In the previous chapter, we took a look at certain phonological constraints and processes. We saw that processes often apply to eliminate ill-formed sound combinations. We also found that some constraints are lexical: they are constraints on word forms; others are post-lexical, i.e., they are constraints on sound sequences. In this chapter, we’ll examine phonological constraints on word forms in detail. The part of phonology which deals with how phonemes can combine within a word is called phonotactics.

Let us start with some interesting observations. The first one concerns what sounds can start or end a word. In English as well as in Hungarian and all other languages, several generalisations can be made.

Firstly, many languages allow words to begin with a vowel or a consonant, e.g., English, Hungarian, French, Japanese, Swahili. There are, however, many languages which require a consonant word-initially, e.g., German, Czech, Standard Arabic. On the other hand, we don’t know any language to require that all words begin with a vowel, i.e., no language forbids word-initial single consonants.

Secondly, the end of the word behaves as a mirror image of the beginning of the word. There are languages — like English, Hungarian, German, Standard French — which have both consonants and vowels in word-final position. Then there are many which require that all words end in a vowel, e.g., Hawaiian, Italian, Southern French — in these languages, no word may end in a consonant. Curiously, there are no languages which forbid word-final vowels!

Thirdly, if a language allows consonant clusters at word edges, a similar “mirror image” situation is found: generally (though not without exception), consonant combinations which are not possible word-initially are found word-finally. No English (or Hungarian) word can begin with nt, lk, mp, rd, for example, but these clusters can occur finally, as shown by English words like count, bulk, lamp. (In RP, rd isn’t found in this position, either, but this is due to an independent constraint operating in many accents of English, incl. RP, Australian, New Zealand, Welsh and New England English, which requires that the phoneme r be followed by a vowel. In accents which lack this constraint — such as GA, Scottish and Irish English as well as most accents of the West of England — rd is fine word-finally, cf. card = GA kard, RP kaud.) On the other hand, many clusters that are well-formed word-initially, are impossible word-finally, e.g., tr, gl, kw, as in try, glue, queen. You are invited to check similar clusters in Hungarian — it’s pretty much the same!

There are many other observations of this kind, but let these three suffice as an appetizer. They clearly show that words in natural language are not simply a random string of phonemes. Earlier, we saw such constraints — for instance, that sb, although pronounceable, is not found in English within words. The great generality of the above observations, however, calls for a generalisation: they are so similar in the sense that all display a mirror-image behaviour with respect word-initial vs. word-final position that this is hardly a coincidence. Instead, we suspect them to be derivable from some more basic principle of sound structure found in all languages of the world.

1 In German and Czech, the spelling doesn’t reflect this, so, for example, the German word acht ‘8’, when written, looks as if it were axt. In fact, it is faxt, that is, it begins with a glottal stop. In this language, no word may start with a vowel, and if there’s no vowel lexically, a glottal stop is inserted by default. The same goes for Czech.

2 In Italian, there are some apparent exceptions, e.g., prepositions such as per ‘for’, in ‘in’, but these words never occur alone: they always attach to the following word, so they aren’t independently pronounced words.
This chapter examines these questions. Let us start our investigation with something that is seemingly unrelated to them. This is **length**.

We saw in Chapter 3 that both English and Hungarian possess long as well as short vowels. In Hungarian, several long vowels are articulatorily identical to short vowels: the vowel of *víz* viz for example is qualitatively the same vowel as that of *vizer* vizet: the two only differ in **quantity**, i.e., length, but both are front high unrounded, CV1. Not all Hungarian short and long vowels can be arranged into qualitatively identical pairs: the vowel of *keze* keze, for example, has no qualitatively identical long vowel phoneme; what is spelt <é> in Hungarian orthography, as in *kéz* kéz, is é, a long CV2, while è is CV3. The same is true for the pair illustrated by *nyara* — *nyár*, i.e., o — a; where the two vowels are, in fact, rather different, the short one being a raised CV13, a rounded raised-low back vowel, the long one being about CV4, i.e., low front unrounded. (You are invited to reflect on why Hungarian orthography “cheats”, pretending that kéz and keze, as well as nyár and nyara, have a qualitatively identical vowel in their first syllable.)

In English, short and long vowels cannot generally be arranged into qualitatively identical pairs, that is, a given quality automatically **implies** a given quantity. The vowel of *law* law, for example, is generally long (= aː), and there’s no short a phoneme. The only exception is é and è, as in *set* and *fairy*, respectively, but this is so in AdvRP only — Cons RP has a diphthong eo in words like *fairy*.³

Nevertheless, length has a very important role in English. Let us consider two phenomena: the first is the occurrence of short and long vowels in stressed monosyllabic words; the second is stress placement in English polysyllabic nouns.

If you consider the following words, a quite interesting picture emerges about the distribution of short and long vowels in English. Note that one English vowel, the schwa (ə) only appears in unstressed syllables, so it can’t appear as the vowel of a stressed monosyllabic word.

(1)

<table>
<thead>
<tr>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
<th>(d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pit</td>
<td>*pt</td>
<td>feed</td>
<td>fid</td>
</tr>
<tr>
<td>push</td>
<td>*puʃ</td>
<td>tool</td>
<td>tuːl</td>
</tr>
<tr>
<td>set</td>
<td>*se</td>
<td>scarce</td>
<td>*skəːs</td>
</tr>
<tr>
<td>lack</td>
<td>*læk</td>
<td>firm</td>
<td>fɜːm</td>
</tr>
<tr>
<td>love</td>
<td>*lʌv</td>
<td>card</td>
<td>kɔːd</td>
</tr>
<tr>
<td>cod</td>
<td>*kɔd</td>
<td>lord</td>
<td>lɔːd</td>
</tr>
</tbody>
</table>

What we find is that if the monosyllabic word ends in a single consonant (please do not be misled by the spellings, especially by the ones with a “silent” <r>: in RP, they contain no /ɹ/, so concentrate on the pronounced forms), the consonant can only be removed if the vowel is long. The forms in (b) are not only nonexistent, but also **ill-formed**: no English monosyllabic word can end in a short vowel. Long vowels are free to occur in this position (in other words, the nonexistence of forms in (b) is not a mere accident; as opposed to this, a form like gaː doesn’t exist either, but this is an accident: as shown by words like fur, cur, stir, etc., aː is perfectly okay at the end of a monosyllabic word). That is, gaː isn’t an English word but it could be; forms in (b) are not words and couldn’t be, either.

³ As opposed to Hungarian and English, Czech, for example, arranges long and short vowels into pairs without exceptions: each short vowel has a qualitatively identical long counterpart, and vice versa.
Let us now see diphthongs. If you check the data in (2), you’ll see that they behave just like long monophthongs. This is strange, since phonetically, long monophthongs are the same type of beast as a short monophthong, having a constant quality, while diphthongs are rather different, being composed of distinct vowel qualities.

(2)

<table>
<thead>
<tr>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>sight</td>
<td>s&quot;ait</td>
<td>sith</td>
</tr>
<tr>
<td>house</td>
<td>hauz</td>
<td>haw</td>
</tr>
<tr>
<td>made</td>
<td>meid</td>
<td>may</td>
</tr>
<tr>
<td>toil</td>
<td>t&quot;ail</td>
<td>taw</td>
</tr>
<tr>
<td>toad</td>
<td>t&quot;oad</td>
<td>tood</td>
</tr>
<tr>
<td>pierce</td>
<td>pil&quot;es</td>
<td>peer</td>
</tr>
<tr>
<td>cured</td>
<td>kju&quot;ed</td>
<td>cure</td>
</tr>
</tbody>
</table>

The picture above suggests that long vowels and diphthongs constitute a common class, as opposed to short vowels. In fact, this pattern is not found solely in monosyllables: instead, short stressed vowels are never found in word-final position. So, for example, allow o"lau, tattoo tae"tu, referee ref"o"ri, restore ri"sto, millionaire ,milja"ne, etc., are okay, forms like *o"la, *tae"to, *ri"sto, etc., would be impossible. We can formulate a general statement to sum up this important constraint of English. Before doing so, however, let us introduce the terms closed syllable, open syllable, and ult. A closed syllable is one that ends in a consonant; an open syllable is one which ends in its vowel. E.g., fe"ee consists of an open syllable, whilst the first syllable of plenti"ty contains a closed syllable. The term ult means “last syllable in a word”, so, for example, the ult of referee ref"o"ri is ri. In monosyllabic words, of course, the ult is the only syllable of the word! Here we go now:

(3) In English, the vowel of a stressed open ult may not be short.

We’ll reformulate this constraint later on. Let me note that this constraint is operative in all Germanic languages, not only in English, and in a number of other languages, too.

Let us now turn our attention to another topic: stress placement in English nouns. Except for nouns ending in certain suffixes, the majority of English nouns follows an interesting pattern as regards stress placement, i.e., which syllable receives the (main) stress of the word. Look at the following data; the dots in the transcriptions indicate syllable boundaries (where the stress mark doesn’t indicate it):

(4)

<table>
<thead>
<tr>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>table</td>
<td>'tei.bal</td>
<td>aroma</td>
</tr>
<tr>
<td>survey</td>
<td>'sa:ver</td>
<td>tomato</td>
</tr>
<tr>
<td>finance</td>
<td>'fa.na:ns</td>
<td>horizon</td>
</tr>
<tr>
<td>weather</td>
<td>‘we.doa</td>
<td>potato</td>
</tr>
<tr>
<td>city</td>
<td>’si:ti</td>
<td>veranda</td>
</tr>
<tr>
<td>habit</td>
<td>’hae.bit</td>
<td>utensil</td>
</tr>
</tbody>
</table>

Let’s now take a look at what characterises these nouns. In column (a), the nouns are disyllabic, i.e., they contain two syllables. In all cases, the stress is on the first syllable. We

4 The symbol , in IPA transcriptions indicates secondary stress. We can neglect it here.
may now state that disyllabic nouns are stressed on their first syllable. If you look at the nouns in (b) and (c), you’ll find that they are all polysyllabic: they contain more than two syllables.\(^5\) You can observe that none of them is stressed on the ult, the last syllable. We can generalise and say that English nouns are never stressed on the ult. This also explains why disyllabic nouns are stressed on their first syllable: as their second syllable is the ult, it’s not stressable, so the stress is placed onto the only remaining syllable — the first.

Let’s now take a look at polysyllabic nouns. The ones in (b) are stressed on their penult, the last but one syllable; the ones in (c) have the stress on the antepenult, the third syllable from the end. We’ll say that polysyllabic nouns have either a penultimate or an ultimate stress. But what does it depend on?

Note that I indicated something in the transcriptions which is not usually shown: syllable boundaries. This helps you to identify the sounds making up the individual syllables. Let’s take a look at the nouns in (b), those with penultimate stress. What can we say about the penult itself in these words? Careful observation shows that their penult is either an open syllable with a long monophthong or a diphthong, or else a closed syllable with a short vowel. Again, we find that diphthongs and long monophthongs behave identically. Furthermore, both behave in the same way as a closed syllable with a short vowel. Schematically, this is what the penult of the words in (b) ends in:

\[
V_1 V_2 = VC
\]

(The formula \(V_1 V_2\), of course, symbolizes diphthongs, while \(V_1\) stands for a long monophthong.)

This is exactly what we find at the end of a stressed monosyllable (= monosyllabic word), as shown by the data in (1) and (2)! This is indeed very curious.

But let us go on and see the nouns in (c). They have antepenultimate stress. What do they have in common? The stressed syllable varies a lot: it can end in a short vowel (as in remedy), a long vowel (as in Abraham) or a short vowel + consonant (as in bungalow). This is hardly what we’re looking for; but if we take a look at their penults, we find that they all end in a short vowel!

The most reasonable explanation for how stress is placed in nouns is to assume that it happens according to the following scenario:

\[
\text{(6) The stress placement procedure in English nouns}
\]

\begin{enumerate}
  \item Start scanning the noun from the end. Neglect the ult (because the ult of nouns is not stressable).
  \item Check the penult. If it ends a (i) long monophthong, or (ii) diphthong, or (iii) a short vowel plus a consonant, stress it.
  \item If the penult ends in a short vowel, stress the antepenult.
\end{enumerate}

Compare now (6c) with what we said about stressed final syllables in English in (3). Altogether, reformulating (3) a bit, we come to the following conclusions:

\(^5\) Strictly speaking, the term polysyllabic means ‘having more than one syllable’, but it is generally used for words of three or more syllables, and we’ll follow this practice, too.
(7)
(a) The **stressed ult of an English word must be** either (i) an open syllable containing a long monophthong or a diphthong, or (ii) a closed syllable. **It may not be** an open syllable with a short vowel.
(b) The **penult of English nouns is stressed** if it is either (i) an open syllable containing a long monophthong or a diphthong, or (ii) a closed syllable. **It is not stressed** if it is an open syllable with a short vowel.

It is impossible not to notice that the parts set in small caps in (7a) and (7b) are identical. This means that English seems to have a preference for stressed syllables **not to be open syllables with a short vowel**. But why are all the other types preferred? What is common in syllables which (i) are closed, (ii) contain a long monophthong, (iii) contain a diphthong? And why are open syllables with a short vowel different?

We can find the answer if we try to concentrate on the fact that a syllable-final VC sequence behaves just like a long vowel or a diphthong. Let’s approach the problem from the diphthongs. A diphthong is a combination of two short vowels in one syllable, as **ai**, for example (= *I ‘én’* or *eye ‘szem*, being monosyllabic). They behave identically to long monophthongs; we could say that long monophthongs are none other than two short vowels, too — **two identical short vowels**, forming one syllable! So we may provide an illustration as in (8), where a single box represents a short segment:

(8)  
<table>
<thead>
<tr>
<th>Short monophthong:</th>
<th>Long monophthong:</th>
<th>Diphthong:</th>
</tr>
</thead>
<tbody>
<tr>
<td>( V_1 )</td>
<td>( V_1 V_2 )</td>
<td>( V_1 V_2 )</td>
</tr>
<tr>
<td>( \varepsilon )</td>
<td>( \varepsilon \varepsilon )</td>
<td>( \varepsilon \varepsilon )</td>
</tr>
</tbody>
</table>

The figures in (8) show that both types of vowel occupy the same amount of “space”; the amount equal to two short vowels. Quantitatively, long monophthongs and diphthongs are equal to two short vowels. Let’s now consider VC combinations. A single consonant, too, is a short segment, so it can be represented as occupying one box in our model. Take, for instance, the VC sequence **en**, ending the stressed syllable of *utensil*.

(9)  
<table>
<thead>
<tr>
<th>Short monophthong + C</th>
<th>( V C )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \varepsilon )</td>
<td>( n )</td>
</tr>
</tbody>
</table>

Let’s make a model now in which the four possibilities are compared:

(10)  
<table>
<thead>
<tr>
<th>Short monophthong:</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \varepsilon )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Long monophthong:</th>
</tr>
</thead>
<tbody>
<tr>
<td>( V_1 V_2 )</td>
</tr>
<tr>
<td>( \varepsilon \varepsilon )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Diphthong:</th>
</tr>
</thead>
<tbody>
<tr>
<td>( V_1 V_2 )</td>
</tr>
<tr>
<td>( a i )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>VC sequence:</th>
</tr>
</thead>
<tbody>
<tr>
<td>( V C )</td>
</tr>
<tr>
<td>( \varepsilon \n )</td>
</tr>
</tbody>
</table>
The picture we have got shows the following: long monophthongs, diphthongs and VC sequences do have something in common — they are quantitatively equal to two short segments. To return to the question of stress placement in nouns, we can say that the penult of a noun is stressed if it ends in two short segments. But this formulation is totally unsatisfactory. What do we mean by “ends in two short segments”? Do we mean to say that it contains (at least) two short segments counted from the end? Alas, this isn’t possible. Consider the penult of the noun America: it is /rI/, so it does have two short segments counted from the end! Okay, you can say, but it doesn’t contain anything else, and maybe that’s the problem! Well, the penult of the word Abraham is /bra/, so it is longer than two segments, and yet, it fails to receive stress.

It appears that it doesn’t really count how many segments there are in the syllable. If you carefully observe the above, you will find that what counts is how many segments there are in the syllable disregarding the consonant or consonants before the vowel. That is, what is counted is the vowel and the following consonant(s) (if there’s any); these are relevant for purposes of stress placement, while the number of consonants before the vowel is totally irrelevant.

This divides the syllable into two parts: (1) the consonant(s) before the vowel, and (2) the vowel and the consonant(s) after the vowel. In phonology, part (1) is called the onset of the syllable, whilst part (2) is the rhyme. We can now reformulate the rule for English noun stress with reference to the rhyme, as in (11) (cf. (7b) above):

(11) The penult of English nouns is stressed if its rhyme contains what is equal to at least two short segments. If not, the antepenult is stressed.

This formulation is okay, though it is a bit clumsy. To make it easier and more elegant, let us introduce some technical terms used in phonology. The amount of segments in the rhyme determines rhyme weight as well as syllable weight. A rhyme with two short segments (remember that this, of course, includes long vowels!) is called heavy; accordingly, a syllable with a heavy rhyme is called a heavy syllable. If the rhyme contains but one short segment (= a short vowel), it is light, and a syllable having a light rhyme is a light syllable. (In other words, open syllables with a short vowel are light, all other syllables are heavy.) Let’s now reformulate the complete English nominal stress rule — this time, thankfully, in a final version:

(12) The stress placement procedure in English nouns
   d) Start scanning the noun from the end. Neglect the ult: it’s not stressable.
   e) Check the penult. If it is heavy, stress it.
   f) If the penult is light, stress the antepenult.

Now it’s time to see what we can say about stressed ults with reference to syllable weight, as in (13) (cf. (7a)):

(13) The stressed ult of an English word must be heavy.

The constraint in (13) is a very strong one: it admits no exceptions. As mentioned earlier, the same constraint is at work in all Germanic languages, not only in English (students of German are invited to check this). The stress rule for nouns, too, shows that English has a general preference for stressed syllables to be heavy: the penult is stressed only if it is heavy, otherwise it isn’t. Note, however, that the weight of the antepenult doesn’t matter: if the

---

6 The word rhyme can also be written rime.
penult isn’t stressed, the antepenult is, whether heavy or not. The stress rule, as it were, doesn’t move further away from the end of the word than the antepenult. Another thing to note is that the nominal stress rule, like all other stress rules in English, are influenced by a variety of factors (such as part of speech, suffixes, etc.), and they also have exceptions.\(^7\) So stress rules are not strict, watertight constraints in English but rather tendencies that hold for the majority of words but not all of them. We’ll come back to the problem of stress in Chapter 7.

Let us now return to the internal division of the syllable. We saw that it contains two parts: the onset and the rhyme. The two parts are quite independent from each other; for example, we have seen that the onset has no influence on the weight of the syllable, that, is syllable weight is none other than rhyme weight. Furthermore, in poetic usage, the two parts behave independently, too. In rhyming poetry, for instance, it is the rhyme that counts: two syllables rhyme if their rhymes are identical, no matter what their onsets are. For example, *lent*, *Kent*, *spent*, *ent* all rhyme with each other. (Still wondering where the term *rhyme* comes from?) In alliterative poetry, on the other hand, it is onsets (or part of them) that agree: the rhymes are completely irrelevant. Another important fact showing the independence of the rhyme and the onset is that many languages (including English and Hungarian) allow onsetless syllables, but no language has syllables without a rhyme: the onset may be optional, but the rhyme is obligatory.

The rhyme, however, can be divided into two parts. If you observe what is obligatory and what is optional within the rhyme, you’ll find that the consonant or consonants in it are optional: all languages, without exception, allow rhymes to consist of a vowel only without a following consonant or consonants: in other words, we find open syllables in all languages. The vowel in the rhyme, on the other hand, is not optional: instead, it is the central part of the syllable, around which optional consonants are gathered. The number of syllables in a word is the same as the number of vowels in it. The vowel and the following consonants seem to form different parts of the syllable, therefore. The consonant or consonants after the vowel are said to constitute the coda of the syllable. For example, the coda of the stressed syllable of *utensil* is *n*. The stressed syllable of *America*, on the other hand, doesn’t have a coda: it’s an open syllable.

Let us now go back to the part that remains if we remove the onset and the coda: the vowel. In fact, to use the term “vowel” to refer to this part misleading. It is true that it tends to consist of a vowel, but not always. In a number of languages, including English, it may consist of a consonant, too. For example, the word *utensil*, transcribed as *juːtɛnsɪl* earlier, can also be pronounced without a schwa, as *juːtɛnsɪl*. Nonetheless, it still consists of three syllables, because the *l* takes over the function of the dropped schwa, and it serves as the central element in the last syllable: the last syllable of *juːtɛnsɪl* is *sl*, with a syllabic *l*. Similarly, the words *even*, *bottom* are pronounceable as *ɪvn*, *bɒtм*, with a syllabic nasal in the second syllable. Generally speaking, the liquids and the nasals can be syllabic in English, but words with a syllabic consonant are always pronounceable with a schwa + consonant sequence, too, e.g., *ɪvən, bɔtəm*; furthermore, syllabic consonants are only found in unstressed syllables in English, just like its sister tongue German.\(^8\) In IPA transcription, syllabic consonants can be indicated by a little subscript vertical line as, for example, in *ɪvən, bɔtəm, ˈsɛntrəl*, etc.

Since the central part of the syllable is not necessarily a vowel, we give it an own name: it’s called the nucleus (plural nuclei).\(^9\) In English, then, vowels as well as certain

\(^7\) For example, some disyllabic nouns are stressed on the ult, such as *event* ˈɪvənt; sometimes the penult is stressed although light, as in *vanilla* ˈvænɪ.lə, or it isn’t stressed although heavy, e.g., *character* ˈkeɪrətər.

\(^8\) But not in all languages: in Czech, for example, syllabic liquids are frequent, but they only occur in stressed syllables. Many languages, such as Hungarian or Polish, have no syllabic consonants at all.

\(^9\) Several linguists use the term peak for nucleus.
sonorant consonants can function as syllable nuclei, while in Hungarian, the nucleus is always a vowel. Altogether, then, the syllable can be conceived of as an arboreal structure, as in (14):

(14)  

```
  Syllable
    /\  
   /  
  Rhyme
    /\  
   /  
Onset  Nucleus  Coda
```

The onset, the nucleus, the coda and the rhyme are the constituents of the syllable. Of these, the onset and the rhyme are immediate constituents, as the syllable is directly decomposable into an onset and a rhyme, then the rhyme is decomposable into a nucleus and a coda. The constituents of the syllable are often given in an abbreviated form as O, N, C, and R — I guess these abbreviations are quite self-explanatory — especially in tree diagrams, where the syllable itself is marked with a lowercase Greek letter sigma, σ.

Recall that our investigation has shown that diphthongs, long monophthongs and VC sequences all make a heavy rhyme, as they occupy as much space as two short segments. Note that whether the segment in the second position is a vowel or a consonant is irrelevant; in other words, two short segments in the nucleus (= a long vowel or a diphthong) make a heavy rhyme just like one short segment in the nucleus (= a short monophthong, or — let’s now add this too — a syllabic consonant10) plus one short segment (= a consonant) in the coda. This shows that quantity is independent of quality: a rhyme is heavy if it has two “filled” segment-sized spaces, no matter what kind of sound is in it. Let us illustrate this equivalence, using as examples the stressed rhymes of arena əˈrɛnə, horizon həˈraɪzən and veranda vəˈrændə (note that i + i = i):

(15)  

```
(15) R  R  R  |  N  N  N  |  C
     \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |   \ |
enter 'entə (= 16a, a closed syllable + a light syllable) and survey 'sərvə (= 16b, with two heavy open syllables, the first containing a diphthong, the second a monophthong)\(^{11}\):

\[
\begin{array}{cccc}
\text{(a)} & \sigma & \sigma & (b) \\
R & O & N & C & R & O & N \\
O & N & C & O & N & O & N \\
X & X & X & X & X & X & X \\
\varepsilon & n & t & \aleph & s & \bar{g} & v & e & i \\
\end{array}
\]

The nucleus of the first syllable is a short monophthong, \(\varepsilon\). This is represented in the diagram by linking the sound’s IPA symbol to one X. Contrast this with \(\varepsilon\)\(\acute{\varepsilon}\), a long monophthong, where the phonetic material (= \(\varepsilon\)) is linked to two X-es. Note that the length mark (\(\acute{\varepsilon}\)) is not used in the tree diagram, since it’s unnecessary: the fact that the vowel is long is shown by the two X-es. In such tree diagrams, quality is clearly separated from quantity, the latter being represented by the line of X-es. This line is generally called the timing tier,\(^{12}\) because it shows units of quantity: timing positions or slots, in the form of X-es. The bottom line, showing phonetic material associated with timing slots is called the melodic tier, because phonetic material (= segments) is often called melody in phonological theory. (This has nothing to do with singing, of course.) The fact that a rhyme containing a long nucleus and a rhyme containing a short nucleus plus a coda are equally heavy ones is beautifully expressed in such tree diagrams: both types of rhyme are the same at the level of the timing tier, occupying two slots. Stress rules, for example, only take the timing tier into consideration: what melody is associated with the relevant slots does not matter.

The difference between two qualitatively identical but quantitatively different phonemes, such as \(\varepsilon\) (as in set) and \(\varepsilon\)\(\acute{\varepsilon}\) (as in fairy) is rather simple to represent, too: the same melody (= phonetic content, so, in this case “mid-low front unrounded vowel”) is associated with one (\(\varepsilon\)) or two (\(\varepsilon\)\(\acute{\varepsilon}\)) timing slots, as in (17a) vs. (17b). The representation of a hiatus (pron. ha\’ret\’\-as) is easy, too. A hiatus is a sequence of two vowels in separate syllables without an intervening consonant, as in Hungarian te.a, ki.oldoz, le.esik. In the last example, leesik le.<jík> two identical vowels sit next to each other, but they remain in different syllables — i.e., different nuclei — rather than forming a long nucleus. Hiatuses can be represented as in (17c):

\[
\begin{array}{cccc}
\text{(17)} & \text{(a)} & \text{(b)} & \text{(c)} \\
N & N & N & N \\
X & X & X & X \\
\varepsilon & \varepsilon & \varepsilon & \varepsilon \\
\end{array}
\]

\(^{11}\) Note that we postulated an empty word-initial Onset in enter. We’ll discuss this later on.

\(^{12}\) The word tier is pronounced tıa. In Hungarian, it’s called tengely.
Recall from Chapter 2 that there is an interesting type of consonant: affricates, such as \( \text{tʃ}, \text{dʒ} \) as in *chain, jam*. They start as plosives but end as a fricative. In fact, you may rightly ask why we consider them to be single consonants? Why don’t we say, for example, that the \( \text{tʃ} \) in *kitchen* \( \text{'kɪtʃn} \) is a consonant cluster, similar to \( \text{kʃ} \), as in *action* \( \text{'ækʃn} \)? After all, both are combinations of a stop plus a fricative!

The answer is that affricates behave as single consonants quantitatively. Consider nouns like *prodigy, astrology, phonology, intelligence*, etc. They are all stressed on the antepenult, which suggests that their penult is light. If \( \text{dʒ} \) counts as a single short segment, this is understandable: *prodigy*, i.e., \( \text{'prədʒi} \), has a light penult just like *remedy* \( \text{'rɛ.mə.dɪ} \). If, however, \( \text{dʒ} \) is none other than a sequence of \( \text{d} \) plus \( \text{ʒ} \), the \( \text{d} \) should belong to the coda of the penult, since a stop + a fricative in English cannot constitute the onset of a syllable; as a result, *prodigy* would be syllabified as \( \text{'prə.dʒi} \), but that would mean that the penult is heavy, and it ought to be stressed. As words like *prodigy* are stressed on the antepenult, this cannot be the case: the affricate must be a single segment, quantitatively speaking. In fact, our model can accomodate this very easily: all we need to say is that melodically, affricates are composed of a stop plus a fricative, but the two parts are linked to a single timing slot:\(^{13}\):

\[
(18) \quad \text{X}
\]
\[
\text{t} \quad \text{ʃ}
\]

As syllable weight is determined on the basis of how many timing slots (X-es) are inside the rhyme, words like *prodigy* will have a light penult:

\[
(19)
\]
\[
\sigma \quad \sigma \quad \sigma
\]
\[
\text{R} \quad \text{R} \quad \text{R}
\]
\[
\text{O} \quad \text{N} \quad \text{O} \quad \text{N} \quad \text{O} \quad \text{N}
\]
\[
\text{X} \quad \text{X} \quad \text{X} \quad \text{X} \quad \text{X} \quad \text{X} \quad \text{X} \quad \text{X}
\]
\[
\text{p} \quad \text{r} \quad \text{d} \quad \text{a} \quad \text{d} \quad \text{ʒ} \quad \text{i}
\]

To conclude the discussion, I would like to take an example from Hungarian, to show you that quantity and quality are treated separately not only in English but in natural language in general. The example involves a widespread phenomenon in the languages of the world: **compensatory lengthening**.

In many accents of Hungarian, the liquid \( \text{l} \) is deleted before a consonant. For example, the word *csinált* \( \text{tʃjínalt} \) is pronounced as \( \text{tʃjínat} \), or *szólsz* \( \text{sɔls} \) is pronounced as \( \text{sɔs} \).\(^{14}\) (In

\(^{13}\) Another argument for affricates being linked to a single X is that other plosive + fricative clusters (such as \( \text{pf}, \text{kʃ}, \text{ks} \), etc. are never found in word-initial position.

\(^{14}\) The latter becomes homophonous with *szősz*, and is often spelt so in informal written discourse such as in emails, e.g., *Mit szősz hozzá?*
several dialects, this is the norm, but in standard Hungarian it is confined to casual styles.)
Now, if the preceding vowel — the nucleus — is short, something interesting happens. For example, the deletion of l in the words voltam, elment results in the forms votom, vement, rather than *votom, *ement — the latter variants simply do not exist. In other words, the short nucleus is lengthened. Why?
If we consider the fact that melody and quantity are separate things, the phenomenon becomes understandable: the l is deleted from the melodic tier, but its timing slot remains. The vowel in the nucleus spreads into the vacated slot, becoming long — to “make up” for the loss of the consonant! This is why such lengthening is called compensatory: it occurs as a kind of compensation. Melody is lost, but quantity is preserved. This is shown in (20), where the first syllable of vol.tam is represented:

(20)  (a)  (b)  (c)

Step (a) is the initial stage, representing the lexical form of the syllable. In step (b), the deletion process applies: the l is removed from its slot, but the slot remains; the o is linked to the empty slot, becoming long. In step (c), the slot becomes associated to the Nucleus. The l is no longer there, but its slot remains: the syllable preserves its heaviness: it now contains two segments on the melodic level, as opposed to three, but the number of timing slots is the same. Of course, if the nucleus is long (as in csinált), no compensatory lengthening takes place, since the vowel would become overlong, occupying three X-es, and that’s impossible. Furthermore, the deletion of l after a long Nucleus doesn’t alter the weight of the Rhyme, anyway, since a long Nucleus makes a heavy Rhyme.

Let us now see the details about the structure of syllabic constituents themselves.

1. The Onset
Assuming for the time being that the initial consonants in a word constitute the Onset of the first syllable of the word, it appears that the Onset in English may contain from zero to 3 consonants, as testified by words such as act, pit, pray, spray. We do not find any English
word to begin with more than 3 consonants. Let’s see if this assumption is correct and what exactly constrains the logically possible consonant combinations as Onsets.

First, what about words beginning with a single consonant? In general, any English consonant can begin a word if immediately followed by a vowel, e.g., *bin, pin, zeal, sit, have, key, node, ray, yet, way*, etc. There’s one exception, however: the velar nasal, i.e., *ŋ*, which never occurs in initial position. We’ll come back to this later.

Second, let us make a jump and check words beginning with 3 consonants. These invariably involve consonant clusters whose first member is *s*, its second member is a voiceless plosive (i.e., *p, t, or k*), and its third member is an approximant, cf. the words in (21):

(21)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>spray</td>
<td>spreɪ</td>
</tr>
<tr>
<td>l</td>
<td>split</td>
<td>split</td>
</tr>
<tr>
<td>j</td>
<td>spew</td>
<td>spjuː</td>
</tr>
<tr>
<td>w</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

As you can see, there are three gaps: there are no words beginning with *spw, *stl, or *stw*. It is not necessarily the case, however, that these initial clusters are non-existent because they are ill-formed. Recall from Chapter 1 that there are accidental gaps in language; e.g., there’s no English word such as *blik*, though it would be perfectly well-formed. It is also possible that we do not find words beginning with *spw, *stl, or *stw for the same reason. It needs further investigation to be able to say something about this.

Consider the words listed in (21). If we remove the initial *s*, and check the remaining 2-consonant clusters, we’ll notice that these clusters are all well-formed in word-initial position:

(22)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>pray</td>
<td>preɪ</td>
</tr>
<tr>
<td>l</td>
<td>play</td>
<td>pleɪ</td>
</tr>
<tr>
<td>j</td>
<td>pure</td>
<td>pjʊə</td>
</tr>
<tr>
<td>w</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

It appears that 3-member clusters are made up of an *s* followed by a well-formed intial 2-member combination. For example, *pr* is a well-formed cluster initially, and so is *spr, etc.*; but *pw* seems to be ill-formed, and so is *spw as a result.

As shown by the word *twin*, marked with a *, tw-initial words do exist. It seems, therefore, that the lack of *stw*-initial words is indeed an accidental gap, but that of *spw or *stl initial ones is not. We can make the following generalisation, using the symbol T for any voiceless plosive and the symbol R for any approximant: If initial TR is well-formed, so is sTR.

The *s*, therefore, appears to be insensitive as to what exactly follows the voiceless plosive: all that matters is that the *s* itself must be followed by a voiceless plosive, and the voiceless plosive and the following approximant must be well-formed in word-initial position.

Needless to say, *s* + a voiceless plosive (without a following approximant) is always well formed intially, cf. *spv, stv, skv.
Let's now examine 2-consonant initial clusters. Here, we find a huge variety of combinations, and the task is to find out if we can make a generalisation about what these combinations have in common.

Remember that we made an observation at the beginning of the chapter. Notably, we seem to find that the mirror image of word-initial clusters does not occur in the same position. This is true for a range of languages, not only for English. For example, initial tr is okay, but *rt is not. We will make use of this observation in our investigation. Let's now see which 2-member clusters are found in initial position in English.

(23)

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>l</td>
<td>j</td>
<td>w</td>
<td>m, n</td>
<td>p, t, k</td>
</tr>
<tr>
<td>1</td>
<td>p</td>
<td>pray</td>
<td>play</td>
<td>pure</td>
<td>(pueblo)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>preɪ</td>
<td>pleɪ</td>
<td>pjʊə</td>
<td>ˈpweɪblou</td>
</tr>
<tr>
<td>2</td>
<td>t</td>
<td>tray</td>
<td>tune</td>
<td>twin</td>
<td>twin</td>
</tr>
<tr>
<td></td>
<td></td>
<td>treɪ</td>
<td>tju:n</td>
<td>ˈtju:n</td>
<td>—</td>
</tr>
<tr>
<td>3</td>
<td>k</td>
<td>cry</td>
<td>clue</td>
<td>cute</td>
<td>queen</td>
</tr>
<tr>
<td></td>
<td></td>
<td>kraɪ</td>
<td>kluː</td>
<td>ˈkluː</td>
<td>ˈkwin</td>
</tr>
<tr>
<td>4</td>
<td>b</td>
<td>brew</td>
<td>blue</td>
<td>'bjuːgl</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>bruː</td>
<td>ˈbruː</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>5</td>
<td>d</td>
<td>drain</td>
<td>dune</td>
<td>dwell</td>
<td>dwell</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dreɪn</td>
<td>djuːn</td>
<td>ˈdjuːn</td>
<td>—</td>
</tr>
<tr>
<td>6</td>
<td>g</td>
<td>green</td>
<td>glue</td>
<td>Gwen</td>
<td>Gwen</td>
</tr>
<tr>
<td></td>
<td></td>
<td>grɪn</td>
<td>gjuː</td>
<td>ˈɡwɛn</td>
<td>—</td>
</tr>
<tr>
<td>7</td>
<td>f</td>
<td>fry</td>
<td>fly</td>
<td>few</td>
<td>thwart</td>
</tr>
<tr>
<td></td>
<td></td>
<td>fraɪ</td>
<td>fliː</td>
<td>ˈfliː</td>
<td>ˈθwɔːt</td>
</tr>
<tr>
<td>8</td>
<td>θ</td>
<td>through</td>
<td>—</td>
<td>(Thule)</td>
<td>thwart</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ˈθruː</td>
<td>(ˈθjuːliː)</td>
<td>ˈθwɔːt</td>
<td>—</td>
</tr>
<tr>
<td>9</td>
<td>v</td>
<td>(vroom)</td>
<td>(Vlad)</td>
<td>view</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(vrʊːm)</td>
<td>(vlɒd)</td>
<td>vjuː</td>
<td>—</td>
</tr>
<tr>
<td>10</td>
<td>z</td>
<td>—</td>
<td>—</td>
<td>(Zeus)</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>—</td>
<td>—</td>
<td>(zjuːs)</td>
<td>—</td>
</tr>
<tr>
<td>11</td>
<td>m</td>
<td>—</td>
<td>—</td>
<td>muse</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>—</td>
<td>—</td>
<td>mjʊz</td>
<td>—</td>
</tr>
<tr>
<td>12</td>
<td>n</td>
<td>—</td>
<td>—</td>
<td>new</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>—</td>
<td>—</td>
<td>njuː</td>
<td>—</td>
</tr>
<tr>
<td>13</td>
<td>l</td>
<td>—</td>
<td>—</td>
<td>(Luke)</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>—</td>
<td>—</td>
<td>(ljuːk)</td>
<td>—</td>
</tr>
<tr>
<td>14</td>
<td>s</td>
<td>—</td>
<td>slim</td>
<td>swine</td>
<td>smile</td>
</tr>
<tr>
<td></td>
<td></td>
<td>—</td>
<td>slɪm</td>
<td>ˈswain</td>
<td>ˈsmɛl</td>
</tr>
<tr>
<td>15</td>
<td>f</td>
<td>shrine</td>
<td>schlep</td>
<td>schwa</td>
<td>schmuck</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ʃraɪn</td>
<td>ʃlep</td>
<td>ʃwəː</td>
<td>ʃmʊk</td>
</tr>
<tr>
<td></td>
<td></td>
<td>—</td>
<td>—</td>
<td>schwa</td>
<td>ʃmʊk</td>
</tr>
<tr>
<td></td>
<td></td>
<td>—</td>
<td>—</td>
<td>ʃwəː</td>
<td>ʃmʊk</td>
</tr>
<tr>
<td></td>
<td></td>
<td>—</td>
<td>—</td>
<td>ʃwəː</td>
<td>ʃmʊk</td>
</tr>
</tbody>
</table>

Let us now check what generalisations we can make based on the above table.
Note that only $s$ and $\mathfrak{j}$ can precede a following nasal (cf. 14-15/EF); note that other consonants cannot. Of the two fricatives, $\mathfrak{j}$ only occurs — except before $r$, cf. 15/A — in a handful of recent German and Yiddish loans. However, as native speakers of English can happily and easily pronounce initial $\mathfrak{j}m$ and $\mathfrak{j}n$ (as well as $\mathfrak{j}l$ and $\mathfrak{j}w$) sequences and no attempt is made to replace the $\mathfrak{j}$ with $s$, we must regard these sequences as well-formed, though rare as yet. The fact that only $s$ and $\mathfrak{j}$ can stand before a nasal stop is rather suspicious, especially because only $s$ (and, of course, in a couple of German loans, $\mathfrak{j}$ — e.g., spiel $\mathfrak{j}p$) can appear initially before an oral stop, i.e., a plosive, too. In other words, only $s$ and $\mathfrak{j}$ are permitted in English before a following stop, either oral or nasal, in word-initial position. No other fricative is (let alone sonorants or plosives). Remember also that 3-member initial clusters always begin with $s$. Therefore, initial clusters beginning $s$ and $\mathfrak{j}$ behave differently from other clusters, so must be treated with caution, and we will indeed come back to them.

Let us, therefore, observe those combinations first which do not begin with $s$ or $\mathfrak{j}$ (rows 9-13 in the table). The following points are to be noted:

1) The mirror image of these clusters is not possible initially. Thus, $tr$, for example is found in word-initial position, but $*rt$ isn’t.
2) The second member of the cluster is always an approximant, i.e., $r$, $l$, $j$, or $w$.
3) There are some gaps, though: initial $*tl$, $*dl$, $*bw$, $*fw$, $*vw$, $*\theta l$ are non-existent. In fact, $pw$ is very rare, too: to the best of my knowledge, it is only found in the Spanish loan pueblo. Interestingly (according to the Longman Pronunciation Dictionary), an alternative pronunciation also exists: $pu\tilde{e}b\acute{l}ou$. This may well reflect the fact that many speakers of English actually find the initial $pw$ unusual and difficult, and replace the $w$ with an $u$: (something that Hungarians will also do). In sum, I regard initial $*pw$ as ill-formed, too; accordingly, I’ll treat pueblo (when pronounced with an initial $pw$) as an exception — an obvious foreignism. Therefore, $*tl$, $*dl$, $*bw$, $*fw$, $*vw$, $*\theta l$ are all considered to be ill-formed initially; their boxes are shaded in the table.

Let us now ask the following questions:

1) Why isn’t the mirror image of the well-formed clusters in rows 1-13 found initially (e.g., $*rt$, $*lp$, $*lf$, etc)? Again, some language do have such combinations (e.g., Russian $rta$ ‘of the mouth’), but again, these languages all have plosive + approximant clusters initially. There is no language to forbid initial $tr$ while permitting initial $rt$, etc.

It would be beyond the scope of this textbook to present detailed arguments from English or other languages, so I will limit the discussion to presenting some generally accepted observations and explanations.

First, let’s use the following abbreviations for the sake of brevity: $T =$ any plosive, $K =$ any plosive different from $T$ (e.g., $TK$ will stand for a cluster of two different plosives), $R =$ any approximant, $N =$ any nasal consonant, $S = s$ or $\mathfrak{j}$, $F =$ any fricative other than $S$.

To answer the question why initial $TR$, $FR$, $NR$ but not $*RT$, $*RF$, $*RN$ clusters are well-formed, we must make reference to sonority. Recall from Chapter 2 that sonority means

---

15 In German, $s$ is not possible in word-initial position, being replaced by $\mathfrak{j}$ before a consonant (and by $z$ before a vowel). The same goes for Yiddish, a close relative of German, as regards words of Germanic origin; however, due to the large amount of Slavonic and Hebrew loans with initial $s$, Yiddish has given up this restriction and allows initial $s$ both pre-consonantally and pre-vocally. (Ádám Nádasdy, p.c.)
proportion of voice vs. noise; and speech sounds can be arranged along a scale according to their sonority level. The most sonorous sounds are vowels and glides, since they contain only voice, no noise. The least sonorous ones are voiceless obstruents, which contain no voice at all. Now, in phonetic terms, our sonority scale is then (where sonority increases from left to right):

\[(24)\quad \text{voiceless obstruents} < \text{voiced obstruents} < \text{nasals} < \text{liquids} < \text{vowels and glides}\]

In fact, this phonetically based sonority scale must be modified for our purposes. Recall that we are doing phonology here, not phonetics; and we already seen that the phonological behaviour of sounds may differ from their phonetic behaviour. Glides, for example, though phonetically vowels, are treated as consonants in phonology, simply because they function that way (e.g., they are unable to function as Nuclei, just like consonants, and unlike vowels). As regards sonority, this manifests itself in that glides behave as less sonorous segments than vowels. This is understandable from a phonetic point of view, too: with glides, there is a narrower closure in the oral cavity than with vowels. The other point of divergence concerns obstruents. Although voiced obstruents contain more voice than voiceless ones, in terms of phonological function, fricatives behave as more sonorous segments than plosives, regardless of voice. That is, even a voiceless fricative is more sonorous phonologically speaking than a voiced plosive. Note that again, plosives are characterised by complete closure (they are non-continuants), while fricatives are produced with a narrow (but not complete) closure; the situation is similar to what we have just seen in the case of glides vs. vowels. Phonological sonority, then, is determined partly by content of voice (phonetic sonority), partly by degree of closure, which is why the above sonority scale needs revision. The final version of our sonority scale, then, is as follows:

\[(25)\quad \text{plosives} < \text{fricatives} < \text{nasals} < \text{liquids} < \text{glides} < \text{vowels}\]

Of course, voiceless plosives are less sonorous than voiced ones, and voiceless fricatives are less sonorous than voiced fricatives, but this is not quite relevant for us. For the sake of clarity, it is useful to give numbers to each sonority level in (25). Let’s take plosives to be level 1, fricatives are level 2, nasals 3, liquids 4, glides 5, and vowels 6. So, for example, the sonority profile of the word trend is 14631, while quiz is 1562:

\[(26)\]

Now, returning to the question of initial TR vs. *RT etc. clusters, observe that TR clusters show a rising sonority profile, as a R is more sonorous than a T. Conversely, RT clusters exhibit a falling sonority profile. TR clusters are fine word-initially; RT clusters are okay word-finally. In monosyllabic words possessing both Onsets and Codas, therefore, sonority
first rises, reaches its maximum in the Nucleus, then it falls. This observation has been called
the Sonority Sequencing Principle:

(27) Within a syllable, sonority falls from the Nucleus towards the edges of the syllable.

Hence, a syllable is none other than a sonority peak, as shown in the diagrams in (26): observe
that in both examples, sonority is level 6 in the Nucleus, and it decreases towards either edge
of the syllable. (The term sonority peak is quite illuminating: the diagrams do indeed resemble
mountain peaks!)

With the help of the Sonority Sequencing Principle (SSP), we can define an Onset
cluster as follows:

(28) A consonant cluster constitutes an Onset iff it exhibits a rising sonority profile\textsuperscript{16}.

It is quite generally accepted by phonologists that the SSP is a universal principle of syllable
structure, which is operative in all languages, and it explains why TR, but not RT clusters are
typical in word-initial position: assuming that a word-initial cluster is the Onset of the word’s
first syllable, RT clusters will contradict the SSP. If you go back to table (23), and check the
two-member initial clusters listed there, you will find that all of them obey the SSP. You can
now play an interesting game: change the order of consonants in each box. For example,
replace $\emptyset r$ with $r\emptyset$, $s m$ with $m s$, etc. In all cases, you will get clusters with a falling sonority
profile. Check if you can find English words beginning with these clusters! (I bet you suspect
by now what the result of your search will be, so I may as well tell you that you won’t find a
single word like that!)

The other question to be answered is, What’s wrong with $*t l$, $*d l$, $*b w$, $*p w$, $*f w$, $*f w$, $*0 l$? These clusters meet the SSP, after all — $*t l$, $*d l$ are 14, just like $k l$
and $g l$, which are good Onsets; similarly, $*b w$ and $*p w$ are 15, just like the well-formed $g w$
and $k w$; $*f w$, $*f w$ are 25, and $*0 l$ is 24.

The problem here is that these ill-formed clusters consist of consonants which have the
same place of articulation, i.e., these are homorganic clusters. In the case of $*b w$, $*p w$, $*f w$, $*v w$, both consonants are labials; as for $*t l$ and $*d l$, both clusters consist of
alveolars. In the case of $*0 l$, the first consonant is dental, the second is alveolar, so they do
not have exactly the same place of articulation, but English seems to treat them as having the
same place of articulation for combinatorial purposes. This may sound strange, but note that
the same is valid for palato-alveolars and palatals, too. Note the lack of initial $*r j$ and $*3 j$, for
example, both being composed of a palato-alveolar C + yod, a palatal glide. These clusters
have the same sonority profile as $l j$ (45) and $z j$ (25), respectively, which are found initially.
The lack of $*r j$ and $*3 j$, then, is possible to explain if we assume that palato-alveolars and
palatals are treated for combinatorial purposes as being homorganic. English Onsets,
therefore, may not consist of two consonants with the same place of articulation. Note,
however, that again, $s$ and $j$ behave specially, witness the possible (and frequent!) homorganic
sequences $s l / s n$ (alveolars) and $f r$ (both palato-alveolar), not to mention the cluster $s t$
alveolars), mentioned earlier. It is time we discussed SC sequences in detail (where S stands
for either s or $j$; of course, C stands for any consonant).

Let us sum up, then, why SC clusters are special.

\textsuperscript{16} The word \textit{iff} is not a misprint: it is an abbreviation of \textit{if and only if}, used in mathematics, too.
1) Observe that ST clusters — i.e., \(sp, st, sk\) — contradict the SSP: each has a falling sonority profile: 21. No other initial cluster does. Note also that no other fricative can stand before a T initially, so there is no word-initial \(*ft, fk, fp, \theta p, \theta k\), etc. Therefore, all other fricatives are perfectly well-behaved with regard to the SSP: we find ST clusters but no FT clusters initially.

2) Only S is possible as the first member of the cluster if the second one is a nasal consonant, so that SN is fine, but FN or TN isn’t. It seems, then, that nasals in general cannot function as the second member of an initial cluster; again, the “bad guys” here are s and fj.

3) As mentioned above, SC sequences can be homorganic ones: sl/sn/st (alveolars) and fr (both palato-alveolar). No other initial cluster can be homorganic.

4) In three-member initial clusters, the first C is always s, followed by a well-formed combination of the TR type. There is no 3-member initial cluster of any other type.

As you can see, SC clusters are special in as many as four ways. In fact, they do not behave like the great majority of initial clusters at all. So far, we tacitly assumed that word-initial consonants constitute the Onset of the first syllable of the word. However, it is not at all obvious that this assumption is correct. Indeed, if we reject this assumption and adopt the view that initial SC clusters are not Onsets, a great number of generalisations can be made about English Onsets. Recall how we defined a possible Onset cluster (see (28)). We can now add the other observations we made to our definition; accordingly we will define an Onset as follows (don’t forget that SC clusters are not taken into account):

(29) English Onsets
a) Any single C save \(nj\) can function as an Onset.

b) The maximum number of consonants in an Onset is 2. In a 2-member Onset, the following conditions hold:
   i) The cluster exhibits a rising sonority profile.
   ii) The second member of the Onset is an approximant.
   iii) The Onset cluster is not homorganic.

Note that if SC clusters — including, of course, STR ones — are taken to be Onsets, all these generalisations are lost. As a result of our analysing initial SC clusters (more precisely, most of them, see the second note below) as non-Onsets, we will have the following 2-member Onsets in English:

(30) Possible English 2-member Onsets

\[
\begin{array}{cccccccc}
pr & pl & pj & tw & tr & tj & kw & kr & kl & kj \\
br & bl & bj & dw & dr & dj & gw & gr & gl & gj \\
fr & fl & fj & \theta r & (\theta j) & sw & (sj) & (lj) & mj & nj \\
(vr) & (vl) & vj & (zj)
\end{array}
\]

Some notes are in order. First, \(vr\) and \(vl\) are quite marginal: the former only occurs in the onomatopoeic word vroom, while the latter is only found in foreign names (e.g., Vlad, Vladimir, Vladivostok). For this reason, I take them to be exceptional. Note also that voiced fricatives are typically missing from English 2-member Onsets.
Second, note that I included two s-initial clusters, viz. sw and sj, as Onsets. This is because of all SC clusters, these two do qualify as Onsets, since they meet the requirements listed in (29). Accordingly, we will treat these two (and only these two) SC clusters as Onsets.

Third, the clusters coronal fricative + yod (= θj, sj, zj) are placed between parentheses because they are very rarely heard in AdvRP. In conservative RP, words such as Thule, suit, Zeus are variably pronounced as θju:li ~ θui:li, sjut ~ sut, zju:s ~ zu:s. In AdvRP, the yodless forms are generally heard. The same is valid for lj, which is so rarely heard in stressed syllables nowadays that it must be considered archaic; in unstressed syllables, on the other hand, it occurs freely (cf. value vælju), which is why it is included in our list of Onsets. General American has gone even further in the simplification of Cj Onsets: in that accent, tj, dj, and nj aren’t found, either, witness the GenAm pronunciation of tune, due, new (tu:n, du:, nu:), but again, these clusters are found in GenAm, too, in unstressed syllables (e.g. venue is pronounced ’venju: in both RP and GenAm).

Fourth, yod is special in another sense, too: it can only be the second member of the Onset if it’s followed by u/o/a — as in cute, cure17; no other nucleus is possible after Cj. This is unusual, since a single j can be followed by any vowel (cf. yet, your, yod, year, yoke, yankee, etc.), and other 2-member Onsets are also found before any vowel. For this reason, several phonologists claim that the yod cannot be the second member of a 2-member Onset at all; instead, they propose that the sequence jju: be analysed as a “complex Nucleus”, a single phoneme. This explains why Cj is not possible before other vowels. Nonetheless, such a solution has its problems, too. If the yod is part of the Nucleus, we would expect that it can be preceded by a complex (complex = 2-member) Onset. Yet, such forms are not found at all (e.g., *flj, *plj, etc.); but if the yod is part of the Onset, the lack of these clusters follows automatically from the fact that they would be 3-member Onsets, which isn’t possible18. As no unambiguous decision can be made, I retain the analysis of Cj clusters as complex Onsets for the sake of simplicity.

Initial SC clusters (exc. sw/sj) are, then, not Onsets. More precisely, the S is not part of the Onset, the other consonant is; similarly, in STR clusters, the TR sequence is an Onset, but the S isn’t. We express this for the time being by placing the S in parentheses, so (s)tay, (s)low, (s)street, (sh)rine, etc. The question is, of course, where the S belongs. In order to find out, however, we need to examine Codas and the syllabification of intervocalic clusters.

2. The Nucleus

The Nucleus represents no particular difficulty in English. Just like the Onset, it can be simplex (containing one X slot) or complex (with 2 X-es). In the latter case, there are two possibilities:
a) A single melody is attached two both nuclear timing positions. Such Nuclei are long monophthongs.
b) Two different melodies are attached to the two X-es. In this case, we have a diphthong. The second element is can only be i, u or a schwa; the first element varies to a greater extent.

It must be added that in English, there is a strong restriction on when the Nucleus can be complex: only if the syllable doesn’t have a Coda. We’ll discuss this soon.

---

17 In unstressed syllables, lj is free to occur before θ/a, which, however, can be regarded as the reduced (unstressed) equivalents of u/o/a; e.g., secular’sek:j:o: ~ ’sek:j:la.
18 It is interesting to note that stj, skw, skj, etc. initial clusters do exist (as in stew, square, skew), again providing evidence that the SC clusters are not Onsets.
A simplex Nucleus may also be filled by a liquid or a nasal, i.e., a syllabic consonant. The analysis of syllabic consonants is not possible at this point, so we postpone it until the necessary background has been presented.

3. The Coda
This is a point of some difficulty, and in order to be able to talk about Codas, three topics must be dealt with: (i) the maximum number of X-es (timing slots) within a constituent, (ii) the behaviour of word-final consonants, (iii) the syllabification of intervocalic clusters, and (iv) superheavy Rhymes.

The maximum size of syllabic constituents
We have seen so far that the maximum number of timing slots (X-es) within both the Onset and the Nucleus is two. It would be desirable if we could make a generalisation and say that this is valid for all syllabic constituents. That is, all syllabic constituents — the Onset, the Nucleus, the Coda, and the Rhyme — may contain a maximum of two X-es. Recall that the syllable is a tree-like structure. In such structures, a constituent may have other constituents below itself. Below the Rhyme, for example, we find two constituents: the Nucleus and the Coda. Furthermore, below those constituents which are at the bottom (there is no other constituent below them — these are the Onset, the Nucleus, and the Coda) we find timing positions (X-es). Any point in the tree below which there are other ones is called a node. So, for example, the Syllable is a node, just like the Rhyme, the Onset, the Nucleus, and the Coda; similarly, the X-es themselves can be considered nodes, since there is melody below them. So, to take an example, the structure of the word handy \( \text{ˈhændi} \) is as follows\(^\text{19}\):

(31)

\[
\begin{array}{cccccccc}
\sigma \\
\mid \\
R \\
\mid \\
O \\
\mid \\
N \\
\mid \\
C \\
\mid \\
X \\
\mid \\
h \\
\end{array}
\begin{array}{cccccccc}
\sigma \\
\mid \\
R \\
\mid \\
O \\
\mid \\
N \\
\mid \\
C \\
\mid \\
x \\
\mid \\
\end{array}
\]

In this tree, both Nuclei are below a Rhyme, and also below the Syllable itself. We say that the Nucleus in any syllable is dominated by the Rhyme as well as the Syllable. If you start out from the Nucleus and go along the branches of the tree upwards, you will first reach the Rhyme, then the Syllable. Therefore, the Nucleus is dominated by the Rhyme as well as the Syllable. The Onset is only dominated by the Syllable: if you move upwards along the branch starting from the Onset, the next (and only) constituent is the Syllable itself. (So, the Rhyme, though visually higher in the tree than the Onset, does not dominate the Onset: in order to reach the Rhyme from the Onset, you’d need to move upwards first to the Syllable, then downwards from the Syllable to the Rhyme.) In general terms:

(32) Domination
Node \( A \) dominates node \( B \) if and only if \( A \) is higher in the tree than \( B \) and there is a unique path along the branches of the tree whereby \( B \) can be reached from \( A \) without any movement upwards in the tree.

\(^{19}\) For why the \( n \) in \textit{handy} is taken to be a Coda, and the \( d \) is treated as an Onset, see shortly.
In sum, the Syllable node dominates every other node; the Rhyme dominates the Nucleus and the Coda; the Onset, the Nucleus and the Rhyme dominate their respective timing positions (X-es), while the timing positions dominate the melody (phonetic material) below them. In case the dominated node is directly below the dominating one, we talk about immediate domination. So, for instance, the Nucleus is dominated by the Rhyme as well as the Syllable, but it is immediately dominated by the Rhyme only, as the Rhyme is the node which one reaches first when moving upwards from the Nucleus.

We have said that the Onset and the Nucleus may not dominate more than two timing positions, and that it would be desirable if we could claim the same for all constituents of the Syllable (for the Rhyme as well as the Coda, too). That is, let’s make the following claim:

\[(33) \text{ Each syllabic constituent may dominate a maximum of two timing positions.}\]

This means, at first sight, that the Onset, the Nucleus, the Coda and the Rhyme may not dominate more than two X-es. Now, take heart: we have a paradoxical situation here. Consider the following syllable structure, where the Onset, the Nucleus and the Coda dominate two X-es each:

\[(34)\]

\[
\begin{array}{c}
\sigma \\
/ \ \\
R \\
/ \\
O \quad N \quad C \\
/ \ \\
X \quad X \quad X \quad X \\
/ \ \\
X \quad X \\
/ \\
X \\
/ \\
X \\
/ \\
X
\end{array}
\]

In this tree, there is a problem: it’s quite true that the Onset, the Nucleus, and the Coda dominate but two X-es each; but the Rhyme dominates as many as four X-es! Now, if you remove one X from either the Nucleus or the Coda, the Rhyme will still dominate 3 X-es. If (33) is correct, this situation isn’t possible. Note, however, that the Nucleus is obligatory; there is no Syllable without a Nucleus. The Nucleus must dominate at least one X (equal to a short monophthong). But then, it follows logically that the Coda may not contain more than one X, because if it did, the Rhyme would dominate more than two X-es, and that’s clearly against the principle stated in (33). As a result, we come to the (apparently) paradoxical conclusion that the Coda may dominate a maximum of one X. This follows from the facts that (i) the Rhyme, as a syllabic constituent, is subject to (33) — i.e., it may not dominate more than two X-es, (ii) the Nucleus is obligatory — i.e., it must dominate a minimum of one X. As a result, according to (33), the maximal Rhyme structures are as follows:
Both structures are heavy Rhymes, as we have already seen. Any structure in which the Rhyme dominates more than two X-es is referred to as a **superheavy** Rhyme. English shows a strong preference to avoid superheavy Rhymes: in most instances, the Nucleus in English can only be long in an open syllable; in closed syllables, only short Nuclei are generally possible. We’ll discuss the exceptions right away, but we must first note something about word-final consonants.

**Word-final consonants**
You can now object to the above claim: what about words like *feed, tide, rope, moon,* etc., with a final consonant and a long Nucleus? There are many such words. Yet, as we’ll see, a word-final consonant, for reasons to be discussed later, fails to count for purposes of syllable weight: it behaves as if it “weren’t there”. The word *feed,* therefore, is in fact $\text{fiic(d)}$, where the bracketed $\text{d}$ — more precisely, the X it is linked to — is not counted, hence this syllable is heavy, the Nucleus dominating 2 X-es rather than three (exactly like *fee*!)! Another important point to make is that the same is valid for the ult of free stems to which a regular suffix, such as $-s, -d,$ etc., is added, as well as the suffix itself. The form *feeds*, therefore, is $\text{fiic(d)(z)}$, where neither the final consonant of the free stem or the suffixal $\text{z}$ counts — again, *feeds* is no heavier than *fee*! The same goes for regular past tense forms (e.g., *peeped $\text{piic(p)(t)}$*), regular plural forms (e.g., *planes $\text{plecl(n)(z)}$*), as well as productive derivational suffixes which attach to free stems, such as $-ness, -ly, etc.$ Therefore, our discussion will be based on monomorphemic (unsuffixed) and irregularly suffixed forms. Regularly inflected forms (such as *peeped, planes,* etc.) will be disregarded. Word-final consonants will be revisited soon; let us now turn to superheavy Rhymes.

Superheavy Rhymes, then, are greatly disfavoured in English. Heavy, i.e., VC (as well as VV) Rhymes are very frequent:

(36) Heavy Rhymes: examples
(a)VC
(i) In ults (= V(C)#): 
$\text{lam(p), lin(k) (\eta(k)!)}, \text{a.dop(t), ac(t), ris(k), was(p), bul(k), hel(p), lif(t), ren(t), lis(t), len(d), mel(t), hel(d), bul(ge) (l(d3)!)}, \text{sin(ce), clean(se) (-en(z)!)}, \text{sel(f), etc};$ in GenAm, also *heart $\text{hardt}, tur(n) \text{tern}, ser(ve) $\text{s3rv, har(p) harp}, etc. In RP, these are

---

$\text{20 But the final consonant of bound roots in suffixed forms does count, so kept is kep(t), rather than *ke(p)(t).}$

This is quite logical: *kep in kept* is not a free word form. So the $\text{t},$ the last consonant in the full word form, doesn’t count, as expected, but the $\text{p}$ is not final in a free form, hence it does. In *peeped $\text{piic(p)(t), pip}$* is a free word form in itself (= *peep*). We’ll discuss this issue in more detail when dealing with morphology and the phonological behaviour of suffixes.
pronounced with a long vowel and without an r (i.e., hat, thin, sw, ha:p) so they belong to group (b) — VV Rhymes — below.

(ii) In non-ults (= VC):  
\[am.ply, an.chor (g.k!), scep.tic, fac.tor, Os.car, whis.per, wel.come, scal.pel, fif.ty, plen.ty, mis.ter, An.dy, anger(g.g!), am.ber, fil.ter, Mul.der, cul.ture (l.t3!), can.ter, fren.zy, al.pha (l.l!), etc; In GenAm, also Car.ter, car.nal, ser.vice, Mar.ple, etc. In RP, these belong to (b) below.

(b) VV  
(i) In ults (= VV(C)#):
\[fee, two, lay, pie, toe, see(d), loa(d), ma(ke), etc; in RP, words like hear(t), tur(n), ser(ve), har(p), etc. also belong here.

(ii) In non-ults (= VV.):
\[po.et (ø!), mu.se.um (i), hu.mor (ur), a.re.na (i), mo.dal (ø0), mi.nor (a1), fa.ther (ur), etc; in RP, Car.ter, car.nal, ser.vice, Mar.ple, etc., too.

As opposed to this, consider words with a Superheavy (VVC) Rhyme:

(37) Superheavy (VVC) Rhymes
(i) In ults (= VVC(C)#)
\[(a) pain(t), fin(d), bol(t) (ø0), fiel(d), laun(ch) (ø1), chan.ge, oun(ce) (ø0)
(b) eas(t), pos(t) (ø0), pas(te) (e1), boost(t), ous(t) (ø0), etc.
(c) (In RP only, all with ø: — GenAm has short ø): las(t), draf(t), tas(k), gras(p), etc.

(ii) In non-ults (= VVC.C)
\[(a) saunter (ø1), laun.dry (ø1), poul.try (ø0), shoul.der (ø0), an.cient (n.ʃ4!), coun.cil, dan.ger (n.ʃ3!)
(b) eas.ter, oys.ter, pas.try (e1), roos.ter, etc.
(c) (In RP only, all with ø: — GenAm has short ø): mas.ter, laugh.ter, bas.ket, exam.ple, etc.

Look at the words in (i/a) and (ii/a). In each instance, the Coda consonant is a coronal sonorant (nasal or liquid), and the following consonant is a coronal obstruent. We can now make the following claim:

(38) A long Nucleus is possible in a closed syllable if it is followed by a homorganic coronal cluster consisting of a sonorant and an obstruent.

Let’s check if this claim is true. Do we find words with a long vowel before a non-coronal cluster? Are there words like blílk, tungk, seimp, faplti, etc? In fact, there aren’t! It seems that (38) is correct.

In (i/b) and (ii/b), the Coda consonant is the coronal fricative s, followed by a coronal plosive, t. Hence this cluster, too, is a homorganic coronal one, but s is not a sonorant (how interesting — it is s which “misbehaves” again!). We can, therefore, modify (38) to include st:

(39) A long Nucleus is possible in a closed syllable if it is followed by a homorganic coronal cluster consisting of (i) a sonorant and an obstruent, or (ii) st.
It is interesting to note that coronals tend to behave in a special way in several languages. In Spanish, for example, words generally end in a vowel; all words with a final consonant end in a single coronal consonant (except plosives). So, for instance, Spanish has words like león ‘lion’, país ‘country’, paz (pron. paθ), amar ‘to love’, sol ‘sun’, Madrid (pron. maθrɪθ), but no other consonant is possible word-finally.\(^{21}\) Finnish, Ancient Greek and Latin are similar. It seems that coronals behave as if they weren’t there — as if the word ended in a vowel!

The words in (i/c) and (ii/c) only contain a superheavy Rhyme in RP, but not in GenAm, where the vowel is short. Here, we always have a in RP. The words last and master can also be assigned to (b), since any long vowel is free to occur before aC clusters, but before the clusters sk, sp, ft, mp we never find any other long vowel, only a. We can, therefore, say that it is a here which behaves specially, inasmuch as it can occur before sC clusters, ft and mp. Note, however, that a doesn’t appear before clusters such as pt, lk, gk, gg, etc., either! That is, it enjoys a wider range of possibilities to occur in a superheavy Rhyme than other vowels, but even so, its possibilities are limited.

Finally, there is a negligible number of words which have a superheavy Rhyme with a vowel other than a before a non-coronal cluster, such as chamber, Cambridge, hoax, traipse, corpse. The last three are irregular with regard to the final consonant cluster, too. We’ll consider these words to be exceptions. Let’s now turn to syllabification.

The syllabification of intervocalic consonants

In polysyllabic words, one can find between 0 and 5 consonants in intervocalic position, cf. museum mju:zɪ:əm (zero), edit ɛdɪt (one), winter ˈwɪntə, betray bɪˈtreɪ (two), amplifying ˈæmplɪʃəst, mistress ˈmɪstrəs, Bentley ˈbɛntliː (three), monstrous ˈmɒnstrəs, extra ˈɛkstrə (four), sempstress ˌsɛmpstrəs (five). The question is that of syllabification: do these consonants belong to the Coda of the preceding syllable, or the Onset of the next syllable? Or are they divided between two syllables? That is, is betray, for example, syllabified as be.tray, bet.ray, or bet.ay?

Now, in case there is no consonant between the two Nuclei (as in museum), the problem does not arise; the adjacent vowels are in different syllables, i.e., mu.se.um. But what about the rest?

The generally accepted view is that as many consonants are gathered into the Onset of the following syllable as possible. This has been called the Onset Maximisation Principle. But what do we mean by “possible”? Well, in the previous section, we described what clusters qualify as possible Onsets in English. According to the Onset Maximisation Principle (OMP), start scanning the intervocalic consonant cluster from the right. The last (rightmost) consonant will inevitably belong to the Onset of the next syllable, as single C’s are well-formed Onsets (except for the velar nasal, to which we’ll return). This means that a single consonant in intervocalic position is always an Onset, so, for example, edit is syllabified as e.dit. In case there are two consonants intervocally, as in winter and betray, both consonants are syllabified as Onsets if they can form a complex Onset. This is true for the tr in betray: it is a well-formed Onset in English. Accordingly, the syllabification is be.tray. The nt cluster in winter, on the other hand, is not a possible Onset, so the syllabification is win.ter: the second C goes to the Onset of the second syllable (according to the OMP), but the n cannot, so it is

\(^{21}\) Foreign words with a non-coronal final consonant are adopted in Spanish with the addition of a final e, e.g. English park appears in Spanish as parque (pron. parke).
syllabified into the Coda of the first syllable. The same will go for words like *sis.ter, fil.thy, fris.by, am.ple, etc.*

The same procedure is at work in the case of intervocalic 3-member clusters. Recall that the Onset may not contain more than two consonants, which means that the first C will belong to the Coda of the first syllable. The last two C’s will, of course, be syllabified as a complex Onset if they qualify for a well-formed Onset. This is what we find in *am.ply* and *mis.tress*: both pl and tr are possible complex Onsets. The cluster tl, on the other hand, as in *Bent.ley*, is not, because the two C’s are homorganic. As a result, only the third C may attach to the Onset of the second syllable, so we seem to have the syllabification *Bent.ley*. In this case, the Coda, apparently, contains 2 consonants. Yet, we proposed that the Coda may dominate a maximum of one X position. The syllabification *Bent.ley* yields a two-member Coda, hence a superheavy Rhyme. The t in *Bentley* seems to be unsyllabifiable — *Bent.ley* isn’t possible because tl is not a possible Onset, and nt can’t be a Coda, either, so *Bent.ley* is out, too! As a result, we have two choices. First, we can abandon the claim that the Coda may not dominate 2 X-es. This, however, is not satisfactory: we have seen that English dislikes superheavy Rhymes. Worse still, we’ll see shortly that this solution will not be applicable in certain cases. Instead, another solution must be sought.

The word *monstrous* is similar. The cluster tr will constitute the Onset of the second syllable, but the ns cannot form the Coda of the first since we assume that the Coda may not contain 2 consonants, so *mons.trous* isn’t possible. Again, where does the s go? Worse still, words like *extra* present an “extra” problem. Recall the Sonority Sequencing Principle in (27), repeated here as (40) for convenience:

(40) Within a syllable, sonority falls from the Nucleus towards the edges of the syllable.

This principle explains why an Onset cluster has a rising sonority slope. Let’s now check *extra*:

(41)

```
  6 * 5 4 3 2 1
  e k s  t r o
```

The first syllable is clearly ill-formed, since the level of sonority may not rise from the Nucleus to the right edge of the syllable! The s cannot go to the second syllable, either, since that solution would also involve a violation of the SSP. Words like extra, therefore, are problematic in two senses. First, as in the case of *Bentley* and *monstrous*, we have a consonant which cannot belong to either syllable. But in addition, we also have a violation of the SSP!

First, let’s discuss the word *Bentley*. There are other words of this type, e.g., handling, empty, Hampton, etc. The underlined consonants cannot be syllabified either as a Coda or an Onset for the same reason as in *Bentley*. In this book, we adopt a solution widely used in current phonological theory, using a category which will be needed to explain other phenomena as well: the notion of empty positions. Specifically, we propose that the apparently unsyllabifiable consonants are in fact Onsets in a syllable whose Nucleus is empty. This is shown in (42) below, taking *Bentley* as an example:
The question is, how is it possible to have a syllable without a Nucleus? After all, we have claimed that the Nucleus is obligatory — indeed, it is (in English at least) the only obligatory syllabic constituent!

Note, however, that we have separated melody and structure in our model. We did this to explain why Rhymes containing a long monophthong, a diphthong as well as a VC sequence are equally heavy: phonetically, they differ, but in a structural sense, they are the same: all of them contain two timing slots, or X-es. When viewed in a structural sense, the problematic syllables do have a Nucleus structurally speaking, but that Nucleus is empty: no phonetic material is attached to its X position.

Furthermore, the use of empty categories is widespread in other branches of linguistics, too — most notably, syntax. To take an example, consider phrases such as the blind ‘a vakok’. It is clearly a Noun Phrase, whose head, the Noun, is missing. The word blind is an Adjective, and Adjectives cannot take a Determiner, so, for instance, you can’t say *a clever arrived, *I like this beautiful, *Come here, my nice — Determiners must be part of an NP. Okay, you might say, what if blind is used here as a Noun? After all, conversion of a word from one part of speech to another without the addition of an affix is frequent in English, e.g., access (from Noun to Verb). Alas, this can’t be the case: if blind were a Noun, it should take a plural suffix (since we talk about all blind people), producing the form *the blinds. Adjectives, on the other hand, cannot be marked for the plural by way of suffixation. The conclusion is that this an NP — with an empty N. In other words, the Noun people is understood to be there, although it is “unpronounced”. A similar situation is found in Hungarian, where the present tense 3rd person forms of be (= van, vannak) is dropped in given syntactic contexts, so that one says A lány szép ‘The girl is beautiful’ (lit. ‘The girl beautiful’) rather than *A lány szép van — although in other tenses, the verb is there (A lány szép volt/lesz), just like in other persons of the present (Szép vagy ‘You are beautiful’). There is no particular reason (except orthodoxy) to believe that phonology is any different, since it is part of the grammar, and it is hardly surprising that some of the organising principles of grammatical structure are found in phonology, morphology, and syntax, too.

Of course, as you can expect, Nuclei are preferably filled, not empty. Recall the Sonority Sequencing Principle (SSP), which states that from the Nucleus towards either edge of the Syllable sonority falls. Syllables with an empty Nucleus clearly violate the SSP: an empty Nucleus is phonetically silent, with no melody attached to it. Now, silence, of course, means lack of any voicing, hence no sonority at all! In (42), in the second Syllable of Bentley, sonority doesn’t rise from the Onset towards the (empty) Nucleus. Therefore, empty Nuclei are — just like superheavy Rhymes — less preferred than filled ones, and we must clearly and explicitly state under which circumstances a Nucleus may remain empty. The situation is similar in Syntax, where phrases with a filled head are clearly the “normal” case, and empty head elements may only occur if specific conditions hold. To return to the
Hungarian syntactic parallel, the Verb (the head of the VP) is only dropped in the 3rd person forms of the Present Indicative if the VP has a subject complement (if it does not, van(nak) cannot be dropped, e.g. *Jani itt van* ‘Johnny is here’). In our case, the condition is as follows; it is called the Empty Category Principle (ECP):

(43) **The phonological Empty Category Principle**
A Nucleus may remain empty if
(a) it dominates only one X position,
(b) it is unstressed,
(c) it is separated from the following Nucleus by a single consonant,
(d) the following Nucleus is not empty,
(e) the Onset before the empty Nucleus is simplex (it dominates one X).

All these conditions hold for Bentley, empty, handling, Hampton. (43a) states that a long Nucleus may never be empty: simply, that would mean a sequence of two empty X positions, and that’s not possible (we’ll see why when discussing word-initial empty positions); that is, a sequence of two empty positions is ill-formed. In accordance with (43b), the Nucleus must be unstressed — this is logical: stress has phonetic properties, a stressed Nucleus being louder, of higher pitch, etc., than an unstressed one; how could these phonetic properties be realised on an unpronounced Nucleus? According to (43c), the empty Nucleus must be followed by a single consonant + a pronounced Nucleus.

The analysis is supported by the fact that some words actually have two possible pronunciations. Handling, for instance, though generally pronounced as *'hændlin*', can also be pronounced as *'hændlin'* . The difference between the two pronunciations, then, is that the schwa is deleted in the former variant, but structurally, the two forms are identical: in the schwa-less variant, the melody is dropped but its X-position remains. In fact, deletion of schwa (sometimes of *i*) is very frequent in English in medial syllables. Such a deletion — when a vowel is deleted in a word-medial syllable — is called syncope (pron. *'sɪŋkəpi*). The following words can all be pronounced with or without a schwa: separate*Adj* 'sep(ə)rət, history*Adj* 'hɪstr(ə)ri, memory*Adj* 'mem(ə)rɪ, definite*Adj* 'def(ə)nət, family*Adj* 'fæm(ə)li, etc. (the last two words can be pronounced with an *i*, too). Compare this to visibly, which can only be pronounced *'vɪzblɪ*; not *'vɪzbli*; because the schwa is followed by a complex Onset *bl*, i.e., it is separated from the following Nucleus by more than one consonant. In secular*Adj* 'sɛkˈdʒala, the underlined schwa cannot drop, either, because it is preceded by a complex Onset *kj*, cf. (43e). (In fact, complex Onsets in English must always be followed by a pronounced vowel, as we will see shortly.)

Let us now turn to words like monstrous and extra. Please note an interesting fact: the segment whose syllabification is problematic is *s*, our old friend. What a pleasant surprise! Recall that there are SC-initial words, such as spy, sky, stay, snore, slow, shrine, etc. We said that these SC clusters are not syllabifiable as Onsets. We delayed the discussion of an important point, though: if SC clusters are not Onsets, where does the S belong in these words? Note that this is exactly the same problem as in extra and monstrous: we have a consonant which cannot be syllabified anywhere.

Note that intervocalic SC clusters — e.g., sis.ter, ves.try, fis.cal, etc. — have been syllabified as Coda-Onset clusters, where the first member, S, is in a Coda, and the second member, C, is in the Onset of the next syllable. This suggests that identical clusters be syllabified in the same way in all cases, even where there is no vowel before the cluster (as in stay or extra).
We have already made use of empty Nuclei, and we also mentioned that employing them in our model will turn out to be useful in the solution to other problems. In this case, however, the situation is somewhat different than with Bentley and similar words: there, we assumed an empty Nucleus after the problematic consonant (as the t in Bentley). As SC sequences behave like Coda-Onset clusters, the obvious solution is to posit an empty Nucleus before the S. This solution removes the problem of word-medial S-es (as in extra), and treats them in the same way as word-initial SC clusters. Let’s see the model in detail. Here’s a representation of stay (44) and extra (45):

(44) $\sigma$ $\sigma$
|   |   |
R   R
|   |   |
N   C   O   N
|   |   |   |   |
X   X   X   X   X
|   |   |   |   |
s   t   e   i

(45) $\sigma$ $\sigma$ $\sigma$
|   |   |   |
R   R   R
|   |   |   |
O   N   O   N   C   O   N
|   |   |   |   |   |   |
X   X   X   X   X   X   X
|   |   |   |   |   |   |
ε   k   s   t   r   ə

As you can see, the k will belong to the Onset of the second syllable, according to the Onset Maximisation Principle. In semipstress, too, the second syllable is (using a _ to indicate an empty Nucleus) p_s. Note that we also postulate an empty Onset word-initially; we’ll come back to this problem in the final part of this chapter. For now, note that in this model, the problematic s is treated in the same manner both word-initially and word-externally.

Now, the problem is this: we said earlier that an empty Nucleus must not be followed by more than two consonants (cf. (43c)). Here, this is not the case. We must, therefore, add a further condition:

(46) A simplex Nucleus may remain empty when followed by a Coda S (S = s/ʃ).

In other words, a syllable whose Coda contains a s or a ʃ, may have a phonetically un realised (= unpronounced) Nucleus. Note that if the following SC cluster is sw or sj, which are complex Onsets, the Nucleus may not be empty.

Not all languages are of this type: Spanish, for instance, does not allow such empty Nuclei. A consequence is that no Spanish word may begin with an SC cluster. Spanish speakers, when trying to pronounce foreign words beginning with such a cluster will automatically insert an e before it, so, for example, they will pronounce the English proper
name Stanley as “Estanley”. (Old Hungarian was of the same type — hence the i inserted in the word *iskola* ‘school’, from Latin *schola*.)

Word-final consonants revisited
We have already paid some attention to word-final consonants, and pointed out that they behave specially: they do not seem to contribute to the weight of the ult. Hence, *keep, soap, soup, cape, carp, thorpe, type*, etc., for example, have a long vowel, though in non-final syllables closed by a Coda we never find long vowels. We simply observed the special behaviour of final C’s, but we didn’t offer an explanation; at any rate, it seems that they are not Codas, since if they were, words like *keep* would contain a superheavy Rhyme. In order to propose an explanation, however, two other phenomena must be examined, both providing evidence that final consonants do not behave like Codas at all.

Evidence No. 1: Verb Stress
We have seen that syllable weight plays a very important role in English when discussing stress placement in English nouns. Recall that the penult of the noun is stressed if it is heavy; otherwise the antepenult is stressed. Now, a similar rule is at work with verbs, but there is a difference: the ult is not neglected. Instead, the verbal stress rule is as follows:

1) Check the ult. If it is heavy, stress it.
2) If the ult is light, stress the penult.

Examples are found in (47). In (47a), the ult is stressed, while in (47b) and (47c), the penult receives stress.

(47)

(a) delay  di'leɪ  answer  'ɑ:n.sə  edit  'ɛ.dɪt
allow  ə'laʊ  discover  dis'kɜr.və  polish  'pʊ.lɪʃ
maintain  mɛn'ten  deliver  di'li.və  cancel  'kæn.ʊl
insist  in'sɪst  protect  prəʊ'tɛkt

In those accents of English (including GenAm) which have a final r, *answer, discover* and *deliver* will also belong to (c).

It seems at first sight that verbs like *edit* are “irregular”: their ult ends in a VC sequence, which ought to be heavy, hence stressed. Note that with Nouns, where the penult’s weight is considered (as a Noun’s ult isn’t stressable), we observed that VC counts as a heavy Rhyme (cf. *verán.da, utén.sil* — the accute accent indicates stress). Why isn’t the same kind of Rhyme heavy word-finally?

Let us observe those verbs which are stressed on their ult. Their ult (i) contains a long Nucleus (*delay, allow, maintain*) or (ii) it has a short Nucleus followed by two C’s (*insist, protect*). That is, word-final -VCC does make a heavy ult!

This phenomenon has long been observed, and phonologists have pointed out that the apparent irregularities (such as *édit*) can be explained if we assume that the very last consonant of the word is not taken into consideration when calculating the weight of the ult. So, *édit*, for instance, is stressed on the penult because the final -t doesn’t count for purposes

---

22 Accute accent = Hu őles őkezet, the one used in Hu to indicate vowel length, as in vőz.
of syllable weight. We can symbolise this by placing the final C between parentheses: \textit{e.lit(t)}, \textit{abo.lif(sh)}, etc. What remains if the bracketed C’s are disregarded is a light rhyme, i.e., \textit{dr, li}, respectively.

In the case of words like \textit{deláy} and \textit{maintáin}, the ult has a complex Nucleus, which is sufficient to make a heavy syllable. No wonder these verbs are stressed on the ult, as their ult is heavy regardless of the number of final consonants. In words like \textit{insís(t)} and \textit{presén(t)}, if we disregard the last consonant (= \textit{insís(t)}, \textit{presén(t)}), we still have a Coda consonant left, which makes a heavy Rhyme.

In sum, the stress pattern of English verbs provides evidence that final consonants do not belong to the Coda of the ult. Instead, they are something else.

\textbf{Evidence No. 2: Word-final consonant clusters}
If you look at the word-final 2-member clusters found in English words (excepting, of course, those which contain a free stem and a regular suffix such as \textit{peeped}, \textit{bags}, etc.), you can notice something quite interesting. Consider the words in (48). \(C_1\) = the first member of the cluster, \(C_2\) = the second member of the cluster.

\begin{center}
\begin{tabular}{cccc}
\hline
 & \textbf{(a)} & \textbf{(b)} & \textbf{(c)} & \textbf{(d) GenAm:} \\
\hline
\hline
\textbf{C}_1 \rightarrow & Obstruent & Nasal & /l/ & /r/ \\
\hline
\textbf{C}_2 \downarrow & & & & \\
Voiceless plosive & apt & lamp & help & carp \\
& act & link (\textit{ŋk}) & melt & sort \\
& risk & wasp & rent & bulk \\
& twist & soft & & park \\
& box (\textit{ks}) & lapse (\textit{ps}) & & \\
& blitz (\textit{ts}) & & & \\
\hline
\textbf{Voiced plosive} & \textit{(*mb#)} & bulb & curb \\
& land & held & card \\
& \textit{(*ŋg#)} & \textit{lg} = \textit{g} & porch \\
& ranch & belch & porch \\
& change & bulge (\textit{ld}\textit{3}) & forge & \\
& triumph (\textit{mf}) & self & serf \\
& \textit{mv} = \textit{f} & solve & serve \\
& plinth (\textit{nθ}) & filth (\textit{lθ}) & forth (\textit{rθ}) \\
& \textit{nð} = \textit{d} & \textit{lð} = \textit{t} & \textit{ð} = \textit{s} & \\
& since & else (\textit{ls}) & purse (\textit{rs}) & \\
& cleanse (\textit{nz}) & Giles (\textit{lz}) & Mars (\textit{rz}) \\
& Thames (\textit{mz}) & & \\
& \textit{nʃ} = \textit{ʃ} & Welsh & harsh \\
& melange (\textit{nʒ}) & \textit{lʒ} = \textit{ʒ} & concierge (\textit{rʒ}) \\
& \textit{(*mn#)} & kiln & barn \\
& \textit{(*nm#)} & film & charm \\
& /l/ & & & \\
\hline
\end{tabular}
\end{center}
Of course, the words in (d) do not end in a consonant cluster in RP, since r is not found in this accent before a consonant (so, carp = kearp, sort = sɔt, etc.)

You can observe that in all cases (exc. for ks/ps/ts, shaded in the table), the cluster exhibits a falling (e.g., 31 as in ns) or a level (= unchanging, e.g., 11, as in kt) sonority pattern. None of these clusters shows a rising sonority profile. Furthermore, note the following:

(i) The clusters *mb and *ŋɡ are not found word-finally in either RP or GenAm. This is because a homorganic cluster of a nasal + a voiced plosive in word-final position may only be coronal (= nd) — again, these tricky coronals! Do not be misled by the spelling: words like bomb and sing, which look like ending in mb and ŋɡ, respectively: they are pronounced without the plosives, i.e., bom and sɪŋ, respectively.

(ii) The clusters *mm and *nm, are also ill-formed word-finally — a combination of two nasals is ill-formed word-finally. Again, do not be misled by spellings such as solemn — the letter <n> here is silent, so, for instance, this word is pronounced 'sɒləm.

(iii) The clusters marked with a  are assumed to be accidental gaps. In the case of mv, nð, lð, rð (GA only), lʃ, note that the voiceless counterpart of each fricative is found in word-final position when preceded by the same sonorant (= mf, nθ, lθ, rθ (GA only), ʃ are okay). You could object, possibly, that voiced fricatives are not possible after a sonorant in word-final position, but as shown by words like solve, Giles, melange as well as — in GA — serve, Mars, concierge, this can’t be the case. Furthermore, note that ʒ occurs in final nʒ (melange), but its voiceless counterpart, ʃ, doesn’t (*nʃ). The unsystematic nature of these gaps suggests that they are accidental, so the clusters marked with a  could be possible (e.g., new words ending in these clusters could be added to the language without any problem).

(iv) Note that ks, ps, ts show a rising sonority profile — no other word-final cluster does! Again, this naughty s! We’ll see later if this can be explained. (It must be added that ts is rare, but nonetheless, it is possible.)

Look at now the words in (49):
(49)

<table>
<thead>
<tr>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
<th>(d) GenAm:</th>
</tr>
</thead>
<tbody>
<tr>
<td>C₁ →</td>
<td>Obstruent</td>
<td>Nasal</td>
<td>/l/</td>
</tr>
<tr>
<td>C₂ ↓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voiceless plosive</td>
<td>scep.tic</td>
<td>am.ply</td>
<td>scal.pel</td>
</tr>
<tr>
<td></td>
<td>fac.tor</td>
<td>an.chor (ŋ.k)</td>
<td>fil.ter</td>
</tr>
<tr>
<td></td>
<td>whis.ker</td>
<td>plen.ty</td>
<td>wel.come</td>
</tr>
<tr>
<td></td>
<td>whis.per</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>mis.ter</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>fif.ty</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voiced plosive</td>
<td></td>
<td>am.ber</td>
<td>el.bow</td>
</tr>
<tr>
<td></td>
<td></td>
<td>An.dy</td>
<td>fol.der</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ang.er (ŋ.g)</td>
<td></td>
</tr>
<tr>
<td>Affricate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>axis (k.s)</td>
<td>com.fort</td>
<td></td>
</tr>
<tr>
<td></td>
<td>sy.nop.sis</td>
<td></td>
<td>al.pha (l.f)</td>
</tr>
<tr>
<td></td>
<td>ac.tion (k.ʃ)</td>
<td></td>
<td>Cal.vin</td>
</tr>
<tr>
<td></td>
<td>op.tion (p.ʃ)</td>
<td></td>
<td>l.θ = ʃ</td>
</tr>
<tr>
<td></td>
<td>exigent (g.z)</td>
<td></td>
<td>l.δ = ʒ</td>
</tr>
<tr>
<td></td>
<td></td>
<td>can.cer</td>
<td>repul.sive (l.s)</td>
</tr>
<tr>
<td>Fricative</td>
<td></td>
<td>fren.zy</td>
<td>pal.sy (l.z)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>sen.sual (n.ʃ)</td>
<td>revul.sion (l.ʃ)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>lin.gerie (n.ʒ)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nasal</td>
<td>arsenic (s.n)</td>
<td>om.nibus</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Den.mark</td>
</tr>
<tr>
<td>/l/</td>
<td>bracelet (s.l)</td>
<td></td>
<td>grem.lin</td>
</tr>
<tr>
<td></td>
<td>Leslie (z.l)</td>
<td></td>
<td>Stan.ley</td>
</tr>
<tr>
<td></td>
<td>mushroom</td>
<td></td>
<td>com.rade</td>
</tr>
<tr>
<td>/r/</td>
<td></td>
<td></td>
<td>Hen.ry</td>
</tr>
</tbody>
</table>

Again, the items in (d) have no rC cluster in RP. Note, moreover, the following:

(i) In neither (48) nor (49) do we find a cluster of two identical consonants, such as ll, i.e., l₁; a long consonant. (Yes, we do find such spellings, as in belly, better, carry, fell, pass, etc., but in such cases, only a single short C is pronounced.) This is because English disallows long consonants (also called geminates) — as opposed to Hungarian, for instance, cf. Hu őtt (pron. ʃʃt).

(ii) There are some clusters — shaded grey in the table — which are not found finally but can occur inside a word. We have already discussed them: such clusters are not possible word-finally.

(iii) Clusters marked with a ʃ in the table do not occur in word-final position, either. Remember, however, that word-final consonant clusters all exhibit a non-rising sonority pattern, and the clusters marked with a ʃ all show a rising sonority pattern, which is why they can’t occur word-finally. (Do not be misled by words ending in a syllabic sonorant, such as channel ʃʃn) — the liquid here sits in the Nucleus, and as the pronunciation variant ʃʃn shows, a schwa is always possible in such words. Hence the n and the l do not form a consonant cluster in these cases. We’ll discuss syllabic liquids in detail in the next chapter.)

(iv) The clusters marked with a ʒ, as earlier, are regarded as accidental gaps.

Now please go back to (48), and check the word-final clusters there. You will find that any cluster which is found finally is also possible inside a word! In word-medial position — that
is, when the cluster is not final — such clusters are syllabified as Coda-Onset clusters. For example, the final clusters in *fact*, *lent*, *belt*, *since* are Coda-Onset clusters in *fac.tor*, *plen.ty*, *fil.ter*, *can.cer*, etc.

Now, we have already pointed out that the status of word-final consonants as Codas is questionable; but we have not decided yet what they are, if not Codas. As the clusters in (48) are easily syllabified as Coda-Onset sequences within a word, it seems more than plausible to suggest that word-final consonants are, in fact, Onsets. So, for example, the clusters in *fac.t*, *len.t*, *bel.t*, *sin.ce* are Coda-Onset sequences — exactly like in *fac.tor*, *plen.ty*, *fil.ter*, *can.cer*, etc.

Recall that final consonants do not contribute to the weight of the ult. So, verbs like *édit* are stressed on the penult, though superficially, they “look like” having a heavy penult. The final C doesn’t count. Similarly, a word-final VVC sequence (as in *beat*, *soap*, *loud*, etc.) doesn’t make a superheavy Rhyme: the final C, again, looks as if it wasn’t there. More precisely, in both cases, the final C doesn’t seem to be part of the Coda. All these phenomena receive an explanation if we adopt the view that final C’s are Onsets: in that case, the fact that they do not make the preceding Rhyme (super)heavy is an automatic consequence of the fact that they do not belong to the Rhyme at all. But if final consonants are Onsets, they ought to be followed by a Nucleus! Yet, there is no vowel after them.

We have already made use of empty Nuclei. If word-final consonants are syllabified as Onsets, they start a syllable — without a pronounced Nucleus. In other words, final C’s are, structurally speaking, not final at all: they are only final on the melodic tier, as shown in the representations below, where the Syllable structure of *edit* (50a), *lamp* (50b), *tempo* (50c) and *keep* (50d) is depicted:

(50) (a) (b) (c) (d)

```
(a) σ σ σ σ σ σ σ
    R R R R R R R
    O N O N O N O N
    X X X X X X X X
    | | | | | | | |
    e d i t

(b) σ σ σ σ σ σ σ
    R R R R R R R
    O N O N O N C O N
    X X X X X X X X
    | | | | | | | |
    l æ m p

(c) σ σ σ σ σ σ σ
    R R R R R R R
    O N C O O N O N
    X X X X X X X X
    | | | | | | | |
    t e m p o v

(d) σ σ σ σ σ σ σ
    R R R R R R R
    O N O N O N O N
    X X X X X X X X
    | | | | | | | |
    k i p
```
In the case of (50a), you can see that the syllable filled by the CV sequence \( \text{dt} \) is actually a light Rhyme, since it has a simplex Nucleus — and no Coda. The \( \text{t} \), final melodically (= it isn’t followed by any other segment), is actually not final structurally, being linked to an Onset followed by an empty Nucleus. As a result, the verb \( \text{edit} \) is stressed on the first syllable.

In (50b), the \( \text{p} \) is, similarly, in an Onset, preceded by a Coda \( \text{m} \), just like in (50c). This beautifully reflects the fact, observed earlier, that word-final consonant clusters behave mostly like word-internal Coda-Onset sequences. Finally, in (50d), there is no superheavy Rhyme at all: \( \text{t} \) is a normal heavy Rhyme.

So far so good, but what makes it possible for the final empty Nucleus (abbr. FEN) to remain empty? After all, we have said that Nuclei are preferably filled, and they can only remain empty under specific conditions. Recall the Empty Category Principle, repeated here as (51) for convenience:

\[
\text{(51) The phonological Empty Category Principle}
\]

\text{A Nucleus may remain empty if}

\begin{itemize}
\item[(a)] it dominates only one X position,
\item[(b)] it is unstressed,
\item[(c)] it is separated from the following Nucleus by a single consonant,
\item[(d)] the following Nucleus is not empty,
\item[(e)] the Onset before the empty Nucleus is simplex (it dominates one X).
\end{itemize}

To this, we added a further condition to explain the special behaviour of SC clusters:

\[
\text{(52) A simplex Nucleus may remain empty when followed by a Coda S (S = s/\text{ʃ}).}
\]

Neither (51) nor (52) can explain why a word-final Nucleus may be empty in English. A further condition (and the last one, finally!) is needed:

\[
\text{(53) A simplex Nucleus may remain empty if it is word-final.}
\]

Remember what we observed about the occurrence of vowels and consonants in word-final position in the world’s languages: some languages have vowel-final words only, while others have both vowel-final and consonant-final words. \text{There is no language in which all words must end in a consonant.} Apart from the fact that our model, employing FENs, explains the special behaviour of final consonants in English, it also achieves a remarkable unification of phonological structure in all languages. In a structural sense, \text{all} words in \text{all} languages end in a Nucleus, but not necessarily a pronounced vowel. Some languages — like English, Hungarian, etc., — allow final Nuclei to remain empty, while others — such as Italian — do not. As a final Nucleus is obligatory in all languages, and no language forbids filled Nuclei in general (= not only finally), our model correctly predicts that there will be no language in which a final vowel is impossible! As we said, Syllables with an empty Nucleus violate the SSP, and the same is valid for Syllables containing a FEN. As a result, it can be expected that no language will make FENs obligatory, i.e., all languages have final filled Nuclei. If you adopt a model in which final C’s are Codas, there’s no violation of the SSP, and the fact that no language requires that all words end in a consonant remains unexplained.

Now, we observed that — apart from the notorious ps/ks/ts clusters — no cluster with a rising sonority profile can occur finally: neither Onset clusters (such as kl, pr, etc.) nor Coda-Onset clusters with a rising sonority pattern (such as sn, tl, etc.) are found in this position. Now, if melodically final C’s are in fact followed by a silent Nucleus, these clusters
could actually be syllabified, as shown in (54), where a representation of a non-existent hypothetical form *lokkr is found:

(54)

\[
\begin{array}{c}
\sigma \\
R \\
O \\
X
\end{array}
\quad\begin{array}{c}
\sigma \\
R \\
O \\
X
\end{array}
\quad\begin{array}{c}
* \\
\text{l} \\
\text{d} \\
\text{k} \\
\text{r}
\end{array}
\]

The question is, if edit and lamp can have a final consonant followed by an empty Nucleus, why is a form like the one in (55) impossible? The answer is that in English, any cluster exhibiting a rising sonority profile must be “supported” by a following filled Nucleus. Hungarian, German, Slovene, Romanian, etc., are of the same type. As already mentioned, Spanish only allows final coronal non-plosives; this means that all other consonants (coronal plosives and any non-coronal consonant) as well as any cluster, including those with a falling sonority pattern, must be supported by a pronounced Nucleus. In Italian, which has no final C’s at all, all Onsets, even simplex ones, must be supported by a filled Nucleus. On the other hand, French, Polish, Russian, etc., can have word-final complex Onsets, cf. the Polish word teatr ‘theatre’, with a final tr\(^23\) or the French word siècle ‘century’, pron. sjekl. That is, languages do not only differ in whether they allow FENs, but also in what types of consonants or consonant clusters they can support. In English, FENs can only support a simplex Onset or a cluster with a falling sonority profile. To sum up, empty Nuclei are capable of supporting a smaller range of preceding consonant clusters than filled ones.

In the light of the abovesaid, the precise meaning of the term ult needs clarification. We said, for example, that edit has a light ult, whereas protect has a heavy one. In structural terms, the ult is identical in both words: a simplex Onset being followed by an empty Nucleus, as the last C — in both words, t — is analysed as belonging to the Onset of a Syllable with a FEN. That is, structurally speaking (cf. (50a)), edit has three syllables, and the one filled with the sequence di is not the last one! While this is true, we’ll retain the use of the word ult in the classical sense, using it to refer to the last syllable in the word with a filled Nucleus, for the sake of simplicity.

Let us revisit the problem of final ps, ks, ts clusters, as in lapse, box, blitz. The s, according to our model, is an Onset followed by a final empty Nucleus. Yet, we have said that final empty Nuclei cannot support a cluster with a rising sonority pattern. We might add a further condition, sounding something like “unless the cluster is ps/ks/ts”. Note, however, that these clusters never occur initially — hence, a pronounced Nucleus cannot support them in word-initial position. We’ve seen that empty Nuclei have a smaller supporting power than filled ones — they cannot support a cluster with a rising sonority pattern, as opposed to filled Nuclei — and it would be strange to say that in the case of these clusters, a FEN is “stronger” than a filled Nucleus. That is, if a FEN can support ps/ks/ts, we would expect that a filled

\(^{23}\) The final liquid in either the Polish or the French word is not syllabic!
Nucleus can do so, too. In sum, we will simply regard final ps/ks/ts clusters as exceptions, in the sense that we can’t explain why they are possible. (Note that models which do not employ empty Nuclei are equally unable to say anything sensible about these final clusters.)

Remember that we posited an empty Onset, too, word-initially, in words which melodically begin with a vowel (cf. (45) and (50a)). You may rightfully ask why we did so. There are four arguments I would like to present. Moreover, we’ll argue that the Onset is always obligatory, not only initially, but word-externally, too. That is, there is an empty Onset between two separate Nuclei, as in *screwing skruːɪŋ,* which contains an empty Onset between *uː* and *ɪ.*

(That is, we are talking about hiatuses here, if you recall.)

First, remember that in the languages of the world, just the opposite situation holds in initial position than in final position. Initially, all languages have consonants; some languages also allow words to begin with a vowel (such as English or Hungarian), while others (like German and Czech) do not. There is no language in which words must begin with a vowel. As in the case of final position, our model unifies word-initial position in all languages: all words in all languages begin with an Onset, but some languages may allow it to remain empty.

Second, in alliterative poetry, consonant-initial words alliterate if they begin with the same consonant: so, for example, *lamp, lord,* and *last* alliterate. Interestingly, vowel-initial words behave differently: any vowel alliterates with any other, so, *entry, east,* and *only* can happily alliterate. Why do vowels behave differently than consonants? If we assume that vowel-initial words actually begin with an empty Onset, we receive an explanation: they can alliterate because they all begin with an identical, empty, Onset.

Third, the behaviour of word-final *r* in many accents of English, including RP, is readily explained with reference to initial empty Onsets. In these accents (referred to as non-rhotic accents), *r* is only possible before a pronounced vowel, so words like *card, course,* and *girl,* etc. are pronounced *kɑːrd, kɔːrs, ɡɜːrl,* respectively, as opposed to those accents (called rhotic — GenAm is such an accent) which permit *r* before a consonant as well as word-finally, so the same words sound as *kɑrd, kɔrs, ɡɜrl,* respectively, in GenAm. Word-finally, however, something interesting happens in (most) non-rhotic accents: a final *r* is pronounced if the following word begins with a vowel. So, there is no *r* in *car* in RP in *This is my car* (absolute word-final) and *My car broke down* (the following word begins with a C); but in *My car is blue,* the *r* is pronounced: *kɑːr ɪz.* How do we explain this? First of all, in terms of structure, we can say the following about the occurrence of *r* in non-rhotic accents:

(55) *Non-rhoticity*

An *r* must be supported by a following filled Nucleus.

This explains why this consonant cannot stand in a Coda (since a Coda is always followed by an Onset) or in final position (where it would be followed by a FEN). It seems, therefore, that the representation of *car* in non-rhotic accents is as follows:
Then, we could say, the \( \sigma \), not being supported by a following filled Nucleus, is dropped. This solution is problematic, though. First, the dropping of the \( \sigma \) would result in a completely empty syllable. Second, if the next word begins with a vowel, we would still expect the \( \sigma \) to drop in case we adopt (56), because — whatever the first sound in following word — it is still followed by the FEN!

The appearance of \( \sigma \) before a following vowel-initial word is easily explained if we assume that (i) the next word, in fact, begins with an empty Onset, (ii) the final \( \sigma \) is not an Onset. Now, of course, it can’t be a Coda, either, since there are no word-final Codas in this model. Note, however, that we have used empty positions: positions without an attached melody. The separation of timing positions and melody also predicts that there can be instances when a melody is unattached to any position (X). It has no host, but simply “floats”. Such segments are called **floating segments** (lebegő szegmentumok). So, the representation of the word *car* is as follows:

![Diagram](image-url)

That is, the word ends in a \( \sigma \) on the melodic tier, but it isn’t linked to any X. Consider what happens if the next word begins with a vowel, i.e., an empty Onset, as in *car is* (I omit the representation of the second syllable of *is* (i.e., \( \kappa + \text{FEN} \)) — simply because it is irrelevant here).
The floating \( r \) is readily pronounceable, since it can attach to an Onset position followed by a filled Nucleus (shown by the dotted line). By contrast, if the following word begins with a C (as in \textit{car broke}), the Onset is already occupied, hence the floating segment cannot attach anywhere. The same happens, of course, if there’s no following word. A floating segment which cannot find a host is called by the funny name \textit{stray segment} (köbor szegmentum) — because it has no \textit{host} (gazda), i.e., an X it can belong to. Hostless, i.e., stray segments cannot be integrated into any syllable, and they are deleted as a result — this deletion is referred to by another funny name: \textit{Stray Erasure} (köbortörlés).

The same analysis can be applied to explain why the indefinite article \textit{an} drops its \( n \) if the following word begins with a consonant. If we analyse the \textit{n} as a floating segment, we can say that it attaches to a following empty Onset position if there’s one available. The same solution can be used for the Hungarian definite article \textit{az}, which behaves identically to English indefinite one. Note, furthermore, that in these cases, we cannot even assume that the final C is in an Onset followed by a FEN, since \( \sigma \) in English and \( \zeta \) in Hungarian are perfectly possible in word-final position (unlike \( \rho \) in non-rhotic accents of English): no other word ending in \( n \) (or \( z \) in Hungarian) drops this consonant before a following C-initial word!

The fourth argument for postulating an empty Onset initially in vowel-initial words as well as between two separate Nuclei word-internally is \textit{hiatus breaking}. Recall that a hiatus is a sequence of two vowels in different syllables, as in English \\textit{screw.ing}, \\textit{Lea skru\textsuperscript{\wedge}n\text{\textdollar}, li\textsuperscript{\wedge}o}, or Hungarian \\textit{te.a, fi.ú}. Now, hiatuses are often broken up, or eliminated, by the insertion of a glide between the two Nuclei. In English, \\textit{Lea} and \\textit{screwing} are often (if not always!) pronounced as \\textit{skru\textsuperscript{\wedge}n\text{\textdollar}, li\textsuperscript{\wedge}o}, where the superscript symbols indicate the inserted hiatus-breaking glide. In standard Hungarian, hiatus-breaking is found after \( \iota \), so \( fiú \) is pronounced as \( fi\text{\textdollar}u \); but in several dialects, it is possible after \( \varepsilon \), too, as in \\textit{tea te\text{\textdollar}o}. In English, the same phenomenon can be observed across word boundaries, too, e.g., \\textit{screw it} pronounced as \\textit{skru\textsuperscript{\wedge}n\text{\textdollar}t}. The assumption that the Onset is always obligatory, that is, syllables which melodically begin with a vowel actually begin with an empty Onset (both word-initially and word-internally), offers an explanation: the empty Onset is available for the hiatus-breaking glide.

Note that we gave the precise conditions under which a Nucleus may remain empty. We did this because syllables with an empty Nucleus violate the SSP, hence they are exceptional. A syllable with an empty Onset, on the other hand, does not: as the Onset is silent, i.e., it has zero sonority, sonority does fall starting from the Nucleus towards the (empty) Onset. Recall, however, how we represented words with an initial SC cluster, cf. the representation of \\textit{stay} in (44), repeated here as (59):
Here, there is no empty Onset. Remember the Empty Category Principle, one point of which says that a complex Nucleus may not be empty. The reason for this is that a sequence of two empty X-es is ill-formed. Positing an empty Onset in words with an initial SC cluster, as in \textit{stay}, would clearly violate this condition. Furthermore, the indefinite article is pronounced without an \textit{n} before such words: \textit{a star} rather than \textit{*an star}; the same goes for the Hungarian definite article, cf. \textit{a sport} rather than \textit{*az sport}. In non-rhotic accents, final floating \textit{r} remains silent, too, before a following SC-initial word: \textit{car stayed has no r}. In sum:

\begin{center}
\textbf{60) An empty Onset is found before any filled Nucleus, but syllables with an empty Nucleus have either a filled Onset or no Onset at all.}
\end{center}

In practical terms, this means that all words begin with an Onset (filled or empty), except SC-initial ones, which begin with an empty Nucleus.

The very last point to clarify is the possibility of having empty Codas. You could, for example, say, that syllables without a pronounced Coda consonant actually have an empty Coda position — in a parallel fashion with Onsets. So, for example, the syllable structure of \textit{America} could possibly be:

\begin{center}
\textbf{(61)}
\end{center}

\begin{center}
\texttt{(59)}\hspace{1cm} (61)
\end{center}

However, there are serious problems with this. First, note that to claim that the Coda is structurally speaking obligatory (but it may remain empty) would predict that there are languages which require the Coda to be filled in all cases — remember, there \textit{are} languages which only have filled Onsets and Nuclei, but no empty ones. But we do not find such languages: all languages of the world have open syllables, and there are many which do not have Codas at all. There is no language which has closed syllables only.

Worse still, the representation in (61) predicts that \textit{America} has a heavy penult. Note that a syllable is heavy if its Rhyme dominates 2 X-es, no matter what is associated with these X-es. In this representation, it actually does! Yet, as the stressing of the word shows (Am\textit{é}rica,
that is with an antepenultimate stress, rather than *America, with stress on the penult), the penult is light. If we adopt (61), there is no way to distinguish the weight of the penult in América (light) vs. veranda (heavy, with a Coda).

In sum, the Coda is not obligatory. Instead, a Coda is only assumed if it is actually filled: there are no empty Codas. Accordingly, (61) must be rejected, and the correct representation of America is as in (62):

(62)

\[
\begin{array}{cccccc}
\sigma & \sigma & \sigma & \sigma & \sigma \\
R & R & R & R & R \\
O & O & O & O & O \\
X & X & X & X & X \\
\emptyset & \emptyset & \emptyset & \emptyset & \emptyset \\
\end{array}
\]

This concludes our discussion of syllable structure, but in later chapters, we will sometimes come back to questions concerning syllabification. In the next chapter, for example, we’ll find further evidence for word-internal empty Nuclei, and we’ll also discuss syllabic consonants.

Many people find the model presented here as “complicated”. In a certain sense, it is: we introduced abstract notions such as empty positions, which, of course, is a complication. But it is worth doing so, because these abstractions make it possible for us to explain a wide range of phenomena found in English and in languages in general. The use of abstractions is widespread and commonly accepted in natural sciences. Newton, for example, observed that (i) the apple falls off the tree, rather than staying floating in the air or going zig-zag, and (ii) the moon orbits the earth, rather than falling down, and so on. He might have chosen the “simple” solution and simply observe these facts (which people had done before), but where would science be if he had chosen this solution? Instead, he proposed the abstract notion of gravity, which, interacting with other forces, explains the phenomena found in the physical world in a unified way. After all, gravity is “invisible”, just like empty positions are inaudible: we can only see the effects of it — just like we can see the effects of empty positions (such as the special behaviour of final C’s, etc.).

Finally, I hope that this chapter has also convinced you that English and Hungarian, as opposed to common (mis)belief, are not at all that different. We have found that in many ways, the two languages behave alike: both have complex Onsets (but only before a filled Nucleus), complex Nuclei, and optional Codas; both languages allow Onsets and Nuclei to remain empty, and, indeed, in almost identical conditions; both languages have an article with a final floating consonant; and so on, and so forth. In spite of the many differences, English and Hungarian show remarkable similarities in the phonological structure of words.
Appendix to Chapter 6: some syllable structure representations

This appendix is meant to help you in clarifying and understanding the main points made in this chapter by presenting the syllable structure of some words in addition to the ones represented in the main body of the chapter. Each representation is followed by a brief explanation.

(a) slow

As sl is an SC cluster not syllabifiable as an Onset, the s belongs to the Coda of the preceding syllable which has an empty Nucleus. There is no initial empty Onset, since that’s only possible before a filled Nucleus.

(b) swine

As opposed to sl in (a), sw is a possible Onset (recall that of all SC clusters, only sw and sj meet the requirements for a complex Onset). The final C, of course, is an Onset followed by a FEN.

(c) sprint

Again, the s is in a Coda, as in slow. The pr can be a complex Onset. Yet again, the very last C of the word occupies the Onset of a syllable with a FEN. The n goes to the Coda of the
second syllable, as nt is not a possible Onset. As nt exhibits a falling sonority profile, it can be followed by a FEN.

(d) Lea

Here, the iː — a long monophthong — occupies two X-es. Note again that the length mark (ː) isn’t used in tree diagrams: in phonetic/phonemic transcriptions, which do not show structure, only melody, the length mark must be used to show that the V is long. In tree diagrams, however, this is expressed by linking the single melody to two positions. The second syllable begins with an empty Onset, as the Onset is obligatory before a filled Nucleus. This empty Onset position may be filled by an inserted j as a hiatus-breaker.

(e) source

As expected, the final s is an Onset, as in (b) and (c). Note that no r is shown in the diagram, as opposed to car (cf. (57) in the chapter). This is because in non-rhotic accents, this word is never pronounced with an r. As a result, it simply has a long vowel + the final C. In fact, it sounds just like sauce. In car, we assumed a final floating r because it is pronounced if the following word begins with a vowel (or, structurally, an empty Onset). In word-internal position, however, the r is never pronounced before a consonant. For non-rhotic speakers, then, source and sauce are homonymous, neither containing an r. (Recall that we are not dealing with spelling, but pronunciation here!) The same goes for all non-final spelt (but unpronounced) r’s (e.g., girl, park, port, beard, scarce, etc.). In rhotic accents, of course, source is sərs, with a Coda r + a final Onset s followed by a FEN.
In (f), the indefinite article — as discussed in the text — ends in a floating n, which is attached to the following empty Onset. If the following word begins with a filled Onset, the n can’t find a host X and will be the victim of Stray Erasure (cf. a lion, not *an lion). In (g), we have a melodically similar situation: the numeral one ends in n, just like an. Yet, one is always pronounced with a final n (one lion). In our model, this is shown by the fact that in one, the n is not floating, but it is attached to an Onset followed by a FEN. As any single final C is possible before a FEN, it will always be pronounced. (All words with a final n behave in the same way as one, except, of course, an.)