V Sounds and the sound system: Phonetics and phonology

Phonetics and phonology (both from Greek *phonè* (φωνή) 'sound') are the branches of linguistics study speech sounds (PHONES). Phonetics concerns itself with the physical properties of speech sounds (sound waves, spectograms, pronunciation, etc.), whereas phonology investigates the behaviour of speech sounds within a SOUND SYSTEM of a particular language and in the universal phonological system underlying all human natural languages more generally — a system of regular (i.e., rule-based) alternations involving speech sounds. For a proper understanding of the phonology of a language, it will always be helpful to have a good sense of the phonetic properties of the sounds of the language. But phonology, which is a branch of theoretical linguistics, often abstracts from the physical properties of sounds to arrive at an understanding of how they interact in the sound system.

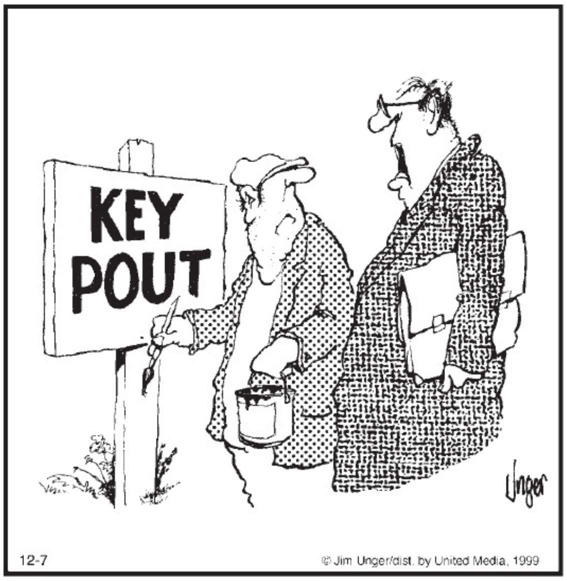
V.1 Spelling and pronunciation

The orthographic systems of human languages are all designed to represent the sounds of these languages. Some are clearly better at this than others. Though Hungarian spelling struggles with the two renditions of the sound [j] (viz., *j* and *ly*), it is otherwise quite a faithful reflection of the sounds of the language. At the other extreme, English spelling system is notoriously ill-equipped for the pronunciation of the language as it is spoken today: it may have been a good approximation of the language of Geoffrey Chaucer (the 14th century author) or William Caxton (who introduced the printing press to England in the 15th century) but lots of sound changes have taken place in the interim while spelling reform has not kept up with these developments. Because *gh* sometimes sounds the same as *f* (as in *rough*), *o* sometimes sounds like *i* (as in *women*), and *ti* is frequently pronounced the same way as *sh* (think of *position*), George Bernard Shaw quipped that we might as well spell *fish* as *ghoti* — which is of course a joke (and a rather lame one at that): *gh* can sound like *f* when it appears in word-final position but never when it is at the beginning of a word; *o* sounds like *i* in one word only; and *ti* sounds like *sh* in the suffix *-tion* but not otherwise. But you get the point: English spelling is a poor reflection of present-day pronunciation practices. To flog this dead horse even further, let us consider the set of words in (1), all ending in the letter combination *ough*. When you say these words out loud, you will discover that each of them pronounces this letter combination in a way different from any of the others. The phonetic transcriptions of these words (using the system introduced by the English phonetician A.C. Gimson, employing the symbols of the International Phonetic Alphabet) are provided in the right-hand column, for the sake of explicitness. (The symbol 'ː' marks a long vowel.)

(1)
a. though [əʊ]
b. through [uː]
c. tough [ʌf]
d. trough [ɒf] or, in American dialects, [ɒθ]
e. bough [aʊ]
f. bought [ɔː]
g. hough [ɔk]
h. hiccough [ʌp]

In light of this, automatic speech-to-spelling or spelling-to-speech conversion would seem a hazardous undertaking for English. But even for languages whose spelling is much less 'chaotic' than that of English, it is difficult to negotiate one's way between spelling and speech. Sounds often 'get lost' in the pronunciation of words in connected speech and sounds are affected by others in their environment, so that in British English [sləː spiːtʃ] could be understood either as *slower speech* or as *slurred speech* (with the *d* of *slurred* elided in front of *sp* in rapid speech because the consonant sequence [dsp] is very hard to produce). Misunderstandings about word boundaries can arise as well: the cartoon in (2) (reproduced from Fromkin *et al*. 2011:230) is based on this.

(2)



As in the case of Shaw's *ghoti*, however, the joke in (2) is a bit of a stretch: *keep out* and *key pout* do certainly have the same vowels and consonants in a broad transcription of these words; but when you say the strings *keep out* and *key pout* one after the other, you will find that one of the consonants sounds distinctly different in each case — the *p* of *pout* is pronounced with a puff of air ('aspiration', annotated with a superscripted 'h') following it ([p^{h}aʊt]); the *p* of *keep* is not, and may instead have a brief interruption of the airflow (a 'glottal stop', annotated as [ʔ]) right before it ([k^{h}iːʔp]). So *keep out* is in fact rather difficult to 'mishear' as *key pout*. Similarly for the famous pair in (3) (from the 1927 song 'I Scream, You Scream, We All Scream for Ice Cream'):

(3) a. I scream
    b. ice cream

Again, the sounds of (3a) and (3b) are broadly speaking the same; but while the [k] of *ice cream* causes the [r] that follows it to 'lose its voice' ([kr̥iːm]; the diacritic '˳' printed below the [r] indicates that the [r] is pronounced without vibration of the vocal cords), the [r] of *scream* is not devoiced. This difference in the voicing of [r] is distinctly audible even in connected speech, and it is an important clue to listeners trying to decipher the incoming speech signal.

Jokes such as *ghoti* as an alternative spelling for *fish* and *I scream* ~ *ice cream* work because English is a written language and most of its users are able to read. But in actual practice, there is obviously no use for *ghoti*, and we are rarely tripped up by segmentation in cases like (2) and (3).

V.2 The International Phonetic Alphabet

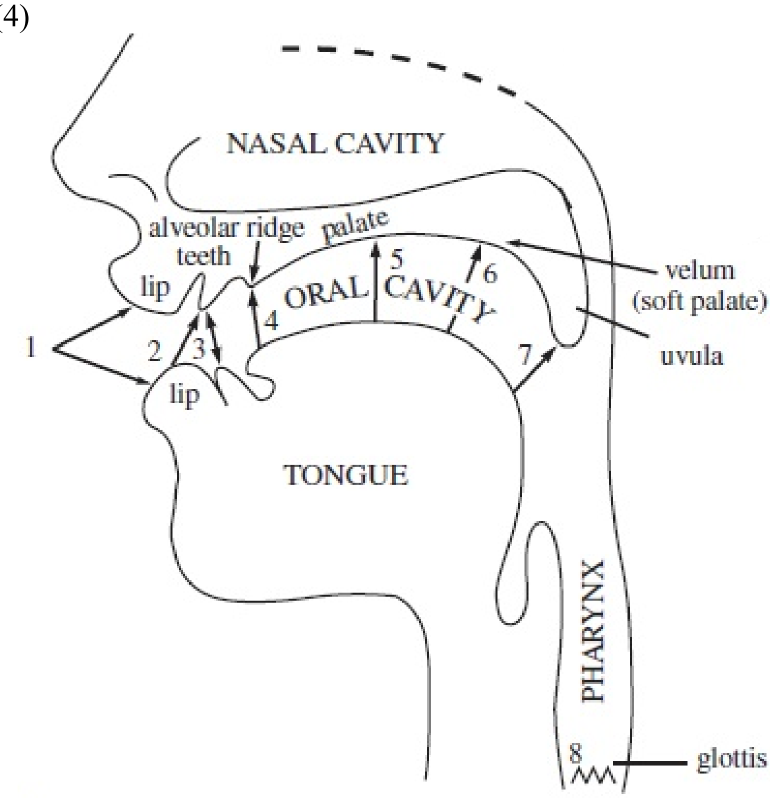
An important message emerging from the previous section is that the writing system of a language is not always a good guide for phonetic analysis. We need an alphabet that is not dependent on the whims of individual spelling systems and which we can use for the transcription of the sounds of all of the world's languages. Thankfully, we do not have to invent such a system anymore: it already exists, in the form of the International Phonetic Alphabet. In the transcriptions provided in the previous section, we availed ourselves of this alphabet already. (There are simpler transcription systems available for English, but because the IPA is the international gold standard, we will be basing ourselves on it here.)

There is no point in learning the entire IPA by heart. You can look up the symbols in a book or on the internet at any time. But it will be helpful to familiarise yourself to some degree with the IPA, so that you will quickly recognise the symbols frequently used in the transcription of English and be able to 'translate' them into sounds in your head.

V.3 The vocal tract

The symbols of the IPA are renditions of all the sounds known to exist in human natural language. Those sounds are all produced somewhere in the vocal tract — the collective name for the parts of the human body involved in the production of speech: the vocal cords, the pharynx (or throat), the mouth, the tongue, the nose and the lips. A picture of the vocal tract (reproduced from Fromkin *et al*. 2011:236) is presented in (4). The main areas of articulation for human speech sounds are marked with numbers, and identified to the right of the picture.

(4)



1. bilabial
2. labiodental
3. interdental
4. alveolar
5. (alveolo)palatal
6. velar
7. uvular
8. glottal

The vocal cords (aka 'vocal folds'), located in the glottis, make the difference between voiced and voiceless sounds. When you allow your vocal cords to come closely together during the production of a speech sound, they will vibrate (which you can feel when you place your fingers on your glottis) and produce a sound that is 'voiced'; when the vocal folds are kept apart, there is no vibration and hence no voicing. The difference between *b* and *p* is all about this: throughout the production of the string [aba], the vocal cords vibrate, so *b* is voiced; by contrast, in the production of [apa] there is an interruption of vocal cord vibration in between the two vowels, precisely at the point when *p* is pronounced.

The velum (aka 'soft palate') plays an active role in making the distinction between an oral sound and a nasal sound. When the velum is lowered, as in the picture in (4), air coming from the lungs can escape through the nose; when it is raised, it blocks off the nasal cavity and air can only escape through the mouth. If in addition to raising the velum we also close our lips, air cannot escape at all and we produce a *b* or a *p*. With the velum down and the lips closed, we instead produce a nasal consonant — *m* or its voiceless counterpart (annotated as [m̥]; although in connected speech we might occasionally come across one, English does not feature voiceless nasals in its phonology).

The active articulators in the production if *b*, *p* and *m* are the lips, which come together in these sounds. We call sounds for which the lips are the active articulator 'labial' sounds (from Latin *labium* 'lip'). The lips are also actively engaged in the production of the consonants [ʃ] (found in *shoot*), [tʃ] (*chunk*), [ʒ] (as in the second *g* of *garage*, if pronounced in a French-like way) and [dʒ] (*junk*), as well as in the semi-vowel [w], and several of the vowels: in particular (for standard British English), the [uː] (as in *food*), [ʊ] (*foot*), [ɔː] (*fought*), [ɒ] (*dot*), [əʊ] (*vote*), and [aʊ] (*out*).

With the exception of bilabial [p, b, m] and labiodental [f, v], all the major consonants of English feature the tongue as an active articulator. Thus, in [tʃ] and [dʒ], apart from the rounding of the lips, the tongue makes contact with the hard palate, while in [ʃ] and [ʒ] the tongue merely approximates the hard palate and allows air to escape through the narrow opening between the two, causing friction (hence the term 'fricative'). The stops [t], [d], [n], [k], [g] and [ŋ] (the final consonant of *hang*) all have the tongue block the airflow out of the lungs through the mouth. In the production of the fricatives [s] and [z], the tongue tip moves very closely towards the ribbed area just behind the upper front teeth (the 'alveolar ridge'), and for [θ] and [ð] (both written as *th*) the tongue tip inserts itself between the front teeth; but in all four cases, air is still allowed to queeze by. The tongue also gestures in the direction of the roof of the mouth without making full contact with it in the production of almost all of the vowels of English (the only exception being the 'bland' vowel of *uh* [ə], where the tongue is 'in rest', in its neutral position). The exact point towards which the tongue is moving and the extent to which it rises makes the difference between the various vowels.

The areas which the tongue gestures towards or makes contact with are called the passive articulators: the upper front teeth (for *oaf* and *oath*), the alveolar ridge (for *oat*, *ode*, and *use*), the hard palate (for *ouch*), and the soft palate (or 'velum', for *oak*). Although the uvula is involved in the pronunciation of *r*-sounds in some dialects of English (the Geordie dialect, spoken in and around Newcastle-upon-Tyne, is famous for this), this extreme of the mouth is not directly implicated in the production of any of the sounds of standard English. Nor does English reach further back into the pharynx to produce sounds. But English does have a (non-orthographic) consonant that involves complete occlusion of the airflow in the glottis: the glottal stop, already encountered in the pronunciation of *keep* as [k^hiːʔp] and of *cat* as [k^hæʔt] . In certain varieties of English (think of Cockney, for instance), the glottal stop regularly replaces oral stops, with *got it!* produced as *goʔ iʔ*. The glottis is also the locus for the pronunciation of one of the fricatives of English: the *h* [h].

This covers the vocal tract as it is exploited by English. It will be useful to have a reasonably keen sense of the shape of the vocal tract, the active articulators, the places in the mouth which the tongue can gesture towards or make contact with, and the difference between voiced and voiceless sounds and between oral and nasal sounds. We will be making reference to these things at various points in the discussion of the phonology of English, to which we now proceed.

V.4 Assimilation and dissimilation

In this segment so far, we have been talking about concrete, physical speech sounds. These distribute in certain ways within the sound system of a particular language, often as a function of the properties of the sounds around them. Consider, for instance, the following pair of cases:

(5)	a.	virus	+ -al	=	viral
	b.	velum	+ -al	=	velar

When we combine *virus* with the adjective-forming suffix *-al* (commonly found in such words as *royal*, *legal*, *physical*, *grammatical* and *(phono)logical*), we get the output *vir-al*. No surprise here. But when we combine *-al* with *velum* (the scientific word for the soft palate), the result is not **vel-al* but *vel-ar*. Here, the consonant of the suffix has changed its form in order to avoid an output in which two *l*'s occur close to one another: it has made itself dissimilar to the *l* of the stem. We see this also in *popul-ar*, *regul-ar*, *spectacul-ar* etc. This phonological process is called DISSIMILATION.

The opposite also exists, and is in fact very common: putting two sounds close together frequently results in the two sounds becoming similar or even identical to one another. This is called ASSIMILATION. We will see this at work in many examples in this segment, including those in (6):

(6)	a.	in- + elegant	=	inelegant	[n]
	b.	in- + credible	=	incredible	[ŋ]
	c.	in- + glorious	=	inglorious	[ŋ]
	d.	in- + proper	=	improper	[m]
	e.	in- + balanced	=	imbalanced	[m]
	f.	in- + moral	=	immoral	[m]
	g.	in- + licit	=	illicit	[l]
	h.	in- + regular	=	irregular	[r]

In (6a), *in-* is realised in its 'pristine' form, with the alveolar nasal [n] (where 'alveolar' refers to the alveolar ridge, the rippled area immediate behind the upper front teeth: in the production of [n], the tip of the tongue makes contact with the alveolar ridge). In (6b) and (6c), although the orthography does not reveal this, the prefix *in-* ends in a nasal consonant involving contact between the back of the tongue and the velum. In (6d–f), the orthography shows us that the place of articulation for the nasal of *im-* is labial: [m]. For (6f), this actually results in a situation in which the nasal consonant of *in-* and the first consonant of the stem are identical. We see such 'chameleon'-like behaviour in an even more extreme form in (6g) and (6h), where the prefix *in-* not only becomes identical with the [l] or [r] of the stem but in the process goes so far as to lose its nasality altogether.

If we state the facts in (6) with reference to entire sounds, we get statements such as those in (7a–d).

(7) a. the nasal consonant of *in-* is pronounced as [ŋ] in front of [k] and [g]
b. the nasal consonant of *in-* is pronounced as [m] in front of [p], [b] and [m]
c. the nasal consonant of *in-* is pronounced as [l] in front of [l]
d. the nasal consonant of *in-* is pronounced as [r] in front of [r]

These are accurate descriptions of (6a–h), but it should be obvious that we are missing a generalisation here: what the nasal of *in-* is doing in all these cases is assimilate in its PLACE OF ARTICULATION and, in the case of (6g,h), also the MANNER OF ARTICULATION to the following consonant — regardless of whether that consonant is voiced or voiceless, or nasal or oral.

The following is a more efficient and insightful way of capturing the place-assimilation facts in (6a–f):

(8) a. the nasal of *in-* is [velar] in front of a consonant that is [velar]
b. the nasal of *in-* is [labial] in front of a consonant that is [labial]

Stating the assimilation rules in terms of articulatory features such as [velar] and [labial], rather than with reference to whole sounds, thus gains us a better insight into the processes involved.

V.5 Distinctive features as the primitives of phonological analysis

Distinctive features are the features in terms of which we can distinguish between speech sounds. These are the true primitives of phonology. There are a wide variety of distinctive features, broadly organisable into three subfamilies (where 'laryngeal' refers to the larynx, where voicing, aspiration, glottalisation and 'creaky voice' are controlled):

(9) DISTINCTIVE FEATURES
a. PLACE OF ARTICULATION: {[labial], [dental], [alveolar], [palatal], [velar], ...}
b. MANNER OF ARTICULATION: {[stop], [fricative], [nasal], [liquid], ...}
c. LARYNGEAL PROPERTIES: {[voiced], [aspirated], ...}

We refer to the sound [ŋ], which we encountered in (6b,c), as a 'voiced velar nasal'; its feature composition contains the features [voiced], [velar] and [nasal]. Similarly, [m] is a 'voiced bilabial nasal', with the specifications [voiced], [labial] and [nasal].

The sound [l], seen in (6g), is called a 'liquid': even though full contact between the tonguetip and the alveolar ridge is made, the air flow is nonetheless uninterrupted (i.e., is liquid) because air can escape via the sides of the tongue. The [r] of (6h) is also a liquid, but its place of articulation is different from that of [l]: in the production of [r] the tongue-tip curls back and upwards towards the hard palate. Both liquids of English share their specification as [voiced] — but this voicing is cancelled when [l] and [r] immediately follow a stop at the beginning of a stressed syllable, as in (10). This is the result of a phonological process that realises the aspiration of the [p], when it is initial in a stressed syllable, on the liquid that follows it — the [p] 'dumps' its aspiration (one of its laryngeal properties) on its next-door neighbour, as it were, and this laryngeal property overrides the natural voicing of the liquid, making it voiceless.

(10) a. play [pl̥eɪ] or [pɬeɪ]
b. pray [pr̥eɪ]

Here again, we see that phonological processes deal in properties of speech sounds, not in speech sounds integrally. In order to truly understand phonological processes, we need to give these properties of speech sounds the theoretical status they deserve. The distinctive features in (9) are the minimal building blocks of phonological analysis. We do not need to learn these lists of features by heart: they and their definitions can readily be looked up in textbooks and on the internet. What matters is that we recognise the importance of couching our phonological analyses in terms of these features, instead of stating them over whole segments.

V.6 Distinctive features and feature classes in the analysis of assimilation processes

Some phonological processes are best understood if we describe them not in terms of *individual* distinctive features but *classes* of them. The three classes of distinctive features recognised in (9) (the PLACE, MANNER and LARYNGEAL classes) are all active in phonological processes. To illustrate the fact that the place-of-articulation features can act *en bloc* in the phonology, we can go back to the facts in (6a–f) and the rules in (8), the latter repeated below:

(8) a. the nasal of *in-* is [velar] in front of a consonant that is [velar]
b. the nasal of *in-* is [labial] in front of a consonant that is [labial]

We can improve our insight into what is going on in (6a–f) if we state the assimilation process at work in these examples as in (11):

(11) the nasal of *in-* is [αPLACE] in front of a consonant that is [αPLACE]

Here α is a variable ranging over all the place-of-articulation features belonging to the list in (9a). So for any particular choice of α for the consonant following the nasal, the statement rule in (11) automatically ensures that the nasal of *in-* will have exactly the same specification for its place of articulation.

From the discussion so far, we may have got the impression that distinctive features of the sort recognised in (9) are relevant for consonants but not for vowels. But that would be a mistake. Although vowels are overwhelmingly voiced and unaspirated (so there is little to say about them with regard to their laryngeal properties: these tend not to be distinctive for vowels), there are differences in the vocalic domain in the realms of manner and place of articulation just as there are in the consonantal domain. Thus, there are both oral and nasal vowels (think here of French *un bon vin blanc* 'a good white wine', where each of the vowels is nasal); and vowels have a wide range of places of articulation. The articulation of the vowel [u] involves lip rounding — a [bilabial] or [round] vowel. In the production of [i], on the other hand, the tongue gestures forwards and upwards to the front of the hard palate — a [palatal] or [front] vowel. And in the case of [ɒ] (the vowel of *hot* in British English), the tongue gestures towards the back of the mouth — a [velar] or [back] vowel.

That it is useful to recognise these different places of articulation in the formal analysis of these vowels becomes apparent when we look at assimilation processes such as VOWEL HARMONY in Hungarian, and UMLAUT in the Germanic languages. Consider first the case of Hungarian vowel harmony, illustrated (very partially) in (12):

(12) a. vet-ett-em ‘I sowed’
a. süt-ött-em ‘I baked’
b. fut-ott-am ‘I ran’

The past-tense and first person singular subject markers of the Hungarian forms in (12) have a vowel in them, but this vowel is not the same throughout: the subject marker alternates between *e* and *a*, and the tense marker even shows a three-way alternation, between *e*, *ö* and *o*. The choice of vowel is not random: we do not have the licence to pick and choose freely (**vet-ött-am*, **süt-ett-em*, etc.). There are two regularities to be noted here: the vowel of the tense marker harmonises with the vowel of the stem in terms of palatality (front/back) and labiality (rounding); and the vowel of the subject marker harmonises with its host for palatality but not for labiality. The precise way of analysing the phonology of Hungarian vowel harmony is quite complex, and we will not develop it in detail here. But from what we have said so far, it is clear that it is essential to be able to make reference to the place features of the vowels (in particular, to [palatal]/[±back] and [labial]/[±round]).

Old English *i*-umlaut is another assimilation process that presents a good illustration of the importance of distinctive features in the analysis of phonological phenomena affecting vowels. Consider the following Old English forms:

(13) a. cuman ‘come-INF’ cymþ ‘come-3SG’
b. dohtor ‘daughter’ dœhter ‘daughter-DAT/PL’
c. faran ‘go’ færþ ‘go-3SG’

There is a vowel quality change in each of these pairs. This vowel change can be traced back to historically reconstructed forms featuring an *i*-suffix (here ‘*’ means ‘historically reconstructed’):

(14) a. cymþ < *cumiði
b. dœhter < *dohtori
c. færþ < *fariði

With this knowledge in mind, we could describe the vowel changes in (13) with the aid of the following rules:

(15) a. /u/ becomes [y] when it precedes */i/
b. /o/ becomes [œ] when it precedes */i/
c. /a/ becomes [æ] when it precedes */i/

But although these rules are descriptively adequate, they are not tremendously helpful if we want to truly understand these vowel changes. The problem with (15) is that it misses a generalisation that emerges when we inspect the Old English vowel system, and analyse it in terms of the binary-valued distinctive features [±back], [±high], [±low] and [±round], as in (16):

(16)

[–back]		[+back]		
i	y	u		[+high, –low]
e	œ	o		[–high, –low]
æ			a	[–high, +low]
[–round]	[+round]		[–round]	

What unites the three processes in (15) is that the underlying [+back] vowel of the stem is fronted (i.e., becomes [–back]) under the influence of the (reconstructed) high front (i.e., [–back, +high]) vowel /i/. This decomposition of the vowels of Old English into feature bundles allows for a reappraisal of the *i*-umlaut phenomena in Old English as in (17):

(17) [+back] becomes [–back] immediately preceding *[–back, +high]

In our discussion of the Old English vowel system so far, we have concentrated just on the role played by the feature [±back], in order to gain an understanding of the changes involved in *i*-umlaut. But it is also worth noting that the system of binary-valued features employed in (16) is very efficient for the analysis of vowel height: with the aid of just two binary features for height, [±high] and [±low], we can describe a three-height vowel system of the Old English type. Of course, the fourth logically possible combination of the binary features [±high] and [±low], viz., [+high, +low], is a logical contradiction, so we do not expect this combination to be found in any natural language.[1]

V.7 Minimal oppositions and the phoneme

The distinctive features which we talked about in the previous sections can participate in phonological processes in groups or on their own on; but they are never realised on the surface all by themselves: it is impossible to pronounce the place feature [velar] without adding manner and laryngeal features into the mix. Distinctive features get grouped together into a speech SEGMENT. A speech segment that is involved in MINIMAL OPPOSITIONS with other segments within the system of which it is a part is called a PHONEME.

Phonemes are language-specific clusterings of distinctive features, awarded phonemic status precisely on the basis of the systematic oppositions that they are engaged in within the sound system of the language. Replacing one phoneme of a particular language with another always delivers a different word from that language. Minimal oppositions are an essential tool in determining whether a particular segment of a particular language is a phoneme of that language or not. It is important to bear in mind that phonemic status is always language-particular: a particular segment may be a phoneme of language X but not of language Y. Whether a segment is or is not a phoneme of a language depends entirely on the way it behaves with respect to *other* segments of the language in question — and such behaviour typically differs from language to language. So it is pointless to look for segments that are phonemes in every human language.

To illustrate the notion of minimal oppositions, the foundation of phonemic analysis, let us consider the following set of words from English:[2]

1 There are, however, languages with a more-than-three-way vowel height system. For these, an additional feature needs to be recognised — one which we will not go into here: for Old English and Hungarian, we do not need such an additional feature, so we are in the clear for these languages.

2 There are two gaps in the set of words in (18). In the column in (18) that has the velar-initial words *core* and *gore* in it, there is no word listed in the bottom row: although English has the velar nasal [ŋ], this phone cannot start a word in English. The other gap, likewise located in the velar column, involves the velar counterpart to *fore* and *sore*. This would have to be a word with a velar fricative, a sound that Modern English does not have in its phoneme inventory. We do not have to look far beyond English to find such a fricative: German has it (think of the well-known composer Johann Sebastian Ba<u>ch</u>); and Scots has it, too (in words such as *lo<u>ch</u>*). English had a velar fricative as recently as a few centuries ago, but not anymore. We are dealing here with an accidental gap in the phoneme inventory of Modern English.

(18)	pore	×	tore	×	core
	×		×		×
	bore	×	door	×	gore
	×		×		
	fore	×	sore		—
	×		×		
	more	×	nor		—

Though the orthography (spelling) is not a perfect reflection of this, the ten words in (18), when pronounced, differ only in their initial consonants (C) — what follows the initial C is consistently the same: [ɔː] (the long vowel found in the exclamation *Aw!*) in British English and other varieties of English that do not pronounce an orthographic <r> when it occurs immediately after a vowel ('non-rhotic' varieties); [ɔr] in American English and other 'rhotic' dialects of English.

The fact that the initial consonants of all of the words in (18) are phonetically quite different in itself is not a reason to identify these consonants as phonemes. What matters for the status of these segments as PHONEMES *of English* is that, *in English*, each of these ten words is different from any of the others: replacing the initial consonant of one of these words with the initial consonant of one of the others results in a different word (a word that means something else: even if they might not be certain what exactly some of these words mean, English speakers will know with certainty that each of the ten words in (18) is different from each of the others).

V.8 Phonologically conditioned alternations, complementary distribution, and allophony

Though there is a clear phonetic similarity between them, the initial consonants of the three words in (19) are also phonetically distinct in English. All three consonants involve the movement of the back of the tongue all the way up to the soft palate (or velum), creating an occlusion which causes the flow of air from the lungs out through the mouth to be completely stopped — hence the name 'velar stop' for all three consonants in (19). But the exact point of contact between the back of the tongue and the soft palate is different in each of the three cases in (19). In (19a), the back of the tongue touches the soft palate at a point right behind the hard palate — the part of the soft palate that is closest to the front of the mouth. This is what the diacritic '+' underneath the consonant signifies: the initial consonant in (19a) is a 'fronted' version of the velar stop. In (19b), by contrast, the point of contact between the back of the tongue and the soft palate lies much farther back, which is what the diacritic '–' expresses. And in (19c), in the course of the production of the velar stop the lips already start taking the shape required for the production of the following vowel. This is indicated with the superscripted 'w', which stands for 'labialised'.

(19)	a.	keep	[k̟]
	b.	carp	[k̠]
	c.	coop	[k^w]

Unlike in the set of words in (18), exchanging the initial consonants among the words in (19) does not result in different words — rather, it results in the same word 'pronounced weirdly'. What we are looking at in (19) is an ALTERNATION between [k̟], [k̠] and [k^w], conditioned by their phonological environment. As we already saw above, it is because of the fact that the velar stop in (19c) is followed by a vowel whose production involves lip rounding that the lips are already getting in position during the production of the stop: the stop is anticipating the coming of the vowel, smoothening the transition from the stop to the vowel — a case of anticipatory assimilation. Similarly, in (19a), the point of contact between the back of the tongue and the soft palate is located towards the very front of the velum in order to anticipate the arrival of the vowel, whose production requires the tongue to gesture forwards and upwards, towards the front of the hard palate. And in (19b), the tongue makes contact with the roof of the mouth at the back of the soft palate because in the production of the following vowel, the tongue wants to gesture towards the back of the mouth as well.

The fact that the alternation between the three forms of the velar stop in (19) is automatically conditioned by the phonological environment in which this stop appears (here, specifically by the following vowel) makes the particular choice of velar stop in the production of the words in (19) something that the language user has no conscious control over. What the user does have control over is whether to pick a velar stop or some other consonant as the initial segment of the word. Replacing the initial *k* of (19) with a *p* results in a new set of three words, each different from any of the ones in (19).

The initial velar consonants of the words in (19) are thus all representatives of the abstract phoneme /k/. Phonemes are conventionally enclosed in slants. The specific realisations of the velar stop that we see in (2a–c), whose distribution is automatically conditioned by the phonetic context, are called ALLOPHONES (‘alternate phones’) of the phoneme /k/. Allophones are conventionally enclosed in square brackets, as all physical speech sounds (phones) are. But allophones are, more specifically, phones that, together with other, phonetically similar phones, are alternative realisations of the phoneme to which they belong. We can think of the connection between phonemes and their allophones in terms of family relations: the phoneme is the mother of all of its allophones. Only the allophones are directly pronounceable; the phoneme is an abstraction, sitting at a higher structural level than the allophones.

(20) /k/ — PHONEMIC level

[k̟] [k̠] [kʷ] — ALLOPHONIC level

Because allophones of a phoneme have their selection automatically determined by their phonetic environment, they are inevitably in COMPLEMENTARY DISTRIBUTION — where one of these sounds can legitimately occur, the other allophones of the same phoneme cannot. Complementary distribution is a very helpful diagnostic test for detecting the allophones of a phoneme.

A little warning regarding complementary distribution as a diagnostic for allophones is in order, though. The phones [h] and [ŋ] are in complementary distribution in English (as in many other languages, though this is by no means universally true): [h] is confined to prevocalic positions while [ŋ] exclusively appears postvocallically (we find words such as *hang* [hæŋ], but not *[ŋæh]). But although [h] and [ŋ] are in complementary distribution in English, it would not be enormously sensible to call [h] and [ŋ] allophones of a single phoneme: the fact that these two phones lack PHONETIC SIMILARITY cautions us to treat each of them as a phoneme in its own right. For two phones to be allophones of the same phoneme, they will need to be phonetically similar to one another. The sounds [h] and [ŋ] are just too different from one another with regard to the ways in which they are articulated or with respect to their acoustics to make it a reasonable strategy to relate them to a single phoneme.

V.9 A challenge for the phonemic~allophonic dichotomy

In the previous section, we found that in order to have the licence to treat two phones as allophones of a single phoneme, they have to meet the requirement of phonetic similarity. One would think that in cases of phonetic *identity*, two sounds would certainly have to be mapped into a single phoneme. But phonemic analysis sometimes leads to paradoxical results in this connection.

English makes a vowel-length distinction between the members of the following set of minimal pairs of words:[3]

(21)	a.	bit	bid
	b.	bet	bed
	c.	bat	bad
	d.	but	bud
	e.	bite	bide
	f.	beat	bead
	g.	bot	bod

The vowels of all the words in the right-hand column in (21) are consistently longer on the surface than the vowels of the corresponding words to their left. This does not seem to be an accident. The length of the vowel in (21) appears to be automatically conditioned by its phonological environment: when a voiced consonant follows it, the vowel is longer than when a voiced consonant follows it. In light of the systematicity of this vowel-length alternation in English, we are naturally inclined to treat the alternation between these short and long vowels as allophonic. Each of the vowel phonemes in (21) has a short and a long allophone. The selection of the long allophone is rule-governed:[4]

(22) VOWEL LENGTHENING
for any phoneme /V/, select the allophone [Vː] before a voiced consonant

The approach to the pairs in (21) based on the allophonic rule in (22) looks extremely simple and attractive. But in certain varieties of American English, the two words in (23) have the same vowel, and differ only in the length of this vowel.

(23)	bomb	balm
	[bɑm]	[bɑːm]

Here we do not seem to be dealing with an allophonic alternation between a short and a long version of the same vowel. After all, the surface phonetic environment is the same in the two words: in both, the vowel is preceded by /b/ and followed by /m/. Allophonic alternations are always conditioned by differences in the environment. The absence of a difference between the surface environments in which [ɑ] and [ɑː] find themselves in (23) would appear to make it impossible to treat this as an allophonic alternation under the purview of the rule in (22).

3 The word *bot* can refer to the larva of the botfly, but it is more familiar as a neologistic abbreviation of *robot*; and *bod* is a slang word for ‘body’ or ‘person’. We can also replace these with *pot* and *pod*.

4 ‘V’ in (22) stands for ‘vowel’, *any* vowel: for any choice of vowel, we get the same effect, so by using the symbol ‘V’ we generalise over all the vowels of the system.

The fact that a vowel phonetically characterisable as [ɑː] participates in both the alternation in (21g) (plausibly treated as allophonic, on a par with (21a–f)) and the alternation in (23) (which, because of the apparent contrastiveness of vowel length here, does not seem treatable as a case of allophony) poses a conundrum. If we decided to call [ɑː] an allophone of the phoneme /ɑ/ in (21g) but a phoneme in its own right in (23), we would no longer be able to associate the phone [ɑː] with just a single phoneme. Rather, [ɑː] would be simultaneously one of the allophones of the phoneme /ɑ/ *and* one of the allophones of the phoneme /ɑː/. This is something our theory should not allow — for if it did, it would quickly become an extremely difficult process for the language learner to associate phones with phonemes. We will return to the problem posed by (21g)~(23) (first identified by the American structuralist phonologist Bernard Bloch) in §V.11, where we will discover that there is a solution for it. But before we can appreciate this solution, we will need to get acquainted with phonological structure above the segment — so-called SUPRASEGMENTAL structure.

V.10 The syllable

The phonological unit that is probably the closest to being a 'household term' is the SYLLABLE. Native speakers and advanced second-language learners of syllable-timed languages such as English generally have a very good sense of how to break up long words into syllables — even nonsense words such as Mary Poppins' *supercalifragilisticexpialidocious*: though nobody knows exactly what this word might mean, everybody who speaks English will agree that it has fourteen syllables. The question of whether the syllable is a *bona fide* structural unit in phonological theory remains a matter of debate. But there are certain regularities of English mono- and polysyllabic words that receive a straightforward descriptive account in terms of the syllable.

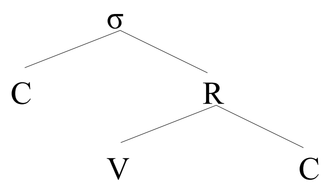
A general restriction on English syllable structure says that stressed syllables are either closed (i.e., have one or more consonants following the vowel) or, if open (i.e., vowel-final), must contain a long (or 'tense') vowel (as in *tea*) or a diphthong or vowel–glide sequence (as in *tie*). Stressed open syllables with a short (or 'lax') vowel are not well-formed. To see this, consider the data in (24) and (25). The vowels in (24) occur in both closed (left-hand column) and open (right-hand column) stressed syllables; the ones in (25) are only legitimate in stressed syllables that have a closing consonant. Note that vowels in (25) do not just want there to be a consonant *near* them: they specifically want this consonant to *follow* them. It does not help these vowels at all that in the right-hand column they are preceded by a consonant.

(24)	a.	/iːt/	*eat*	/tiː/	*tea*
	b.	/eːt/	*eight*	/teː/	*Tay* (river name)
	c.	/aːt/	*art*	/taː/	*ta(r)*
	d.	/oːt/	*oat*	/toː/	*toe*
	e.	/uːt/	*hoot*	/tuː/	*too*
	f.	/aɪt/	*height*	/taɪ/	*tie*

(25)	a.	/ɪt/	*it*	*/tɪ/
	b.	/ɛt/	*ate*	*/tɛ/
	c.	/æt/	*at*	*/tæ/
	d.	/ɒt/	*hot*	*/tɒ/
	e.	/ʊt/	*foot*	*/tʊ/

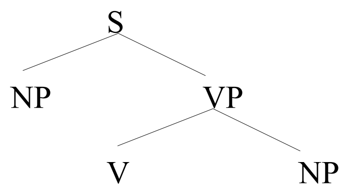
The fact that there is an important distinction to be made between pre- and postvocalic consonants in the structure of the syllable with regard to their relationship with the vowel of the syllable (while postvocalic consonants can help certain vowels occur in stressed syllables, prevocalic consonants provide no assistance of this sort) should not be surprising: the familiar system of 'end rhyme' in poetry and song is based precisely on this distinction. The distinction calls for the recognition, in our phonological analysis of the syllable, of a CONSTITUENT that groups the vowel and the postvocalic consonant(s) together to the exclusion of the prevocalic consonant(s). The name that has been given to this constituent, conveniently, is the RHYME (or 'rime' in American orthography). We will use the letter 'R' to make reference to this constituent in the structure assigned to the syllable. And to the syllable as a whole, we will refer with the symbol 'σ' (the lower-case non-final sigma of the Greek alphabet, chosen because *(a)* the Greek word that is the ancestor to *syllable* starts with 'σ', and *(b)* the letter 'S' from the Roman alphabet had already been taken, in grammatical analysis, by the sentence; see (27)). This gives us the following structure for a syllable with both a prevocalic and a postvocalic consonant alongside the vowel:

(26)



There is a remarkable (and very welcome) similarity between th structure of the syllable in (26) and the structure of the sentence, discussed in segment III, and reproduced in (27):

(27)



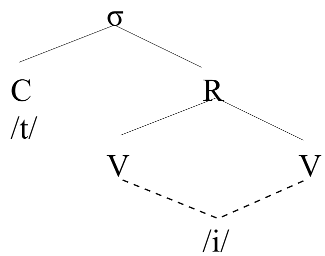
And just as in the structure of the sentence, the NP within the VP is not always there (there are, after all, both transitive and intransitive verbs, the latter not taking an NP-complement), so the structure of the syllable does not necessarily feature a consonant in the Rhyme: there are closed as well as open syllables, the the latter not having a Rhyme-internal consonant (called the CODA).

When we now return to the English stressed syllables in (24) and (25), and look at them against the background of the structure in (26), the generalisation in (28) can be advanced to account for them:

(28) an English stressed syllable is well-formed only if the Rhyme has two dependents

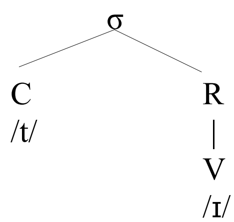
When there is a coda consonant alongside the vowel, the R-node has two dependents: the vowel nucleus and the coda consonant. When the syllable has a *long* vowel, it is standard to assume that the R-node has two branches as well: a long vowel is a single feature bundle associated with two V-slots, as in (29a) (the representation for the syllable *tea*).

(29) a.



A *short* vowel, by contrast, is associated with just a single V-slot. If a short vowel fails to co-occur with a consonant to its right, this gives us a syllable without a branching R-node, as in (29b):

(29) b.



Though /tɪ/ frequently occurs as a syllable in English (as in *plenty* and many other words), it cannot be stressed. That is because stressed syllables in English require a branching Rhyme, as per (28).

That (28) is unconsciously active in the phonology of English native speakers can be shown by looking at the variation among speakers of English with respect to the pronunciation of the vowel of the first syllable in the words in (30):[5]

(30)	a.	Kenya	/kɛnjə/	~	/kiːnjə/
	b.	zebra	/zɛbrə/	~	/ziːbrə/

This variation can be understood against the background of the preceding discussion as being rooted in syllabification. Speakers pronouncing either or both of (30a,b) with the short vowel [ɛ] must syllabify these words as in (31) (where '.' marks the syllable boundary): */kɛ́/ (where '´' marks stress) is ill-formed since stressed syllables in English must have a branching Rhyme, and the short vowel [ɛ] cannot fill two positions in the Rhyme all by itself. On the other hand, speakers pronouncing either or both of (30a,b) with the long vowel [iː] syllabify these words as in (32), with the long vowel single-handedly making up the branching Rhyme of the stressed first syllable.

(31)	a.	Ken.ya	(32)	a.	Ke.nya
	b.	zeb.ra		b.	ze.bra

5 An early demonstration of the psychological reality of (28) is the following observation by the American structuralist Edward Sapir. Sapir taught a group of non-linguists the so-called glottal stop [ʔ] (recall the discussion of *keep* and Cockney *goʔ iʔ* towards the end of §V.3), and discovered that once equipped with the knowledge of the glottal stop, they tended to transcribe the right-hand forms in (25), when they were read out to them, with a glottal stop at the end. This makes sense from the perspective of (28): the glottal stop [ʔ] is a consonant, so if [ʔ] is postulated in postvocalic position in the words in the right-hand column in (25), these words can be mapped onto a syllable structure with a branching R-node, in compliance with the constraint on stressed syllables in (28).

Now that we know this, it will also become understandable why English sometimes hyphenates words in ways that, to non-English speakers, seem baffling. Take, for instance, the fact that the nouns *phonology* and *analysis* are hyphenated after the *l*, whereas the adjective *phonological* and the verb *to analyse* have the hyphen inserted before this *l*. Once we pronounce these words, it becomes apparent that this hyphenation convention follows the constraint on English stressed syllables postulated in (28). In the nouns *phonólogy* and *análysis*, stress falls on the second syllable, which has a short vowel (/ɒ, æ/) as its nucleus. To make this syllable comply with (28), a coda consonant must be provided. So the postvocalic /l/ must be mapped into the second syllable as its coda (rather than into the third syllable as its onset), for compliance with (28). The placement of the hyphen after the *l* (*phonol-ogy*, *anal-ysis*) reflects this. In the adjective *phonológical* and the verb *to ánalyse*, by contrast, stress is not placed on the vowel preceding this /l/, making it perfectly legitimate for /l/ to serve as the onset to the third syllable. In general, natural language prefers to treat intervocalic consonants as onsets of the syllable headed by the vowel following them ('onset maximisation'), rather than as codas of the syllable headed by the vowel preceding them. The hyphenation of *phono-logical* and *to ana-lyse* is a perfect reflection of this syllabification of the word.

V.11 Sylllable weight, compensatory lengthening, and the mora

From the discussion in the previous section, we learnt that English stressed syllables are subject to a constraint that says that, in one way or another, their Rhyme must branch (i.e., have two dependents). This is a weight restriction on stressed syllables. Many languages are QUANTITY SENSITIVE in the sense that they impose such weight restrictions on the syllables that are assigned prominence (stress, pitch accent). And we see that in the need to meet these weight restrictions, languages sometimes go out of their way to compensate for the lack of a coda consonant by lengthening the vowel: so-called COMPENSATORY LENGTHENING.

Lumasaaba (a Bantu language spoken in Uganda) will serve as a good illustration of compensatory lengthening being productively at work in the phonology of one of the dialects. Consider the facts in (33) (with Hungarian spelling adopted for the palatal consonants in (33e,f)).

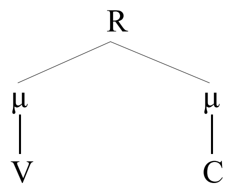
(33)		*dialect A*	*dialect B*	*gloss*
	a.	in-dali	in-dali	'beer'
	b.	in-temu	iː-temu	'snake'
	c.	im-beba	im-beba	'rat'
	d.	im-piso	iː-piso	'needle'
	e.	iny-gyoːla	iny-gyoːla	'I grow'
	f.	iny-tyese	iː-tyese	'sheep'

In both dialects, the nasal of *in-* assimilates in place of articulation to the following consonant — something we remember from English negative *in-*. But in dialect B, the nasal of *in-* does not always show up: this dialect has a phonological rule that deletes the nasal when it precedes a voiceless consonant (a rule that we could state as in (34)). And interestingly, whenever the nasal is forced to drop out by the deletion rule in (34), we see the vowel that precedes it *lengthening*, apparently in an effort to 'make up' or compensate for the loss of the nasal. This is what is called COMPENSATORY LENGTHENING. The pattern is systematic in dialect B of Lumasaaba.

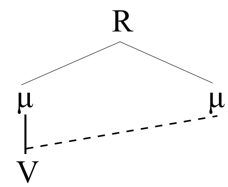
(34) a consonant that is [+nasal] is deleted in front of a consonant that is [–voice]

It is difficult to understand compensatory lengthening well in traditional syllable theory: in what sense can the length of the syllable nucleus (a vowel) be seen to be a function of the presence or absence of a coda consonant? We would like the structure of the syllable to give formal recognition to the concept of syllable weight in such a way that compensatory lengthening can be understood more transparently. A representation of the syllable that, instead of Vs and Cs, takes the dependents of the Rhyme to be MORAE (the plural of MORA) does this very elegantly. This representation (which dates back to the phonologists of the Prague School, in particular Nikolai Trubetzkoy) says, quite simply, that a heavy syllable is a syllable that has two morae, whereas a light syllable is a syllable that has just one mora. The mora is thus the unit in terms of which we measure the weight of a syllable (like a gram is the unit in terms of which we measure the weight of an object). The mora is that which a heavy syllable has two of, so to speak. With this in mind, we can understand compensatory lengthening as in (35):

(35) a.



b.



The syllable of *in-* in Lumasaaba is a heavy syllable, which means that it is represented in terms of two weight units or morae (the symbol 'µ', the lower-case Greek letter corresponding to the Latin alphabet's 'm', is used to reference the mora). Loss of the postvocalic nasal (as a consequence of rule (34)) in dialect B automatically triggers a repair strategy: when the segmental content of the second mora is deleted by (34), the vowel steps in and spreads its features to this second mora. The association of a vowel with two morae is the representation of a *long* vowel. So spreading of the vowel's features over to the second mora leads to lengthening of the vowel.

It is important to note that morae are not 'privileged' for consonantal or vocalic material: anything that contributes to the weight of a syllable can associate with a mora. (Even if a syllable has just a single mora, it can have consonantal rather than vocalic feature content: we see this in so-called 'syllabic nasals' in English words such as *button*, whose second syllable contains just [n̩], associated with the mora.) In (35a), the coda consonant is what makes the syllable heavy; in (35b), where there is no such consonant, the vowel 'makes up' for it by associating with the second mora. It is precisely the fact that morae accept both vocalic and consonantal feature content that makes mora theory ideally placed to account for compensatory lengthening processes.

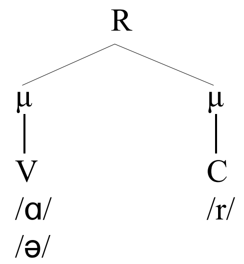
You might at this time ask why we are discussing compensatory lengthening at all, if this is a phenomenon found only in 'exotic' languages. But as a matter of fact, compensatory lengthening is also found in English. We will go through three cases that arguably involve this process. The last of these allows us to loop back to a puzzle left open in the discussion in §V.9: the vowel-length distinction in the American English pair *bomb~balm*.

First, consider the difference between British and American English with respect to the pronunciation of words such as those in (36):

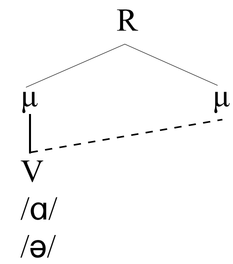
(36) a. far
b. fur

Where American English produces a vowel–consonant sequence (with /r/ as the consonant), British English produces just a vowel. But this is not the only difference between rhotic and non-rhotic varieties of English when it comes to the pronunciation of these words. It has been observed that the duration of the vowel in the British English forms is noticeably greater than that of the corresponding vowel in the American English forms — apparently to compensate for the non-pronuciation of the postvocalic /r/. This compensatory lengthening can be straightforwardly understood within the mora-based approach to the syllable:

(37) a.



b.



Very much as in the case of dialect B of Lumasaaba, the loss of the final consonant /r/, with preservation of the moraic structure, leads to vowel lengthening, representable as spreading, as depicted by the dotted line in (37b).

The second case of compensatory lengthening in English comes from pairs such as the ones in (38):

(38)	a.	sign	signal
	b.	malign	malignant
	c.	design	designate

In the words in the right-hand column, we hear a voiced velar stop, [g], being pronounced right before the nasal, [n]. But in the words in the left-hand column, this velar stop is not present. English phonology has a rule that deletes an underlying [g] in front of a nasal in a word-final syllable. Now, this [g], before it was deleted, occupied the coda position of the syllable in which it occurred. And that syllable is a stressed syllable, so it is beholden to the well-formedness constraint in (28). If we were to produce the short lax vowel [ɪ] (the vowel preceding [g] in the words in the right-hand column) in this stressed open syllable, we would be trespassing against this constraint: [ɪ] can only associate with a single mora. The problem is solved if, as compensation for the loss of [g], we ‘spread the vowel out’ across the two morae — which delivers the diphthong [aɪ], precisely as desired. Of course no lengthening of [ɪ] to [aɪ] will happen in the words in the right-hand column because here [g] is actually pronounced, filling the second mora of the syllable.[6]

6 While it makes good sense to treat the facts in (38) in terms of compensatory lengthening, a wrinkle emerges when we consider that the diphthongisation of the vowel also happens in the case of *paradigm*, alternating with *paradigmatic*. Though [g]-loss is in evidence in *paradigm*, there is nonetheless no cogent motive for compensatory lengthening: the syllable from which [g] is deleted is not stressed in this case, so it should, by the logic of the analysis in the text, be able to survive with a short lax [ɪ]. One suspicion that one may have in this connection is that the diphthongisation of the vowel in *paradigm* happens on the ANALOGY of the phonologically forced diphthongisation of the vowel in *sign*, *malign* and *design*. There can be no doubt that analogy is found in natural language, not just in phonology but elsewhere as well. But it is often difficult to provide concrete and cogent evidence for it. So although appeals to analogy are frequently made in the linguistics literature, they rarely amount to more than hand-waving.

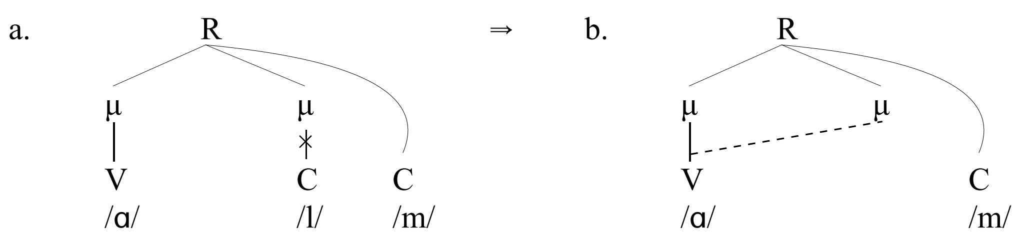
In closing, let us return to the pair of words in (23), repeated below, which in varieties of American English are distinguishable only by the length of the vowel — the quality of the vowel and its consonantal entourage are the same in the two words.

(23)	bomb	balm
	[bɑm]	[bɑːm]

The vowel-length distinction in (23) previously seemed to throw a wrench into the notion that such vowel-length alternations are generally the product of an allophonic rule in English (as they plausibly are in (21)). But now that we have encountered compensatory lengthening, we have a tool to argue this headache away. What we observe in *balm* is that the orthography of the word has the vowel followed by an *l* and an *m*, while the pronunciation of the word only has an /m/ in postvocalic position. If we assume that there is in fact an /l/ in the phonological representation underlying *balm*, but that this /l/ is prevented from associating to the second mora of the Rhyme (e.g., by rule (39)), we arrive at (40b).

(39) delete /l/ immediately before /m/ if the two consonants are part of the same Rhyme

(40)



The /m/ in the representation of *balm* does not itself occupy a mora slot: it is associated directly to the Rhyme, as an adjunct. This will be irrelevant here. What matters is the fate of the second mora of the Rhyme of *balm* after /l/ has been deleted as a function of rule (39). The 'orphaned' mora left behind by this deletion process needs to be rescued. The vowel (already associated with the first mora) comes to the rescue of the second mora by spreading its feature content to it. The result is a long vowel — a long vowel that is derived by phonological rule, not one that is part of the underlying phonological representation of the word *balm*. The underlying representation of *balm*, given in (40a), has a short /ɑ/, associated with just a single mora. In this respect, the /ɑ/ of *balm* is no different from the /ɑ/ of *bomb*. There is no 'deep' length distinction, therefore, between the vowels of *balm* and *bomb* in the relevant varieties of American English — the length of the vowel in *balm* is the product of a phonological process set in motion automatically by the application of the rule in (39) (which, by the way, is not specific to American English: it applies to all varieties of English). This eliminates the threat that (23) seemed to pose to the allophonic rule of vowel lengthening conditioned by voiced coda consonants. So all is well.

V.12 Rule ordering in phonology

But (21g)~(23) is not the only context in which we are confronted with a challenge for the phonemic~allophonic dichotomy. The following example (again from American English, and made famous by Chomsky in his critique of American structuralist phonemics) poses another such challenge:

(41)	a.	write	[rajt]
	b.	ride	[raːjd]
(42)	a.	writer	[rajɾər]
	b.	rider	[raːjɾər]

The alternation in (41) is a straightforward case of vowel lengthening, conditioned by voicing: the final consonant of *write* is voiceless whereas that of *ride* is voiced; the vowel is lengthened before a voiced consonant — so the phonological context of the vowel determines whether it is pronounced short or long. But the vowel appears to be in exactly the same environment in the two examples in (42): the vowel occurs immediately before the alveolar flap [ɾ] in both cases. Yet even though the contexts are identical, a vowel length distinction is made between the two examples. As in the *bomb~balm* case, the surface identity of the environments for the long and short vowel seems to preclude an analysis of the vowel-length alternation between (42a) and (42b) as a phonologically conditioned allophonic alternation.

However, we can actually understand (42) as an allophonic alternation if we look 'under the hood', and state the vowel-length alternation not with reference to the *surface* context in the narrow phonetic transcriptions given above , but over the more *abstract* representations in (41′) and (42′), below, featuring phonemic transcriptions. The alveolar flap [ɾ] in (42a) is a surface realisation of the phoneme /t/, while the [ɾ] in (42b) is a surface form of /d/. If we apply the vowel lengthening rule in (22) *before* the application of the phonological rule producing [ɾ], at the point at which we still have the contrast between /t/ and /d/ (which flapping eventually 'washes out'), the fact that vowel lengthening occurs in (42′b) (with /d/) but not in (42′a) (with /t/) follows.[7] (43) sums this up.

(41′)	a.	write	/rajt/
	b.	ride	/rajd/
(42′)	a.	writer	/rajtər/
	b.	rider	/rajdər/

(43)		INPUT ⇒⇒⇒	VOWEL LENGTHENING ⇒⇒⇒	FLAPPING ⇒⇒⇒	OUTPUT
	a.	/rajter/	n/a	✓	[rajɾər]
	b.	/rajder/	✓	✓	[raːjɾər]

This treatment of the facts in (42) highlights the importance of RULE ORDERING in phonology. We already see in the discussion of morphology in segment II that phonology is not alone in applying its rules in a particular order.

On this note of *rapprochment* between phonology and morphology, we close our discussion of phonology.[8]

7 Although basing ourselves on (42′) handily solves the problem involving the vowel-length alternation, we are not entirely out of the woods: the fact that the alveolar flap [ɾ] serves as a surface realisation of both /t/ and /d/ still continues to present trouble for the hypothesis that a phone can never be an allophone of two different phonemes.

8 Though we have not looked beyond the syllable, there is further suprasegmental phonological structure (the FOOT, most immediately): syllables can be grouped together into feet (such as the iamb and the trochee), which we are familiar with from the metres of poetic verse. But since the analysis of the foot does not bring any major structural insights to the table that we could not already glean from the discussion of the syllable, we will leave a discussion of metrical structure above the syllable aside.