

## Chapter Two

### The melody and the skeleton

#### 2.1 Introduction

As we have seen, segments can be pronounced in somewhat different ways depending on the context in which they appear; they also vary in their ability to combine with other segments. There are other properties of phonological units which deserve attention and which should be reflected in a description. Some of these properties are language-specific, while others characterise a great many or perhaps all languages; in the latter case such properties are not part of individual language systems and their phonology but rather belong to the theory of language and its phonology. The theory of phonology reflects our current understanding of the organisation and the working of the sound system of languages. A general problem in this theory is the question of whether the segments that we identify in languages are indivisible units or whether they have an internal structure.

In what follows we shall argue that segments are complex structures. Each segment comprises a slot, or a position, linked to a group of phonetic properties. The phonetic properties are called **the melody**, while the slots with which they are associated make up what is known as **the skeleton**. The slots which make up the skeleton can be thought of as a sequence of x's so that a word like *America* can be represented as a two-layered structure with melodic units attached to skeletal positions in the following way:

[1]

x	x	x	x	x	x	x	x
ə	m	e	r	i	k	ə	

This representation offers an extremely simple case since there is a one-to-one relationship between melodic units and skeletal positions. If this were always the case, the skeletal level - in addition to the melodic one - would be superfluous; there is quite a lot of evidence, however, which shows the need for the skeleton apart from the melodic representation. In what follows we shall review some of the evidence and relate it to specific phonological phenomena.

#### 2.2 Equivalence of long vowels and diphthongs in English

The vowels or **vocalic nuclei** of Modern English comprise at least three types of objects:

vowels which are short and lax, e.g.: [ɪ, ʊ, ɒ] *bitter* ['bɪtə], *soot* [sʊt], *folly* ['fɒli]

vowels which are long and tense, e.g.: [i:, u:, ɔ:] *lead* [li:d], *stool* [stu:l], *lawn* [lɔ:n]

diphthongs, e.g.: [aɪ, ʊə, əʊ] *revile* [rə'vaɪl], *allure* [ə'ljuə], *notice* ['nəʊtɪs].

The phonetic tradition recognises an additional category of triphthongs for some varieties of English, illustrated by [aɪə] in *Messiah* [məs'aɪə] or [eɪə] in *layer* [leɪə]. We need not be concerned with this complication here - suffice it to say that such complex phonetic units could be interpreted as sequences of a diphthong followed by a vowel, i.e. as [aɪ] or [eɪ] plus [ə].

Another issue which deserves mention is the question of vowel quantity. As is well-known, vocalic length depends to some extent on the environment so that the vowel in *bit*

[bit] is longer than the first one in *bitter* ['bitə], but the nucleus in *bid* [bid] is longer than either of these. Such contextually determined length distinctions should be carefully distinguished from the more basic contrast where the vowel of *bit* or *bid* is distinct from that of *beet/beat* [bi:t] or *bead* [bi:d]. The somewhat varying quantitative differences accompany a basic difference in quality often referred to as a **tense-lax** vowel opposition. Tense vowels are said to require a greater articulatory effort and a more significant departure from the neutral position than lax vowels. Thus tense vowels are both higher and longer as compared to the lax ones. Thus the English short or long vowels are in reality quality-quantity complexes of lower or more open and relatively shorter nuclei as against closer and relatively longer ones. The difference between short and long nuclei, then, involves not only duration but also tenseness. Long vowels which might be said to be of the same duration as diphthongs differ from them in maintaining their articulatory configuration throughout. Diphthongs modify the configuration so that [aʊ], for example, starts as an open front vowel and ends as half close back. The identification of three vocalic categories - short vowels, long vowels, and diphthongs - is thus phonetically motivated. There are, however, serious considerations which force us to look at the three classes in a different way.

Consider the vocalic nuclei which can appear before the velar nasal and a sequence of a labial nasal and another consonant within English morphemes. There are words like *wrong* [rɒŋ], *link* [lɪŋk], *lump* [lʌmp], *timber* ['tɪmbə] in which the vowel preceding the nasal is short; words with either a long vowel or a diphthong are impossible in this position, i.e. something like \*[reɪŋ], \*[laɪŋk], \*[laʊmp] or \*[təʊmbə] are not possible words of English. There is a phonological regularity which only allows a short vowel in such positions; in other words, long vowels **and** diphthongs are banned from this environment. We see that these two distinct nuclei groups - long vowels and diphthongs - are treated in a uniform way in English.

The non-appearance of long vowels and diphthongs before the velar nasal and certain nasal consonant sequences is part of a more general regularity. Specifically, before a sequence of two consonants a short vowel is the preferred option as shown in [2a]; long vowels and diphthongs, while not totally excluded, are seriously restricted [2b].

[2]

- |    |                 |                  |
|----|-----------------|------------------|
| a. | nymph [nɪmf]    | fact [fækt]      |
|    | sombre ['sɒmbə] | sceptre ['septə] |
|    | bump [bʌmp]     |                  |
| b. | salt [sɔ:lt]    | hind [haɪnd]     |
|    | child [tʃaɪld]  | paint [peɪnt]    |
|    | fiend [fi:nd]   |                  |

Without going into details (we return to this question later on), short vowels are tolerated in a broader range of contexts than long vowels and diphthongs. The latter are most frequently found before a sequence of **coronal** consonants, i.e. those whose primary articulation is located in the dental and alveolar region. This means that although short nuclei are possible before a coronal context

[3]

- |                |             |
|----------------|-------------|
| melt [melt]    | bond [bɒnd] |
| Hilda ['hɪldə] | mint [mɪnt] |
| hand [hænd]    |             |

long ones and diphthongs are normally not tolerated before noncoronal clusters, i.e. \*[paʊmf], \*[klaʊkt], \*[seɪptə] etc. These two groups of vocalic elements, which we shall call **complex nuclei**, are subject to the same restriction.

A very different case where complex nuclei pattern together concerns stressed word final vowels. In English short stressed vowels are impossible in the final position, i.e. there are no words ending in any of the following stressed segments [ɪ, e, æ, ɒ, ʌ, ʊ]. Combinations of segments such as \*[rə' bæ], \*[ɪn'tɒ], \*[æblə'me] are not potential English words. The vowels we find regularly in stressed word-final positions are complex nuclei, i.e. either long vowels or diphthongs.

[4]

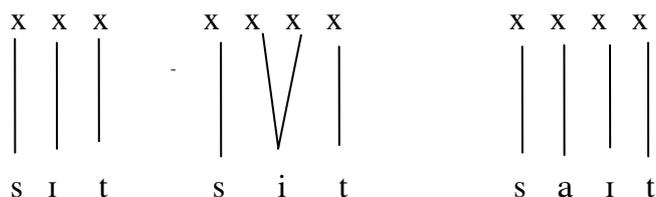
- |    |                 |                         |
|----|-----------------|-------------------------|
| a. | see [si:]       | absentee [æbsən'ti:]    |
|    | accrue [ə'kru:] | hullabaloo [hʌləbə'lu:] |
|    | flaw [flɔ:]     | macaw [mækɔ:]           |
|    | spa [spa:]      | bizarre [br'zɑ:]        |
| b. | sty [stɑɪ]      | apply [ə'plɑɪ]          |
|    | ploy [plɔɪ]     | destroy [də'strɔɪ]      |
|    | glow [gləʊ]     | below [bə'ləʊ].         |

Although we may not be able to find examples for every single diphthong, the general pattern is unmistakable: short vowels are strictly excluded from the position in question while there is no problem about either long vowels or diphthongs appearing there. English phonology treats complex nuclei one way, and simplex ones differently.

As a final argument let us note that there is variation between long vowels and diphthongs in dialects of English. In some dialects the nucleus which we have represented as [eɪ], e.g. *tale* [teɪl] can be more adequately transcribed by means of a long vowel, i.e. [te:l]; conversely, the long vowels [i:, u:], in e.g.: *beam* [bi:m], *boom* [bu:m], in certain other dialects or even varieties of the same dialect should be represented as diphthongs [iɪ, uɪ] (or [ij, uw]). Quite obviously, apart from the monophthongal or diphthongal pronunciations of certain vocalic nuclei, the phonological systems of the dialects treat them uniformly. This goes to show that English interprets long nuclei as equivalent to diphthongs.

We have seen a few instances of the split among English vocalic elements into two classes: simplex or short vowels and complex or long vowels and diphthongs. The phonetic distinction into three groups we noted at the outset reduces to a phonological division into just two classes. The complex nuclei form a single group and hence we need a mechanism for treating them as a unit: referring to a disjunction of long nuclei **and** diphthongs fails to achieve this aim, since a class of two objects could, in principle, comprise any two groups. Thus, theoretically, we could have a class comprising long vowels and short rounded vowels, or long vowels and diphthongs ending in [ɪ] etc. What we need is to separate what we have called simple nuclei from complex ones. This is where the skeletal representation proves useful. A short vowel corresponds to a single skeletal point while both a long one and a diphthong represent a two-point structure. The skeletal and melodic representation of the words *sit*, *seat* and *sight* take the shape in [5].

[5]



Within a representation such as [5], a short or simplex vowel is defined as a melodic unit - in this case [ɪ] - associated with a single skeletal position, while a long vowel or a

diphthong is a melodic unit associated to two such positions. Our phonological observations referring to short vowels or long vowels and diphthongs can be translated into statements about single or double skeletal slots with associated vocalic melodies:

- before a labial or velar nasal followed by a plosive only a single nuclear position is possible;
- two vocalic positions are only possible before coronal clusters;
- a stressed final nucleus must contain two positions.

In no case do we need to refer to a disjunction of either a long vowel or a diphthong since from the phonological point of view they form a unity. This unity is reflected in their being attached to two skeletal slots. The phonological regularities invoked above refer not so much to any vocalic melodic properties but to the duration of the melodies, which can be either short or non-short; whether the latter type has a stable melody, yielding a traditional long vowel, or a changeable one, yielding a diphthong, is immaterial from the point of view of these regularities. Hence we only need to refer to the skeletal structure of words. The split of segments into a two-layered representation is thus justified: we need both the melodic level and the skeletal tier. An important condition on the elements of the two tiers is that they must be linked: the units of the melodic tier must be **associated** with skeletal positions, as it is only together that they form complete structures. Neither melodic elements on their own nor skeletal points without associated melodies can be pronounced, i.e. they do not constitute pronounceable phonological expressions. This we shall refer to as the **Association Condition**. Below we present other arguments attesting to the reality of the tier split and the need for association between the levels.

### 2.3 Germanic and Finnish nuclear simplifications

A different instance illustrating a similarity in phonological behaviour that is puzzling at first blush comes from developments in early Germanic dialects. It is generally recognised that in unaccented syllables long vowels tended to become short and short ones to disappear. Similar, more detailed claims can be made for the grammars of individual languages: in Old English short vowels were lost and long vowels were shortened. As an example, and leaving aside various details which would be required in an exhaustive description of the phenomena, consider the reconstructed form *\*pēodænæs* ‘prince, gen. sg’, pronounced presumably something like [θe:odænæs], which yielded Old English *þēodnes* [θe:odnes]: here the middle vowel is lost; on the other hand, the earlier *\*stānǣ* [sta:næ:] ‘stone, dat. sg.’ and *\*blindūst* [blindu:st] ‘blindest’ became first *\*stānǣ* [sta:næ] and *\*blindust* [blindust], before emerging as *stāne* [sta:ne] and *blindost* [blindost] in classical Old English. Additionally, diphthongs become monophthongs, as in the second syllable of *\*arbaiþ-* [arβaiθ] ‘trouble’ which became the historical form *earfeþ* [æarveθ].

We seem to be dealing not with two but three independent regularities: vowel loss, vowel shortening, and finally with monophthongisation of diphthongs; the only thing the regularities have in common is that the nuclei they affect all appear in unstressed positions. In principle, the three phonological mechanisms could be independent of each other, and the fact that they apply in the same context of an unstressed syllable could be an insignificant accident. However, the recurrence of such accidents in the languages of the world suggests that this interpretation should be ruled out and a more fundamental phonological reason should be sought. If we look at the Germanic phenomena from the point of view of the skeleton, we detect a very simple pattern: unstressed nuclei lose one skeletal position. The monophthongisation of diphthongs is the loss of one position, as is the shortening of long vowels. The loss of a position singly associated with a vowel melody means that the vocalic

melody will be unassociated to the skeleton and hence, in accordance with the Association Condition, will not be pronounced. Thus the superficially three distinct operations are merely different manifestations of one and the same mechanism: the removal of one position dominating an unaccented vocalic melody. We will now look at some of the complex vocalic alternations in modern Finnish which seem to mirror closely the situation in early Germanic.

The superlative degree of adjectives in the nominative singular is formed by adding the suffix *-in* to the adjectival base, which ends in a vowel. Note what happens to the final vowel of the base in the following examples:

[6]

- |    |                               |                      |
|----|-------------------------------|----------------------|
| a. | vanha ['vanha] 'old'          | vanhin ['vanhin]     |
|    | köyhä ['køyhæ] 'poor'         | köyhin ['køyhin]     |
|    | suure- ['su:re] 'great'       | suurin ['su:rɪn]     |
| b. | tärkeä ['tærkeæ] 'important'  | tärkein ['tærkein]   |
|    | lyhye- ['lyhye] 'short'       | lyhyin ['lyhyin]     |
|    | pimeä ['pimeæ] 'dark'         | pimein ['pimein]     |
| c. | tervee- ['terve:] 'healthy'   | tervein ['tervein]   |
|    | rakkaa- ['rak:a:] 'beloved'   | rakkain ['rak:ain]   |
|    | oppinee- ['op:ine:] 'learned' | oppinein ['op:inein] |

The stems in [6a] lose the final vowel before the superlative suffix. In [6b] the adjectival base contains a sequence of two vowels but, just as in [6a], the final vowel is lost and the remaining two vowels appear to yield a diphthong. The stems in [6c] shorten their final long vowel and the resulting combination of two vowels again looks like a diphthong. We are, then, dealing with vowel loss and vowel shortening, a configuration that is largely parallel to what we found in the early Germanic examples above. The contexts of the changes are different but the virtual identity in the treatment of long and short nuclei is something that should be captured in phonological terms.

Let us approach the data with the skeleton-melody distinction in mind.

Loss of a vowel, it will be recalled, means that a given vocalic melody has no skeletal position associated with it and as such is not pronounced. In other words, vowel loss is skeletal position loss. In Finnish this happens to the stem-final vowel, no matter whether this melody is itself preceded by a consonantal or a vocalic melody. The mechanical addition of the suffix to the stem would result in the representation in [7a] and [7c] below, while what we actually find is depicted in [7b] and [7d].

[7]

- |    |   |   |   |   |   |   |   |
|----|---|---|---|---|---|---|---|
| a. | x | x | x | x | x | x | x |
|    |   |   |   |   |   |   |   |
|    | v | a | n | h | a | i | n |
| b. | x | x | x | x | x | x |   |
|    |   |   |   |   |   |   |   |
|    | v | a | n | h | i | n |   |
| c. | x | x | x | x | x | x | x |
|    |   |   |   |   |   |   |   |
|    | t | æ | r | k | e | æ | n |

d.     x       x       x       x       x       x       x  
        |       |       |       |       |       |       |  
        t       æ       r       k       e       i       n

Consider now the shortening of long vowels. Within the skeleton-melody model, a traditional long vowel is a single melody attached to two skeletal positions (see [5]). Given this we need to say nothing new about the Finnish data - vowel shortening is another instantiation of the same nuclear simplification mechanism following the removal of the skeletal position preceding the ending *-in*. Consider the pre-loss configuration in [8a] and the attested representation in [8b]:

[8]

a.     x       x       x       x       x       x       x       x  
        |       |       |       |        \       /       |       |        |  
        t       e       r       v        e       i       n

b.     x       x       x       x       x       x       x  
        |       |       |       |       |       |       |  
        t       e       r       v       e       i       n

Vowel loss and vowel shortening turn out to be mechanical consequences of the suppression of a single skeletal position. The same is obviously true about the early Germanic simplifications which we briefly illustrated above: short vowels are lost when the one and only skeletal point associated with them is removed. Long vowels and diphthongs, when one skeletal position is taken away, become short vowels. Thus the same mechanism is involved in what are phonetically different effects. The regularities can be formulated as a single operation only when the skeletal level is recognised as distinct from the sequence of melodic units.

We thus conclude that the representation of melodic elements which is linguistically relevant consists in the association of the melodic unit with one or two skeletal slots or positions. If one melodic unit corresponds to one skeletal point, we are dealing with what is traditionally called a short vowel or a short consonant [9a]. If a single unit is associated with two slots we end up with a long vowel or a long consonant, normally called a **geminate** [9b]. The Finish adjectives *rakkain* ‘most beloved’ and *oppinein* ‘most learned’ in [6b] contain geminate plosives which in traditional phonetic transcription one denotes by means of a colon, i.e. [k:], [p:]; such a transcription should be seen as a shortcut or replacement for the phonologically more adequate representation as a two-tiered structure. Diphthongs within this system denote a situation where a complex melodic unit is attached to two consecutive skeletal positions [9c].

[9]

a.     x                               x  
        |                               |  
        a                               p

b.     x   x                           x   x  
        \   /                          \   /  
        a                               p

c.     x       x  
        |       |  
        a       i

The fact that a single melodic unit can be associated with two skeletal positions making up a long vowel or a geminate consonant means that the segments in question are at the same time single entities (melodies), and complex structures (skeletal sequences). The fundamental function of the skeletal tier is to capture the purely quantitative or temporal properties of linguistic forms, while their qualitative properties are located on the melodic level. It should be stressed that the representation of long segments in [9b] makes the specific claim that melodically the segments are single and hence they are expected to function in a unitary fashion. On the other hand, it should be possible to have a sequence of two identical slot-melody associations, which would provide evidence against the conflation of the melody to a single unit. Consider in this context the English negative prefix *un-*, e.g.: *unstable* [ʌn'steɪbl], *unkind* [ʌn'kaɪnd] and the present participle suffix *-ing*, e.g.: *building* ['bɪldɪŋ], *dashing* ['dæʃɪŋ]. In forms such as *unnecessary* [ʌn'nəsəsəri], *unnatural* [ʌn'nætʃərəl] we have a sequence of two nasals which happen to be next to each other in much the same way as they are neighbours in *ten names* [tenneɪmz] or *tin knife* [tɪnnaɪf]. The very clear morphological boundary falling “in the middle” of the long nasal in forms like *unnatural* argues for two melodic nasal segments not unlike the two melodic segments which are recognised at the boundary in *unstable* for example. Similarly in *studying* ['stʌdɪŋ] or *carrying* ['kæərɪŋ], we do not want to talk about a long vowel [ɪ:] but rather of a sequence of two melodic units and their associated skeletal positions which happen to occur together. Such sequences of ostensibly long consonants are often called spurious or **fake geminates**. Fake geminate consonants, like the nasal of *unnecessary* or pseudo-long vowels such as the vowel of *carrying*, will be represented as sequences of simplex structures:

[10]



The system separating temporal from qualitative properties predicts the possibility that a complex vocalic or consonantal melody could be attached to a single slot. This is borne out by language data: in the former case we encounter the so-called short diphthongs found, for example, in Modern Icelandic; their consonantal equivalent are affricates, combinations of plosives and spirants within a single segment found in numerous languages. The Icelandic word *hætta* ['haihta] ‘stop’ with the short diphthong [ai] and the German word *zehn* [tse:n] ‘ten’ with the affricate [ts] can be represented in the following ways:

[11]



Further implications of the two-level representation will become evident in subsequent discussion. For the moment we recognise four possibilities for the melody-to-skeleton association:

- one melody - one skeletal position (traditional short vowels and consonants)
- one melody - two skeletal positions (long vowels, geminate consonants)
- two melodies - one skeletal position (short diphthongs, affricates)
- two melodies - two skeletal positions (diphthongs, vowel and consonant sequences).

## 2.4 Compensatory lengthening in Germanic and Turkish

The view of the phonological structure of words presented so far entails the claim that the two levels of representation - the skeletal tier and the melodic one - are independent of but associated with each other. To be pronounced a melodic unit must be connected with one or two skeletal positions and likewise, a skeletal point without an attached melody is silent. A further confirmation of the existence of the skeletal level of representation comes from a phenomenon called **compensatory lengthening**. This notion, functioning both in synchronic and diachronic studies, refers to a situation where a skeletal position exists independently of the melody to which it was originally attached. In other words, a skeletal slot may be shown to persist independently of its melody, thereby strengthening the argument for the skeletal level of representation. We will first look at a historical example involving old Germanic, and then consider a case taken from Modern Turkish.

When certain Old English (OE) forms are compared with those of a related Germanic language such as Old High German (OHG) or Gothic (G), it can be seen that a long vowel in English corresponds to a sequence of a short vowel and a consonant in the other language(s). Consider the pairs of words in [12], concentrating on the quantitative differences:

[12]

[o:] - [an]	OE gōs ‘goose’	OHG gans
	OE oþer ‘other’	G anþer, OHG ander
[o:] - [am]	OE sōfte ‘softly’	OHG samfto
[i:] - [in]	OE līþe ‘gentle’	OHG lindi
[i:] - [im]	OE fif ‘five’	G fimf
[u:] - [un]	OE ūs ‘us’	OHG, G uns
	OE dūst ‘dust’	OHG tunst

The examples show clearly that a long vowel in Old English must historically come from a sequence of a short vowel and a nasal consonant before a spirant. This is normally explained as being due to the loss of the nasal, which is compensated for by the lengthening of the preceding vowel. Note that in our terms a long vowel is a two-slot segment; a short vowel followed by a nasal consonant likewise embraces two slots in the skeletal representation. Thus the development of compensatory lengthening can be represented as a case of re-association where a vowel and a nasal, each associated to single skeletal positions, is replaced by a representation where the vowel is attached to two positions while the nasal remains unattached. The unattached nasal might have remained in the phonology of the speakers who introduced the change, but it was bound to disappear from later representations since subsequent generations of speakers would have no base for assuming any nasal in such words at all. Once the sound change was complete, a word such as *ūs* [u:s] would contain a two-skeletal vocalic melody followed by a fricative. A possible scenario for the change is suggested in [13].

[13]

x	x	x		x	x	x
			=== >	\	/	
u	n	s		u	n	s

What is important here is that if phonological representations consisted solely of segmental melodies, the lengthening of a vowel attending the loss of a consonant before another consonant would have to be an accident. The Germanic case shows that the skeletal representation remains stable so that a short vowel followed by a consonant at some stage corresponds to a long vowel at a later stage. The lengthening of a vowel entails the loss of a consonant; this is possible since the skeletal point associated with the consonant, rather than

being lost like its melody, is combined with the melody of the preceding single slot. What the operation yields is a phonetically long vowel.

In Modern Turkish we find a very similar case. Rather than being an instance of historically related forms, however, this case involves phonological **optionality**, i.e. the ability of some forms to appear in two or more phonetic shapes. This is rather like having the English words *prints* pronounced both [prints] and [prɪns] or *French* as [frentʃ] and [frenʃ], depending perhaps upon the tempo and style of speaking. In Turkish we find numerous words where short vowels are followed by the consonants [h], [j], [v] and another consonant. The first consonant can be optionally deleted; if this happens the preceding vowel is lengthened in a way which is strikingly similar to the Germanic change. Both can be said to represent the process of compensatory lengthening. In Germanic the change is prehistoric and has to be reconstructed; Turkish is caught in the act with speakers free to choose either the form with the short vowel and two following consonants the first of which is one of [h, j, v], or a form with a long vowel and the first consonant of the cluster lost. Consider these examples of alternations resulting from the existence of both options in the language.

[14]

a.	kahya [kah'ja] 'steward'	[ka:'ja]
	tahsil [tah'sil] 'education'	[ta:'sil]
	Ahmed [ah'met] 'name'	[a:'met]
b.	eylül [ej'lyl] 'September'	[e:l'yl]
	düğme [dyj'me] 'button'	[dy:'me]
	seyret [sej'ret] 'watch'	[se:'ret]
c.	sevmek [sev'mek] 'love'	[se:'mek]
	övmek [œv'mek] 'praise'	[œ:'mek]
	savmak [sav'mak] 'repel'	[sa:'mak]

The alternations in [14] show that speakers of Turkish have at their disposal two different pronunciations of the same lexical items. The optionality is phonological and phonetic since it amounts to variant realisations of the same forms: long vowels accompany the suppression of the immediately following consonantal melody. This is a synchronically motivated case of compensatory lengthening and documents the perseverance of the skeletal representation, hence its independence of the melodic structure. Consider the alternative representations of the pronunciation of the word *sevmek* 'love':

[15] a.	x	x	x	x	x	x	b.	x	x	x	x	x	x
									\	/			
	s	e	v	m	e	k		s	e	v	m	e	k

As the representations show, the two different phonetic realisations are not due to different melodic or skeletal structures: the word consists of the same number of skeletal points and the same melodic units arranged sequentially, one after the other. The alternative pronunciations emerge as the result of different associations between the units on the two tiers; specifically, the spirant [v] is unassociated in [15b], hence it is not phonetically audible. The skeletal position to which it is attached in the careful or monitored pronunciation does not disappear but is linked to the preceding vowel. The vowel is rendered phonetically long. Turkish compensatory lengthening presents a situation where a unit present in the melody may be realised directly when associated with its slot, or may remain unrealised; alternatively we could say that the unit is realised indirectly through the association of its slot with a different - neighbouring - melody.

As noted above, compensatory lengthening in early Germanic is a historical mechanism. It is assumed in order to account for certain correspondences between related languages: one of the languages deriving from the same hypothetical proto-language, in this case the predecessor of Anglo-Saxon, compensatorily lengthens some vowels, whereas other languages preserve the original situation. The crucial point is that at the Old English stage, a word like *gōs* 'goose' containing the long vowel [o:] which is the result of a historical compensatory lengthening (cp. OHG *gans*), is not phonologically different from a word such as *dōn* 'do' having the same vowel [o:] which continues an earlier long vowel. There is nothing to make these two vowels different. In other words, one cannot talk about any compensatory lengthening in synchronic Old English phonology. It may be legitimate to involve compensatory lengthening as a historical mechanism, but synchronically Old English *gōs* contains no traces of a short vowel and a following nasal. The Turkish situation is drastically different since the long vowel alternates with a sequence of a short vowel and a consonant; thus compensatory lengthening in Turkish is a synchronically **productive** phonological regularity.

The discussion so far indicates that a segment must be seen as a skeletal-melodic complex whose interpretation depends on the nature of the association between its component parts. In the most straightforward case each unit of the melody is attached to one or two skeletal positions. The Turkish case shows that this simple relation can be disrupted by leaving melodic units stranded. A somewhat more complex case is provided by the phonological aspects of English inflectional morphology, to which we now turn. This new material will allow us to study a different factor conditioning the association between the melody and the skeleton.

## 2.5 The phonology of English inflectional morphology

The inflectional morphology of English is extremely simple. In the present-day language we have in the nominal system the regular ending of the plural, e.g. *bat* - *bats*, the regular ending of the genitive singular, e.g. *taylor* - *taylor's*, and the genitive plural of irregular nouns, e.g. *children* - *children's*. In the verbal system there is the regular past tense, e.g. *walk* - *walked* which happens to coincide with the past participle, e.g. (*I have*) *walked*, and the ending of the 3<sup>rd</sup> person singular present tense of most verbs, e.g. (*s*)*he walks*. As far as the actual endings go, the English system is frugal in the extreme; the phonological properties of the endings and their representations are what we are interested in. We will start with the regular plural morpheme, denoted in spelling by *-(e)s*.

The regular *-(e)s* ending corresponds to three phonetic forms, namely [s], [z], and [ɪz]. The distribution of the phonetic variants is strictly determined by the character of the final segment of the noun. Thus we find [ɪz] when the noun ends in one of the hissing fricatives or affricates [s, z, ʃ, ʒ, tʃ, dʒ], e.g.:

[16]

buses ['bʌsɪz]	roses ['rəʊzɪz]
ashes [æʃɪz]	corteges [kɔ:'teʒɪz]
entourages ['ɒntʊrɑ:ʒɪz]	garages [gæɪrɑ:ʒɪz/gæɪrɪdʒɪz]
watches ['wɒtʃɪz]	sausages ['sɒsɪdʒɪz]

The voiced spirant is the most liberally distributed variant in terms of the number of environments as it occurs after vowels and diphthongs, sonorants, voiced plosives and the voiced fricatives [v, ð], e.g.:

[17]

laws [lɔ:z]	remedies ['remədɪz]
flies [flaɪz]	toys [tɔɪz]

games [geɪmz]	drills [drɪlz]
songs [sɒŋz]	ribs [rɪbz]
beds [bedz]	rogues [rəʊgɪz]
loves [lʌvz]	scythes [saɪðz]

Finally, [s] appears after voiceless plosives and the voiceless fricatives [f, θ], e.g.:

[18]

lamps [læmps]	tickets ['tɪkɪts]
sticks [stɪks]	roofs [ru:fs]
myths [mɪθs]	

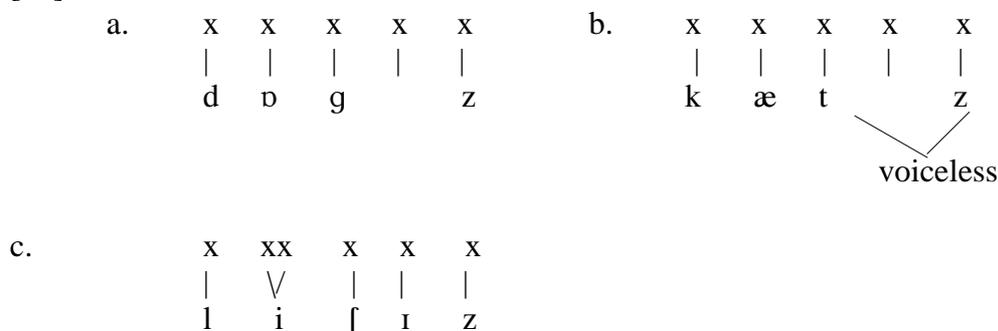
It is evident that the contexts where the variants occur are **mutually exclusive** in the sense that in a given environment one and only one variant can appear; alternatively, wherever one variant appears, none of the others is allowed. Thus, for example, after [g] we can only have [z], and if we have [z] in some word, this is the only possibility in that context, hence [pɪgz] rules out \*[pɪgs] or \*[pɪgɪz]. The presentation of the variants as illustrated in [16-18] gives an account of the facts by **listing** the contexts where each of them appears. Listing forms is basically a way of recording exceptional or **unpredictable** properties of forms: we list the comparative of *bad* as *worse* or the past tense of *sing* as *sang* but the regular comparative of, say, *small* or the regular past tense of *wait* require no individual listing. Listing individual contexts where each of the plural variants appears makes the implicit claim that their distribution is to a larger or smaller extent unpredictable and, in fact, could be very different. Thus our account would not be significantly altered if the facts of English were altogether very different; if, say, we had the ending [s] after plosives and vowels, the ending [ɪz] after diphthongs and voiceless spirants and [z] in all remaining cases, we would find it just as easy to produce a list-like solution, along the lines of [16-18]. All that would differ would be the specific contexts where each variant occurs. In other words, the list-like solution implies that the context where a given variant occurs is an accident, historical or otherwise, just like the comparative of *bad*. This implication is false: the comparative of *bad* could be *badder* just as the comparative of *sad* is *sadder*, but we will argue below that the plural of *lamp* could not be \*[læmpz] or \*[læmpɪz]. The distribution of the plural variants strictly depends on the final segment of the base - we can, in fact, say that the shape of the plural marker is conditioned by what ends the singular noun. A phonological account is interested in capturing this dependence.

Turning to the three variants we observe that one of them is pronounced with a vowel, i.e. it is vocalic ([ɪz]), while the other two are non-vocalic as they consist of just the single consonants [s] and [z]. Note also that the voiceless variant [s] can appear exclusively after another voiceless consonant, whereas the voiced one can appear only after a voiced segment, be it a consonant or a vowel. Thus the spirant of the plural marker has the same voicing as the final segment of the stem. In this sense one can make the tentative observation that the two shapes are merely a manifestation of the requirement of voice agreement between consecutive obstruents that English seems to possess.

The variant with the vowel (i.e. [ɪz]) appears only when the stem ends in a hissing fricative or affricate; the consonants which make up this group constitute the class of hissing obstruents. The consonant of the plural ending is also a hissing obstruent, hence the vocalic variant occurs between two hissing obstruents. In all other contexts the non-vocalic variant is present. We can generalise these observations as follows: the plural marker in English contains two skeletal positions, of which the first is vocalic and the second is the voiced coronal spirant. The vocalic element is pronounced [ɪ] only when attached to a stem ending in a hissing coronal obstruent; if added to a different segment, it is only the coronal obstruent of

the ending that is pronounced and, furthermore, it is realised as voiceless after a voiceless obstruent. A representation for the words *dogs*, *cats* and *leashes* is suggested in [19]; note that the melody of the ending should be specified in terms of properties such as [voicing], [hissing], [coronality] but the simplified representation is adequate for our immediate concerns.

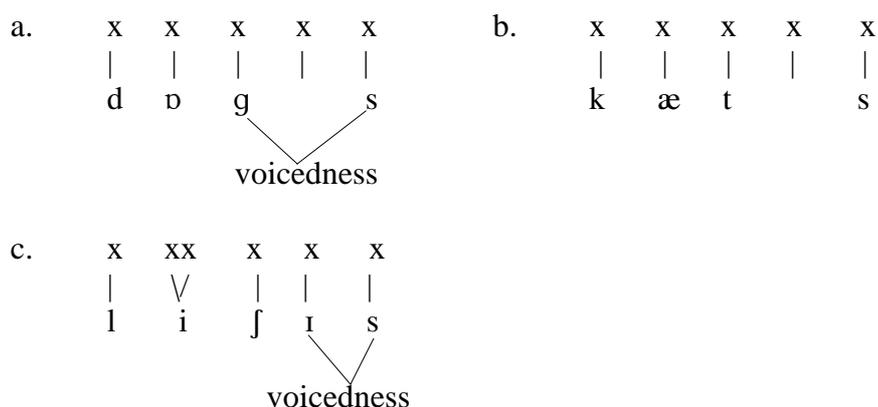
[19]



In [19a] and [19b] the skeletal position preceding the final consonant has no melody attached to it - it is **an empty position**; the melody [ɪ] is attached only when the flanking consonants both belong to the same class of hissing obstruents, as in [19c]. Additionally, in [19b] the final consonant of the ending is specified as *voiceless* in agreement with the voicelessness of the stem final plosive; a [z] which is specified as *voiceless* is, of course, nothing else than, phonetically speaking [s]. We will say that voicelessness is **shared** by the two final consonants.

Since the final consonant of the ending varies between voiced [z] and voiceless [s] we might legitimately ask why it is the voiced consonant which appears in the representations in [19]. Our list-like interpretation in [16-17] makes it clear that after a voiced segment, be it vowel or consonant, the hissing obstruent of the ending must be voiced, while after a voiceless consonant, it must be [s]. Consequently we could adopt a different interpretation and claim that the final consonant is voiceless, and acquires its voicedness from the preceding vowel or consonant. The words of [19] would then be represented in a slightly different way, namely as [20]:

[20]



As we can see, the facts of the English plural ending can be described in two ways. Which is correct?

The two interpretations make different claims or predictions. The analysis embodied in [19] says in effect that a sequence of a voiceless and a voiced obstruent is not acceptable in English; nor is a sequence of two hissing coronals. The analysis in [20] says that a sequence of a vowel and a voiceless hissing coronal is not tolerated in English and that is why the

obstruent shares its voicing with the preceding vowel. This latter claim is factually incorrect since there is no shortage of words which end in a vowel and [s], words which are either morphologically simplex as in [21a], or morphologically complex as in [21b].

[21]

- |    |                      |                          |
|----|----------------------|--------------------------|
| a. | miss [mɪs]           | bogus [bəʊgəs]           |
|    | loss [lɒs]           | mice [maɪs]              |
| b. | happiness ['hæpɪnəs] | manageress ['mænɪdʒərəs] |
|    | hostess [həʊstəs]    | hopeless ['həʊpləs]      |

Furthermore, pairs of words such as those in [22] indicate that the voicing of the coronal spirant after a vowel is not phonologically determined, hence cannot be predicted.

[22]

- |             |             |
|-------------|-------------|
| rice [raɪs] | rise [raɪz] |
| bus [bʌs]   | buzz [bʌz]  |
| miss [mɪs]  | Ms [mɪz]    |

Thus the claim embodied in [20] is seen to be false: the voiceless spirant [s] is perfectly acceptable after a vowel in English. In fact, this analysis encounters other obstacles, since it would also need to voice the final spirant not only after vowels but also after sonorants in words like *bills* [bɪlz], *lambs* [læmz], *rains* [reɪnz]. As there are words where the voiceless spirant [s] freely follows voiced sonorants, both within [23a] and across morphemes [23b], we conclude that just as in the case of vowels, there is no requirement that sonorants must be followed by a voiced hissing coronal.

[23]

- |    |                              |                          |
|----|------------------------------|--------------------------|
| a. | pulse [pʌls]                 | false [fɔ:ls]            |
|    | ransom ['rænsəm]             | dance [dɑ:ns]            |
|    | Samson ['sæmsən]             |                          |
| b. | troublesome ['trʌblsəm]      | minstrelsy ['mɪnstrəlsi] |
|    | insoluble [ɪn'sɒljʊbl]       | youngster ['jʌŋstə]      |
|    | circumstance ['sɜ:kəmstɑ:ns] |                          |

For these reasons the representations in [19] can be said to be phonologically motivated, while the alternative ones would be either at odds with the facts of the language or would require additional complications, e.g. we would need in some way to explain the voicing in *bills* as against its absence in *pulse*.

The analysis of the English regular plural ending reveals the existence of two phonological conditions of the language: 1. sequences of hissing coronals are not acceptable, and 2. sequences of obstruents must agree in voicing. By a sequence we understand consecutive or directly adjacent skeletal positions with their melodic associations. Conditions of this sort will be referred to as phonological **constraints** of the language. We have seen independent evidence showing that it is the voiced coronal spirant in English that adjusts itself in voicing to the preceding voiceless consonant; also, the analysis calls for a skeletal position which is filled by a vocalic melody when a certain phonological constraint needs to be observed, while otherwise the slot remains empty.

The material discussed so far covers the regular plural ending in English. As we noted at the outset, the *-(e)s* ending also marks the 3<sup>rd</sup> person present tense of most verbs, hence side by side with the examples of nouns in [16-19] we find verbs with a homophonous ending. In [24] there are verbs whose stem ends in a hissing coronal.

[24]

(s)he misses ['mɪsɪz]	rises ['raɪzɪz]
bashes ['bæʃɪz]	garages ['gærɑ:ʒɪz] (or ['gærɪdʒɪz])
watches ['wɒtʃɪz]	savages ['sævɪdʒɪz]

The voiced spirant occurs after vowels and diphthongs, after sonorants, after voiced plosives and the voiced fricatives [v, ð].

[25]

(s)he soars [sɔ:z]	remedies ['remədɪz]
flies [flaɪz]	toys [tɔɪz]
names [neɪmz]	drills [drɪlz]
sings [sɪŋz]	robs [rɒbz]
beds [bedz]	begs [begz]
loves [lʌvz]	clothes [kəʊðz]

We are left with [s], which appears after voiceless plosives and the voiceless fricatives [f, θ].

[26]

dumps [dʌmps]	bites ['baɪts]
sticks [stɪks]	spoofs [spu:fs]
sleuths [slu:θs]	

On the face of it there is nothing surprising that what looks the same is pronounced the same way. However, the “look” of sameness is highly misleading: a moment’s reflection will convince us that although the *-(e)s* endings look identical, i.e. they are spelt in the same way, they are completely different entities. They represent different morphemes - plurality vs. 3<sup>rd</sup> person sg. present tense - and are attached to nominal or verbal bases. It is entirely possible that their phonetic realisations could be different - the verbal ending could have just two variants: [ɪz] in the contexts of [24] and [s] elsewhere, or transitive verbs could have one and intransitive verbs the other variant, to take just one possibility. If the variation were morphological or morphologically-conditioned, any arrangement other than the one actually attested would be equally plausible. In actual fact, the two endings are morphologically distinct but the phonology treats them in the same way. In other words, they do not differ as phonological objects, they are subject to the same constraints and consequently display the same phonetic variation. At best we can talk about **morphological homophony** where different morphological categories are expressed by the same phonological means; another instance of such homophony in the area of derivational morphology is the suffix *-er* [ə]: it can denote agentive nouns such as *singer* ['sɪŋə], *player* ['pleɪə] or the comparative degree of adjectives, e.g.: *broader* ['brɔ:də], *darker* ['dɑ:kə].

The morphological homophony in English involving the *-(e)s* ending is even broader. Apart from the two categories just discussed, a same-sounding ending also denotes the genitive singular, e.g.: *George’s* ['dʒɔ:dʒɪz], *child’s* [tʃaɪldz], *cat’s* [kæts] and the genitive plural. The latter case is restricted to plurals formed in an irregular fashion, hence the examples are not very numerous, but they are completely unambiguous: *oxen’s* ['ɒksənz], *sheep’s* [ʃi:ps], *geese’s* ['gi:si:z]. In all these cases the phonological distribution is governed by the same constraints and the effects they induce are identical: voice agreement in obstruent clusters and the filling of the empty position by the vowel [ɪ] when the obstruents happen to be coronal hissing consonants. The identity of the phonological consequences means that the

representation of all these different endings is the same: an empty position followed by the voiced coronal spirant.

Let us now turn to the other ending which practically completes the English inflectional inventory, namely *-(e)d* marking the regular past tense and the past participle. That these two categories are distinct can be seen in the irregular or semi-regular verb groups: thus the past of (*I*) *went*, (*I*) *showed* is distinct from the past participle of (*I have*) *gone* and (*I have*) *shown*. In the regular group of verbs no formal distinction is observed, hence in what follows, while keeping in mind the dual function of the ending, we will speak about the past only.

The past tense ending appears in three shapes whose distribution, illustrated in [27], can be formulated as follows: a. [ɪd] with verbs ending in [t, d]; b. [d] with verbs ending in vowels, sonorants and voiced obstruents except for [d]; c. [t] with verbs ending in voiceless obstruents except for [t].

[27]

a.	wait [weɪt]	waited [ˈweɪtɪd]
	remind [rɪˈmaɪnd]	reminded [rɪˈmaɪndɪd]
b.	score [skɔː]	scored [skɔːd]
	supply [səˈplaɪ]	supplied [səˈplaɪd]
	repel [rɪˈpeɪl]	repelled [rɪˈpeɪld]
	wrong [rɒŋ]	wronged [rɒŋd]
	grab [græb]	grabbed [græbd]
	behave [brɪˈheɪv]	behaved [brɪˈheɪvd]
	clothe [kləʊð]	clothed [kləʊðd]
	advise [ədˈvaɪz]	advised [ədˈvaɪzd]
	charge [tʃɑːdʒ]	charged [tʃɑːdʒd]
	beg [beg]	begged [begd]
c.	sip [sɪp]	sipped [sɪpt]
	laugh [lɑːf]	laughed [lɑːft]
	bath [bɑːθ]	bathed [bɑːθt]
	pass [pɑːs]	passed [pɑːst]
	vanish [ˈvænɪʃ]	vanished [ˈvænɪʃt]
	coach [kəʊtʃ]	coached [kəʊtʃt]
	stroke [strəʊk]	stroked [strəʊkt]

As with the *-(e)s* ending, the distribution of the *-(e)d* variants is complementary but the contexts for variation are partially different. It is true that the voiced/voiceless variants can appear after a voiced/voiceless segment only. This can be regarded as due to the regularity established above which requires that obstruent sequences should be uniform in voicing. However, the vocalic variant appears after hissing sonorants in the plural and 3<sup>rd</sup> singular present, e.g. *hisses* [ˈhɪsɪz], and after a coronal plosive in the past tense, e.g. *waited* [ˈweɪtɪd]. As a result, a verb ending in a hissing sibilant takes the vocalic variant in the 3<sup>rd</sup> person singular present tense but the non-vocalic one in the past: *miss* [mɪs] - *misses* [ˈmɪsɪz] - *missed* [mɪst]. Likewise a verb ending in a coronal plosive takes the non-vocalic present tense but the vocalic past tense variant: *fade* [feɪd] - *fades* [feɪdz] - *faded* [ˈfeɪdɪd]. A closer look at the data reveals a certain similarity between the two groups: the variant containing a vowel appears when the consonant of the ending is similar to the final consonant of the base to which the ending is attached. Thus the hissing coronal spirant of the *-(e)s* ending is separated by a vowel from a hissing coronal obstruent terminating a base, and likewise the coronal plosive of the -

(e)d ending is separated from a coronal plosive in the base - in some sense, then, English disallows sequences of too similar consonants and requires that the slot separating such consonants should be filled by a vocalic melody.

Viewed in this way, the three variants which we find in our endings are all governed by the same two constraints: 1. very similar consonants may not form a sequence but must have a vocalic melody between them and 2. obstruent sequences must be uniformly voiced or voiceless. There is no need to list the individual variants or specify contexts for the distribution. Note specifically that the different contexts for the vocalic variants of the two endings (*misses* - *faded*) follow from the same, more general constraint disallowing sequences of similar obstruents and need not be specified separately for each of them. The representation of the endings is simply as follows:

[28]

x	x	x	x
	z		d

Given these representations, we can provide an account for their variants by means of the two constraints. The variant phonetic forms constitute **an interpretation** of the representations in [28]. In other words, the phonetically attested variants are interpreted representations of linguistic forms.

We have been assuming so far - and this is reflected in the representations in [28] - that the empty vocalic position is filled under specified conditions, i.e. between similar obstruents. It is perfectly possible to imagine an alternative analysis, namely one where the vocalic melody is present in the representation and gets de-associated from the skeletal position when **not** surrounded by similar obstruents; with the association severed, the melody cannot be pronounced. Our main concern in this chapter is the separation of the melody from the skeleton; we also entertain the possibility that there may exist skeletal positions without any melody attached to them. From this point of view we do not have to make up our minds which of the two potential interpretations is to be selected - this is something that would belong to a comprehensive account of English phonology. For our immediate purposes we note that slots without attached melodies will figure in both of the accounts: in the interpreted representations there are going to be slots with unassociated melodies, i.e. empty slots in one analysis, or slots with severed, hence also unassociated, melodies in the alternative one. It is worth pointing out, however, that there is some evidence which argues in support of the analysis which severs the association between the melody and the skeleton as against the empty position analysis presented above.

There are adjectives in English ending in *-ed* such as e.g. *witted* ['wɪtɪd] in *quick-witted*, which are related to the participle forms of verbs. In this case the *-ed* ending is pronounced in the same way as in the verb (*I have been out*)*witted* [aʊt'wɪtɪd], with the slot preceding the final consonant filled by a melody. In other cases, however, the ending has distinct pronunciations in the past tense and in the adjectival form:

[29]

learned [lɜ:nd]	learned ['lɜ:nɪd], <i>a learned person</i>
aged [eɪdʒd]	aged ['eɪdʒɪd], <i>an aged eagle</i>
loved [lʌvd]	beloved [br'lʌvɪd], <i>the beloved country</i>

Similarly, when participles are turned into adverbs, there is a difference in the way the *-ed* ending is pronounced:

[30]

advised [əd'vaɪzd]	advisedly [əd'vaɪzɪdli]
deserved [dɪ'zɜ:vɪd]	deservedly [dɪ'zɜ:vɪdli]
assured [ə'ʃʊəd]	assuredly [ə'ʃʊərɪdli]
fixed [fɪkst]	fixedly ['fɪksɪdli]
supposed [sə'pəʊzd]	supposedly [sə'pəʊzɪdli]
pronounced [prə'naʊnst]	pronouncedly [prə'naʊnsɪdli]

The empty slot solution would need to fill the slot with the melody in the participles when used as adjectives or adverbs. Thus, in addition to the phonologically motivated slot filling when in the environment of similar consonants in nouns and verbs, this solution would also need to fill all slots irrespective of the context in adjectives and adverbs. The melody severing analysis does not need to say anything about adjectives and adverbs, as it merely severs the association in nouns and verbs when the vowel slot is not surrounded by similar consonants. This second solution is simpler and more direct and as such is perhaps preferable.

A more general argument in support of the second solution may derive from the fact that severing association lines between slots and melodies is well-attested in other areas of English phonology. An illustration comes from the so-called **strong** and **weak forms**. We will consider just one instance of this phenomenon here, namely the verb *has*, pronounced [hæz] in its strong form, and [həz, əz, z, s] in the weak version.

Consider the forms [əz, z, s] which are highly reminiscent of the *-(e)s* variants found in the inflectional endings we have just discussed. That the distribution of the variants is identical in both groups can be illustrated by the following examples:

[31]

- |    |                         |                               |
|----|-------------------------|-------------------------------|
| a. | Tom's [tɒmz] car        | Tom's [tɒmz] been here        |
| b. | Jack's [dʒæks] car      | Jack's [dʒæks] been here      |
| c. | George's [dʒɔ:dʒɪz] car | George's [dʒɔ:dʒəz] been here |

If we assume that the weak form of *has* is [əz], then the realisations found in [31] are due to the severing of the association between the slot and the vocalic melody and the observation of the two constraints formulated above. Voice agreement accounts for [31b], while the ban on consecutive similar obstruents accounts for [31c]. As [31c] shows, the form *George's* has two possible pronunciations in RP. The vocalic melody [ə] is the melody which has not been severed; by the same reasoning we may say that [ɪz] in *George's car* is the melody which has not been delinked in observance of the ban of adjacent similar consonants. This brings us again to the conclusion that the representations of the inflectional endings contain vocalic melodies. Thus, rather than [27] we may postulate [32] as the representation of the inflectional morphemes.

[32]

x	x	x	x
ɪ	z	ɪ	d

We have considered two different analyses of the English data. They embrace a situation where a melody is de-associated from its skeletal position, or where a skeletal position has no melody. In other words we envisage the possibility of a phonologically interpreted representation containing slots and melodies which remain phonetically inaudible. Another such case is discussed directly below.

## 2.6 English 'linking' *r* and the unassociated melody

One of the differences among English dialects concerns the distribution of the sonorant [r]. There are dialects, called **rhotic**, where [r] occurs not only before vowels but also before consonants and word-finally. These dialects are to be found in large areas of the British Isles, in particular in Scotland and Ireland, and also in most of the USA. The **non-rhotic** dialects, which disallow preconsonantal and word-final [r], are, among others, those of southern Britain and the eastern United States. This means, in effect, that the distribution of the sonorant [r] within non-rhotic dialects is restricted to the position before a vowel and marginally also the semi-vowel [j] in unstressed positions, e.g.: *bright* [brɪt], *berry* ['berɪ], *caress* [kə'res], *garrulous* ['gærjʊləs]. The distribution of [r], just like that of the dark [ɹ], can be described as requiring the presence of a following vocalic element. Rhotic dialects are not restricted in this way since the sonorant can appear in non-vocalic contexts as well, e.g.: *bark* [bɑ:k], *bar* [bɑ:].

If the differences were restricted to the presence as against the absence of a segment in a specified position, we would be dealing with a partially different distribution of a segment in some dialects of the language. However, the non-rhotic dialects which disallow word-final [r] admit it there in certain cases. Specifically, a final [r] is not possible when the next word begins with a consonant or when there is a pause; if the next word starts with a vowel, however, the final [r] must be pronounced. This final pronounced [r] is called **linking r** in the phonetic tradition. We shall consider the phonological significance of this phenomenon now, starting with a list of relevant examples.

[33]

fear [fɪə]	fear of flying ['fɪərəv 'flaɪɪŋ]
sure [ʃʊə]	sure of himself ['ʃʊərəv him'self]
share [ʃeə]	share of it ['ʃeərəvɪt]
far [fɑ:]	far above ['fɑ:rə'bʌv]
bore [bɔ:]	bore us ['bɔ:rʌs]
for [fə]	for example [fə ɪg'zɑ:mpəl]

The left hand column words all end in a vowel, while the same words end in [r] if followed by another word beginning with a vowel. The same regularity can be observed when a vowel-initial suffix is attached to a word that ends in a vowel. Consider the examples in [34].

[29]

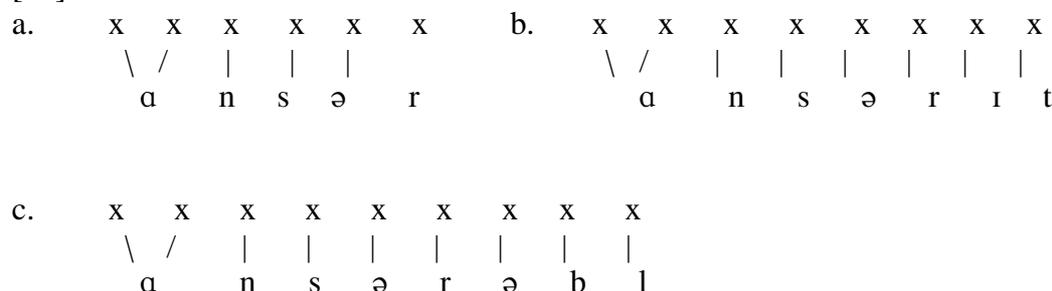
answer ['ɑ:nsə]	answerable ['ɑ:nsərəbl]
cover ['kʌvə]	coverage ['kʌvərɪdʒ]
appear [ə'pɪə]	appearance [ə'pɪərəns]
declare [drɪkleə]	declarative [drɪklærətɪv]
consider [kən'sɪdə]	consideration [kən'sɪdə'reɪʃn]
secure [sə'kjʊə]	security [sə'kjʊərətɪ]
murder ['mɜ:də]	murderer ['mɜ:dərə]

The vowel-initial suffixes of [34] behave in exactly the same way as words beginning with a vowel: [r] appears as the final consonant of the preceding word or morpheme. As noted above, within non-rhotic dialects the sonorant [r] must be followed by a vocalic element. In other words, it is only the presence of a following vowel that allows a preceding [r] to be pronounced - we shall use the term **licensing** to refer to this sort of situation, and say that [r] must be licensed by a following vowel in non-rhotic dialects. In rhotic dialects no licensing relation of this type is required.

The question might be asked as to how the alternations in [33-34] should be represented. If in absolute word-final position [r] cannot be licensed since no vowel follows it,

the melody must remain unattached. If, however, a vowel follows - either as part of the following suffix in the same lexical item or in the following word - then the sonorant is licensed and retains the association between its melody and skeletal point. Consider the verb *answer* in *answer it* and *answerable*.

[35]



In [35a] the final [r] is not licensed by a vowel, hence the association between the two constituent parts of the segment - the skeletal position and the melody - is not present (or has been severed) with the result that the segment is not pronounced. In both [35b] and [35c] the vowel following the sonorant in question licenses it, thereby maintaining the segment's internal association - consequently the sonorant is pronounced in this position. The representation [35a] is the one we are concerned with, as it offers another illustration of the absence of association between the skeleton and the melody resulting in the segment not being pronounced. In other words, the association is part of interpreted representations, and thus without it, the melody is only latently present. The alternations between [r] and zero in [33-34], just as the earlier alternations between [ɹ] and zero in inflectional endings, show that different factors determine the presence of association between the two tiers of representation: the nature of the surrounding consonants in one case, and the presence of a following vocalic melody in the other. As we will see in a number of cases below, phonological regularities result from the relations between neighbouring segments just as much as from the internal organisation of segments.

Before concluding this brief account of [r]-zero alternations in non-rhotic dialects of English, we would like to mention a phenomenon which often appears in the context of linking *r*, namely the so-called **intrusive r**. This is found in non-rhotic dialects and consists in the appearance of [r] at the end of a word before the vowel of the next word or a following suffix in forms which have no [r] in rhotic dialects (or in the spelling). The word *withdraw* [wɪð'drɔ:] in non-rhotic dialects can be pronounced with [r] in *withdraw it* [wɪð'drɔ:ɹɪt] or *withdrawal* [wɪð'drɔ:ɹəl], while rhotic dialects tend to maintain the usual pronunciation of the infinitive in the other forms as well: [wɪð'drɔ:ɪt, wɪð'drɔ:əl]. It seems that the non-rhotic dialects treat words like *withdraw* in the same way that they treat words like *answer*, i.e. with a final [r]-melody unassociated to a skeletal position. If a vowel follows, it licenses the association and the segment is pronounced. In rhotic dialects, the intrusive *r* phenomenon does not exist and we are dealing with representations without unassociated segments: words like *withdraw* simply end in a vowel.

## 2.7 Summary

This chapter has introduced the need for a layered view of phonological representations. Contrary to everyday intuitions, the linear sequence of segments provides only a small portion of the phonologically relevant information. Segments do follow each other in a linear sequence, but they also enter into such closely-knit relationships with one another that it often becomes impossible to separate them, since certain properties belong to more than one

segment at the same time. To reconcile the segmentability with the inseparability, phonological segments are viewed as consisting of skeletal positions and associated melodies.

Positions are basically temporal slots appearing in a sequence and thus they reflect the segmentability intuition. Melodies, while attached to skeletal positions, need not be singly attached but may at the same time span two or more such positions. The phonetic effect of such double - or multiple - attachment is the simultaneous presence of a given property in consecutive timing slots, in other words the inseparability of sounds.

A fundamental insight behind the skeleton-melody distinction is the independence of units at each level. Pronounceability requires that every melodic unit should be attached to some skeletal position and, conversely, that skeletal positions without any melody attached should remain silent. Thus, for example, a unit of vocalic melody attached to two skeletal positions corresponds to what is usually called a long vowel. Since, however, the two levels are independent, operations at one level do not influence the other. The removal of one skeletal position from a doubly attached melody will result in the shortening of that melody without affecting it in any other way. Similarly, the establishment of a melodic connection between consecutive skeletal positions will in no way influence the number or order of the positions. Phonological operations may affect either of the two levels and also the associations connecting them.

An important implication of such a model of organisation is the recognition of two types of phonological objects. On the one hand we can have skeletal positions without any attached melody (or with a severed melody) - these are the so-called empty positions. In this chapter we have come across one case of such a possibility, in the discussion of English inflectional morphology. Another object predicted by the model are unassociated melodies; obviously, if they are unassociated, they will remain inaudible but their existence must be justified by phonological evidence. In other words, the phonological evidence must call for the recognition of such unassociated melodies. This is the case of the linking *r* in non-rhotic dialects of English.

In this chapter we have seen several instances, most of them quite simple, of what might be termed phonological generalisations. Such generalisations, also called constraints, may affect the melodic tier, the skeletal tier or the association between the two. Phonological regularities will occupy us for the rest of this book. To see how they can be established and justified we will look at a number of individual cases in some detail.

## 2.8 Suggested further reading

The relation between quantity and quality in English is covered by phonetic descriptions such as Gimson and Cruttenden (1994) or Jones (1975).

The Germanic lengthenings are described in grammars of older Germanic languages, in particular Meillet (1970) and Hogg (1992).

For the Finnish data consult Whitney (1959) and Morozova (1972); a more comprehensive account is to be found in Kiparsky and Keyer (1984).

Geminates have been subjected to numerous phonological analyses, e.g.: Schein and Steriade (1986), Hayes (1986), the relevant parts in Kenstowicz (1994).

For the skeleton and compensatory lengthening see Prince (1984), Goldsmith (1990, Chap. 2), Kenstowicz (1994, Chap. 8), Harris (1994, Chap. 2.2), Bickmore (1995), Perlmutter (1995).

The presentation of Turkish is based on the data in Sezer (1986).

For partially different accounts of the phonological regularities displayed by English inflectional morphology see Bloomfield (1933, Chap. 13) Anderson (1974, Chap.4). Linking and intrusive *r*'s are reviewed and interpreted in Wells (1982), Giegerich (1992, 1999) and Harris (1994, Chap. 5).