Chapter 3: The Sounds of English. Consonants and Vowels. An Articulatory Classification and Description. Acoustic Correlates

3.1. Consonants and Vowels. Traditional distinctions. Chomsky and Halle’s SPE definition

3.2. Criteria for consonant classification. Vocal cord vibration. Sonority


3.4. Sonorants. The Approximants: glides and liquids

3.5. Oral and nasal articulation

3.6. Force of articulation

3.7. Place of articulation

3.8. The Description of English consonants

A. The Approximants

1. The Glides. There are two sounds in English, [w] and [j], having vowel-like features as far as their articulation is concerned, but which differ from their vowel counterparts [u] and [i] respectively through their distribution, force of articulation and length. When we articulate a glide the articulatory organs start by producing a vowel-like sound, but then they immediately change their position to produce another sound. It is to the gliding that accompanies their articulation that these sounds owe their name. As we have seen earlier, precisely because of their ambiguous nature they are also called semivowels or semiconsonants. Unlike vowels, they cannot occur in syllable-final position, can never precede a consonant and are always followed by a genuine vocalic sound.

a. [w] is a labio-velar, rounded sound. At the beginning, its articulation is similar to that of the vowel [u], but then the speech organs shift to a different position to utter a different vocalic sound. The distribution of the sound includes syllable-initial position before almost any English vowel (e.g. win [wʌn], weed [wi:d], wet [wet], wag [wæg], work [wɜ:k], won [wʌn], woo [wu:], wood [wʊd], walk, [wɔ:k] wander [wɔ:ndər]) or a diphthong (e.g.
Before [r'], (e.g. write) the sound is no longer pronounced. [w] can also occur after a plosive (e.g. twin, queen) or a fricative consonant (e.g. swine). It can be rendered graphically either by the letter w (the most common case) (e.g. sweet) or by u (e.g. quite).

b. [j] is an unrounded palatal semivowel. The initial stage of its pronunciation is quite similar to that of the short vowel [i], but then the sound glides to a different vocalic value. Like [w], [j] cannot occur in final position (as a quite similar palatal sound very often does in Romanian), is never followed by a consonant and occurs in front of back, central and front vowels. (e.g. yes, young, youth). It can be preceded by a plosive (e.g. tune) or a fricative (e.g. fume). The sound may be spelt y (as in year) while in words spelt with u, ue, ui, ew, eu and eau read as the long vowel [u:]. The palatal sound is often inserted. The insertion is obligatory if the preceding consonant is: an oral plosive (p, b, t, d, k, g), a nasal stop (m, n), a labio-dental fricative (f, v) or a glottal one (h). A word like beauty can only be read [bju:t] and not [bu:t]. Cf. also: pure, bureau, tulip, deuce, queue, argue, mule, neutral, furious, revue, huge. The palatal sound is not inserted after affricates or after [r] or [l] preceded by a consonant: chew, June, rude, clue. When [l] is not preceded by a consonant or when the sound preceding [u:] is an alveolar fricative [s, z] or a dental one, the usage varies: cf. suit [sju:t], but also [su:t]. In words like unite, unique, university, etc, where u forms the syllable alone the vowel is always preceded by the semivowel: [ju:nai].

2. The Liquids. These are approximant sounds, produced in the alveolar and postalveolar region and include several variants of the lateral [l] and of the rhotic [r].

a. The lateral [l]. The main variants of [l] are a so-called “clear” [l] and a “dark” [ł]. The clear [l] is distributed in prevocalic positions. When this sound is articulated, the tip of the tongue touches the alveolar ridge and the air is released either unilaterally or on both sides of the active articulator. The front part of the tongue also raises towards the hard palate. Words like lake [leik], look [luk], flute [flu:t], lurid [ljurid] delight [dilaut] illustrate the distribution of the consonant in syllable-initial position or after a plosive plot [plot], Blake [bleik], clean [kli:n], glue [glu:] or a fricative slot [slot], fly [flai] and in front of a vowel or the glide [j].

The dark [ł] is distributed in word-final position or before a consonant. As in the case of the clear [l] the tip of the tongue touches the alveolar ridge and the air is released laterally, but now it is the body of the tongue that raises against the soft palate, modifying the resonance of the sound and giving it a more “stifled” character. Words like kill [kil], rule [ru:l], belfry [belfri], belt [belt], silk [silk] illustrate the distribution of the sound either
at the end of the word (syllable) or before a consonant. The phoneme is spelt either l or ll in words like link or call, for instance. In many words, however, before plosive sounds like [k] or [d] – cf. chalk, could; or before nasals like [m] or [n] – cf. calm, Lincoln; the labio-dental fricatives [f] and [v] – cf. calf, calves; the lateral sound is not pronounced.

b. The rhotic [r]. The class includes several variants which are pretty different both in articulatory terms and in auditory effect.

The RP [ɹ] is a frictionless continuant, articulated very much like a fricative, but friction does not accompany the production of the sound. The tip of the tongue slightly touches the back of the alveolar ridge, while the body of the tongue is low in the mouth.

A flapped [r] is used by many speakers of English, especially when it occurs at the beginning of unstressed syllables. The tongue rapidly touches the alveolar ridge with a tap.
A rolled [r] is common in northern dialects and in Scotland. It is produced by a quick succession of flaps, the tongue repeatedly and rapidly touching the alveolar ridge and vibrating against it. This sound is not characteristic for RP.

The letter r or double rr reproduces the sound graphically: right, barren. In postvocalic word- or syllable-final position the sound is not pronounced in standard English – cf. car, party. If the word is, however, followed by a vowel, [r] is reinserted: the car is mine. The same insertion takes place when an affix is attached to a base ending in a (normally) silent [r]: hear [hɪə] / hearing [hɪərɪŋ]; Moor [mʊə] / Moorish [mʊərɪʃ]. This type of [r] is called “linking r”.

B. The English Stops

1. The oral plosives. In terms of their place of articulation they are bilabial, alveolar and velar.

   a. [p] is a voiceless, bilabial, fortis plosive. Its variants include an aspirated plosive if the consonant is followed by a stressed vowel and occurs in syllable-initial position. Being a bilabial stop, [p] is produced by completely blocking the airstream at the level of the lips and by suddenly releasing the air with an explosion. Except for the aspirated variant, the phoneme is pretty similar to its Romanian counterpart. It is distributed in initial, medial and final position: pane, appear, lip. It is spelt p: plane or
pp\textsuperscript{5}: opposite and only exceptionally gh in hiccough. The letter p is silent when followed by another obstruent or a nasal in word-initial position: psalm, pterodactyl, pneumatic.

b. [b] is the voiced, lenis counterpart of [p]. Voicing and force of articulation are the features that contrast the two phonemes, [b] being like [p] a bilabial sound. It is distributed in all three basic positions; initial, medial and final: bet, above, cab. It is spelt b: about or bb: abbot. The letter is silent in final position after m: limb, crumb, dumb and in front of t in words of Latin origin where the sound has long been lost: debt, doubt, subtle. The variants of [b] include partially devoiced allophones in initial position: big, blow, bring and laterally or nasally released allophones when [b] is followed by the lateral l: bless or by a nasal consonant: ribbon. It is not audibly released in final position: rib.

c. [t] is a voiceless, apico-alveolar, fortis plosive. Like [p], it has an aspirated variant that occurs before stressed vowels when the phoneme is distributed in syllable-initial position: tube. If preceded by s, however, [t] is unaspirated: stain. Its distribution includes all basic positions: try, attain, pit. It is laterally or nasally released if followed by [l] or by a nasal consonant, respectively: little, written, utmost. The English phoneme is more retracted than its Romanian counterpart which is rather a dental sound. It is spelt with t: toe, with tt: cutter or with th: Thomas, Thames.

d. [d] is the voiced, lenis counterpart of [t], voicing and force of articulation differentiating between the two sounds that share the same place of articulation in the alveolar region. Both [t] and [d] can become dentalized in the vicinity of the dental fricatives, in words like eighth and breadth. The sound is distributed in initial, medial and final position: dime, addition, pad. It is partially devoiced in initial position: duke and devoiced in final position: road. It is laterally released if followed by [l]: riddle and nasally released if followed by [m] or [n]: admit, sudden. It is spelt d: read or dd: adder.

e. [k] is a voiceless, dorso-velar, fortis, plosive sound, articulated with the dorsum of the tongue against the soft palate. Like the other voiceless plosives described above, it has an aspirated variant if the sound is distributed in syllable-initial position, in front of a stressed vowel: cat. [k] is distributed in initial, medial and final position: coat, accuse, sack. It can be followed by a nasal consonant and be consequently nasally released:

\textsuperscript{5} Doubling the consonants is strictly a problem of spelling convention in English. The English language does not have double (geminate) consonants. Compare English innate [mɛnt] to Romanian înăscut [înăskut].
thicken or by the lateral liquid and be laterally released: fickle. In spelling, the sound can be represented by the letter c (e.g. comb) or by cc (e.g. accuse), by k (e.g. kill), by ck (e.g. pick), by ch (e.g. architect), by qu (e.g. queen). As in Romanian, the sequence [ks] can be rendered by the letter x (e.g. extreme). In words like muscle and knave the letters c an k are silent.

f  [g] is the voiced, lenis pair of [k] and it has basically the same features as its Romanian counterpart. It is distributed in initial, medial and final position: game, begin, rag. Its allophones include partially devoiced variants in initial position: gain, devoiced variants in final position: dog, laterally released, when followed by [l]: giggle and nasally released when followed by [m]: dogmatic. In spelling, the consonant can be rendered by g: get by gg: begged, or by g followed by h, as in ghastly, by ua, ue or ui, as in guarantee, guess or linguist, respectively. The voiced counterpart of [ks], [gz] can also be rendered by x in words like example.

g. The glottal stop [?] is a glottal, voiceless, fortis sound produced in the glottal region by bringing the vocal cords together and then separating them, thus completely blocking and then suddenly releasing the airstream. It is a sound that has been compared with a slight cough. It appears in syllable-final position especially when it separates two adjacent vowels that are not part of the same syllable (in a hiatus): geography [ʤɪ'?əgrəfɪ] or between a vowel and a syllable-final voiceless stop or affricate that it reinforces. In some accents (notably Cockney), it replaces voiceless plosives like [k] and [t] at the end of a syllable. E.g. sick guy [sɪ?ɡa] or quite right [kwa?rɪt].

Acoustically, English voiced plosives can be distinguished from their voiceless counterparts by having a low frequency component determined by the feature voice. The release stages of the three classes of stops in terms of place of articulation: bilabial, alveolar and velar, respectively, differ as regards the noise burst they produce. Alveolar plosives display higher frequencies (3000-4000 cps) than the bilabial (around 360 cps) and velar ones (around 700 cps).

2. The Nasal stops.
a. [m] is a bilabial, voiced, lenis, nasal stop. As in the case of all nasal sonorants, when we articulate this sound the velum is lowered, blocking the oral cavity and letting the air escape through the nose. There are no differences between the English sound and its Romanian counterpart. [m] is distributed in all basic positions: initial, medial and final: make, remote, dim. It can be spelt with m or mm: come, common. It should be said, however, that English does not accept a sequence of two nasal sounds in the same syllable, words like solemn and hymn differing from their Romanian counterparts as the last nasal sound is not pronounced. If an
affix is added, nevertheless, that begins with a vowel, the second consonant is recovered. Compare solemn [solemn] to solemnity [solemniti].

b. [n] is an alveolar, voiced, lenis, nasal stop. The place of articulation is similar to that of [t] and [d], but [n] is a nasal sound, so the air is released through the nose and not through the mouth. It is similar to its Romanian counterpart. It is distributed in all three basic positions – initial, medial and final: name, renown, can. It is spelt n or nn: dean, annual. The sound is elided in final position after [m], but recovered in derived words: damn, damnation. (See also solemn and solemnity above).

c. [g] is a velar, voiced, lenis, nasal stop. It occurs in the vicinity of the velar oral plosives in words like link or wrong. It is to be noted that in present-day English the velar oral plosive in the last word is no longer pronounced, but we can find the velar nasal in front of [g] in connected speech in sequences like I can get it. A similar sound can be found in Romanian, in words like bancă, rangă, but in our language it does not have a phonemic, contrastive value. As pointed out above, this phoneme has a limited distribution: it always precedes the voiceless velar plosive or occurs in syllable-final position in front of an elided [g].

C. The English Fricatives

Fricatives are, as we remember, sounds that are produced by narrowing the speech tract and letting the air out, a process which is accompanied by friction and in some cases by a hissing sound.

[f] is a labio-dental, voiceless, fortis consonant. It is produced by pressing the lower lip against the upper teeth and forcing the air out between them. The sound is similar to its Romanian counterpart. The sound can be spelt f – as in fine, flare, fringe, feud, loaf, stifle, ff – as in effort, snuff, ph – as in physics, graph, or even gh – as in enough, tough. The word lieutenant [lɛftˈɛnənt] is a particular case.

[v] is the voiced, lenis pair of [f] with which it shares the place (labio-dental) and manner (fricative) of articulation. It is important to remember that the English sound is a labio-dental and not a bilabial fricative (as its Spanish counterpart, for instance). It has exactly the same characteristics as the Romanian sound. It is spelt with the letter v. (Exceptionally, by ph in Stephen, nephew and f in of). Certain English nouns voice their labio-dental final fricative when they pluralize displaying the alternance f/v: e.g. leaf / leaves, wife/wives. Derivational affixes can also voice the final consonant: life/liven.
\[\theta\] is an interdental, voiceless, fortis fricative. The phoneme does not have any distributional variants. It occurs in word-initial, medial and final position. It is produced with the tip of the tongue between the teeth, the air escaping through the passage in between. It is a sound difficult to pronounce for Romanian speakers who often mistake it for [s] or even [t]. The sound exists in other European languages too, such as Spanish or Greek, the symbol used in the IPA alphabet being in fact borrowed from the Greek alphabet. The sound is rendered graphically by \texttt{h}: e.g. thin, method, path. The sound often occurs in clusters difficult to pronounce: eightths [\texttt{eit\theta{s}}], depths [\texttt{dep\theta{s}}], lengths [\texttt{len\theta{s}}].

\[\theta\] is the voiced pair of \[\theta\] being an interdental, voiced, lenis fricative. In initial position it is only distributed in grammatical words such as demonstratives: this, that, these, those, there; articles: the; adverbs: thus. It occurs freely in medial position: brother, bother, rather, heathen. In final position it often represents the voicing of \[\theta\] in plurals like mouths [\texttt{mau\theta{z}}], wreaths [\texttt{ri:z\theta{z}}] which may prove difficult to pronounce, or in derived words like bath \[\texttt{ba: \theta}\] (noun)/bathe \[\texttt{be\theta{a}}\] (verb) or breath \[\texttt{bre\theta}\] (n.)/breathe \[\texttt{bri:\theta}\] (v.). The sound is always spelt \texttt{th}, like its voiceless counterpart.

\[s\] is an alveolar, voiceless, fortis fricative, produced with the blade of the tongue against the alveolar ridge, a sound quite similar to its Romanian counterpart. It is a hissing sound distributed in all major positions: at the beginning, within and at the end of a word. It is in fact the only obstruent sound in English that can occur in front of another obstruent, provided the latter is voiceless: e.g. spot, stop, skin. \[s\] is the plural allomorph for nouns ending in a voiceless consonant as well as the allomorph of the 3rd person singular present indicative morpheme. It is spelt \texttt{s}, \texttt{ss} or \texttt{c} in front of \texttt{e}, \texttt{i} or \texttt{y}: e.g. sour, say, hiss, assign, ceiling, cigarette, cypress, bicycle. Sometimes the spelling can be \texttt{sce}, \texttt{sci} or \texttt{scy} (e.g. science, scent, scene, scythe). \texttt{s} is silent in words like corps, island, viscount.

\[z\] is the voiced, lenis, alveolar fricative that corresponds to the voiceless \[s\]. It is quite similar to its Romanian counterpart, but it plays a more important role in English as it is one of the main allomorphs of the plural morpheme (distributed after a voiced consonant or a vowel). Like its voiceless counterpart, \[z\] is a hissing sound, produced with a high-pitched friction. Because when these sounds are articulated the air is expelled through a narrow groove along the middle of the blade they are also called \texttt{grooved} fricatives. Together with the more retracted, alveo-palatal fricatives and with the affricate sounds they are called \texttt{sibilants}. The sound is spelt \texttt{z}. It is often spelt \texttt{s} when the sound does not occur in initial position e.g. nose, easy, desire), and, exceptionally, \texttt{tz} in \texttt{tzar}. Similarly, when it marks the plural of nouns
ending in a voiced sound (e.g. boys, balls, ribs) or when it is the voiced allomorph of the 3rd person singular present indicative of verbs ending in a voiced sound (e.g. plays, calls, adds) the spelling is s. Exceptionally, the sound can be spelt double ss in words like dissolve, possess.

[ʃ] is an alveopalatal, voiceless, fortis fricative consonant. The uttering of this sound should not raise any particular problems for Romanians as its articulatory features are similar to those of its counterpart in Romanian. The blade of the tongue is raised against the region behind the alveolar ridge and the air is forced out through a groove a little wider than in the case of [s], its more fronted counterpart. [ʃ] is distributed in all three main positions in the word. It is often spelt sh in words like shoe, cushion or push. It can also be spelt s (e.g. sure, sugar) or ss (e.g. pressure, mission) or ci (ancient, delicious), sci (conscious) ce (ocean), si (pension, mansion), ti (tuition, retribution). It is a variant of [sj] in words like issue, tissue. In words of French origin the sound is spelt ch: champagne, charade, chargé, moustache, attaché. The same spelling is used in proper names like Charlotte, Chicago, Chicoutimi, Michigan.

[ʒ] is the voiced counterpart of [ʃ]. It is an alveopalatal, voiced, lenis fricative and is pronounced very much like the corresponding sound in Romanian. It is not, however, a very common sound in English as it occurs mainly in loan (particularly French) words. It is never distributed in initial position, but it can occur in medial (pleasure, treasure, measure) or final position (garage, prestige). It can be spelt either s when followed by u (visual) or i (decision), or z if followed by u (seizure) or ge (massage, espionage). In words like casual the alternative pronunciation [zj] is possible, while in other cases the fricative is replaced by the affricate [dʒ] (e.g. garage).

[h] is a glottal fricative in English, a voiceless, fortis sound produced by letting the air pass freely through the mouth during expiration. Thus, its place of articulation in the glottal region is more retracted than in the case of the Romanian sound which is rather a velar sound, closer to the variant occurring in Scottish English: loch [lox]. A palatalized version is used when the sound is followed by a palatal: humane [hjumen]. Unlike in most Romance languages h freely occurs in initial position in English: home, hiss, hut. “Dropping the h’s “ is even considered a sign of lack of education. In a small number of words the sound is, however, dropped even in standard English in both in initial and medial position: hour, heir, honour, honest, vehicle, annihilate. It is also common (even for educated people to drop the initial h in unstressed (weak) forms of the personal pronouns (he, him) possessives (his, her) or the verb have h is also silent in final position in the interjection ah or in words like shah. The conservative spelling of English has preserved the
letter h after r in words of Greek origin where no h sound or aspiration is heard nowadays (rhapsody, rhetoric, rheumatism, rhinal, rhinoceros, rhombus, rhyme, rhythm).

D. **The English Affricates**

The affricate phonemes of English are [tʃ] and [dʒ]. They differ from their Romanian counterparts as they can be distributed in all three basic positions (including the word-final one) and can be followed by any vowel. Therefore, they are far less palatalized than the corresponding Romanian sounds that must be followed by either e or i. Even when they are followed by i and e the English affricates differ considerably from the corresponding sounds in Romanian. In order to realize the difference between the English sounds and their Romanian counterparts it is enough to compare the English word chin to the Romanian cin or the English gem to the Romanian gem.

[tʃ] is a voiceless, fortis, alveo-palatal sound produced with the blade of the tongue raised against the region just behind the alveolar ridge. As in the case of any affricate sound, its articulation starts like that of a plosive – in our case [t] – by completely blocking the outgoing airstream and then continues by a gradual release of the air, as for a fricative [ʃ]. The very symbol used in the IPA alphabet for the notation of the sound suggests the mixed nature of the affricate. We should make a difference, however, between the affricate proper (pitch [ptʃ]) and the sequence of the plosive and the fricative [t] + [ʃ] (courtship [kɔ:tʃip], right shoe [ræʧu:]). The phoneme is represented graphically by ch: (charm, chinchilla, rich) or tch (kitchen, bitch) or by t followed by u (creature, culture) when the plosive is palatalized. In words like habitual, sanctuary the pronunciation with an affricate is a variant of [tʃ]. Exceptionally, we can have ce or cz as graphic representations of the sound in (violon) cello or Czech.

[dʒ] is the voiced counterpart of [tʃ], being an alveo-palatal, voiced, lenis, affricate consonant. It can be rendered graphically by j in either initial or medial position in words like justice, John, rejoice, pyjamas, by ge in all basic positions: gesture, agent, sage, by gi in initial and medial position: giraffe, rigid; and gy in initial position: gymnastics. In certain words it can be spelt d followed by u: gradual, individual, procedure/al. In all these cases, however, there is an alternative pronunciation [dj]. In a number of proper names or common nouns originating in proper names ch is read [dʒ]: Norwich, Greenwich, S/sandwich. Another spelling can be dg in words like ridge or edge.

The following table summarizes the information about the articulatory features of the English consonants:
<table>
<thead>
<tr>
<th></th>
<th>Bilabial</th>
<th>Labiodental</th>
<th>Dental</th>
<th>Alveolar</th>
<th>Alveopalatal</th>
<th>Palatal</th>
<th>Velarar</th>
<th>Uvular</th>
<th>Pharyngeal</th>
<th>Glotal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plosive</td>
<td>p, b</td>
<td>t, d</td>
<td>k, g</td>
<td>?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasal</td>
<td>m</td>
<td>n</td>
<td>η</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trill</td>
<td></td>
<td></td>
<td>r</td>
<td></td>
<td>R</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flap</td>
<td></td>
<td></td>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fricative</td>
<td>f, v</td>
<td>θ, δ</td>
<td>s, z</td>
<td>s, ʒ</td>
<td>h</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Approximant</td>
<td>w</td>
<td>Ʉ</td>
<td>j</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>l</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Approximant</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affricate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ʒ, ɸ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>