• 2 •

Diagnosing Stress Patterns

This chapter addresses a question that seems logically prior to any other in developing a theory of stress, namely: When we talk about stress, what is the physical phenomenon we are talking about? The definition of stress is one of the perennially debated and unsolved problems of phonetics. A body of careful experimental work has established that no one physical correlate can serve as a direct reflection of linguistic stress levels; see Lehiste 1970; Lea 1977; Berinstein 1979; Ladd 1980; Beckman 1986; and references cited there.

The absence of a clear physical definition of stress means that any theory of stress is in an indirect relation with the facts that support it. In most areas of phonology, it is not too difficult to ascertain when the observed facts confirm or falsify a hypothesis. This is because the distinctive features which form the core of phonological representations have relatively clear acoustic or articulatory correlates.

In the case of stress, however, even the facts are often in doubt. For example, mutually inconsistent systems exist for recording English stress contours (compare, for instance, Jones 1956 with Kenyon and Knott 1944). Our inquiry would obviously be served by the discovery of a clear and unambiguous phonetic correlate of stress. But the phonetic literature suggests that no such thing exists.

In this chapter, I first review the evidence against an invariant phonetic correlate for stress. Second, I try to show that the absence of such a correlate does not exclude the possibility of defining and investigating stress in a rigorous way.

2.1 THE PHONETIC CORRELATES OF STRESS

The nicest available definition of stress, if only it were true, would be the proposal made by Stetson 1928. Stetson believed that the prosodic organization of speech is reflected directly in the actions of the muscles that control expiration. Roughly, Stetson claimed that every syllable in an utterance is accompanied by a **breath pulse:** a contraction of the muscles of the chest wall, which gives the syllable peak its increased sonority over the segments of the syllable margin.

Further, Stetson believed that *stressed* syllables carry an extra-strong breath pulse, executed by abdominal rather than chest-wall muscles.

The work of Ladefoged (1967, and work cited there) refuted many of Stetson's claims. In particular, individual syllables do not have separate breath pulses, and the division of labor among the expiratory muscles is not governed phonologically as Stetson thought. However, Ladefoged did endorse Stetson's claim that stress can always be diagnosed by examining expiratory effort.

Other work, however (Lieberman et al. 1967; Lieberman 1968; van Katwijk 1974; J. Ohala 1977), has strongly argued against the Stetson/Ladefoged view of stress. In the observations made by these researchers, speakers produced "breath pulses" for a *subset* of stressed syllables (especially syllables receiving emphatic stress), but, crucially, not for all of them. That is, one often finds syllables that by other criteria must clearly be counted as stressed which nevertheless are not accompanied by a breath pulse.

Peterson 1958, 402–4, describes a paralyzed patient who had normal control of the larynx but who breathed entirely under the control of a respirator. Although this patient was "unable to produce strongly stressed speech," she could nevertheless speak in a way that "closely resembles normal conversational speech in all aspects." It appears that breath pulses (which this patient could not produce) are not obligatorily present in normal speech.

The research above can be summed up as follows: if a breath pulse is present, it is probable that stress is present, but there are many stressed syllables produced without a breath pulse.

Work in the perceptual domain has further advanced our understanding of the phonetic correlates for stress. Fry 1955, 1958 developed an experimental paradigm that used synthetic speech stimuli. The idea was to vary certain physical parameters systematically, keeping segmental content the same, and let listeners judge if the stress contour was affected. Schematic stimuli are as in (1):

(1) a.	**Loudness**	*PERmit*	vs.	*perMIT*
b.	**Duration**	[pr̩ːmɪt]	vs.	[pr̩mɪːt]

c.	**Pitch**	*per*
		    mit
			vs.	*m*
				         i
				per *t*

d.	**Listener's judgment**	*pérmit*	or	*permít*?

Fry found that loudness had the least effect on stress perception, despite its intuitive status as the most natural correlate of stress. Duration changes had a greater effect, with longer syllables more likely to be perceived as stressed. The strongest effects on stress perception were achieved by altering the pitch contours, as shown. Thus pitch and duration, rather than loudness, seem to be the principal perceptual cues for stress. Later work, such as Bolinger 1958; Morton and Jassem 1965; and Nakatani and Aston 1978, supported Fry's findings. Loudness does have some effect, since if one integrates loudness over the duration of the syllable, this yields a measure that correlates with stress judgments slightly better than pitch does (see Beckman and Pierrehumbert 1986, 272, and references cited there).

The multiple phonetic cues for stress, and the subordinate role of loudness, are particularly interesting when one considers that languages use duration and pitch in their phonological systems for entirely different purposes. Duration is the phonetic cue for vowel length, which is phonemic in many languages. Duration is also widely used to mark phonological phrasing: the right edges of major phrases typically receive extra duration (Klatt 1975; Wightman et al. 1992). Further, pitch is the phonetic cue for tone, in languages with phonemic tone systems, and also is the phonetic basis of intonation. The basic point is this: aside from the marginal role of loudness, stress is **parasitic,** in the sense that it invokes phonetic resources that serve other phonological ends.

Partly as a consequence of this, stress is phonetically realized on a language-specific basis. For example, languages with phonemic vowel length contrasts have been shown to avoid using duration as a correlate for stress; see Berinstein 1979. This makes sense, since using duration to mark stress in these languages would obscure the phonemic vowel length contrast.

A more dramatic example of this type can be found in Finnish. Carlson 1978 points out that Finnish emphatic stress can involve lengthening of unstressed, rather than stressed, syllables. This occurs when the emphasized word begins with a short-voweled open stressed syllable. For example, an emphatic utterance of the word *vítut* would involve a lengthened (but unstressed) second syllable: [vítuːt]. For discussion of the motivation for this pattern, see Carlson 1978 and Prince 1980.

The use of pitch to realize stress is likewise dependent on the character of a language's intonational system. For example, Chung 1983, 38, notes that in Chamorro, the standard declarative intonational contour places the lowest rather than the highest pitch on what is perceived to be the stressed syllable of a word. In English, it is also clear that high pitch does not invariantly signal stress, since in the intonational pattern for yes/no questions the stressed syllable receives the lowest pitch:

(2) *Pennsy l vá n i a ?*

It will not do either to say that stress is signaled by a "pitch excursion" or some similar notion. English has an intonational contour called "scooped" by Ladd 1980, in which the main pitch excursion comes one or two syllables after the stressed syllable:

(3) a. *Ídon't know*

b. *But Mánitowoc has a library!*

In fact, it is possible to construct examples from which pitch information is excluded, with the stress pattern emerging intact. Huss 1975 compared contrasting forms like *tórment* ~ *tormént,* placed after the main stress of a declarative utterance, so that the pitch contours were neutralized to a uniform low.

(4) a. *We FIRST tormént, he said.* b. *His FIRST tórment, he said.*

The stress distinction in pairs like *tórment* ~ *tormént* was systematically reflected in relatively greater duration assigned to the syllable bearing the word stress. Listeners showed only a marginal ability to recover the intended stress contour out of context, but the fact that the phonetic outputs were different shows that the speaker's own internalized representation does include a stress distinction here.

To sum up the point of these examples: it is certainly true that pitch and duration are both intimately linked with stress. However, because the relation between stress and pitch/duration is both indirect and language-specific, it is impossible to "read off" stress contours from the phonetic record.

At this point, the situation may seem rather unpleasant: How can we investigate a phenomenon that has no consistent physical correlates? There is a reasonable answer to this, which has tacitly or overtly governed the practice of many phonologists. I present this answer in two parts, providing first a definition of stress, then discussing how it can be investigated rigorously.

2.2 A DEFINITION OF STRESS

The central claim of metrical stress theory, argued in Liberman 1975 and Liberman and Prince 1977, is that stress is the linguistic manifestation of rhythmic structure. That is, in stress languages, every utterance has a rhythmic structure which serves as an *organizing framework* for that utterance's phonological and phonetic realization. One reason for supposing that stress is linguistic rhythm is that stress patterns exhibit substantial formal parallels with extra-linguistic rhythmic structures, such as those found in music and verse; see § 3.1–2, § 4.5.

If the equation of stress and rhythmic structure is valid, then we automatically account for why there is no invariant physical realization for stress. The reason is that rhythm in general is not tied to any particular physical realization;

one can detect and recognize the same rhythm irrespective of whether it is realized by (for example) drumbeats, musical notes, or speech. Because of this independence, we are not bound to the prediction that any particular phonetic correlate will invariably realize stress in any particular language.

Naturally, certain phonetic correlates serve more readily as cues for stress than others: it is only natural that strong rhythmic beats should coincide with breath pulses, with greater duration, and with raising of pitch. But these are only tendencies, and since phonology serves many ends other than rhythmic ones, they sometimes override the natural correlation between strong rhythmic beats and particular phonetic phenomena. The prediction corresponds to what we observe: there are general tendencies in how stress is manifested across languages, but nothing is iron-clad.

This conception gives rise to a practical difficulty: given that rhythm is an abstract notion that cannot be directly observed, how can we give our observations about stress a solid empirical basis? Chomsky and Halle (1968, hereafter *SPE,* 24–26) were aware of this problem, and took an approach fairly similar to that taken in syntax: if a large group of native speakers who understand the task can agree on a linguistic observation (such as a grammaticality judgment, or the presence of a particular ambiguity, or a stress contour), then we are justified in taking that agreement as a datum. This seems reasonable, and my own judgments about English stress are in broad agreement with the stress contours claimed in *SPE.* However, not everyone's judgments are, and *SPE* was criticized in the literature (e.g. Lehiste 1970; Stockwell 1972; Vanderslice and Ladefoged 1972) for its view of the data.

Part of the difficulty is that prosodic phenomena are among the least accessible to consciousness. For example, teachers of beginning phonetics often encounter students who, although native speakers of English, simply cannot hear where the main stress of an English word falls.

2.3 AN EMPIRICAL BASE FOR THE PHONOLOGY OF STRESS

The difficulties that arise from using intuitive judgments of stress are to some extent avoidable. By carefully examining and comparing various phonological diagnostics for stress, it is possible to study the stress system, at least in English, with little recourse to intuition. The basis for this is the fact that the stress phonology of English forms a tightly organized system: segmental rules, intonational patterns, and phonotactic constraints agree with each other in diagnosing a particular stress pattern. The underlying stress pattern is most clearly demonstrable in that it unifies in a coherent way a broad set of phenomena. The following discussion is intended to support this claim.

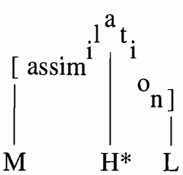
Before proceeding, a note on the data: in a number of areas to be discussed, there is variation across dialects of English. We will be considering patterns that are common in American English. Dialectal differences will affect individual arguments, but the basic points could probably be made in other dialects as well.

2.3.1. Diagnostic I: Attraction of Nuclear Intonational Tones

Most generative work on English stress has maintained that every intonational phrase (roughly speaking, every phrase of approximately sentence length; Selkirk 1984; Nespor and Vogel 1986) has one and only one primary stress. The best evidence for this comes from the English intonational system. Research on intonation, e.g. Liberman 1975 and Pierrehumbert 1980, indicates that the English intonational system can be analyzed in part as an inventory of **tunes.** These are abstract tonal sequences, each paired with a somewhat elusive meaning, which may be assigned to any given text. I present an encapsulated and simplified version of four of these tunes; see Pierrehumbert 1980 for a more serious account.

The **declarative** tune consists of the tonal sequence M H* L (Mid–High–Low), where H* is linked to a specified syllable and M and L are linked to the initial and final edges. Aligning this declarative tune with the word *assimilation* gives (5) (for association of tones to edges, see Liberman 1975; Pierrehumbert and Beckman 1988; Hayes and Lahiri 1991a):

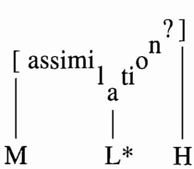
(5)



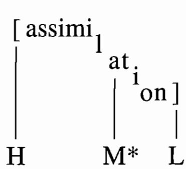
Note that between the phonological tones, pitch is in continuous movement up or down. This is analyzed by Pierrehumbert as interpolation, carried out within the phonetic component; phonologically, such sequences are toneless.

The same text as in (5) can be aligned with different tunes. For example, the **question** tune (6a) has the tonal sequence M L* H; it is used for yes–no questions and other purposes. The **downstepping** tune (6b) is H M* L; it implies that what is being said is in some sense predictable from context (Pierrehumbert and Hirschberg 1990).

(6) a.



b.



The **delayed rise** or **scooped** tune begins low, places a L*+H (rising) sequence starting on a particular syllable, and ends low. (For its meaning, see Ward and Hirschberg 1985; Pierrehumbert and Hirschberg 1990.)

```
                   i
              t  o
(7)  [assimila      n !]
      |         |     |
      L        L*+H   L
```

What is of interest here is the way in which the various tunes line up against different texts. Consider the three examples in (8), each lined up with all four tunes:

```
(8) a. [ assimilation    ]    b. [ preliminary ]    c. [ pontoon      ]
         |        |      |        |     |       |        |     |        |
         M        H*     L        M     H*      L        M     H*       L
         M        L*     H        M     L*      H        M     L*       H
         H        M*     L        H     M*      L        H     M*       L
         L        L*+H   L        L     L*+H    L        L     L*+H     L
```

The point is that the starred tone of each tune consistently docks to the same syllable in each word, namely the one marked in dictionaries (and agreed upon by most native speakers) as the main stressed syllable: *assimilátion, prelími-nary, pontóon.* Thus English intonation can be used as a diagnostic for the position of main stress.

These examples illustrate the view adopted here concerning pitch and stress (cf. Ladd 1980, chaps. 1–2): pitch is directly determined by the intonational system, but the rules linking tones to texts refer to the position of stress. As a result, pitch can serve as a powerful phonetic cue for stress location. However, in locations where the intonational system places no tones, pitch cannot serve as a cue for stress, and other cues such as duration take over; cf. (4) above, as well as Nakatani and Aston 1978.

The diagnostic of intonation can be used to bring evidence to bear on stress patterns that are not so clear as main word stress. For example, there is disagreement in the literature about whether simple two-word phrases, such as *tall trees* or *divine right,* bear a weak stress followed by a strong one, or two equal stresses. The intonational evidence argues for the first interpretation, because in a neutral reading the starred tone of an intonation contour must fall on the second word, as in (9a).

```
                       i                          i
(9) a.               r | g          b.        .v | n
                   e   |                     i   |   e
        [ divin        |  h              [ d     |
         |             |    t]            |      |     r
         |             |     |            |      |      ight ]
         |             |     |            |      |          |
         M             H*    L            M      H*         L
```

The alignment in (9b) would be well-formed only in a context in which it was being emphasized just what sort of right was at issue; that is, in a context with contrastive stress on *divine.*

2.3.2 Diagnostic II: Vowel Quality and Segmental Rules

It is a fairly uncontroversial assumption that a syllable of English is completely stressless if its vowel is schwa; examples are the boldfaced vowels in *about* [əbáwt], *comet* [kámət], *medicine* [mɛ́dəsən], *connect* [kənɛ́kt], and *August* [ɔ́gəst]. By schwa is meant not the mid back unrounded vowel of *cup* [kʌ́p], but rather a reduced vowel, which is shorter, higher, and perceptually less distinct than [ʌ]. Some dialects have two reduced vowel phonemes, /ɨ/ and /ə/; as in *American* [əmɛ́rɨkən]. I assume that both of these vowels are stressless.

If schwa is always stressless, we would expect that it would never bear main stress. The evidence of Diagnostic I from the previous section bears this out: a schwa is never assigned the starred tone of an intonational contour. If one constructs examples in which it is, they sound odd because of the phonological contradiction:

```
(10) a. *medícine   [ mɛd ə sən ]        b. *comét   [ kam ət  ]
                      |    |   |                       |     |  |
                      M    H*  L                       M     H* L
                      M    L*  H                       M     L* H

                      etc.                             etc.
```

Thus there is reason to believe that all schwas are stressless vowels. However, it can be argued that not all stressless vowels are schwa. Stress and vowel quality in English are in a fairly intricate relationship, carefully analyzed in *SPE* and earlier work. I will not go into the *SPE* analysis but will only try to establish the existence of stressless vowels other than schwa on a rigorous basis.

The best way to do this is to use the segmental phonological rules of English as diagnostics for stress. Numerous rules of English refer to the distinction between completely stressless vowels and other vowels and thus can serve as diagnostics. I will describe four such rules here, drawing on Kahn 1976 and others. The rules are formulated in purely descriptive terms; more insightful versions would rely on syllable structure (Kahn 1976; Selkirk 1982) or on foot structure (Kiparsky 1979; Hayes 1982a).

(a) The phonemes /t, d/ may be realized as flap [ɾ] in certain environments. Word-internally, they become flap only when preceded by a vowel or glide and followed by a completely stressless vowel. I use /t/ examples here, since flapping is easier to hear for them.

(11) a. **Flapping** $t, d \rightarrow ɾ \; / \; [-\text{consonantal}] \_\_\_\_ \begin{bmatrix} \text{V} \\ -\text{stress} \end{bmatrix}$

b. *data* [déyɾə] vs. *attain* [ətéyn]

(b) In many dialects the sequence /ns/ optionally receives a brief transitional epenthetic stop when a stressless vowel follows. I refer to this rule as **/t/ Insertion,** without intending to claim that a full /t/ is actually inserted (see Hayes 1986b; Clements 1987 for autosegmental analyses).

(12) a. **/t/ Insertion** $\emptyset \rightarrow \breve{t} \;/\; n \_\_\_\_ s \begin{bmatrix} \text{V} \\ -\text{stress} \end{bmatrix}$

b. *Mensa* [mɛ́**nt͡sə**] vs. *insane* [ɪ**ns**éyn]

(c) The phoneme /l/ optionally becomes voiceless through most of its duration when it occurs between /s/ and a stressless vowel (Kiparsky 1979):

(13) a. **/l/ Devoicing** $l \rightarrow [-\text{voice}] \;/\; s \_\_\_\_ \begin{bmatrix} \text{V} \\ -\text{stress} \end{bmatrix}$

b. *Iceland* [áys**l̥**ənd] vs. *Icelandic* [àys**l**ǽndɪk]

There is some devoicing in other contexts as well, but it is not as severe as in the context of (13).

(d) Word-medial voiceless stops are aspirated provided they are in the onset of a stressed syllable and are not preceded by a strident:

(14) a. **Medial Aspiration**

$$\begin{bmatrix} -\text{son} \\ -\text{cont} \\ -\text{voice} \end{bmatrix} \rightarrow \begin{bmatrix} +\text{spread} \\ \text{glottis} \end{bmatrix} \;/\; [-\text{strid}] \_\_\_ ([+\text{son}]) \begin{bmatrix} \text{V} \\ +\text{stress} \end{bmatrix}$$

b. *append* [ə**pʰ**ɛ́nd] vs. *campus* [kʰǽm**pə**s]

accost [ə**kʰ**ɔ́st] vs. *chicken* [čʰɪ́**kə**n]

Other rules could be added to this list. In general, the consonant allophones found before schwa in English differ from those found before stressed vowels.

Consider now the vowels that are not schwa, but do not bear main stress either. The chart in (15) illustrates the behavior of some of these vowels with respect to the four rules above.

(15)	/ey/	/ɛ/	/ay/	/a/
Flapping	*imitate* [ɪ́mə**tʰ**eyt], *[ɪ́mə**ɾ**eyt]	*protest*$_{\text{Noun}}$ [prów**tʰ**ɛst], *[prów**ɾ**ɛst]	*maritime* [mǽrə**tʰ**aym], *[mǽrə**ɾ**aym]	*proton* [prów**tʰ**an], *[prów**ɾ**an]
/t/ Insertion	*compensate* [kʰámpə**ns**eyt], *[kʰámpə**nt͡s**eyt]	*incest* [ɪ́**ns**ɛst], *[ɪ́**nt͡s**ɛst]	*insight* [ɪ́**ns**ayt], *[ɪ́**nt͡s**ayt]	*consolidation* [kə**ns**alədéyšən], *[kə**nt͡s**alədéyšən]
/l/ Devoicing	*legislate* [lɛ́ǰə**sl**eyt], *[lɛ́ǰə**sl̥**eyt]	(no cases found)	(no cases found)	*Islam* [ɪ́**sl**am], *[ɪ́**sl̥**am]
Medial Aspiration	*octane* [ák**tʰ**eyn], *[ák**t**eyn]	*appendicitis* [ə**pʰ**ɛndəsáyɾəs], *[ə**p**ɛndəsáyɾəs]	*Malachi* [mǽlə**kʰ**ay], *[mǽlə**k**ay]	*coupon* [kʰú**pʰ**an], *[kʰú**p**an]

In these data, /ey/, /ɛ/, /ay/, and /a/ that lack main stress behave just like main-stressed vowels. For example, the word *imitate* shows that a /ey/ that lacks main stress blocks Flapping of a preceding /t/, just like the main-stressed /ey/ of *attáin.*

These observations lead to the following tentative conclusion: English has at least three degrees of stress. Complete stresslessness is diagnosed by the presence of schwa or by segmental rules including (11)–(14). Main stress is diagnosed by its attraction of the starred tone of the intonational contours in (8). Finally, secondary stress is diagnosed by having none of these characteristics. Henceforth secondary stress is marked with / ̀/, as in *ímitàte* or *appèndicítis.* Note that by "degrees of stress" I do not mean three absolute levels, but simply that syllable X has more stress than syllable Y, which in turn has more than syllable Z. I continue this use of the term "degree" below.

An alternative to positing three degrees of stress is to suppose that the segmental rules of (11)–(14) are triggered by the actual vowel [ə], rather than by stress. For example, Flapping would look like (16):

(16) **Flapping** *t, d* → *ɾ* / [−consonantal] ____ ə

This position is untenable, however, because schwa is not the only stressless vowel of English. A subset of the English vowels are able to appear in both main-stressed and completely stressless positions, as we have diagnosed them with intonational and segmental evidence. Although there is dialect variation, the vowels in (17) typically show this behavior:

(17) **Vowel**		**Main Stressed**		**Stressless**	
/r̩/		*burr*	[bŕ̩]	(a) *butter*	[bʌ́**ɾ**r̩]
		refer	[rəfŕ̩]	(b) *cancer*	[kʰǽ**nt͡s**r̩]
				(c) *wrestler*	[rɛ́**sl̥**r̩]
				(d) *upper*	[ʌ́**p**r̩]
/i/	in word-final position	*bee*	[bí]	(a) *pity*	[pí**ɾ**i]
				(b) *fancy*	[fǽ**nt͡s**i]
				(c) *Leslie*	[lɛ́**sl̥**i]
				(d) *hockey*	[há**k**i]
	before a vowel	*Ian*	[íən]	(a) *Whittier*	[hwí**ɾ**ir̩]
		museum	[myuzíəm]	(b) *fanciest*	[fǽ**nt͡s**iəst]
				(c) *Wes-leyan*	[wɛ́**sl̥**iən]
				(d) *Gropius*	[gró**wp**iəs]
/ow/	in word-final position	*go*	[gów]	(a) *motto*	[má**ɾ**ow]
		Trudeau	[trudów]	(b) *Rinso*	[rí**nt͡s**ow]
				(c) *Oslo*	[á**sl̥**ow]
				(d) *Harpo*	[hár**p**ow]

	before a vowel	*boa*	[bówə]	(a)	*Ottawa*	[áɾowə]
				(b)	—	
				(c)	—	
				(d)	—	
/ɪ/	before /ŋ/	*ring*	[ríŋ]	(a)	*Keating*	[kʰíɾɪŋ]
				(b)	*tensing*	[tʰɛ́**n͡ts**ɪŋ]
				(c)	*whistling*	[hwísl̥ɪŋ]
				(d)	*hoping*	[hó**wp**ɪŋ]

In the right column of (17) are examples showing that these vowels can behave just like schwa in triggering the four phonological rules noted above under (11)–(14). For example, syllabic [r̩] behaves just like schwa in triggering Flapping of a preceding /t/, as shown by *butter* [bʌ́ɾr̩]. Using such data, one can show that in a full taxonomy of English vowels we must distinguish three categories: always stressed (e.g., (15)); variably stressed or stressless (17), and never stressed (schwa). Since some vowels can be either stressed or stressless, and it is the stress that determines the outcome, phonological rules like (11)–(14) cannot refer to schwa vowel quality rather than stress.

To complete the picture, here is the full taxonomy of vowels as found in my own speech. Naturally, other dialects will differ (sometimes greatly) in their vowel inventories and the ways in which vowel quality can be used to diagnose stress.

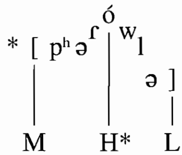
(18) a. **Never Stressed:** [ə, n̩, m̩]

b. **Variable:**
[r̩]
[l̩] (*pull* [pʰĺ̩] vs. *apple* [ǽ**pl̩**])
[i, ow] / ____ #, / ____ V, and in prefixes
(*comprehend* = [kʰàm**pr**ihɛ́nd, kʰàm**pr**əhɛ́nd])
[ɪ] / ____ ŋ
[yu] ~ [yə] (*occupy* = [á**ky**upʰày], [á**ky**əpʰày])

c. **Always Stressed:** [ey, ɛ, æ, a, ɔ, ʌ, ʊ, u], and [i, ow, ɪ] when not in the contexts of (b)

I conclude from this that a distinction of at least three degrees of stress in English is well motivated. Note that the most important aspect of the evidence is its highly systematic nature; a large number of independent criteria agree with each other. For example, the fact that the agentive suffix *-er,* with syllabic [r̩], fails to trigger Medial Aspiration in *hiker* [háy**kr̩**] shows that it is stressless, which in turn explains why the same suffix triggers Flapping in *fighter* [fáyɾr̩] and /l/ Devoicing in *wrestler* [rɛ́sl̥r̩]. The fact that Flapping cannot apply before stressed syllables in the same word, together with the restriction of starred tones in intonational contours, means that a main intonation peak can never occur on a medial syllable beginning in a flap, as in the hypothetical form (19).

(19) *petola*



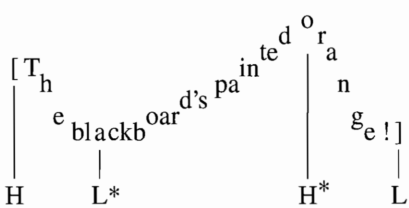
With patience, one could make a very large number of empirical connections of this sort. It is these connections that demonstrate most clearly that the stress system constitutes a unifying framework for English phonology.

2.3.3 Diagnostic III: Non-Nuclear Intonational Tones

The discussion so far has motivated three degrees of stress in English: stressless, secondary stress, and main stress. In fact, it can be argued that additional intermediate degrees exist. To do this, it will be necessary to examine more diagnostics.

All the English intonational contours diagnose the place of main stress. In addition, some of them diagnose weaker stresses as well. As an example, consider the "surprise-redundancy" contour (Sag and Liberman 1975; Liberman 1975; Pierrehumbert 1980), illustrated in (20):

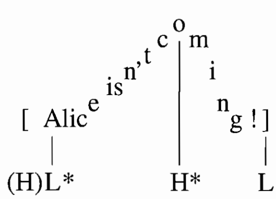
(20)



Roughly, this contour is used to express surprise or the view that one's interlocutor should already know what one is saying. (Readers not experienced in performing out-of-context intonation might try placing the examples in the context "Well, I'll be damned," _____, as a reviewer has aptly suggested.)

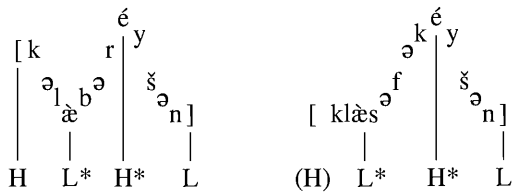
Note that in texts with initial stress, there is no room available for the initial H tone, and it goes unrealized:

(21)



For present purposes, the crucial question is how the tune is aligned with a given text. It is straightforward to establish that the H* tone is aligned with the strongest stress. More interesting is how the L* is aligned: it is associated with the strongest stress to the left of the main stress. This can be argued for in a number of ways. First, consider utterances that contain a single long word, in which only one syllable to the left of the main stress has a non-schwa vowel:

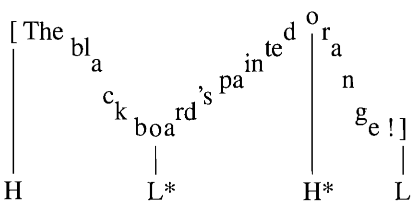
(22) a. *collaboration!* b. *classification!*



Apparently, the L* seeks out the vowel with secondary stress in preference to stressless vowels. Each word of (22) sounds odd if pronounced with the contour of the other.

Another argument that the L* tone seeks out the strongest stress to the left of the main stress is as follows. Note that the compound *blackboard* bears initial stress and that the phrase *black board* bears final stress. These statements can be justified straightforwardly by Diagnostic I (§ 2.3.1). If we place these words in an utterance to the left of the main stress, then we get a corresponding difference in the placement of the L* tone. Compare (20) with (23):

(23)

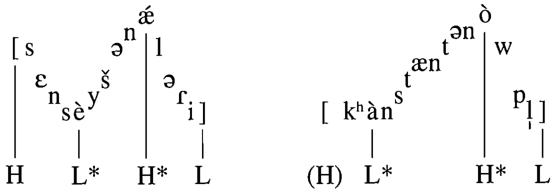


Under the plausible assumption that the stress contrast between *bláckboard* and *black bóard* persists even when neither one bears the main stress, the intonational difference follows, since the L* tone seeks out the strongest stress available to it.

At this point I have at least indirectly motivated four degrees of stress. In (23), the main stress is the first syllable of *orange;* it is stronger than the stress on *board* (Diagnostic I). *Board* bears stronger stress than *black,* since it attracts the L* tone (Diagnostic III). Finally, *black* must bear a stronger stress than the preceding word *the,* since by Diagnostic II syllables containing /æ/ have a higher degree of stress than syllables containing /ə/.

At least in some dialects, four degrees of stress can be observed in a single word. Consider the contrast in (24):

(24) a. *sensationality!* b. *Constantinople!*

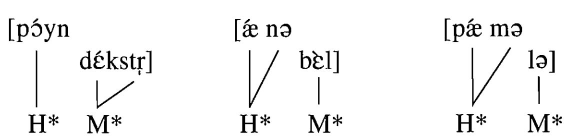


These two words begin with two weak stresses (the vowels /ɛ, ey, a, æ/ always bear some stress), followed by a stressless syllable and the main stress. Yet for many speakers the alignment with the surprise-redundancy contour differs, with the L* falling on the first syllable in *sensationality* but on the second syllable in *Constantinople* (for why this is so, see Kiparsky 1979). By the same reasoning as above, one can argue that these words show four levels of stress.

Not all dialects of English have this contrast; I have found speakers who level out the distinction in the direction of (24b). However, even these speakers arguably have at least four levels in multi-word utterances.

Another English intonation contour that can diagnose secondary stresses is the so-called "chanted vocative," which differs from unchanted contours in having sequences of level pitch, assigned by tonal spreading. Here, a H* tone falls on the main stress, and a M* tone on the strongest stress that follows the H*. This results in a different alignment when the position of the secondary stress varies, as (25a, b) show. Where there is no stress after the main stress, M* docks onto the final syllable by default, as in (25c).

(25) a. Poindexter! b. Annabel! c. Pamela!



For discussion and analysis, see Liberman 1975; Ladd 1978; and Hayes and Lahiri 1991b. A similar Dutch case is analyzed in Gussenhoven 1993.

2.3.4 Diagnostic IV: The Rhythm Rule

Consider now one final diagnostic for stress in English. Like many languages, English has a rule that shifts a stress leftward when a stronger stress follows, creating alternations like *thirtéen* ~ *thìrteen mén* (rather than *thirtèen mén*). Informally, the rule can be stated as in (26):

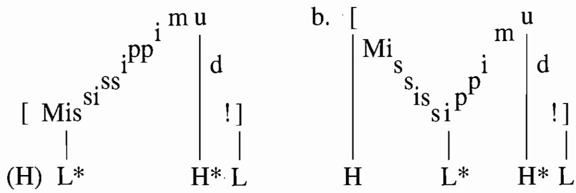
(26) **Rhythm Rule** 3 . . . 2 . . . 1 → 2 . . . 3 . . . 1

where 1, 2, and 3 indicate the relative rankings of the stresses of a phrase. For references and full discussion of this rule, see § 3.4.2, § 9.5.

The existence of the Rhythm Rule is easily motivated by the diagnostics already presented. The various stress contrasts can be justified as follows. The word *Mississíppi* has penultimate main stress, as can be shown by Diagnostic I. From the evidence of vowel quality (Diagnostic II), the first syllable of this word bears secondary stress: [mìsəsípi]. Suppose we embed *Mississippi* in the phrase *Mississippi múd.* By Diagnostic I, the main stress falls on *mud.* We can now inquire where the second strongest stress falls, using Diagnostic III, the surprise-redundancy contour.

Interestingly, there are two possible outcomes, shown in (27):

(27) a. b.



Case (27a), the more normal reading, is the result of applying the Rhythm Rule: the stress levels of the first and third syllables of *Mìssissíppi* have reversed, so that the L* of the surprise-redundancy contour seeks out the first syllable.

```
(28)  3    2    1           2    3       1
   Mississippi mud →  [ Mississippi  mud ]
                          |            |  |
                     (H)  L*           H* L
```

Case (27b) shows that the Rhythm Rule is optional, at least in certain contexts: the L* falls on *-sip-*, showing that it has retained its prominence over *Miss-*.

It would be incorrect to interpret these data as simply free assignment of the L* to any stressed syllable. To see why, consider the alignments in (29):

```
(29) a.  2    3    1            b.   2    3    1
       [alligator alley! ]         * [ alligator alley! ]
        |         |     |            |     |     |     |
        L*        H*    L            H     L*    H*    L
```

The word *álligàtor* has a falling stress contour, and must retain this contour in all contexts, since there is no "Anti-Rhythm Rule" that takes 2 . . . 3 . . . 1 to 3 . . . 2 . . . 1. Accordingly, there is only one option for docking the L* tone of the surprise-redundancy contour. This shows that alignment of this tone is not free in general. The varying alignment of the surprise-redundancy contour with *Mississippi mud* must be due to varying stress.

The Rhythm Rule can be used to shed more light on the English stress system. First, it is a diagnostic for stresslessness. The reason is that the rule is unable to retract stress onto a completely stressless syllable. The inventory of names in (30) illustrates this:

(30) a.
3 2 1 → 2 3 1
Christine Smith → *Christine Smith* ([krìstín])

3 2 1 → 2 3 1
Dawnelle Allen → *Dawnelle Allen* ([dɔ̀nɛ́l])

3 2 1 → 2 3 1
Maureen Reagan → *Maureen Reagan* ([mɔ̀rín])

b.
2 1
Lamont Cranston → no change ([ləmánt])

2 1
Collette Craig → no change ([kəlɛ́t])

2 1
Patrice French → no change ([pətrís])

Again, it is possible to show that the relevant constraint applies to stressless vowels in general, not just to schwa:

(31)
2 1
Jerome Smith → no change ([ǰr̩ówm])

2 1
Pierre Jones → no change ([piɛ́r])

The Rhythm Rule can also serve as a confirming diagnostic for other stress levels. The crucial point is that the retracted secondary stress seeks out the most prominent "landing site" on its left to move to. Consider the following example. In *Sùnset Párk,* familiar diagnostics show that the main stress falls on *Park* and the second strongest stress on *Sun-*. However, since the vowel of *-set* is [ɛ], it too must bear a weak degree of stress.

Observe now that when the Rhythm Rule applies to *Sunset Park,* the stress on *Park* is retracted all the way to *Sun-,* and cannot land on *-set.* This can be shown by Diagnostic III, the surprise-redundancy contour.

(32)
3 4 2 1 → 2 4 3 1 ; 3 2 4 1
Sunset Park Zoo → [Sunset Park Zoo], not * [Sunset Park Zoo]
(H) L* H* L ; H L* H* L

That is, the fact that *Sun-* has stronger stress than *-set* makes it the more qualified landing site when the stress on *Sunset Park* is retracted. I take this as evidence that the stress contrast between the two syllables of *Sunset* is preserved even when the word is embedded in a more complex expression. Notice that all four syllables in the input to (32) bear some degree of stress; if we add in a stressless syllable (e.g. *Sunset Park Gardens*), we have evidence for five levels of stress in English.

The "landing site" property of the Rhythm Rule confirms some conclusions arrived at earlier. For example, I argued that the difference between schwa and most other vowels of English involves stress as well as vowel quality. This is confirmed by the fact that secondarily stressed vowels take precedence over schwa in receiving a retracted Rhythm Rule stress:

```
(33) a. the Macassarèse Émbassy → [ the Macàssarese Émbassy ]
             ([məkǽsəríz])          |        |        |      |
                                    H        L*       H*     L

     b. the Kalamazòo Blúes → [ the Kàlamazoo blúes ]
           ([kæ̀ləməzú])        |     |           |  |
                               H     L*          H* L
```

Further, the Rhythm Rule can be used to verify the contrast of secondary versus tertiary stress within the same word. Those speakers who get a stress contrast between *totàlitárian* and *Cònstantinóple* can, if pushed, apply the Rhythm Rule to these words. The result comes out as expected:

```
            4 3  2     1                 4 2  3     1
(34) a. totalitarian tendencies → [ totalitarian tendencies ]
                                    |  |        |         |
                                    H  L*       H*        L

             3   4   2     1              2   4    3     1
     b. Constantinople trains → [ Constantinople trains ]
                                     |               |   |
                                (H)  L*              H*  L
```

This is probably pushing the method as far as it can go, as the stress contours of (34) are somewhat strained. This is arguably due to the difficulty of applying the Rhythm Rule to sequences of this length; the most natural outcome here is not to apply the Rhythm Rule at all. However, to the extent that it is possible to apply the Rhythm Rule, the landing sites seem to be the predicted ones.

2.4 CONCLUSIONS

I have now covered several phonological diagnostics for stress in English. The overall picture is summed up as follows.

(a) The diagnostics invariably agree with each other. The evidence is strong that underlying all of these diverse phenomena is a single organizing system, namely, the stress system.

(b) Although the exposition above was devoted to the gradual discovery of more stress levels, I would not want to claim that there is a specified number of levels of stress available in English. The maximum found was five, but we looked only at utterances that are syntactically fairly simple. Slightly longer utterances may include more levels; for example (35a) includes six (the five numbered ones, plus stressless), and in (35b) it seems possible to intuit seven.

2 5 4 3 1
(35) a. *non-totalitarian tendencies*

3 6 5 4 2 1
b. *Non-totalitarian tendencies appall him.*

With utterances of increasing complexity, it gets more difficult to judge the number of stress levels; they become "bleached out" in a complex wash of pitch, duration, and rhythm.

It is not clear how such utterances should be categorized. My own view (borrowed from *SPE*) is to let the number of stress levels be determined by the theory, that is, by the most general and explanatory rules that one can devise on the basis of the simple and clear cases. As we will see, rule systems of this sort do generate a large number of levels when applied to complex utterances. Some of these levels may be essentially undetectable in the phonetic record; after all, although the phonetic resources of English for revealing stress contours are substantial, they are not infinite.

An alternative view is to limit the number of stress levels in the phonological system to the number that (it is claimed) can be clearly realized in the phonetics. The postulated number varies from author to author; for instance three (Stockwell 1972), four (Vanderslice and Ladefoged 1972), or five (Gussenhoven 1991). In such a view, the phonological rules simply stop generating more levels once the maximum is reached.

A reason to be skeptical of this view is that there is no clearly defined cutoff point where, as these accounts would predict, the phonology stops creating new levels. The data rather suggest that the phonology can create an unbounded number of levels, and that these gradually blur out as they increase in number and the phonetic cues signal them with progressively less clarity.

(c) The kind of evidence examined here does not require instruments to gather. The relevant facts of segmental phonetics and pitch contours are clear to anyone with a reasonably good ear and a little practice. Since the auditory data involve discrete patterns rather than physical quantities, they are more reliable and easier to interpret.

This is not to say that experimental work on the phonetics of stress is uninteresting; quite to the contrary. But such work is more likely to lead to insights concerning the nature of phonetic rules and speech perception, rather than to any kind of automatic diagnostic for stress.

(d) I believe a similar approach would prove fruitful in the first-hand study of stress in other languages; that is, many of the diagnostics for stress in English carry over directly, especially those involving attraction of intonational tones.

The diagnostic of schwa vowel quality can be stated in more general terms. What seems crucial concerning English schwa is not so much its actual quality (though centralization often is a characteristic of stresslessness), but the fact that the number of phonemic vowel contrasts in English is greatly reduced in stressless position. It is the reduction of contrasts that generalizes most readily across languages. For example, Italian (Vincent 1987) contrasts seven vowels in stressed position, but only five in stressless; Russian reduces its five vowel phonemes to three in stressless position (Jones and Ward 1969); and Catalan (Mascaró 1975) reduces seven vowels to three (or five in borrowed words). In Mandarin, stresslessness is marked by loss of all tonemic contrasts, in syllables bearing the "neutral tone" (Yip 1980).

(e) Lastly, one must consider how to analyze stress in other languages when one is relying on evidence from secondary sources. Such data are not necessarily unreliable. In particular, linguists describing stress in other languages often provide descriptions of its phonetic correlates and of phonological rules that diagnose stress. Where the facts reflect purely auditory intuition, we are on shakier ground. Still, it seems that most people have fairly good intuitions about stress. The studies of Lea 1977 and Thompson 1980 showed good inter-speaker reliability concerning where stress was judged to fall, even though the subjects had no phonetic training. Further, people's intuitions about stress agreed fairly well with what emerges from phonological argumentation.

The situation with regard to the data thus seems short of optimal, but still good enough to justify at least preliminary cross-linguistic investigations of the sort conducted here. The crucial point made in this chapter is that by intensive study of the intonational and segmental phonology of a language, it is possible to make the investigation of stress patterns more rigorous. It is to be hoped that further study of this nature will be carried out in the future, strengthening the empirical basis for theories of stress.

METRICAL S•T•R•E•S•S THEORY

Principles and Case Studies

BRUCE HAYES

THE UNIVERSITY OF CHICAGO PRESS
Chicago and London

Bruce Hayes is professor of linguistics at the University of California, Los Angeles.

The University of Chicago Press, Chicago 60637
The University of Chicago Press, Ltd., London

Printed in the United States of America
03 02 01 00 99 98 97 96 95 94 12345
ISBN: 0-226-32103-7 (cloth)
0-226-32104-5 (paper)

Library of Congress Cataloging-in-Publication Data

Hayes, Bruce, 1955–
Metrical stress theory : principles and case studies / Bruce Hayes.
p. cm.
Includes bibliographical references and indexes.
1. Accents and accentuation. 2. Metrical phonology.
I. Title.
P231.H38 1995
414′.6—dc20 93-34063
CIP

The paper used in this publication meets the minimum requirements of the American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.