

Bruce Hayes

The Phonology of Rhythm in English

Much recent research in the phonology of stress has centered on the English Rhythm Rule, which is responsible for stress alternations such as those in (1):

- (1) *fourtéen* – *fòurteen wómen*
Mississíppi – *Mississippi législature*
seventy-séven – *sèventy-seven séals*

The rule is of interest as a purely analytic problem, because it is difficult to formalize in a way that does justice to the facts. But it is of even greater interest because of the theoretical problems it raises. For example, the Rhythm Rule has figured heavily in the controversy over the representation of stress. There is a fair consensus in the field that the segmental approach to stress proposed in SPE (Chomsky and Halle (1968)) is inadequate, and that stress requires some kind of suprasegmental representation. But the form of this representation is very much in doubt: various researchers advocate “metrical trees,” “metrical grids,” and other representations, as well as theories that mix the two. Another question the rule raises is that of defining the role of rhythm in stress. Pre-theoretic descriptions of the phenomenon typically invoke rhythmic pressures, attributing it to a need to alleviate “clashing” stresses or achieve “alternating rhythm.” If these intuitions are right, one is led to wonder just what rhythm is, what formal model could describe its structure, and through what mechanism it influences the patterning of stress.

This article is an attempt to resolve these questions with a fairly detailed theory of rhythmic phonology. My analysis uses mostly English data, and builds on core ideas presented in Liberman and Prince (1977). It leads to the following specific theoretical conclusions:

(a) Separate representations are required for rhythmic structure and for linguistic stress. These representations should be identified with Liberman and Prince’s metrical grids and metrical trees, respectively. The two representations will be shown to play sharply distinct roles in rhythmic phonology.

(b) It is accordingly mistaken to suppose that a theory incorporating both grids and

I would like to thank Morris Halle, Michael Hammond, Donald Hayes, Marie Huffman, Patricia Keating, Alan Prince, Stanislaw Puppel, an anonymous LI reviewer, and many others for their comments and help.

trees is redundant. The argument here will be not just conceptual, but also empirical: there exist cases that pose serious problems for theories that propose to eliminate the grid (for example, Kiparsky (1979)), as well as for theories that would eliminate metrical trees (notably Prince (1983), Selkirk (forthcoming)).

(c) The notion of "stress clash" proposed in Liberman and Prince's work plays no role in rhythmic phonology. Instead, the rules of the phonology refer to specific rhythmic targets, encoded in "rules of eurhythmty." Some tentative work suggests that these rules may be universal or subject only to minor variation across languages.

(d) Arboreal rules are subject to a general constraint on how they may analyze tree structure. Like similar constraints in syntax, the constraint I will propose allows idiosyncratic information to be factored out of individual rules, thus avoiding overgeneration and increasing the explanatory adequacy of the theory.

The exposition is organized as follows. I first review the essentials of Liberman and Prince's analysis, showing how it fails in a fair number of cases. I then propose a revised conception of the rhythmic target of the stress rules, showing how it alleviates the earlier difficulties. In the second half of the article, I address the issue of whether a grid-and-tree theory is redundant, developing a fairly elaborate argument in favor of a theory in which trees represent stress; grids rhythmic structure. In the final section, I summarize the results.

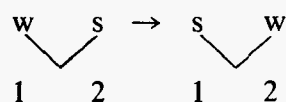
1. Stress and Rhythmic Targets

1.1. Liberman and Prince's Account

Liberman and Prince (1977, hereafter LP) propose a system of rules constructing metrical grids from metrical trees, along with an explicit formulation of the Rhythm Rule. Taken at face value, their system performs two functions: it accounts for native intuitions of syllable prominence more accurately than the n -ary [stress] feature of SPE, and it predicts when the Rhythm Rule will apply. LP programmatically suggest additional functions of the grid, but their suggestions are not precise enough for empirical testing. In this section, I will review LP's rules, showing that, as they stand, they cannot account for the full range of facts. I will then propose a more articulated approach to the problem.

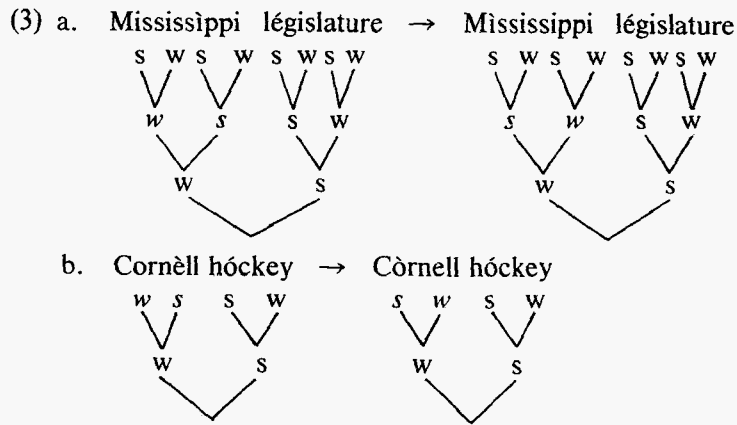
LP propose that a metrical derivation begins with the construction and labeling of metrical trees. Later work (e.g. Selkirk (1980), Hayes (1982)) has suggested substantial modifications of these tree construction rules. But for purposes of rhythmic phonology, the various versions are largely equivalent, so that in what follows I will mostly employ the notationally concise LP-style representations. Metrical trees permit a very simple statement of the Rhythm Rule, as shown in (2):

(2) *Rhythm Rule* (LP version)



where 2 is not the strongest element of its phrase

The rule applies to forms like *Mississippi legislature* and *Cornell hockey* as follows:

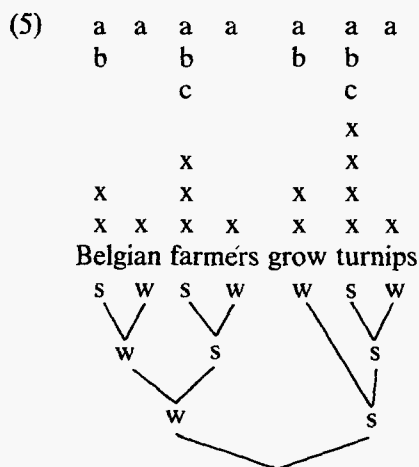


However, the simple arboreal formulation of (2) fails to do justice to the complexity of the facts. Although applying the rule would be the norm in examples like (3), cases like *Mississippi législat^uion*, *Cornèll athlét^uics* would usually not undergo it, even though they contain the appropriate trees. It is to capture distinctions of this sort that the metrical grid comes in. Grids are derived from tree structure by a set of rules that may be stated explicitly as follows:

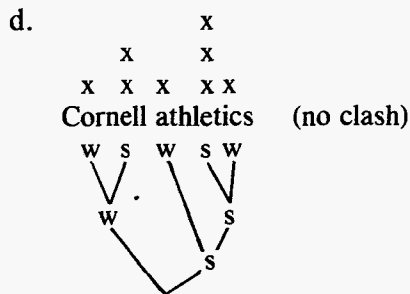
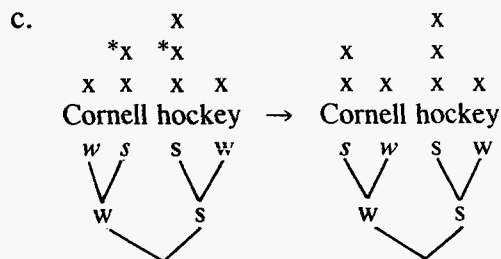
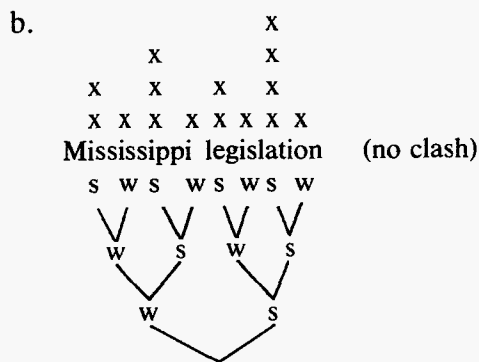
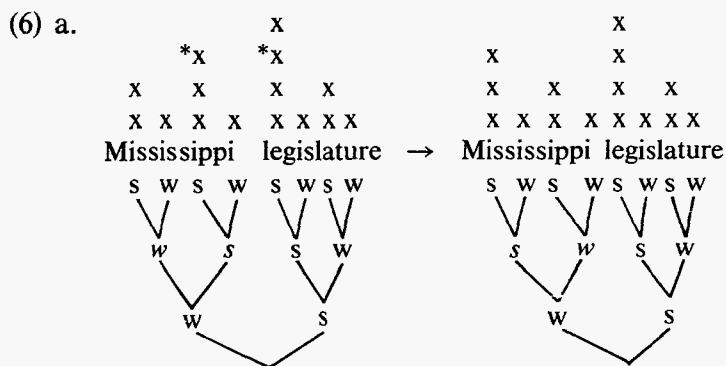
(4) *Grid Construction* (from LP, 315–316, 322)

- a. As a place marker, assign every syllable a mark on the lowest level of the grid.
- b. Assign a mark at level two to the strongest syllable of every phonological word.
- c. Assign sufficient additional marks so that the strongest syllable of every constituent labeled S has a higher grid column than the strongest syllable of its weak sister.

These rules are illustrated in the following sentence, where I have indicated with letters the subrules that have applied in creating each column.

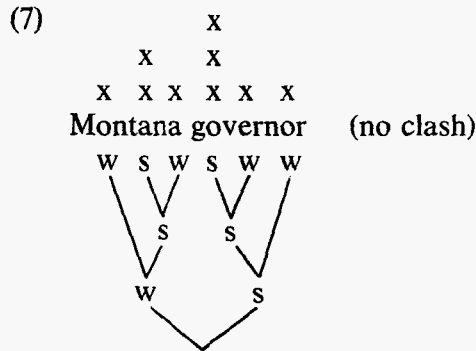


The crucial additional condition on the Rhythm Rule is based on the grid: the rule applies preferentially when it alleviates a *stress clash*, defined as two marks adjacent on their row, with no intervening mark on the immediately lower row. This condition distinguishes the relevant cases. In the examples that follow, stress clashes are marked with asterisks, and relabeled nodes are italicized.

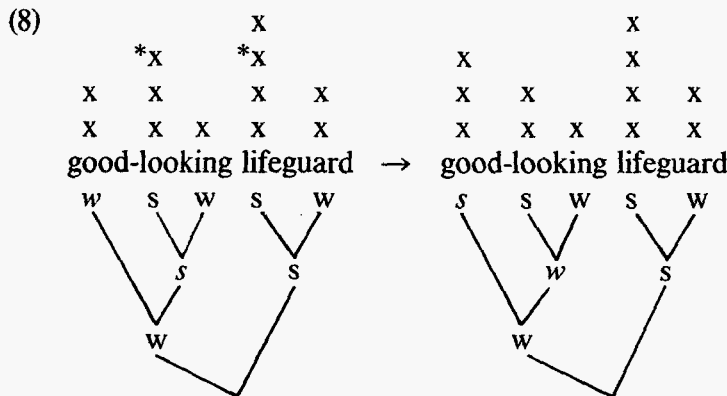


1: the rule
s adjacent
condition
e marked

The notion of stress clash also correctly predicts that words with the stress pattern of *Montana* will not ordinarily undergo relabeling:

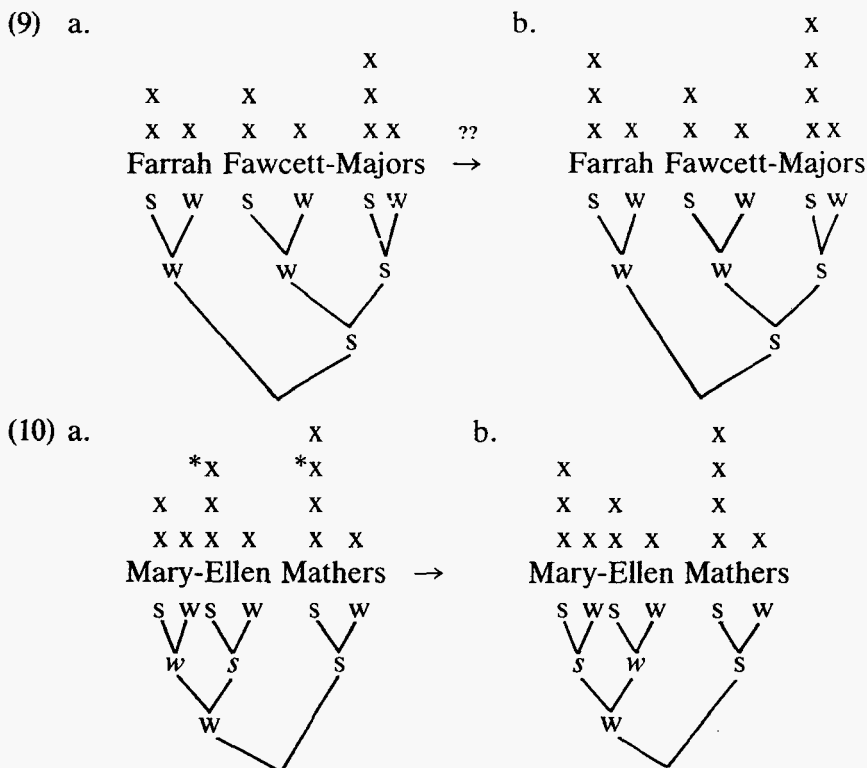


In addition, the stress clash theory provides motivation (beyond mere rhythmic intuitions) for rule (4b) of the grid construction algorithm, as shown by examples like *good-looking lifeguard*. In (8), the extra mark that rule (4b) places on the syllable *good* causes rule (4c) to promote the column of *look* into a position that clashes with *life*. The Rhythm Rule accordingly applies, even though *good-looking* has the same tree shape as *Montana*, under (7).



Further arguments for rule (4b) may be found in Hayes (1983).

LP's work contains a second rule of rhythmic adjustment, although they do not state it explicitly. This rule is based on the fact that in right-branching structures that are labeled *wws*, an alternating prominence pattern usually develops that cannot be derived by relabeling the tree. For example, the stress pattern of *Farrak Fawcett-Majors*, under (9), displays the same sort of rhythmic alternation found in the left-branching form *Mary-Ellen Mathers*, under (10). But it is only in the left-branching case that alternation can be achieved by the Rhythm Rule. The rules stated so far provide no way of converting the predicted form (9a) into the correct output, (9b).



In fact, this problem is quite general. As (11) shows, it is easy to construct pairs of examples that differ in labeling and direction of branching just like (9) and (10), yet are homophonous:¹

- (11) a. $\begin{matrix} & 2 & 3 & & 1 \\ \text{[sea-green]} & \text{soup} & & & \\ & 2 & 3 & & 1 \end{matrix} = \begin{matrix} & 2 & 3 & & 1 \\ \text{see} & \text{[green soup]} & & & \\ & 2 & 3 & & 1 \end{matrix}$
- b. $\begin{matrix} & 2 & 3 & & 1 \\ \text{[twenty-eight]} & \text{steaks} & & & \\ & 2 & 3 & & 1 \end{matrix} = \begin{matrix} & 2 & 3 & & 1 \\ \text{twenty} & \text{[ate steaks]} & & & \\ & 2 & 3 & & 1 \end{matrix}$
- c. $\begin{matrix} & 2 & 3 & & 1 \\ \text{[twenty women's]} & \text{jackets} & & & \\ & 2 & 3 & & 1 \end{matrix} = \begin{matrix} & 2 & 3 & & 1 \\ \text{twenty} & \text{[women's jackets]} & & & \\ & 2 & 3 & & 1 \end{matrix}$

An additional rule is thus needed to induce rhythmic alternation in the right-branching cases. LP do not formalize this rule, but their article implies something like (12), which I will call *Beat Addition*, adapting a term from Selkirk (forthcoming).

(12) *Beat Addition*

Freely add additional marks to the grid columns, provided the relative prominence relations specified in the tree are preserved.

By virtue of this rule, grid (9a) may be amplified to produce (9b), since the additional marks that (9b) contains do not violate any of the tree labelings. Notice that the rule achieves "prosodic homophony" with *Mary-Ellen Mathers*. Similar results hold for the examples of (11).

The *Beat Addition* rule (12) seems precise enough as far as its structural *change*

¹ No theoretical status should be attributed to the numbers in the examples; they are only a shorthand notation for prominence rank.

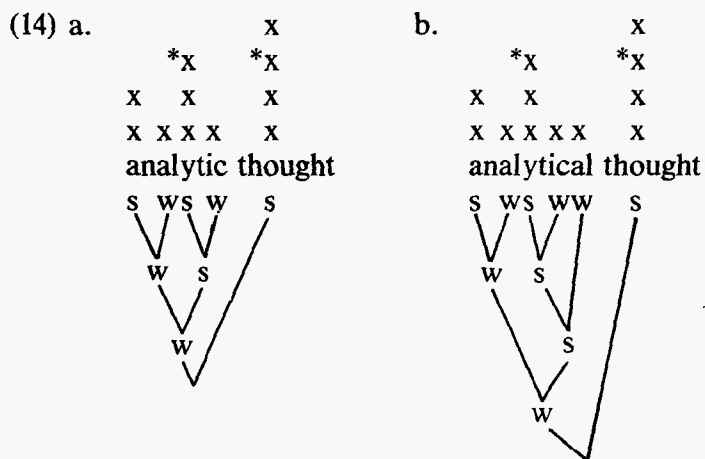
goes, but it includes no structural *description*; that is, it does not say which syllables should be amplified. LP's remarks on this matter (p. 327) are interesting but programmatic. I will try to provide an explicit account below.

1.2. *Some Arguments against LP's Rules*

LP's system predicts that phrases eligible for the Rhythm Rule will fall into two classes: those containing stress clashes, in which relabeling is preferred; and those lacking them, in which the basic stress contour is normally retained. This dichotomy is appealingly simple, and it would be a significant result if true. But a fair range of examples suggests that stress clash is neither a necessary nor a sufficient basis for predicting when the Rhythm Rule will apply. A simple case of this sort involves comparing utterances having originally disyllabic interstress intervals with examples having originally trisyllabic intervals:

(13) a.	$\begin{matrix} 2 & 3 & 1 \\ \text{Mississippi Mabel} \\ 2 & 3 & 1 \\ \text{Punxatawny Pete} \\ 2 & 3 & 1 \\ \text{analytic thought} \\ 2 & 3 & 1 \\ \text{diacritic markings} \\ 2 & 3 & 1 \\ \text{the Passamaquoddy verb} \\ 2 & 3 & 1 \\ \text{Alabama relatives} \\ 2 & 3 & 1 \\ \text{European history} \\ 2 & 3 & 1 \\ \text{Oklahoma congressman} \\ 2 & 3 & 1 \\ \text{two thousand one} \end{matrix}$	b.	$\begin{matrix} 2 & 3 & 1 \\ \text{?Minneapolis Mike} \\ 2 & 3 & 1 \\ \text{?Passaconaway Pete} \\ 2 & 3 & 1 \\ \text{?analytical thought} \\ 2 & 3 & 1 \\ \text{?diacritical markings} \\ 2 & 3 & 1 \\ \text{?the Potawatomi verb} \\ 2 & 3 & 1 \\ \text{?Alabama connections} \\ 2 & 3 & 1 \\ \text{?European historian} \\ 2 & 3 & 1 \\ \text{?Oklahoma congressional district} \\ 2 & 3 & 1 \\ \text{?two thousand and one} \end{matrix}$
---------	---	----	---

It seems clear that the trisyllabic cases are more resistant to relabeling. By LP's theory, there should be no difference between the two cases, since both involve a stress clash as their rules define it:



airs of
yet are

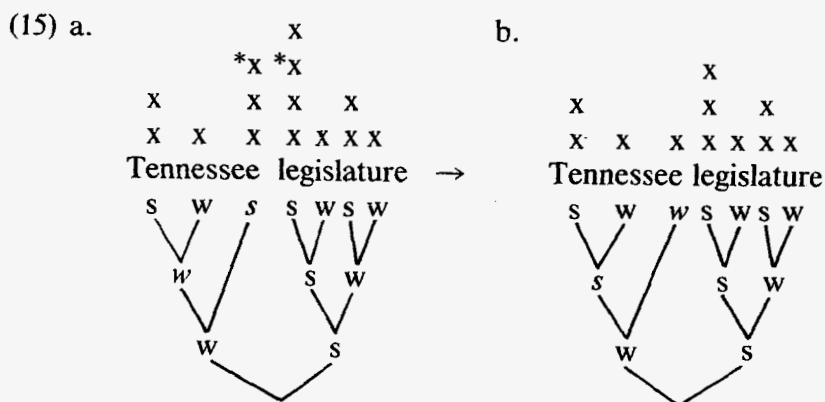
aching
which

promi-

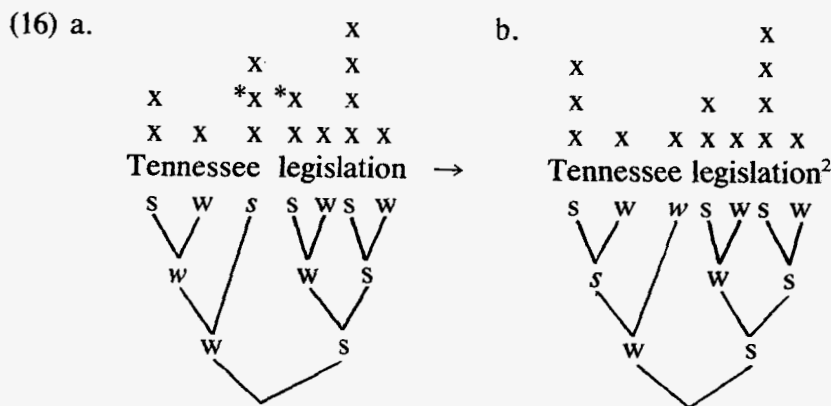
itional
e rule
or the

change
orthand

There are other forms that contain a clash in LP's sense, yet are resistant to re-labeling. Consider the contrast between *Tennessee legislature*, where rhythm is the norm, and *Tennessee legislation*, where the original stressing is usually retained. The first example is a straightforward case of clash resolution through tree relabeling:



However, under LP's theory there is no reason why the same adjustment should not take place in *Tennessee legislation*, which also suffers from a stress clash:

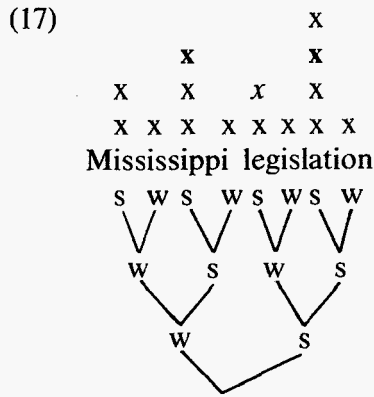


One remedy would be to restrict the definition of stress clash to cases in which the second column involved is higher than the first. As we shall see, this proposal is untenable.

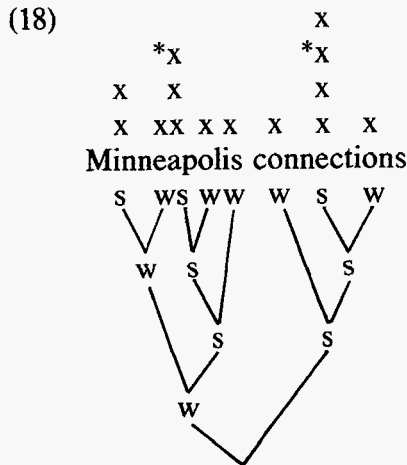
Another class of cases in which stress clash fails to induce relabeling can be con-

² The additional grid mark on *Tennessee* would be derived by Beat Addition.

structured as follows. Consider first the grid of *Mississippi legislation*:



By LP's account, what keeps the boldface grid marks in (17) from clashing is the italicized "intervener" stress on the immediately lower level. By a judicious shift of word boundary, we can remove this intervener while keeping all else constant, as in (18):



This introduces clash, but surely no greater propensity to undergo the Rhythm Rule. Other pairs of examples behave the same way: cf. *Punxatàwny celebrations* vs. *Pasacònaway convèntions*, *achromàtic aberràtion* vs. *mathemàtical invèntions*, *sympathètic Presbytérians* vs. *evangèlical Repúblicans*.

Words like *Ticonderoga* (also *Ompompanoosuc*, *Dodecanesus*) pose yet another problem for the theory of stress clashes. These words have two possible stress contours, 32010 and 23010. I follow Kiparsky (1979) in assuming that this results from a property of English tree construction: above the foot level, the direction of branching is free when not dictated by morphological structure. This derives the 3-2-1 version straightfor-

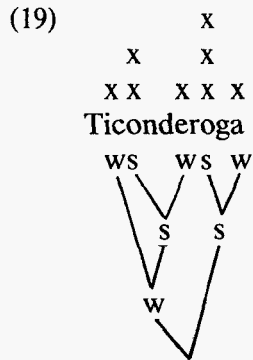
nt to re-
ie norm,
The first

ould not

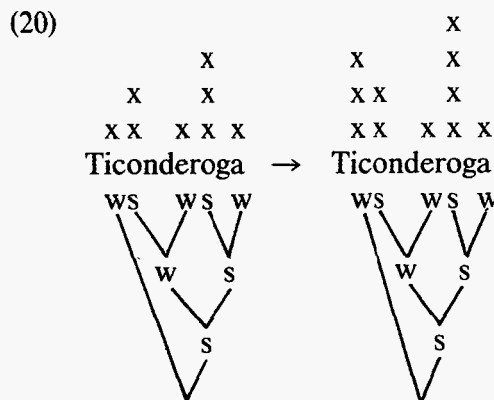
hich the
al is un-

be con-

wardly, given the rules for English word tree labeling (for details, see Hayes (1982)):



The difficulty lies in how the right-branching tree is to be interpreted as a 2-3-1 pattern. Kiparsky assumes that this is to follow from LP's algorithm for translating trees into SPE stress numbers, in which the less deeply embedded of two weak constituents is taken to be more prominent (see LP, 259). However, this algorithm has been shown to be wrong both by LP and by Selkirk (forthcoming). More plausibly, the 2-3-1 contour derives from Beat Addition, as in (20):

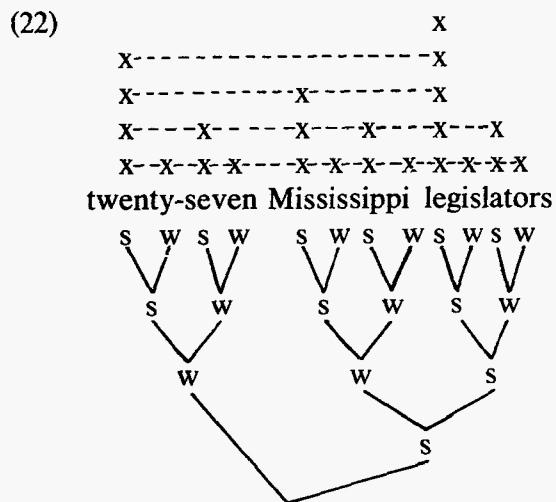


The crucial question is why the shift of (20) should count in any way as a rhythmic improvement. Under LP's theory, it clearly does not, since it only adds a stress clash.

Further evidence can be found in the work of Bolinger (1965b). Bolinger's argument is based on the well-known inability of adjectives beginning in stressless *a-* to occur preminimally, as in **alivè péople*, **aslèep stúdents*. The restriction is due historically to the origin of these adjectives as prepositional phrases, but can be shown to be synchronically productive by its extension to *a-* adjectives not descended from PPs, as in *afraid*, *aghastr*, *averse*. The interesting point is that in adjectival compounds, such as *hàlf-awake péople*, *sòund-asleep stúdents*, the restriction is abandoned. Bolinger argues that the historical pressure to retain the restriction is prosodic. His assumption is that rhythmic pressures act as a shaping force in linguistic change, disfavoring innovations

rather than of linguistic stress. I propose that certain rhythmic structures are designated as highly valued, or "eurhythmic," and that the propensity of the Rhythm Rule to apply is determined by how much it increases the eurhythmy of the text.

There is good reason to suppose that grids embody rhythmic structure, beyond its strictly linguistic manifestation. For example, Jackendoff and Lerdahl (1980) argue that a formal theory of music should include grids as the representation of rhythmic structure. In Hayes (1983) I propose that the rhythmic structure of poetic meters should be represented with grids as well. Grids also neatly capture our intuitive notions of rhythm. The most basic feature of rhythm is repetition of beats at roughly equal intervals. For example, in *twenty-seven Mississippi legislators*, the most basic beat is set by a rapid sequence of twelve evenly spaced syllables. What makes rhythm interesting, however, is that several interbeat intervals may occur simultaneously. In the example just cited, the hearer may perceive disyllabic intervals separating the beats *twen . . . sev . . . Mis . . . sip . . . leg . . . lat*. Quadrisyllabic intervals also occur, separating the beats *twen . . . Mis . . .*, and *leg*; and many hearers will perceive an octosyllabic level encompassing the beats *twen . . . leg*. The intervals obey the law that any beat taking part in longer intervals necessarily participates in all shorter intervals—to use LP's apt phrase, rhythm involves a "hierarchy of intersecting periodicities." Observe now that all of these intuitions are reflected quite clearly in the grid derived for this text by LP's rules:

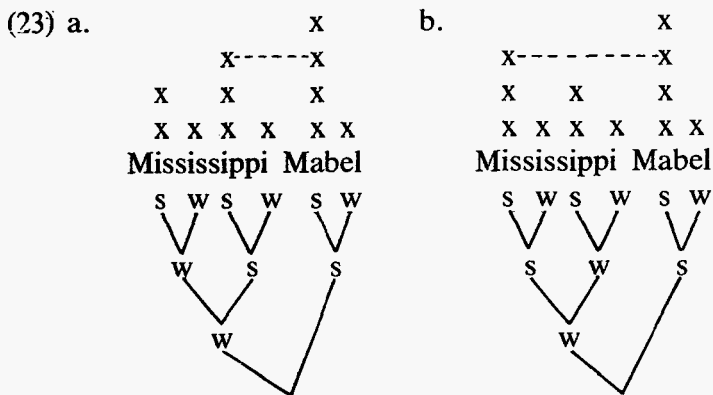


The rows of the grid correspond to the interbeat intervals described above, and the height of each column reflects stress prominence, as a result of the grid construction rules. Under grid theory, the law that stronger beats participate in longer intervals is reflected in the formalism.

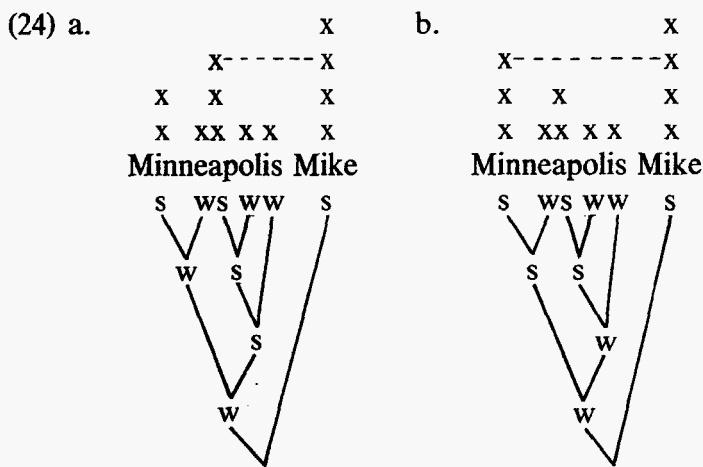
The strategy pursued here, then, is as follows: I am assuming that grids represent rhythmic structures, and that they are projected from arboreal stress representations by rules (4a-c). The Rhythm Rule applies more readily when as a result the text receives a more highly valued rhythmic structure. The value of a rhythmic structure is computed

from its grid by a set of rules I will call *rules of eurhythm*. From this perspective, the problem of predicting the likelihood that the Rhythm Rule will apply consists in defining the eurhythmy rules, and it is to this task that I now turn.

From the preceding discussion, a reasonable hypothesis would be that the eurhythmy rules require equal spacing of grid marks at all levels. This comes close to the truth, I believe, but other factors enter in as well. In particular, eurhythmy requires a *particular spacing of marks* to be found at some level of the grid. This interval appears to center around four syllables, with greater dysrhythmy at greater divergence. We can see this by comparing a number of cases. In *Mississippi Mabel*, relabeling is clearly the norm—a quadrisyllabic interval is preferred to a disyllabic one, as (23) shows.



But in *Minneapolis Mike*, shown with its grid in (24), there is little pressure to relabel—a pentasyllabic interval is no improvement over a trisyllabic one if four syllables is the target distance.



These cases in fact form part of a larger continuum: as the interstress interval of the input form is decreased from the ideal of four syllables, the propensity to apply the Rhythm Rule correspondingly increases. This can be perceived by reading the chains of examples below consecutively.

- | | | |
|---------|-------------------------|-------------------------|
| (25) a. | Tenessèe abbreviàtions | Tènessee abbreviàtions |
| | Tenessèe législàtion | Tènessee législàtion |
| | Tenessèe connèctions | Tènessee connèctions |
| | Tenessèe relatives | Tènessee relatives |
| b. | Mississippi législàtion | Mississippi législàtion |
| | Mississippi connèctions | Mississippi connèctions |
| | Mississippi relatives | Mississippi relatives |

Further, when the interstress interval of the input is kept constant, the propensity to relabel decreases as the interval of the output is increased beyond four syllables:

- | | | |
|------|------------------------|-------------------------------------|
| (26) | Àlabama relatives | Alabàma relatives |
| | Àlamogordo relatives | Alamogòrdo relatives |
| | Àpalachicola relatives | Apalachicòla relatives ⁴ |

These examples motivate the following rule of eurhythmy:

(27) *Quadrisyllabic Rule*

A grid is eurhythmic when it contains a row whose marks are spaced close to four syllables apart.

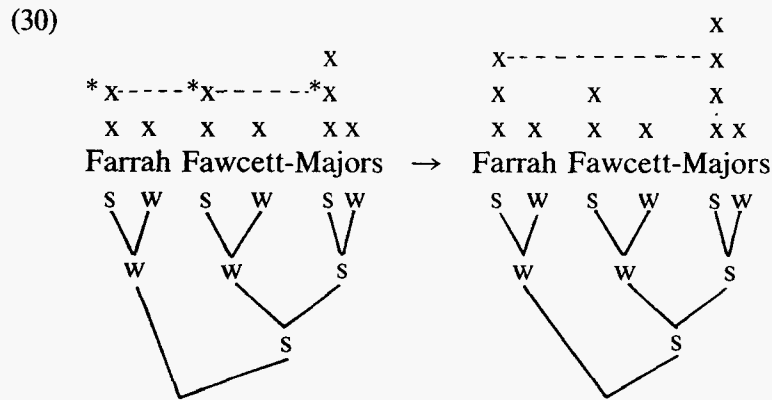
The Quadrisyllabic Rule should be taken as a principle of grid scansion: a grid is searched for the row that best fits the rule's description, and eurhythmy is proportional to how close the marks come to the desired four-syllable interval. Although judgments are delicate, my impression is that the scale of dysrhythmy is nonlinear, at least at the lower end: each successive syllable deducted from the quadrisyllabic norm results in a greater loss of eurhythmy.

In what follows I will refer to the level selected by the Quadrisyllabic Rule as the *level of scansion*.

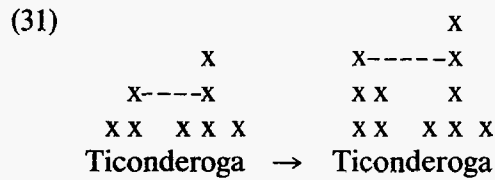
Returning to the earlier data, we find that the differences in propensity to relabel attributed by LP to the presence or absence of a clash are equally well explained in terms of greater or lesser gains in eurhythmy, with one exception to be dealt with below. For example, the interval shifts at the level of scansion found in *Mississippi legislature* and *Mississippi legislation* ((6a,b)) are from two to four and from four to six syllables, corresponding respectively to eagerness and reluctance to undergo the Rhythm Rule. Further, all the cases that proved problematic for the stress clash theory are accounted

⁴ It is quite difficult to retract stress just two syllables in this example, as in **Apalàchicola relatives*. This follows from a precedence principle requiring smaller syntactic domains to receive rhythmic adjustment prior to larger ones—note the contrast between [*Alabàma Street*] *bús route* and [*Àlabama Road*] *bús route* (see also Prince (1983)). A full discussion of this phenomenon goes beyond the scope of this article.

eurhythmy as the Quadrisyllabic Rule defines it:



The application of Beat Addition to *Ticonderoga*, under (20), is similarly motivated by an interval shift from two to three syllables:



That this creates a stress clash in LP's sense again appears not to matter.

The Quadrisyllabic Rule thus accounts for a fair range of data. However, it is not the only principle that dictates rhythmic adjustment. As mentioned above, rhythm involves a whole set of isochronous intervals enacted in a simultaneous hierarchy. It should not be surprising if the rhythmic goals of the phonology should involve more than one level. In particular, I would argue for another principle of eurhythmy, as follows:

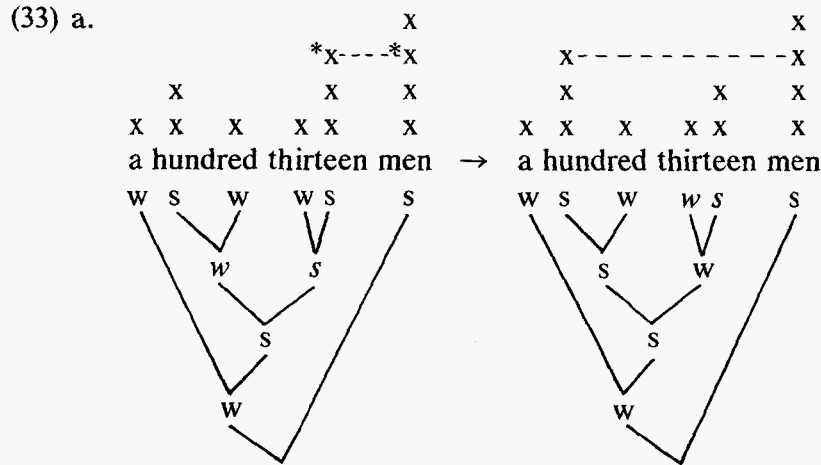
(32) *Disyllabic Rule*

The domains delimited on the level of scansion should be divided evenly by a mark on the next lower grid level.

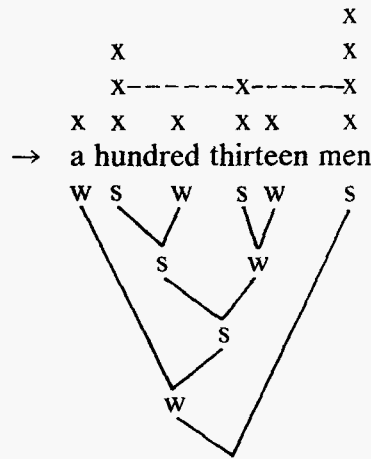
According to (32), the level of scansion is primary, but the general principle that rhythmic intervals should be equal holds both for the level of scansion and for the level below it as well.

Several arguments support this claim. First, there are cases in which the Rhythm

Rule applies twice within a four-syllable phrase, as in (33):



ivated by



, it is not
hythm in-
It should
than one
ws:

b. 2 3 1
an almost hard-boiled egg

c. 2 3 1
an extremely unkind comment

In (33), the “outer” application of the Rhythm Rule (to *hundred thirteen*) is clearly motivated by the Quadrisyllabic Rule. But the rule provides no explanation for the “internal” rhythm found on *thirteen*.⁶ This relabeling would follow directly from a pressure for even spacing on the subscansion level, as the Disyllabic Rule requires. If the rule is

⁶ Prince (1983) claims that internal rhythm is impossible, even though he presents examples of it (claiming that they are special cases). His examples that purport to show the impossibility of internal rhythm are for the most part ruled out independently, since they do not involve an increase in eurhythmy as defined here. Selkirk (forthcoming) admits internal rhythm, supporting her claim with examples that conform to the eurhythmy rules proposed here. Further examples of internal rhythm appear below.

venly by

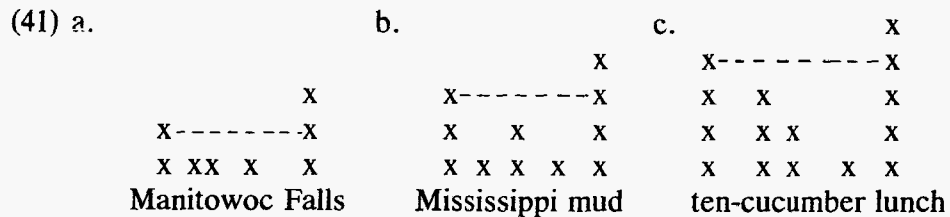
hythmic
below it

Rhythm

1.4. The Scope of the Eurhythmy Rules

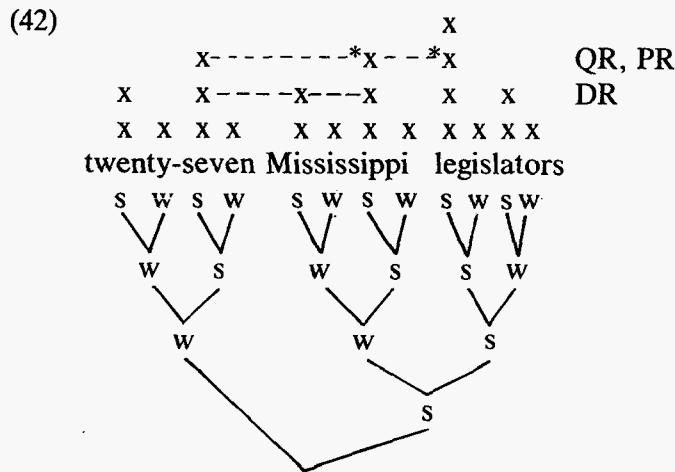
There is evidence, then, for three distinct rhythmic targets in English: (a) a level of grid marks spaced about four syllables apart (the level of scansion); (b) a lower level that divides the level of scansion in half where possible; and (c) a higher level that gives the initial beat of the level of scansion precedence over any other beats at that level. These targets are formalized in the Quadrisyllabic Rule, (27), the Disyllabic Rule, (32), and the Phrasal Rule, (40). In this section I will present further evidence for the validity and generality of these rules.

One important point is that no one level of the grid always constitutes the level of scansion. For example, in (41a), (41b), and (41c) the level of scansion is found on the second, third, and fourth levels of the grid, respectively:



Examples (10) and (54) similarly satisfy the eurhythmy rules at different levels. For this reason, the eurhythmy rules should be interpreted as a "scansion" procedure, which searches the grid for the best available match-up.

A second point is that the Rhythm Rule and Beat Addition may apply several places in a string when eurhythmy is increased. For example, the phrase *twenty-seven Mississippi legislators* is markedly dysrhythmic before the rhythmic adjustment rules apply to it:



Specifically, the marks with asterisks are spaced too closely to conform to the Quadrisyllabic Rule, the Disyllabic Rule is only halfway satisfied, and the Phrasal Rule is not

rnation:

not really
is placed
This was

s, spaced

e in (6b),
to apply
ed more

