Chapter 2
Rhythmic Patterns in Language

2.1 Generalities

Speech – like music, that other culturally codified production of sound – is a rhythmically organized activity. Just as the rhythmic schemes of the different musical traditions are a codification presumably reflecting both universal properties of rhythmic organization and the conventionalized choices about particulars made by the culture (from the range of choices that are available within the confines of a (universally defined) rhythmic organization), so too, we will argue, do the diverse rhythmic patterns of prominence found in language reflect both universal and particular aspects of rhythmic organization. The rhythmic patterning of every language is codified, and the linguistic description of speech rhythm, therefore, is a description of that code. A successful theory of rhythmic structure in language will involve sorting out what is language-particular and what is universal in that code.

Rhythm, as we said in chapter 1, is founded upon a recurrence of pulses, some of which emerge as accented with respect to each other, forming patterns on different levels. Various aspects of rhythmic organization inhere in the metrical grid representation. A grid position is a pulse, a point in (abstract) time. The notion of pattern requires a differentiation among the pulses, and this is provided by ascribing to them a place on different levels. Sequences of rhythmically equivalent pulses are represented as the sequence of grid positions on the different metrical levels. To every level, then, there is a patterning. Thus, claiming that speech is rhythmically organized – that in the linguistic representation of an utterance syllables are aligned with a metrical grid – amounts to claiming that (a) there are discernable pulses in speech, (b) some of these syllables emerge as accented with respect to each other, forming recurrent patterns, and (c) there may be a hierarchy of accented syllables, corresponding to patterns on the different metrical levels. It may be argued, and has been, that all these statements hold true for speech.

What does not inhere in the grid, in our conception of it, is the precise nature of the patterns themselves. An accented (strong) position could in principle be followed, or preceded, by any number of non-accented (weak) positions. This, we believe, is entirely appropriate: the grid itself simply *represents* patterns, as the organization of pulses of abstract time into levels. The patterns are *characterized* by principles or rules that govern the particulars of the metrical grid organization.

There is arguably a universal rhythmic ideal, one that favors a strict alternation of strong and weak beats. Following Sweet 1875-76, we will call this, as in chapter 1, the *Principle of Rhythmic Alternation* (PRA). That the rhythmic organization of language or music aspires to such an ideal state is indicated by a number of general tendencies that are attested in rhythmic patterning: (i) binary patterns (successions of *ws* or *sw*), at various levels, are by far the most prevalent in either domain of activity, (ii) ternary patterns (*wws* or *sww*) are not (usually) basic in the rhythmic patterning at any level, but rather exist alongside binary patterns and constitute a special departure from them, and (iii) quaternary patterns simply do not exist, inasmuch as they may be interpreted as two binary patterns. This is Sweet's point, one also implicit in Liberman 1975. In his work on rhythm in French, Dell (to appear) also argues for a universal "ideal" rhythmic organization, following *le principe d'eurythmie*, according to which an ideal metrical grid contains no adjacent strong positions and is maximally alternating.

Our position is not that the PRA plays a direct role in the linguistic description of a language's patterns of prominence, but rather that the rules of a grammar that define the possible metrical grid alignments of the sentences of that language conspire in approximating that

ideal, on every level of organization. At the end of this chapter, we will outline our theory of metrical grid alignment (or theory of the notion "possible rhythmic 'score' " for language) and will show how we think the rhythmic ideal is embodied in the rules for constructing a metrical grid on the basis of a text. In the immediately following section, we will examine the evidence for viewing speech as a rhythmically organized activity, one that both respects an abstract rhythmic structure having the character of a metrical grid alignment and shows patterns reflecting the operation of the PRA.

2.2 The Rhythmic Nature of Speech

Scholars working on the phonetics of English have long recognized that rhythmic organization at what we will call the "lower levels" plays an important role in the description of speech.[1] A major insight of these investigations is that the quantities (durations) of the syllables of a sentence are determined by the sentence's rhythmic properties. According to D. Jones 1964:§886, for example,

Vowel length depends to a considerable extent [in English] on the rhythm of the sentence. There is a strong tendency in connected speech to make stressed syllables follow each other as nearly as possible at equal distances.

For D. Jones 1964:§888-890, the musical notation in (2.1a) and (2.1b) represents the fact that in (2.1a) "*ei, ai* are something like twice as long in the second sequence as they are in the first" and that in (2.1b) "the *i:* in *scene* is considerably longer than the *i:* in *scenery.*"

(2.1)

a. 　♫　　　♫　　　♫　　vs.　♪　♪　♪

　　'eiti:n　　'nainti:n　'twenti　vs.　'eit　'nain　'ten
　　eighteen　nineteen　twenty　　　*eight　nine　ten*

b. 　♪　♪.　♪　♬　　vs.　♪　♬♬　♬♬

　　ðə　'si:n　wəz　'bju:təfl　vs.　ðə　'si:nəri　wəz　'bju:təfl
　　The　scene　was　beautiful　　*The　scenery　was　beautiful*

Pike 1945:34 makes a not dissimilar observation about American English. He points out that the space of time elapsing between the stressed syllables *man* and *here* is roughly the same in *The man is here* and *The manager's here*, with the result that in the latter the syllables are "crushed together" (pronounced very rapidly), in order to be "fitted in."

Building on the insights of Jones and others, Abercrombie 1964 laid the foundation for further work on the topic by positing a unit of analysis, the *foot* (not to be confused with the foot of metrical phonology), in terms of which, he argued, the rhythmic properties of syllables may be explained:

English utterances may be considered as being divided by the isochronous beat of the stress pulse into feet of (approximately) even length. Each foot starts with a stress and contains everything that follows that stress up to, but not including, the next stress. '*This is the* '*house that* '*Jack* '*built* has therefore four feet, and they can be most conveniently represented by the use of vertical lines:

[1] Steele 1775, Sweet 1875-76, D. Jones 1964, Classe 1939, Jassem 1949, 1952, W. Allen 1954, Kingdon 1958a,b, Abercrombie 1964, 1967, 1968, Catford 1966, 1977, Halliday 1967a, Gimson 1970, Pike 1945, Bolinger 1965b, G. Allen 1972, 1975, Lehiste 1980 (and references cited therein), and Thompson 1980 are to be cited as among the more notable who have worked on problems of speech rhythm in English.

This is the | house that | Jack | built

The quantity of any syllable is a proportion of the total length of the foot within which the syllable occurs, and it is relative to the quantity of any other syllable in the foot. (1964:217)

The claim being made, then, is that English utterances can be characterized in rhythmic terms, as consisting of a sequence of isochronous pulses. This claim about isochrony in English speech has engendered considerable debate (see Lehiste 1980 for a review).[2] We feel that much of this debate, and the apparent lack of resolution on the question, stems from a failure to understand rhythm, and the isochrony of its basic pulses, in sufficiently abstract terms. With Liberman, Lehiste, and others, we claim a psychological reality for rhythmic organization in both the production and the perception of speech, but acknowledge that the ideal isochronicity this organization presupposes may not always reveal itself in easily measurable terms in the acoustic signal.

We may liken the metrical grid alignment of a sentence – the linguistic representation of the sentence's rhythmic structure – to a musical score. It must be recognized that a musical score, or a metrical grid alignment, though grounded in its very conception by the rhythmic capacities of the human organism (the rules for defining possible scores or metrical grid alignments embodying, as they do, universals of rhythmic organization), is but an abstract scheme. The same score may be interpreted in many different ways by the same or different performers; though of course there are limits within which variation in interpretation is confined, if the score is to remain recognizable, that is, if the score is to be considered to have been "realized" in the performance. A distinction must therefore be made between the score and its interpretation. This is the distinction between *langue* and *parole*, or between competence and performance (using the latter term now in its technical sense). For linguistic patterns of prominence, this is the distinction between the metrical grid alignment of the sentence and its phonetic implementation. In measuring for isochrony, one is measuring only the performance of the score, and not the abstract patterning that makes it up. We give full credence, therefore, to the *impression* of isochrony – to the impression of rhythm – as revealing something about how the mind grasps the organization of speech in time.[3] It is what the mind grasps that is of interest to us here, for that is presumably what is embodied in the metrical grid alignment of a sentence. Indeed, isochrony should in no way be considered as the *sine qua non* of a truly rhythmic system. The notion of *pattern*, as consisting of a regular recurrence of motifs defined in strong-weak terms, may be just as important as isochrony in establishing the rhythmic character of speech.

The thrust of the observations about isochrony in English is that not all syllables are rhythmically equivalent to each other; only some are. Moreover, only some are in the relation of (ideal) isochrony with respect to each other. This means that all syllables do not qualify as pulses, in Cooper and Meyer's sense. Rather, it is only the "stressed," Abercrombian foot-initial syllables that align with the pulses, or beats, of the rhythmic organization of English speech. The others have in fact a somewhat variable realization in time, depending in part on how many of them there are between the basic pulses. What does this mean for the alignment of these syllables with the metrical grid? We could contemplate giving Pike's second example the grid alignment below, where the sequence of *x*'s represents the lowest level of the metrical grid (and the higher levels of beats are not represented):

---

[2] A variety of experimental studies have failed to show an exact isochrony of stressed syllables in speech production (Classe 1939, Bolinger 1965b, Uldall 1971, 1972, Lehiste 1973b, 1975a, Lea 1974), though most show a tendency toward isochrony. See Thompson 1980 for a recent study disputing the presence of isochrony in production.

[3] There is a fair amount of evidence favoring the psychological reality of isochrony as a perceptual phenomenon (e.g., G. Allen 1972, 1973, 1975, Lehiste 1979a, Donovan and Darwin 1979).

(2.2)

```
      x       x
The manager's here
```

Here only the syllables that coincide with a pulse, an *x*, are specified for an alignment. The others are in limbo, so to speak. Alternatively, we could postulate a yet lower level on the metrical grid, one with which all syllables would be aligned, but not one where the points are taken to mark out ideally isochronous pulses. Given this approach, the pulses, or beats, would be represented only on the second level and above. *The manager's here* would have the following grid alignment (still incomplete at the higher levels):

(2.3)

```
      x     x
  x   x x   x
The manager's here
```

As will become clear, it is extremely useful to assume that all syllables of the utterance are integrated into the rhythmic organization of the utterance, in the sense of having a specified alignment with the metrical grid. In particular, this assumption makes possible a straightforward treatment of patterns of word stress, a description of rather fine details of timing within feet, and a representation of subtle aspects of syntactic timing. We will therefore adopt the latter sort of metrical grid alignment for language. Within the metrical grids of speech rhythm, we will thus distinguish a *first metrical level* (the lowest), a *second metrical level*, and any number of levels (in principle, infinitely many) above that. The positions on the second metrical level and above will be referred to as *beats*. The positions on the second metrical level will be referred to more specifically as *basic beats*. Finally, the positions on the first level will be referred to as *demibeats*, this term being chosen to reflect the only quasi-pulse status of the positions on the lowest level.

It may be argued that the difference between languages that Pike 1945 has called *stress-timed*, like English, and those that are *syllable-timed* can, and indeed must, be represented at the basic beat level of rhythmic organization. A syllable-timed language is one in which it is said that there is a tendency towards isochrony of all syllables, a relative constancy in the duration of syllables in the utterance, and a (relative) lack of vowel reduction (Pike 1945:35-37, Catford 1977:85-88, Abercrombie 1967:96-98). French, Italian, and Spanish are often cited as examples of syllable-timed languages.[4] For example, in the Italian *il popolo*, the syllables are pronounced as a staccato progression of evenly spaced beats. Viewed in metrical terms, a syllable-timed language is one in which each syllable is aligned with a (basic) beat in the metrical grid. Given the assumption that beats appear only at the second metrical level, then, the minimal well-formed syllable-to-grid alignment for *il popolo* will be as follows:

(2.4)

```
x  x xx
x  x xx
il popolo
```

Evidence for this particular conception of the grid alignment for syllable-timed languages is presented in Selkirk (in preparation).

So far, then, we have demonstrated some reason for positing two levels of rhythmic organization in speech: the level of basic beats and a lower level of demibeats. It is around these levels that the major part of the discussion concerning isochrony in speech has re-

---

[4] In Selkirk 1978a and Selkirk (in preparation) we have argued that French is not entirely syllable-timed, in that syllables containing schwa are not associated with a beat in the initial alignment of syllables with basic beats.

volved.[5] It is also at these levels that the most copious evidence concerning rhythmic patterning has accumulated, as, for example, in recent work on word stress, where it has been shown that in stress-timed languages the number of weak "stressless" syllables intervening between "stressed" syllables is usually one or two, when there is a pattern at all (see Halle and Vergnaud 1979, Safir 1979, Hayes 1980). In other words, in the patterns of stress that are attested in natural language, it appears that at most two weak demibeats intervene between strong ones. This sort of syllable count gives obvious support to our formulation of the Principle of Rhythmic Alternation, and more generally to the idea that speech is a rhythmically organized activity.

In studies subsequent to Abercrombie 1964 such as Catford 1966 and Halliday 1967a, the rhythmic *foot* has been seen as the central, and in fact unique, unit in the analysis of sentential rhythm.[6] Translated into the terms in which we have been discussing the problem of rhythm, this amounts to saying that only two metrical levels are involved in speech rhythm: syllables are organized into basic beats, and that's it. If this were indeed a correct assessment of the facts, the motivation for invoking the complexly hierarchized representation of rhythm that is embodied in the metrical grid would be weak. However, as we will demonstrate, speech does show a greater hierarchy of rhythmic arrangements. Indeed, the term *foot*, as defined by Abercrombie, is relevant only to these lower levels of rhythmic organization, where the difference between stress-timing and syllable-timing is represented.

What of the *degrees* of rhythmic prominence (often referred to, in the tradition of Trager and Smith 1951, as *degrees of stress*) – that is, the distinctions between strong and weak beats on various levels? It seems that these distinctions have been recognized, but that in the notational representation of rhythm they have been obscured. Catford, for example, allows the sentence *John bought two books last week* to have any of the following arrangements into feet:

(2.5)
| John | bought | two | books | last | week
| John | bought two | books last | week
| John bought | two books | last week
| John bought two | books last week
etc.

But the notion "foot" employed here cannot be the same one defined by Abercrombie, for each of the monosyllabic words in these examples is stressed, and so constitutes a foot on its own. The vertical marks must in fact be taken as indicating strong-weak relations on a level above that of the foot. The point is made clear when multisyllabic words consisting of a stressed syllable followed by unstressed syllables are substituted for the monosyllables in (2.5): for example, *Mary purchased twenty pamphlets yesterday morning*. The intuition is that the same placements of vertical lines are appropriate, that is, that the same rhythmic groupings are possible for the sentence. Yet within the spans of the utterance flanked by the lines still further rhythmic distinctions are made, as shown below, where the italics indicate

---

[5] Most research has concerned itself with the isochrony of stressed (vs. stressless) syllables. See Lehiste 1980 for discussion.

[6] Catford 1966 and Halliday 1967a, for example, posit a hierarchy of units of the phonological organization of the utterance in which the foot has a place: tone group, foot, syllable, phoneme. But since the elements of any given level in the hierarchy are not recursive, this proposal implies that just one level of grouping intervenes between the syllable and the tone group (intonational phrase).

In fact, it is a mistake to view the foot as a constitutive unit of phonological representation arranged hierarchically between syllables and phrases. It can be shown that the latter form part of a hierarchy that is qualitatively distinct from the hierarchy involved in rhythmic organization. On this point, see section 1.2.

local rhythmic prominences (beats):

(2.6)
| Mary | purchased | twenty | pamphlets | yesterday | morning
| Mary | purchased twenty | pamphlets yesterday | morning
| Mary purchased | twenty pamphlets | yesterday morning
| Mary purchased twenty | pamphlets yesterday morning
etc.

The intuition, then, is that rhythmic groupings are made at ore than one level.[7] This, along with other evidence (to be reviewed directly) that there are degrees (or levels) of rhythmic prominence in speech, shows that a hierarchical representation of speech rhythm such as the metrical grid is necessary.

The necessity for distinguishing a minimum of two metrical levels above the basic beat level in the metrical grid with which a sentence is aligned is quite common. Consider the English sentence *Abernàthy gesticulàted*. Here some syllables are aligned with beats (they are marked with accents) and some are not; and among the beat-aligned syllables some are aligned with a strong beat (they bear acute accents). Thus, in this example a strong-weak alternation of beats is clearly perceived. Any representation of speech rhythm requires some means of denoting this strong-weak contrast. The alignment of the words of the sentence with the metrical grid would minimally involve three metrical levels:

(2.7)
```
x             x
x     x       x   x
x  x  x  x  x  x x x x
```
Abernathy gesticulated

The syllables aligned with strong beats here (coinciding with points on the third metrical level) are often referred to as syllables bearing word stress, or more specifically, main word stress. And indeed, in the words of many languages, whether they are stress-timed like English or syllable-timed like Italian, there will be a locus of rhythmic prominence, a beat reliably stronger than the others. The existence of (main) word stress, then, indicates rhythmic organization above the basic beat level.

The strong beat of main word stress does not mark the highest level of rhythmic organization in an English sentence, or in the sentences of many other languages. In the normal English sentence, there exists a strong-weak distinction between beats at more than one level: a beat that is strong on one level may coincide with a beat on a higher level that is either weak or strong. This is the case in the sentence *Abernathy gesticulated*, where the syllable -ti- is the most prominent of all, either in the "neutral" pronunciation of the sentence with a pitch accent only on the verb, or in the "nonneutral" pronunciation in which both words bear pitch accents. The syllable -ti- may be said to bear main sentence or phrase stress. The full alignment of the sentence with the grid would be as follows:

---

[7] Actually, Catford 1966:612 acknowledges the existence of something like Trager-Smith degrees of stress, and claims that the sort of representation he proposes is enough to convey the distinctions, in conjunction with other independently required notions. His proposals concerning degrees of nonnuclear stress are somewhat programmatic, but apparently do not rely on a *rhythmic* characterization of these stresses; hence, they strike us as not making possible a unified treatment of stress and rhythm, as the metrical grid does.

(2.8)

```
                  x
x                 x
x     x           x   x
x  x  x  x  x  x x x x
```

Abernathy gesticulated

Thus the existence of just one degree of sentence or phrase stress alongside main word stress shows that at least two levels of rhythmic organization must be distinguished above the level of basic beats. Again, it is quite common among languages for the main-stressed syllables of the words making up a sentence to differ in their degrees of rhythmic prominence. Partly the existence and location of phrasal rhythmic prominence are to be attributed to the operation of rules of grammar that are sensitive to syntactic structure, such as the Nuclear Stress Rule of English, and partly the appearance of rhythmic prominence on the phrase is to be attributed simply to the demands of rhythmic organization per se, and in particular to the PRA. The different contributions to phrasal prominence will be discussed in section 2.3 and in chapters 3 and 4.

One potentially very telling sort of "nonintuitional" evidence for degrees of rhythmic prominence in speech is hinted at by Pierrehumbert 1980. She shows that the phonetic values for the tones (i.e., pitch accents) composing intonational contours in English are in part a function of the relative prominence of the syllables with which they are associated, where "prominence," for Pierrehumbert, is largely (though not entirely) a matter of stress, i.e., rhythmic prominence.[8] Pierrehumbert offers a convincing case for representing intonational contours as consisting (largely) of a sequence of atomic pitch accents. She argues that these pitch accents, which associate with the main stresses of words, are to be characterized either as one of two single level tones, high (H) and low (L), or as binary combinations of these (see section 5.3). Moreover, she shows that in the same intonational contour, given a pair of pitch accents consisting, say, of one high tone each, the high tone that is associated with the rhythmically most prominent syllable will consistently have the higher frequency (when the effects of declination are factored out, of course).[9] This relation holds not only between a nuclear pitch accent and a prenuclear pitch accent, but also between two prenuclear pitch accents with different prominence values, as shown in figure 2.1. In this example, *In November, the region's weather was unusually dry*, each of the principal words bears a high tone pitch accent. The sentence has a "declarative contour." It corresponds to two intonational phrases, demarcated here with the % symbol. (For the particulars of such contours, see section 5.3.)
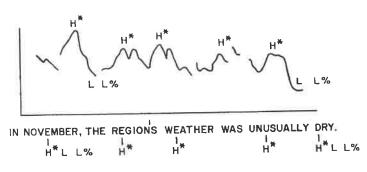

Figure 2.1

---

[8] Assigning a higher F_0 target level within the current pitch range to an H pitch accent on a syllable with higher stress gives very successful results in synthesizing the intonation of neutral declarative contours (Pierrehumbert 1981).

[9] In speaking of relative pitch height here, we assume Pierrehumbert's theory, according to which values for height must be calculated relative to a (declining) F_0 baseline (see Pierrehumbert 1979, 1980).

The English Nuclear Stress Rule determines that *weather* will be more prominent than *region's*, that *dry* will be more prominent than *unusually*, and that the most prominent element of the last phrase (i.e., *dry*) will be more prominent than any other element of the sentence. The alignment of the sentence with the metrical grid would thus be, minimally, as shown in figure 2.2.[10] A comparison of the two figures shows them to be consistent with Pierrehumbert's suggested generalization: that F_0 height is a reflection of relative rhythmic prominence. We will pursue this question of the relation between intonation and rhythm in chapters 4 and 5. For now, we seek only to point out that there exists evidence beyond intuitions about timing that points to the sort of rhythmic organization being claimed to exist.
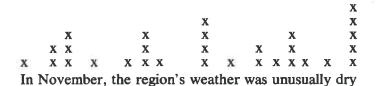
```
                                              X
                              X               X
          X           X       X       X       X
        X X           X       X       X   X   X
  X     X X   X     X X X      X   X   X X X X   X   X
In November, the region's weather was unusually dry
```

Figure 2.2

Liberman 1975 and Liberman and Prince 1977 have offered another important sort of evidence from English for degrees of rhythmic prominence and, specifically, for representing these prominence relations in terms of levels on the metrical grid. They argue that to properly characterize the conditions under which the pervasive phenomenon of *stress shift* (often referred to as the "Rhythm Rule") takes place crucially requires making a distinction in levels of rhythmic organization. Consider, for example, the location of stress prominences in the normal pronunciation of the phrases below:[11]

(2.9)

Dúndèe mármalàde	(the) thírtèenth of Máy
Wéstmìnster Ábbey	(the) únknòwn sóldier
(a) góod-lòoking lífegùard	ánaphòric réference
thírtèen mén	áchromàtic léns

The acute accent indicates "primary word stress" and the grave indicates "secondary word stress"; in the terms used here, strong vs. weak beats.) Of significance is the fact that main stress in the first word is not located where it would be if that word were pronounced in isolation or in other phrasal contexts: *Dùndée, thìrtéen, ànaphóric, more gòodlóoking than you*, etc. The main stress has been "shifted backward" in the examples above. Liberman and Prince's approach to this matter is to say, roughly, that a surface representation like *áchromàtic léns* is derived by a rule of stress shift from an underlying representation in which the first word has its "normal" lexical word stress pattern, e.g., *àchromátic léns*. Liberman and Prince argue that the stress shift rule applies only when two stresses of the *same level* are *adjacent*, in the relevant sense, and hence constitute a "stress clash." The stress clash is eliminated by shifting the first of the clashing stresses backward.
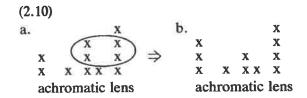
Clearly, strict adjacency of stressed syllables is not required as a precondition for stress shift, for the rule applies in the phrases with *Westminster*, *good-looking*, and *anaphoric*, as it

---

[10] In what follows we will have reason to modify some of the particulars of the metrical grid alignment illustrated in figure 2.2, though the relative prominences, which are at issue here, will remain the same. See chapter 4 on the Nuclear Stress Rule.

[11] D. Jones 1964:§931, Kingdon 1958:164ff., and Gimson 1970:§11.03 discuss this stress shift phenomenon as well.

does in *(the) thirteenth (of May)*, where a stressless syllable intervenes between the two apparently clashing stressed syllables. Thus, as Liberman and Prince point out, the offending adjacency must be represented as occurring at a higher level; it cannot be explained purely in terms of the linear arrangement of syllables. This notion of higher level can of course be characterized, in terms of the metrical grid or in other not strictly linear representations of stress.[12]

Assume that "before" stress shift the metrical grid alignment of *achromatic lens* is as shown in (2.10a) and that stress shift converts it into (2.10b).

(2.10)

```
a.              x          b.                      x
           x    x                 x               x
     x     x    x       ⇒         x       x       x
     x  x  x x  x                 x    x  x x      x
     achromatic lens             achromatic lens
```

(The alignment at the first three levels in (2.10a) is determined by the principles of word stress in English, and the fourth level beat on *lens* corresponds to the effects of the Nuclear Stress Rule.) It is easy to explain in terms of the grid why the concatenation of rhythmic units in (2.10a) is ill formed and therefore subject to modification by stress shift. In general, a grid configuration of the following sort will be judged deviant, regardless of what metrical levels are involved:

(2.11)

```
..........
...X X...
...X X...
..........
```

This is because the grid (i.e., the rhythmic structure) fails to "alternate" in such a configuration, in the sense that two strong beats are not separated by a weak beat. This offending configuration appears in the representation (2.10a) of *achromatic lens*, where it is circled. Thus, though the stressed syllables in (2.10a) are not themselves adjacent, the strong beats with which they are aligned are, in the metrical grid, and this is what constitutes the stress clash. Thus the clash can be defined here only in terms of the second metrical level and above. Supporting this characterization of the stress clash is the fact that stress shift does not take place as readily, or at all, when the main stress is some distance away and thus not involved in a clash. We will see examples of this in section 4.3, where the "Rhythm Rule" of English is discussed in some detail.

As Liberman and Prince point out, of course, stress shift is not the only means available for avoiding a stress clash in the grid. Alternatively, the final syllable of the first word may be lengthened, or a pause may be placed between the words. It may be argued that the possibilities of lengthening and pausing are attributable to the presence of additional "silent" beats, whose appearance is governed by syntactic factors (see section 4.3 and chapter 6) and which themselves prevent stress clash from arising.

What is important to the more general point at hand is the fact that characterizing the ill-formed grid configuration referred to as a "clash" requires reference to two metrical levels. In the particular cases under discussion, it means that reference is made to both the basic beat level and one level above; and this in turn means that stress clash and its consequent stress

---

[12] Proposals have been made for characterizing stress clash in terms of metrical trees, e.g., Kiparsky 1979 and Selkirk 1978c. We will see in chapter 4 why there is nothing to be gained by this approach.

34

shift provide evidence for the existence of at least three levels of rhythmic organization in English words.

In discussing rhythmic organization in speech and in particular the metrical grid, we have suggested that language exhibits a quite general *Principle of Rhythmic Alternation* (PRA), which determines the patterns of alternation appearing in the grid. A provisional formulation of the PRA is as follows:

(2.12)
Between two successive strong beats on a metrical level n there must intervene at least one (and at most two) weak beat(s) of the metrical level n.

Note now that in (2.11) the two points on the lower metrical level are both strong beats, in that they both coincide with beats of the next metrical level up, and that because the two strong beats are adjacent, this configuration does not conform to the PRA. It seems reasonable, then, to view the PRA as in some way responsible for ruling the stress clash in (2.10a) ill formed, and this is in fact the view we will take (see section 2.3). Note that although the PRA itself does not refer to more than one metrical level (the strong and weak beats to which it refers are all of one level), the very notions of "strong beat" and "weak beat" that it invokes require reference to two metrical levels. Thus the characterization of a stress clash as a configuration ruled ill formed by the PRA does involve an appeal to more than one metrical level.

The characterization of stress clash given by Liberman 1975 and Liberman and Prince 1977:312-313 does not invoke a general principle of alternation. Instead the offending ill-formedness is defined directly, as that grid configuration in which two positions on level m are not separated by a position on the next lower level $m-1$. Either approach makes the case for higher levels of rhythmic organization, and the two are equivalent for the cases examined so far. As we will show, however, there is quite independent motivation for the PRA – it apparently operates in numerous other circumstances where the notion of stress clash is simply irrelevant. The value of characterizing the ill-formedness of a stress clash in the way advocated here, then, is that this analysis invokes a principle of some generality in the grammar, not one relevant only to the problem at hand.

Another sort of arhythmicity may arise in the derivation of the phonological representation of a word or sentence, one that, like the stress clash, is often enough done away with in the surface phonological representation of a sentence. This other sort of grid configuration is an overlong sequence of weak beats that is not punctuated by any strong, as depicted in (2.13).

(2.13)

```
. . . . . . . . . . .
.... o o o ....
.... x x x ....
. . . . . . . . . . .
```

(The symbol o indicates the absence of any beat at the point that o occupies in the grid.) Here, too, we find an absence of rhythmic alternation. We might call this a rhythmic lapse, to distinguish it from the clash. It will be desirable to (re)formulate the PRA so that it can rule out both types of nonalternating configuration.

Considerable evidence shows that the rhythmic organization of speech abhors a lapse as much as it does a stress clash. We can see this in the rhythmic organization of phrases. For example, for a normal "neutral" pronunciation of sentence (2.14), where there is a pitch accent only on the final word, which also bears main phrase stress (see sections 4.2.2 and 5.5), there must be a rhythmic prominence before the final main stress.

(2.14)
(I know quite well that) it's organized on the model of a gallon of worms.

The metrical grid alignment (2.15) is avoided in favor of some alternating pattern, be it that of (2.16) or (2.17).[13] (To simplify matters, the lowest metrical level noted here corresponds to primary word stress (i.e., the third metrical level).)

(2.15)

```
                              x
   x              x     x     x
. . . it's organized on the model of a gallon of worms
```

(2.16)

```
        x
x       x
x   x   x   x
```

(2.17)

```
        x
    x   x
x   x   x   x
```

(2.15) is the minimal grid alignment consistent with the Nuclear Stress Rule (NSR) and the assignment of a pitch accent to the final word (see chapter 4). But consistency with the NSR and pitch accent assignment is not enough to make the metrical grid alignment well formed, if the grid does not otherwise exhibit the appropriate alternation in rhythmic organization. Something like the PRA would seem to be at work.

Another set of facts indicating that the rhythmic organization of speech avoids lapses involves secondary word stress. In both stress-timed and syllable-timed languages, the rules of the grammar will define the alignment of syllables with basic beats and often enough will also pick out which of these beats is the most prominent (main-stressed) in the word. For example, the rules for basic beat alignment and main word stress in syllable-timed Italian will give the grid alignment (2.18) for the penultimately stressed word *generativa*:

(2.18)

```
   x
x x x x x
x x x x x
generativa
```

But this alignment, which involves a lapse preceding the main stress, is not actually attested. What is found instead is a secondary stress preceding the main stress at a distance of either one or two syllables (Malagoli 1946, Nespor and Vogel 1979, Chierchia 1982b, Vogel and Scalise 1982), as illustrated in (2.19) and (2.20).

(2.19)   (2.20)

```
      x                  x
x     x              x   x
x x x x x            x x x x x
x x x x x            x x x x x
generativa           generativa
```

---

[13] In discussing the same sentence, Pierrehumbert 1980 allows for an even greater variety of metrical grid arrangements, but these, we submit, would only be available either under conditions of nonneutral intonation, where some pretonic word bore a pitch accent as well, or under conditions where the contributions of syntactic timing introduced lengthening or pausing. We will defend this interpretation of the facts in chapters 4 and 5.

We will also be proposing (section 2.3 and chapter 3) that the very existence of alternation at the lowest (basic beat and demibeat) levels is to be ascribed to the PRA in some way. A sequence of syllables each of which is aligned only with a (weak) demibeat at the first metrical level is but one very long rhythmic lapse and is obviously avoided.

In general, the tendency to avoid lapses, at all levels of rhythmic organization, is overwhelming. We will reformulate the PRA to reflect this. As provisionally stated in (2.12), the PRA rules out any instance of clash, as well as those instances of lapse where strong beats (or demibeats) flank a sequence of more than two weaks. But there are other instances of rhythmic lapse where the overlong sequence of weaks is not flanked on both sides by strongs, e.g., (2.15) and (2.18). If the failure of these configurations to appear in surface phonetic representation is to be attributed to the PRA, then it must be reformulated so as not to require the presence of the flanking strongs. Suppose the PRA specified instead that any weak beat or demibeat may be preceded by at most one other weak. This says, in effect, that there may be at most two weaks in sequence. It has the effect of ruling out overlong weak sequences between strong beats or demibeats, and such sequences not flanked on both sides as well. Let us call this the *anti-lapse provision* of the PRA.

We might now consider reformulating the *anti-clash provision* of the PRA, which currently states that at least one weak beat or demibeat must intervene between two strongs. In fact, one alternative to this appeal to the flanking strongs appears to make certain correct predictions about the facts that the earlier formulation leaves unexplained. This alternative consists in specifying simply that at least one weak beat (or demibeat) must follow a strong.[14] Like the first, this formulation ensures that there will be at least one weak between two strongs; unlike the first, however, it ensures in addition that a weak beat (or demibeat) will always follow a final strong beat. (The notion "final" will be defined with respect to some domain – word, phrase, etc. – and thus is not limited to sentence-final position.) For the time being, then, let us entertain the formulation in (2.21) as an alternative to (2.12):

(2.21)
The Principle of Rhythmic Alternation

a.   Every strong position on a metrical level n should be followed by at least one weak position on that level.

b.   Any weak position on a metrical level n may be preceded by at most one weak position on that level.

The consequences of formulating the principle in this way will be examined more fully in the following chapters.

To sum up, then, there appears to be considerable evidence for the rhythmic organization of speech, and for representing that rhythmic organization as an alignment of the syllables of a sentence with positions in a metrical grid. Like a musical score, the alignment of a sentence with a grid represents the (ideal) isochrony of the pulses of speech, the relative durations of individual syllables, and their degrees of relative prominence. Moreover, the grid permits an understanding of the alternations in the rhythmic realizations of words in the sentence. This is an extremely important point. The metrical grid allows the generalization to be expressed that the same rhythmic ideal, the Principle of Rhythmic Alternation, governs (in a way yet to be defined) the patterns to be found *at all levels*. The PRA, formulated with respect to levels of the metrical grid, "expresses" the propensity to alternation at the lower levels in the same terms as the propensity to avoid clash, or to introduce alternation, at the higher levels. Thus, viewing stress patterns in terms of the metrical grid rationalizes the properties of stress sys-

---

[14] Logically speaking, another alternative would be to specify that at least one weak *precede* every strong. We have no overriding reason for choosing one alternative over the other here.

tems even further than the metrical tree theory of stress patterns, for it rationalizes the pattern question itself (see sections 1.2.2 and 4.2). It is for this reason that we view the metrical grid as crucial to a theory of stress patterns and thereby more than worthy of supplanting the now otiose trees in this function.

2.3 Building the Grid

2.3.1 The Framework

We are interested here in developing a theory of patterns of rhythmic prominence in language, a theory of the notion "possible metrical grid alignment." We believe that that theory must incorporate universals of rhythmic organization, and must delineate the range of choices that are available to grammars of individual languages for the codification of particular rhythmic patterns. What we are aiming for is a core theory of rhythmic patterns (see Chomsky 1981 and references cited therein). Given such a core theory, the grammar of an individual language will specify not language-particular rules, but which among the (universally defined) rules made available by the theory are actually at play in a particular language. Following Chomsky, we will call this a language-particular specification of the *parameters* delineated by the theory.

In recent years considerable research has been done in metrical phonology with a view toward developing a core theory of word stress patterns, based on a metrical (prosodic) tree representation of prominence relations. This effort has been advanced most notably by the work of Halle and Vergnaud 1979 and developed more recently in Hayes 1980. As will become clear, many of the fundamental insights into the parameters of stress theory gained in these works, and others, are readily and perspicuously characterized in a metrical grid framework. With Prince 1981, 1983, we argue that a theory of stress patterns is better expressed when the metrical grid is assumed to be the basic representation of stress. Many of the basic lines of thinking pursued here find their inspiration in Prince's work, though the particular articulations of the ideas are different in certain respects, having been developed independently.

As stated earlier, we propose that the relation between the text (a syntactic representation) and its metrical grid alignment is to be expressed as a set of rules that "construct" a metrical grid alignment for the text according to the principle of the cycle. On this theory, these rules progressively build up the metrical grid alignment of a sentence (a rhythmic "score"), from the lower levels to the higher, on successively larger cyclic domains. An alternative theory is entirely conceivable, according to which the text is in one way or another "matched up" with a full-fledged grid somehow independently defined, a grid that may undergo modification once aligned with the text. This is the approach implicit in Liberman 1975, Liberman and Prince 1977, and Dell (to appear). There, the text, which includes the abstract stress pattern of the sentence, is matched up with a "preexistent" grid, without appeal to the cycle. Let us consider this alternative theory, stripped of its assumption that the text includes metrical tree patterns of stress.

Two issues must be distinguished here: whether the text-grid relation is established by "construction" or "matching," and whether the relation is established cyclically or not. We submit that the principle of the cycle is essential to the proper characterization of this relation, and we will argue for this position in several ways in this chapter and in chapters 3 and 7. The fact that the text-grid relation must be cyclically defined removes any possibility of entertaining the matching theory, we believe. Fragments of a metrical grid, corresponding to the lowest cyclic domains, could perhaps be considered to be preexistent and matched with the text. But there could no longer be a preexistent metrical grid of the entire sentence, in any interesting sense, given that the grid matched on lower cycles would be modified cyclically. And to