exists in English (and is roughly equivalent to the final sound in Bach).

To advance the study of speech, the International Phonetic Association was founded in 1886 with the goal of developing a phonetic alphabet known as the International Phonetic Alphabet (or IPA). This alphabet went through several revisions, with the most recent version established in 2005 (see www.arts.gla.ac.uk/ipa/ipachart.html, accessed April 30, 2008). In this alphabet, each symbol corresponds to a single sound, making it possible to describe the sounds of any language in the world. The IPA allows for two types of transcription: a broad transcription or a narrow transcription. A broad transcription is focused on individual phonemes. For instance, the words pat and spat would be transcribed, respectively, as /pæt/ and /spæt/. A narrow transcription, in contrast, would capture the phonetic differences between the sounds in these words, specifically that in pat the /p/ is aspirated, as indicated by a superscript \(h\) following /p/: \([p hæt]\). There are two conventions for transcribing phonemes and allophones. Transcriptions of phonemes are placed within slashes: //. Transcriptions of allophones of a phoneme are included within brackets: [].

To provide a more detailed description of the IPA, the next two sections describe the symbols used in the IPA as they apply to consonants and vowels in the English language. For all consonant and vowel sounds described in these sections, simple English words containing the sounds will be provided to aid in matching symbols with the sounds that they describe.

**English consonants**

English consonants are classified along three parameters: voicing, place of articulation, and manner of articulation. To introduce these notions, one type of consonant – plosives – will be described before the entire range of consonants in English is presented.
**Plosives.** Table 7.1 lists all of the consonant phonemes in English. Across the top of the table are the places of articulation: the parts of the mouth involved in the articulation of each phoneme. The left-hand column classifies the consonants according to their manner of articulation: where the air flows in the mouth while each consonant is articulated, and the degree to which the air flows freely or is subject to varying degrees of constriction.

<table>
<thead>
<tr>
<th>English</th>
<th>bilabial</th>
<th>labiodental</th>
<th>dental</th>
<th>alveolar</th>
<th>postalveolar</th>
<th>palatal</th>
<th>velar</th>
<th>glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>/p/ pat</td>
<td>p</td>
<td></td>
<td></td>
<td>t</td>
<td></td>
<td>k</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/b/ bat</td>
<td>b</td>
<td></td>
<td></td>
<td>d</td>
<td></td>
<td>g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/m/</td>
<td>m</td>
<td></td>
<td></td>
<td>n</td>
<td></td>
<td></td>
<td>η</td>
<td></td>
</tr>
<tr>
<td>/f/</td>
<td>f</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>h</td>
</tr>
<tr>
<td>/v/</td>
<td>v</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/θ/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/s/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/ʃ/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/z/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/ʒ/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/tʃ/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/dʒ/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/w/</td>
<td>(w)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/j/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/w/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/l/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

English has six plosives, which are found at the start of each word in the list below:

/p/ pat  /t/ tack  /k/ kite
/b/ bat  /d/ dark  /g/ get

Three of the plosives are voiceless (or unvoiced): /p/, /t/, and /k/. Three are voiced: /b/, /d/, and /g/. Voicing is a property of the vocal folds, which are located in Figure 7.1 at the bottom of the vocal tract. When the vocal folds vibrate during the articulation of a consonant, the consonant is voiced; if the vocal folds do not vibrate, the consonant is voiceless. Thus, /p/ and /b/ contrast because the latter is voiced but the former is not. It is easy to recognize the absence or presence of voicing in these sounds by feeling the larynx when pronouncing syllables such as pa and ba: pa will produce much less vibration than ba.

While these two consonants differ in voicing, they are identical in terms of their place of articulation. Both sounds are bilabial because their articulation involves the lips, which move together as /p/ and /b/ are produced. The remaining four plosives have different places of articulation. The consonants /t/ and /d/ are alveolar because when they are articulated, the tip or blade of the tongue touches the alveolar ridge in the mouth (again, see Figure 7.1 for these and other places of articulation). The remaining two stops, /k/ and /g/, are velar because during their articulation,
the back of the tongue touches the soft palate (also known as the velum). All six consonants are considered plosives because of their manner of articulation. That is, when these sounds are made, the flow of air in the mouth is initially blocked and then subsequently released. When /p/ and /b/ are articulated, for instance, the lips will initially be pursed, causing blockage of air in the oral cavity. The lips are then opened, allowing the air to flow outside the mouth. With the alveolar and velar stops, a similar effect is achieved with the tongue, which first touches the alveolar ridge or soft palate, blocking the flow of air. The tongue then pulls away, allowing the air to flow outside the mouth.

**Nasals.** English has three nasal consonants, each of which is voiced:

/m/ make  
/n/ nice  
/ŋ/ long

These sounds have the same place of articulation as the six plosives in English. The manner of articulation is likewise identical, except that the air flows through the nasal cavity rather than the oral cavity. With the exception of the nasals, all consonants in English are oral: the soft palate moves back in the mouth and touches the pharynx, preventing air from flowing through the nasal cavity and forcing it instead through the oral cavity. With nasal consonants, however, there is no such movement of the soft palate. Consequently, the air is allowed to flow through the nasal cavity.

**Fricatives.** English has nine fricatives, each of which, except for voiceless /h/, has a voiceless and voiced counterpart. The fricatives in the top
row are voiceless; those in the bottom row are voiced:

| /f/ five | /θ/ thin | /s/ sip | /ʃ/ ship | /h/ hip |
| /v/ vice | /ð/ this | /z/ zip | /ȝ/ measure |

The phonemes /f/ and /v/ are labiodental. When they are articulated, the upper teeth touch the lower lip. The phonemes /θ/ and /ð/ are dental. To produce these sounds, the tip or blade of the tongue touches the upper teeth. The phonemes /s/ and /z/ are alveolar. The phonemes /ʃ/ and /ȝ/ are postalveolar: the articulation of these sounds involves the tip or blade of the tongue touching the back of the alveolar ridge. The phoneme /h/ is glottal. Because this sound originates at the glottis, its articulation is independent of the other articulators (e.g. the tongue or lips). Instead, it is produced “by bringing the vocal folds close enough to produce a hissy sound” (Hewlett and Beck 2006: 36).

With fricatives, there is some constriction of the airflow in the oral cavity. For instance, when /θ/ and /ð/ are articulated, the tongue does not completely block the flow of air as it touches the upper teeth. Instead, it creates a narrow opening through which the air flows, resulting in turbulence in the oral cavity. In fact, with some fricatives, such as /s/ and /z/, one can even hear a hissing sound as the sound is produced.

**Affricates.** English has two affricates, one voiceless and the other voiced:

/ʈʃ/ church  
/ɖʒ/ judge

Both of these sounds are palatal. Their articulation involves the front of the tongue touching the hard palate. As the IPA symbols for these sounds suggest, an affricate is the combination of a stop and a fricative. That is, when these sounds are produced, the tongue causes complete blockage of air. However, after the air is released, the tongue creates enough obstruction of the flow of air to create the kind of turbulence associated with a fricative.

**Approximants.** There are two types of approximants: central and lateral. English has three central approximants:

/ɹ/ ripe  
/j/ yet  
/w/ wet

and one lateral approximant:

/l/ like

All four approximants are voiced. The phonemes /ɹ/ and /l/ are alveolar, while /j/ is palatal. The phoneme /w/ is different from the other approximants in that its articulation is bilabial/velar. That is, as this sound is articulated, the lips narrow and the back of the tongue touches the soft palate. Thus, the production of this sound involves two places of articulation. In
some dialects of English, it is possible to find a voiceless equivalent of /w/, which is transcribed as /ʍ/ and is typically found in words beginning with the orthographic characters wh. For speakers who make the distinction, the word *witch* would begin with /w/ and *which* with /ʍ/. While common in earlier periods of English, the distinction between /w/ and /ʍ/ is currently maintained by relatively few speakers of English. With all four approximants, there is relatively little obstruction of the air in the oral cavity—certainly much less than is found with fricatives. For this reason, the approximants are sometimes regarded as semi-vowels, since the articulation of vowels involves very little obstruction of the flow of air.

**Vowels in English**

While the number of consonants across dialects of English is relatively constant, with vowels there is considerably more dialectal variation. For this reason, the discussion in this section will be restricted to the vowels found in the standard varieties of American and British English. These varieties are sometimes referred to, respectively, as General American (GA) and Received Pronunciation (RP) or BBC English. It must be emphasized, however, that while these varieties have a certain amount of prestige in the United States and Great Britain, they are not spoken by everyone in these countries, and variations from these standards do not necessarily result in non-standard pronunciation. In the United States, people residing in Nashville, Tennessee, for instance, do not desire to speak like people in Chicago, Illinois. Each part of the United States has its own regional standard of pronunciation, which may or may not resemble GA. The same situation holds true in Great Britain, where only a small fraction of individuals (10 percent of population) actually speaks RP.

Table 7.2 lists the vowels found in General American and RP. As this table illustrates, vowels are classified according to the relative position of the tongue in the mouth: how high or low the tongue is positioned (the vertical axis on the left) and the degree to which the tongue is placed in the front or back of the mouth (the horizontal axis on top). The vowel /i/ is considered a close front vowel because when this vowel is articulated, the front of the tongue is at a very high position in the very front of the
mouth. The vowel /u/ is a close back vowel because like /i/ the tongue is high in the mouth. However, rather than the tongue being positioned in the front of the mouth, the back of the tongue is high in the back of the mouth. The vowel /a/ is also a back vowel but it is open because unlike /u/ the back of the tongue is low in the mouth. The vowel /æ/, known as the schwa, is pronounced with the tongue at a height mid-way between the top and bottom of the mouth and positioned in the center of the mouth. The vowel /ɛ/ is pronounced with the tongue positioned slightly lower in the mouth than the schwa but not as close to the front of the mouth as /i/. The remaining vowels exhibit varying degrees of height and frontness and backness.

It is important to realize that the system of vowel classification illustrated in Table 7.2 is somewhat of an idealization. Ladefoged (2001: 71) characterizes notions such as height and frontness and backness as “labels that describe how vowels sound in relation to one another. They are not absolute descriptions of the position of the body of the tongue.” Unlike consonants, he continues, “there are no distinct boundaries between one type of vowel and another.” But while the categories in Table 7.2 may indeed be merely “labels,” they are a convenient way for linguists to describe and compare vowels.

Vowel differences between GA and RP. To illustrate how vowels differ between GA and RP, it is first of all necessary to see how each vowel in Table 7.2 is pronounced in actual words:

<table>
<thead>
<tr>
<th>GA</th>
<th>RP</th>
</tr>
</thead>
<tbody>
<tr>
<td>/i/ feet</td>
<td>/ɪ/ fit</td>
</tr>
<tr>
<td>/eɪ/ fate</td>
<td>/ɛ/ pet</td>
</tr>
<tr>
<td>/u/ suit</td>
<td>/æ/ sack</td>
</tr>
<tr>
<td>/ou/ soak (GA)</td>
<td>/u/ book</td>
</tr>
<tr>
<td>/ɔ/ fought</td>
<td>/ɔ/ arrive [vowel in first syllable]</td>
</tr>
<tr>
<td>/a/ cot (GA)</td>
<td>/æ/ fun</td>
</tr>
<tr>
<td></td>
<td>park (RP)</td>
</tr>
<tr>
<td>/ɔ/ /cot (RP)</td>
<td></td>
</tr>
</tbody>
</table>

Although many of the words in the list are pronounced similarly, there are some notable differences. The vowels in boat and cot, for instance, are pronounced differently in GA and RP. Within the United States itself, many speakers would not pronounce cot with /ɑ/ but with /ɔ/ because the vowels /ɑ/ and /ɔ/ are merging, with /ɔ/ replacing /ɑ/ in many words. But the most notable difference between GA and RP is the pronunciation of the word park, a word that would be pronounced as /pɑːk/ in GA but /pɑk/ in RP. GA is rhotic because /r/ occurs following vowels within a syllable. RP is non-rhotic because /r/ does not occur following vowels in the same context.

Because RP is non-rhotic, it has many more diphthongs than GA. All of the vowels discussed thus far have been monophthongs: vowels whose “quality remains relatively constant” during their articulation (Ashby and Maidment 2005: 75). If, however, a vowel occurs within a syllable and its
quality changes during its articulation, the vowel becomes a diphthong. GA and RP share three diphthongs:

/ai/ fight
/au/ house
/ɔi/ boy

In each of the diphthongs above, the tongue changes position as each part of the diphthong is articulated. In the case of /ɔi/, for instance, the tongue is initially positioned in the lower back part of the mouth and then “glides” to the upper front of the mouth. This feature of diphthongs explains why in the American tradition of transcription, the three diphthongs above are transcribed, respectively, as /ay/, /aw/, and /ɔy/. The sounds /y/ and /w/, sometimes referred to as glides (or semi-vowels), are used to reflect the gradual transition between vowels inherent in diphthongs.

While GA and RP share three diphthongs, RP has four additional diphthongs occurring in syllables where GA would have a vowel + /ɹ/ sequence, a sequence leading to an /ɹ/-colored vowel (Ladefoged 2005: 30). Thus, the four words below would have different pronunciations in GA and RP:

<table>
<thead>
<tr>
<th>GA</th>
<th>RP</th>
</tr>
</thead>
<tbody>
<tr>
<td>fear</td>
<td>/fiɹ/</td>
</tr>
<tr>
<td>fair</td>
<td>/feɹ/</td>
</tr>
<tr>
<td>tire</td>
<td>/taɹ/</td>
</tr>
<tr>
<td>four</td>
<td>/foɹ/</td>
</tr>
</tbody>
</table>

In GA, any time a vowel occurs before an /ɹ/ in a single syllable, the /ɹ/ “colors” the vowel, creating in a sense a single sound. In RP, in contrast, no such sequences of vowel + /ɹ/ exist, resulting instead in a diphthong.

Other classifications of vowels. Vowels have additional classifications. In Table 7.2, all of the vowels outside the rectangular box are considered tense, those within the box lax. In English, tense vowels tend to be long, while lax vowels are short. For instance, the tense vowel /i/ in a word such as feet has a longer duration than the lax vowel /ɪ/ in fit. But tense and lax vowels have other differences too. While tense vowels can occur in both open and closed syllables, lax vowels are restricted to closed syllables. A closed syllable is a syllable ending with a consonant. In such syllables, one finds either tense or lax vowels:

Tense vowels: read /rid/, suit /sut/, hate /heɪt/, talk /tɔk/
Lax vowels: fit /fɪt/, sat /sæt/, help /hɛlp/, took /tʊk/

Open syllables, in contrast, end in a vowel. Only tense vowels can occur in such syllables:

bee /bi/, sue /su/, pay /peɪ/, law /lɔ/

Tense and lax vowels also have different distributions on the vowel chart in Table 7.2, with tense vowels on the “periphery” and lax
Vowels (enclosed in the square box) “in the central regions” (Hammond 1999: 6).

Vowels can also be rounded or spread. Most vowels in English are spread; that is, when a vowel such as /i/ is articulated, the lips are spread apart. In contrast, when /u/ is pronounced, the lips are rounded. The rounded vowels in English are the back vowels /u/, /u:/, /o:t/, /o:/, and /o/. All other vowels are spread.

**Allophones and allomorphs**

The discussion thus far has focused on phonemes: the distinctive sounds (consonants and vowels) that exist in the English language. As was noted earlier, however, phonemes have different pronunciations in specific contexts. For instance, the phoneme /p/ is aspirated initially in a syllable but unaspirated medially or finally. Thus, the phoneme /p/ has two allophones that are in complementary distribution: aspirated [pʰ] in one context and unaspirated [p] in another. Similar kinds of variation occur with morphemes. The plural marker -s has three different pronunciations that are in complementary distribution: [s], [z], and [əz]. The specific environments in which each pronunciation occurs will be described later in this section. But because these variations in pronunciation involve morphemes, they are considered allomorphs of the plural morpheme in English.

This section contains a description of some examples of allophonic and allomorphic variation in English, with an emphasis on one more general process, assimilation, as well as several other processes. Because the focus is on allophones and allomorphs rather than phonemes, a narrow system of transcription, enclosed in brackets [], will be used.

**Assimilation.** Assimilation is a general process whereby adjacent sounds with differing properties become more similar in terms of their voicing, place of articulation, or manner of articulation:

**Nasal plosion**

In a word such as flatten [flætnn], the alveolar plosive /t/ undergoes nasal plosion. First of all, /t/ occurs at the end of a syllable before the alveolar nasal /n/. Both of these consonants are therefore homorganic: they share the same place of articulation. Second, /n/ is syllabic: it occupies a single syllable with no accompanying vowel and is thus transcribed narrowly as [n]. In this environment, when the plosive is released, the air flows through the nasal cavity rather than the oral cavity in anticipation of the articulation of the nasal consonant /n/. In this environment, the plosive is transcribed narrowly with a superscript n: [tn]. Other words in which this process occurs include hidden [hidn] and written [writn]. The process also applies, as Yavas (2005: 59) notes, before /n/ occurring at the beginning of either a subsequent syllable, as in madness [madnɛs], or a subsequent word, as in sad news [sædnuz].

**Other assimilatory processes involving nasals**

When the bilabial nasal /m/ and alveolar nasal /n/ occur before the labiodental fricatives /f/ and /v/, the nasals become labiodentals [n]. Thus, comfort
would be pronounced [kæŋfə:t] or [kæŋfət] and convince as [kənvəns]. When /n/ occurs before a dental consonant such as /θ/, its articulation becomes more dental [ŋ]. Words such as tenth would therefore be pronounced as [tenθ]. Vowels occurring before nasal consonants will themselves be nasalized: ran [rən], room [rəm], ring [rəŋ].

**Devoicing of consonants**

When plosives, fricatives, and affricates occur before unvoiced consonants, they will be devoiced. For instance, in isolation, his ends with the voiced fricative /ʒ/: /hɪz/. However, if his is followed by a word beginning with a voiceless consonant, the /ʒ/ becomes devoiced [z]. Thus, his fist would be pronounced as [hɪz fɪst]. Other examples include:

had to [hæd tu]
with sympathy [wɪθ sɪmpəθi]
yours truly [jɔz tjuəli]

**Voicing assimilation with English -s:**
The inflection -s in English occurs in three contexts: as a third person singular present tense verb form (takes), as a plural marker on nouns (dishes), and as a marker of possession (man’s). But while each inflection is spelled as -s or -es, the inflections have three different pronunciations.

Two of the pronunciations, [s] and [z], are the result of voicing assimilation between the consonant ending the base to which these inflections are attached and the inflections themselves. If the base ends with a voiced consonant or a vowel (all vowels in English are voiced), then the inflection will be voiced [z]:

hose [hoʊz] or [həʊz]
feels [fɪəlz]
child’s [tʃaɪldz]
doors [dɔʊz] or [duəz]
gangs [gæŋz]

If the stem ends with a voiceless consonant, the inflection will be voiceless:

walks [wɔks]
fights [faɪts]
fifths [fɪfθs]
words [wɜdz]
huffs [hʌfs]

With stems ending with the consonants /s/, /z/, /ʃ/, /ʒ, /tʃ/, and /dʒ/, neither [s] nor [z] is possible, since a consonant cluster such as [ss] or [dʒz] is not possible in English. As a result, it is necessary to insert an [i] or [ə] between the consonant ending the stem and the inflection [z]:

hisses [hɪsɪz] or [hɪsəz]
fizzes [fɪzɪz]
washes [wʌʃz]
George’s [dʒɔʊzdʒz]
churches [tʃɜːtfjiz]

Place of articulation assimilation with negative prefixes
The negative prefixes im-, in-, ir-, and ir- all mean ‘not’ in words such as impertinent or illegal. Exactly which of these three prefixes is used depends upon the place of articulation of the consonant beginning the base to which they are affixed. If the base begins with a bilabial sound, then [ɪm] will be used:

immodest [ɪmˈmədɪst]
imperfect [ɪmˈpɛrkɪt]
immobile [ɪmˈmoʊbəl]
implausible [ɪmpləˈzəbəl]

If, on the other hand, the stem begins with an alveolar or velar consonant, then [ɪn] or [ɪŋ] will be used:

intangible [ɪnˈtæŋɡəbəl]
incomplete [ɪnˈkəmplɪt]
inauspicious [ɪnəˈspɪʃəs]
indefensible [ɪnˈdɛfnəsəbəl]

If the stem begins with [l] or [ɹ], the prefixes [ɪl] and [ɪɹ] are used, respectively:

illegal [ɪlɪɡəl]
illicit [ɪlɪˈsɪt]
irrelevant [ɪrɪˈleɪvənt]
irredeemable [ɪrɪˈdɪməbəl]

Miscellaneous processes. Other processes occur in English as well:

Alveolar flapping
In American English, the words writer and rider can be homophones because the alveolar plosives /t/ and /d/ can undergo a process known as alveolar flapping, resulting in both words being pronounced as [ɹaɪrʃ]. For this process to apply, /t/ or /d/ must occur between vowels, and the primary stress needs to be placed on the syllable preceding /t/ or /d/. Both of these conditions apply in writer and rider, as both /t/ and /d/ occur between vowels, and the primary stress occurs on the syllables preceding both consonants. Other words in which this process can apply include butter [ˈbaɪrʃ], literature [ˈlɪrətərɪʃ], and fatter [ˈfætər].

Lengthening of vowels
Vowels will differ in length depending upon whether they occur before a voiced or unvoiced consonant. In the list below, the vowels in the second column are lengthier because they occur before voiced consonants. The
vowels are shorter in the first column because they occur before voiceless consonants:

<table>
<thead>
<tr>
<th>Word (first column)</th>
<th>Word (second column)</th>
</tr>
</thead>
<tbody>
<tr>
<td>bit [bɪt]</td>
<td>bid [bɪ:d]</td>
</tr>
<tr>
<td>feet [fɪt]</td>
<td>feed [fɪ:d]</td>
</tr>
<tr>
<td>hiss [hɪs]</td>
<td>his [hɪz]</td>
</tr>
<tr>
<td>leaf [lɪf]</td>
<td>leave [lɪv]</td>
</tr>
</tbody>
</table>

Intrusive /ɹ/

Some speakers of English will insert an /ɹ/ in contexts where one is not ordinarily found. In the 1960s, the former US President John F. Kennedy made famous the pronunciation of Cuba as [kjubə] rather than [kjuba]. This process occurs when a word ends in a schwa [ə] and precedes either a pause or a word beginning with a vowel. Therefore, President Kennedy would have said *Cuba* [kjubə] *is a threat* but *Cuba* [kjuba] *threatens us*. Other words in which intrusive /ɹ/ can occur include *idea*, *Toyota*, and *Rebecca*.

**Suprasegmentals**

The study of suprasegmentals extends the focus of inquiry to units that are larger than individual segments – syllables, words, phrases, and clauses – and to the features of sound that describe these units, specifically stress and intonation. Of key importance to both stress and intonation is the notion of the syllable.

**Syllables**

A syllable consists of three parts: an onset, a nucleus, and a coda. The nucleus typically consists of a vowel that is preceded by the onset and followed by the coda. In a simple word such as *hat*, the nucleus would be the vowel /æ/, the onset /h/, and the coda /t/. While the nucleus is usually a vowel, it is also possible, as suggested earlier, for the nasal consonants /m/ and /n/ to be syllabic in words such as *bottom* and *button* and for the approximant /l/ to be syllabic in words such as *bottle* /batʃ/ and *little* /lɪt/.

Most people have an intuitive sense of what a syllable is. For instance, if presented with a word such as *happiness*, even a non-linguist would recognize this word as containing three syllables (separated by periods) with the primary stress (marked by ’) falling on the first syllable: /ˈhæ.pi.nəs/.

More formally, Ashby and Maidment (2005: 7) define a syllable as:

one pulse of speech. It always contains one loud or prominent part (almost always a vowel sound), and may optionally have consonant sounds preceding or following the vowel.

Because many words in English have alternate pronunciations, differing pronunciations will sometimes lead to varying numbers of syllables. The word *smile* can be pronounced with one syllable /smaɪl/ or two syllables /smaɪ.yal/. Syllable boundaries can also vary. In the word *ketchup*, it is possible to place the syllable boundary in two places, depending upon how the word is pronounced: /kε.tʃ.əp/ or /kε.ɪ.tʃ.əp/.
Exactly which consonants can occur in the onset and coda is determined by a series of **phonotactic constraints**. The assignment of the primary stress to a particular syllable in a word is dictated by a fairly complex set of rules for stress assignment in English.

**Phonotactics.** English has a series of constraints, known as phonotactics, that specify the permissible sequences of consonants in the onset and coda. For instance, /p/ can occur singly in the onset /pæt/ or the coda /tæp/. It can also follow /s/ in the onset /spæt/. However, it cannot precede /s/ in the onset (e.g. */psæt/). The other consonants in English have similar constraints: other consonants before or after which they can or cannot occur. In total, English allows up to three consonants in the onset and three consonants in the coda, with different sequences of consonants permissible in each.

With the exception of /ŋ/, all consonants in English can occur singly in the onset. The velar nasal /ŋ/ is somewhat exceptional in that it is usually found only singly in the coda in words such as ring /rɪŋ/ and sing /sɪŋ/. It can also occur before /θ/ in one-syllable words such as strength /stɹɛθ/. However, in this and other similar words (e.g. length), some speakers of English will substitute /n/ for /ŋ/ (e.g. /lɛnθ/) and use /ŋ/ only singly in the coda.

As consonants are added to the onset, more restrictions are placed on the possible combinations that are permitted. For instance, if the onset contains two consonants, all plosives and voiceless fricatives can occur prior to the approximant /ɹ/ in words such as pray /prɛ/, trip /trɪ/, crime /krɛ/, bring /brɪ/, drink /drɪ/, greed /grɛ/, free /frɛ/, through /θɹ/, and shrill /ʃɹ/. Many (though not all) of these same consonants can precede /l/: play /pla/, clasp /kl/, blank /bl/, glad /gl/, and slap /sl/. The approximants /j/ and /w/ permit a wide range of consonants to precede them in many syllables. Some examples with /j/ include the first syllables in words such as puny /pju/, furious /fju/, and humid /hj/. Some syllables can contain an optional /j/. For instance, the first syllable in coupon can be pronounced as either /kju/ or /ku/. Some examples with /w/ include twice /twɛ/, quick /kwɛ/, and the first syllable of dwindle /dwɛ/. If the onset contains three consonants, the options are very limited. Some examples include spring /spɛ/, splash /spl/, strike /stɪ/, and scratch /skɛ/.

In the coda, all consonants are permitted to occur singly with the exception of the fricative /h/ and the approximants /j/ and /w/. Whether /a/ is allowed, as in harm /hæm/ or in the second syllable of never /nɛvrɛ/, depends upon whether the speaker’s dialect is rhotic or non-rhotic. In codas with two consonants, the approximants /l/ and (in rhotic dialects) /ɾ/ can precede many different consonants, as in, for instance, help /hpɛ/, bold /bdɛ/, self /slɛ/, elm /ɛlm/, harsh /hɛʃ/, and church /tʃɛ/. The three nasals too can precede a range of different consonants, as in lamp /lɛmp/, hand /hænd/, and link /lɛŋk/, as can the plosives /k/ and /p/ in words such as sixth /sθɪ/, ax /aks/, and taps /ps/. Sometimes, whether a coda can be said to contain two or three consonants depends upon whether the speaker actually articulates all of the consonants possible. For instance, all speakers would articulate a word such as unkempt as
containing three consonants in the coda: /mpt/. However, in other cases, such as twelfth, some speakers might include three consonants in the coda /lfθ/, while others might pronounce only two /lfθ/.

**Stress.** All English words will contain one syllable that has primary stress: a syllable that is more prominent than the other syllables in the word. In the word happy, for instance, the primary stress falls on the first syllable because of the two syllables in this word, the first syllable is more prominent than the second syllable: 'ha.ppy. Some words contain syllables with varying degrees of stress. As was noted in the last chapter, compound words in English are marked by a specific stress pattern: primary stress on the first element and secondary stress on the second element, as in head.light. However, the extent to which English words consistently exhibit varying degrees of stress is a controversial notion. As a result, most discussions of word stress in English focus mainly on primary stress.

Because English is a language with variable stress rather than fixed stress, determining which syllable in a word receives the primary stress can be a very complicated process. A language such as French exhibits fixed stress because the primary stress falls most frequently on the last syllable of a word, and less frequently on the second to last syllable. English, however, has variable stress. A survey of words in English reveals that the primary stress falls on many different syllables. In the examples below, the primary stress falls on three different syllables:

- 'con.cert (penultimate [i.e. second to last] syllable)
- re.'place.a.ble (antepenultimate [i.e. third to last] syllable)
- re.'ceive (ultimate [i.e. last] syllable)

The variability of stress placement in these and other words in English is on one level a consequence of the history of the language, specifically the fact that English vocabulary is a mixture of words of Germanic and non-Germanic origin. And two of the languages from which English has borrowed extensively – Latin and French – have different conventions than English for the assignment of primary stress.

Words of Germanic origin in English are subject to the Germanic stress rule, which stipulates that primary stress is placed on the first syllable of the base of a word, as in the examples below:

- 'ba.by
- 'hun.gry
- 'mo.ther
- 'friend.ly

Even if derivational affixes are added to words of Germanic origin, the primary stress remains on the first syllable of the base:

- 'bel.ly
- 'pre.tty
- un.'friend.ly

- 'fa.ther
- 'pre.tti.ness
- un.'ha.ppiness
During the Middle English period, however, English received an extensive number of borrowings, particularly from French and Latin. As Italic (or Romance) languages, French and Latin had, as Fournier (2007) comments, very different systems of stress. In Germanic languages, as demonstrated above, primary stress is on the initial syllable of the base. If a single prefix is added, as in the case of *unfriendly*, the stress moves rightward to the first syllable of the base. In contrast, Italic languages such as Latin and French “share a feature which is the exact opposite of the stress systems of Germanic languages: stress is determined from the end of words (or tone units) rather than from their beginning” (Fournier 2007: 228). In the word *de'vout*, a French borrowing into Middle English, the stress is on the final (ultimate) syllable, a common stress pattern for disyllabic words in Romance languages. When a suffix is added, as in *de'vo.tion*, the primary stress moves leftward to the penultimate syllable.

As a result, during the Middle English period, two very different systems of stress placement co-existed. One consequence, as Dresher and Lahiri (2005: 78) note, were “doublets,” words with two different patterns of stress: one Germanic, the other Latinate or French. Commenting on the list of words below, Dresher and Lahiri (2005) remark that while Chaucer would have employed the French system of stress, the Germanic system would have existed in English as well:

<table>
<thead>
<tr>
<th>French Stress</th>
<th>Germanic Stress</th>
<th>Modern English Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>ci.'tee</td>
<td>'ci.tee</td>
<td>‘city’</td>
</tr>
<tr>
<td>com.'fort</td>
<td>'com.fort</td>
<td>‘comfort’</td>
</tr>
<tr>
<td>di.'vers</td>
<td>'di.vers</td>
<td>‘diverse’</td>
</tr>
<tr>
<td>ge.'aunt</td>
<td>'ge.aunt</td>
<td>‘giant’</td>
</tr>
<tr>
<td>Pla.'to</td>
<td>'Pla.to</td>
<td>‘Plato’</td>
</tr>
<tr>
<td>pre.'sent</td>
<td>'pre.sent</td>
<td>‘present’</td>
</tr>
</tbody>
</table>

Early (pre-1500) disyllabic borrowings such as these have almost entirely been assimilated into the Germanic system of stress. In a survey of 200 disyllabic French loanwords from this period in Wells’ (2000) *Longman Pronunciation Dictionary*, Svensson and Hering (2005: 123–4) found that only six words in Modern English had stress on the final syllable (as in *di.'verse*). However, later borrowings (post-1700) revealed a greater mixture of the Germanic and French systems, and also some notable differences between Modern British and American English. For instance, while words such as *bro.chure* and *ca.fe* carry primary stress on the first syllable in British English, in American English the primary stress is on the second syllable (*bro.'chure* and *ca.'fe*). Later borrowings had an additional effect on stress placement in English: they introduced into the language words and affixes that resulted in polysyllabic words that were often three syllables in length or longer. The resultant words brought about significant changes in the assignment of primary stress, so that in Modern English, assignment of primary stress is dependent upon (1) the type of affix added, (2) the number of syllables a word contains, and (3) which syllables in the word are...
light (incapable of taking primary stress) or heavy (capable of taking primary stress).


<table>
<thead>
<tr>
<th>Word Pair</th>
<th>Stress Pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>'hope.ful'</td>
<td>'hope.ful.ness'</td>
</tr>
<tr>
<td>'pro.fit'</td>
<td>pro.fi.'teer</td>
</tr>
</tbody>
</table>

The words hopeful and hopefulness contain suffixes, -ful and -ness, that are stress-neutral: their addition to hope has no effect on the placement of primary stress, which remains on hope. However, in the words profit and profi-teer, the suffix -er is stress-demanding: its addition to profit causes the stress to shift to the end of the word. Other suffixes that Stockwell and Minkova (2001) classify as stress-neutral include -ess (e.g. 'host.,'hos.tess), -man (e.g. po.'lice, po.'lice.man), and -ist (e.g. 'fet.ish, 'fet.ish.ist); those they classify as stress-demanding include -naire (e.g. 'ques.tion, ques.tion.'naire), -esce (e.g. in.can.'des.cent, in.cand.'esce), and -tee (e.g. 'am.pu.tate, am.pu.'tee).

The word pairs below illustrate the effects of the number of syllables in a word on stress placement as well as whether a potential syllable for stress is light or heavy:

<table>
<thead>
<tr>
<th>Word Pair</th>
<th>Stress Pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>'mul.ti.ply'</td>
<td>mul.ti.pli.'ca.tion</td>
</tr>
<tr>
<td>'mys.ti fy'</td>
<td>mys.ti.fi.'ca.tion</td>
</tr>
</tbody>
</table>

In each of the pairs above, the second examples are derived from the first examples through the process of affixation, a process that in each example adds additional syllables and shifts the placement of primary stress. In 'mul.ti.ply', the primary stress is on the antepenultimate (third to last) syllable, whereas in mul.ti.pli.'ca.tion, the primary stress is on the penultimate (second to last) syllable. The difference in stress placement in these words is directly attributable to the kind of syllable occurring in the penultimate position.

In mul.ti.pli.'ca.tion, the penultimate syllable (spelled ca but pronounced /kei/) is heavy: it contains a tense vowel. If the penultimate syllable is heavy, it will carry the primary stress (heavy syllables can also contain any vowel, tense or lax, followed by a consonant, which is why a word such as con.'tent.ment has primary stress on the penultimate syllable, even though the vowel in tent, /ε/, is not tense but lax). In contrast, in 'mul.ti.ply, the penultimate syllable is light: it contains only a lax vowel (spelled i but pronounced as /a/). With light syllables in the penultimate position, stress is moved to the antepenultimate syllable.

Additional examples are given below, with heavy syllables in the first column and light syllables in the second column:

<table>
<thead>
<tr>
<th>Penultimate stress</th>
<th>Antepenultimate stress</th>
</tr>
</thead>
<tbody>
<tr>
<td>res.ti.'tu.tion</td>
<td>mul.ti.'fac.e.ted</td>
</tr>
<tr>
<td>dis.con.'tent.ment</td>
<td>rep.re.'hens.i.ble</td>
</tr>
<tr>
<td>ad.jec.'ti.val</td>
<td>de.'riv.a.tive</td>
</tr>
<tr>
<td>fun.da.'men.tal</td>
<td>pre.'var.i.cate</td>
</tr>
</tbody>
</table>
While derivational affixes affect stress placement, inflectional affixes do not. In all of the word pairs below, even though the addition of a suffix adds an additional syllable to each word, the placement of stress does not change:

pre.'var.i.cate   pre.'var.i.ca.ted
es.'tab.lish    es.'tab.lish.ing
fa.'mil.iar.ize  fa.'mil.iar.iz.es

Of course, there are exceptions to many of the patterns of stress discussed thus far. Words such as ex.'pa.tri.ate and 'con.cen.trate have primary stress on the antepenultimate syllables, even though the penultimate syllables in these words are heavy and should therefore carry the primary stress. Just the opposite is true as well: words such as re.'vi.sion, fru.'ti.on, and con.'fe.sion in which the penultimate syllable is stressed despite the fact that it is light. The noun and verb forms of certain words are distinguished simply by changes in stress. When the stress is on the first syllable, as in 'con.tract, the word is a noun. When the stress is on the second syllable, the word is a verb: con.'tract. Other examples include 're.cord/re.'cord, 'con.vict/con.'vict, and 'im.port/im.'port. And as noted earlier, there are a number of disyllabic words in English (e.g. ga.'rage) that contain primary stress on the final rather than the first syllable, contrary to the more common patterns of stress assignment in English. These and other exceptions indicate that stress assignment in English is, at best, only partially predictable.

Intonation

The study of intonation involves the investigation of pitch and stress across groups of words occurring within a tone unit. In a typical (i.e. unmarked) tone unit, the pitch will begin rising at the start of a tone unit, peak on one particular syllable occurring towards the end of the tone, and then fall before rising again at the start of the next tone unit. For instance, the excerpt below contains two tone units:

I couldn’t REAlly || let my company DO this ||

(London-Lund S.12.6.758-759)

In this excerpt, the tone unit boundaries are indicated by the two vertical lines || and the syllables receiving the highest pitch and the greatest stress are marked with upper-case letters and bold-face type. (To enhance readability, the annotation used above and elsewhere in this section has been changed from the annotation used in the London-Lund Corpus.) In the first tone unit, the pitch rises, peaks on the first syllable of REAlly, which receives the greatest stress of any word in the tone unit, and then falls. In the second tone unit, this same pattern is repeated, except that after the pitch peaks on DO, the tone unit does not end until after the following word, this. To understand why this pattern exists, it is important first of all to distinguish word stress from what is sometimes referred to as sentence stress, and to discuss the role that intonation plays in highlighting new information, a point introduced earlier in the section of Chapter 4 dealing with information structure.
In an unmarked tone unit, the pitch will peak on the tonic syllable. As the last section demonstrated, in isolation all words will have one syllable carrying primary stress: the preposition *be*tween has primary stress on the second syllable, the adverb *rea*lly in the above example on the first syllable. However, when words occur together in an unmarked tone unit, one syllable of one word will receive greater stress (called sentence stress) than the other words. This stress will occur on the syllable, called the tonic syllable, carrying primary stress in the last content word of the tone unit. In unmarked tone units, only content words will contain the tonic syllable, not function words. This is why in the second tone unit in the example above, the lexical verb *do* (a content word) contains the tonic syllable rather than the last word in the tone unit, the pronoun *this* (a function word).

The reason that content words rather than function words contain the tonic syllable results from the fact that the goal of intonation in spoken discourse is to highlight new information. And since content words are more meaningful than function words, it is only natural that content words would receive the greatest stress in a tone unit. In fact, function words are often so lightly stressed in rapid speech that the vowels they contain become subject to *vowel reduction*, and the consonants ending function words are sometimes deleted. Vowel reduction occurs when the vowel in a lightly stressed syllable changes to a schwa [ə]. For instance, if carefully articulated, the articles *a* and *the* in English can be pronounced as [ei] and [əi], respectively. However, if the articles are lightly stressed, they will be pronounced as [ə] and [əə]. The conjunction *and* [ænd] can have its vowel reduced too, and one or both of its final consonants deleted, resulting in pronunciations of [ən] or simply [ə].

While function words do not ordinarily contain the tonic syllable, with some kind of contextual motivation, they and potentially any word in a tone unit can contain the tonic syllable. For instance, if the invented tone unit below is unmarked, the tonic syllable will occur on the second syllable of *tomorrow*:

```
I will call you toMORrow /dblpipe
```

It is possible, however, to imagine contexts in which just about any word in the above tone unit could contain the tonic syllable, resulting in a marked tone unit, a tone unit in which the tonic syllable occurs somewhere other than in the last content word of the tone unit. For instance, if several people are speaking and one person asks two of the conversants which one will call her, one of the conversants could reply by placing the tonic syllable on *I* to emphasize that she rather than the other person will be doing the calling:

```
I will call you tomorrow /dblpipe
```

If one of the conversants is worried that one of the others will not call her tomorrow, the person who will be making the call can place the tonic syllable on *will* to reassure the person being called that she will indeed be called:

```
I WILL call you tomorrow /dblpipe
```

One could imagine other contexts in which other words in the tone unit could receive the tonic syllable, but the point to remember is that
potentially any word in a tone unit can be heavily stressed for purposes of emphasis and meaning.

In addition to highlighting new or important information, tone units help to segment spoken language into grammatical units. The extent to which tone unit boundaries correspond to major grammatical boundaries depends upon the particular grammatical boundary being considered and also whether the spoken text is planned or spontaneous. In an analysis of a spoken monologue, a more carefully planned type of speech, Altenberg (1990) found that while certain kinds of structures were very commonly associated with tone unit boundaries, others showed a weaker correspondence. For instance, Altenberg (1990: 279) discovered that 150 of 153 (98%) coordinated main clauses were separated by a tone unit boundary, whereas only 19 of 32 (59%) of nominal that-clauses were set off by a tone unit boundary. These trends can be observed in the excerpt from a monologue below:

your Provost has SAID / that I was going to talk about the ARTS / and indeed I had inTENded / to TALK about that / but hearing President NIXon / MOUthing about the DEATH penalty / and about the permissive soCiety / I decided that I would talk inSTEAD / about something which conCERNS me / in the THEatre /

(London-Lund S.12.7.2–10)

There are two instances of coordinated main clauses separated by a tone unit boundary: the coordinator and in the early part of the excerpt (and indeed I had) is preceded by a tone unit boundary, as is the coordinator but (but hearing President NIXon). In contrast, the one instance of a nominal that-clause is not set off by a boundary (I decided that I would talk).

If, however, a spontaneous dialogue is examined, there is less of a correspondence between grammatical and tone unit boundaries, largely because the spontaneous structure of such speech-types often results in fewer grammatically well-formed structures. In addition, because the speaker is planning what to say as he or she is speaking, there are more hesitations, reformulations, and so forth that interrupt the flow of speech. In the excerpt below from a spontaneous dialogue, one finds boundaries before the two clauses introduced by so (...so I said FINE... so I got a perEMPtory), but also boundaries not separating major grammatical categories, such as the boundary between the two repetitions of I’m (I’m || I’m just hanging ON) and between an adjective and noun within a noun phrase (a perEMPtory || command)

oh well you KNOW || might get TERRibly || you KNOW || I’m || I’m just hanging ON now || and could take you on PERmanently || we may need YOU || to do some work in the EVEning || so I said FINE || being oBLiging || so I got a perEMPtory || coMMAND || over the PHONE || RIGHT ||

(London-Lund S.1.5.223–234)

There are many other features of intonation that can be described. Speech is filled with pauses, which are distinct from tone unit boundaries. The tone and tempo of speaking is also important: speakers can vary the loudness of what they say, and the tempo (fast, slow) at which they speak. In declarative sentences, the pitch falls after the tonic syllable is reached: