

Chapter 3. What is a Mismatch?

3.0 Introduction

Among the examples discussed in the previous chapter was the socialist anthem the Internationale. Following its appearance in French in 1888, the song would quickly appear in translation in numerous languages¹, many of which would become well known during the first half of the 20th century.

While it may seem obvious, it should be noted that what we call translations of songs involve an alteration only of the text with the tune generally remaining largely unchanged from the original. It follows, therefore, that the differences which are found in these translated variants are of two types. One of these is trivial: the presence of a text in a new language implies that the sounds a listener hears and the singer produces will be drawn from the phonemic inventory of the target language-e.g. bilabial fricatives in Japanese (but not English), retroflex liquids in Hindi, vowel harmony in Turkish etc (Ladefoged and Maddieson 1996). More interesting for our purposes is a second characteristic of textsetting which becomes apparent in the comparison of translated texts: identical phonological units shared by diverse languages will be treated differently in different textsetting traditions.

We have already made note of one such phonological feature, namely stress. As we observed previously, stress and beat tend to be aligned in English textsetting whereas in other textsetting traditions (most notably French), stress to beat matching is less rigidly

¹ A collection of mp3 files of the Internationale in more than 40 languages is available at <http://www.hymn.ru/internationale/index-en.html>.

The normative arrangement of English songs can be seen in the first verse and chorus of the American English version of the Internationale shown below.

(2) The Internationale: American English version: (Charles Hope Kerr (1900))

			x						x		
			x			x			x		x
	x	x	x	x	x	x	x	x	x	x	x
1	x	x	x	x	x	x	x	x	x	x	x
		Ar	ise	ye	pris' ners	of	star	va	tion		
2		Ar	ise	ye	wret ched	of	the	earth			
3		For	ju	stice	thun ders	con	dem	na	tion		
4	A	-	bet	ter	world's		in	birth!			
5	No	-	more	tra	di tion's	chains	shall	bind	us		
6	Ar	ise,	ye	slaves,	no	more	in	thrall;			
7	The		earth	shall	rise	on	new	foun	da	tions	
8	We	have	been	naught	we	shall	be		all.		
9	'Tis	the	fi		nal	-	con	flict			
10	Let	-	each	stand	in		his	place			
11	The		In	ter	na	tional	Un	ion			
12		shall	be	the	hu	man		race.			

The above transcriptions allow us to simply count the mismatches in bold with the results obtained (39 for French, 22 for English) providing a rough statistical indication of the differences in the textsetting practice in the two languages. That said, this approach is clearly unsatisfactory in two respects. First, as we shall see, the definition of mismatch appealed to is insufficiently descriptive in that there are particular stress/beat configurations which very rarely occur in attested settings but which the definition does not denote as a mismatch. Conversely, numerous attested configurations violate the definition for unitary mismatch.

A second problem, related to the first, is that taking the unitary definition of mismatch as predictive is clearly at variance with one's intuitive judgments in that what are experienced as acceptably matched configurations often meet the conditions for unitary mismatch. Thus, to take one example from line 8 in (2), "shall", a relatively

unstressed word, appears on a strong position and qualifies as a mismatch. But as this is surely not experienced as an unnatural relationship of a note and syllable, an analysis which defines it as such needs to be revised.

3.0.1 Chapter Organization

The above provides a brief introduction to the purpose of this chapter which is to identify as precisely as possible the underlying mechanisms involved in listeners' judgments as to what constitutes a mismatch in English (and to a lesser degree French) textsettings. It turns out that these are somewhat complex matters which require developing several distinct analytical and explanatory threads. These include a clarification of the musical and linguistic elements involved in the configuration, a precise description of the interactive structure in which these separate elements are implicated and finally, a discussion of the empirical requirements imposed on a theory claiming to account for the facts coupled with an analysis of the extent to which the proposed solution meets these requirements.

Implementing this framework results in the following organization of the chapter: In section 3.1 we formulate a definition of the independent elements (musical and linguistic) which are involved in the mismatched configuration. In sections 3.2 we will introduce three plausible candidates for mismatched text/tune configurations with a view towards determining which is most likely to best describe textsetting practice and our intuitions with respect to it. In section 3.3, having provisionally adopted the binary mismatch as the most likely candidate, we examine some of its properties. Sections 3.4 and 3.5 are devoted to the discussion of certain linguistic forms which require partial revision of the binary mismatch criterion proposed in 3.3. Section 3.6 raises the question

of the empirical requirements for verifying the proposed mismatch criteria and subsequently proposes a formal mechanism for determining whether these requirements are met. Section 3.7-3.9 implement the evaluative scheme proposed in 3.6 in the course of which questions having to do with the effect of prosodic constituency on text setting are addressed. Section 3.10 discusses how the binary mismatch criterion needs to be construed such that it can be made relevant to the external corpus-i.e. the composed vocal literature-generally and applied specifically to particular songs, including the International. 3.11 shows that the binary mismatch criterion does indeed provide significantly better empirical results than the informal mismatch just discussed and that the differences between French and English textsetting practice are revealed much more starkly than in our initial informal description. Finally, 3.12 discusses the relevance of these results for how metrical form (in a broad sense) is understood within some of the disciplines in which it is taken as operative.

3.1.0 Independent Elements: Terminological Preliminaries

As discussed above, the mismatched configurations we are dealing with in this chapter result from the interaction of two independent characteristics of music and language. Before we are able to discuss the nature of this interaction we must first provide a sufficiently rigorous definition of the independent elements which are implicated in these configurations. These are linguistic stress (or accent) and (musical) metrical accent respectively; the following two sections will be concerned with laying this foundation.

3.1.1 Definition of Independent Elements (Linguistic): Stress/Accent

With respect to the relevant linguistic characteristic implicated in mismatched configurations, namely stress, two essential facts need to be taken into account. The first is that the informal understanding of words as having a single accented syllable is insufficient. It is well known (e.g. Halle-Vergnaud 1987) that in addition to carrying primary stress on the fourth syllable, a word such as “Ticonderoga” carries secondary and tertiary stress on the first two syllables as shown in the stress grid below.

(3)

```

          x
    x      x
  x x     x
x x  x x x
Ticonderoga
```

Secondly, when words are combined within phrases, the absolute levels applicable to words considered in isolation become subordinated to phrasal stress patterns (Hayes 1995). Thus, as shown in example (4), if the word “farmer” is appended to “Ticonderoga” what was previously a two level grid column associated with the stressed syllable must be augmented with three additional x’s in order to represent its receiving main phrasal, as opposed to word level, stress.

(4)

a.

```

  x
x  x
farmer
```

b.

```

          x
    x      x
  x      x  x
x x     x  x
x x  x x x  x  x
Ticonderoga farmer
```

For our purposes, the existence of multiple degrees of stress creates a combinatorial explosion of possible arrangements which make analysis of the problem of text to tune alignment unwieldy. A complete treatment would need to itemize all possible phrasal stress patterns in relationship to the rhythmic structure to which they can in principle be assigned and develop an explanation for how each behaves in context.

Fortunately, for reasons discussed in Halle Lerdahl (1997) (hereafter HL) the problem can be simplified considerably without significantly misconstruing the essential nature of the facts. The key to this simplification involves recognizing that musical settings of texts frequently manifest significant departures from the normative stress pattern associated with a given utterance. Thus, for example, the phrase “drunken sailor”, in a normal pronunciation will require greater stress assigned to “sailor” than “drunken”.

(5) stress grid

```

           x
        x   x
       x x  x x
drunken sailor.
```

In the context of the song “The Drunken Sailor” this weak-strong stress pattern is assigned to a strong-weak metrical sequence contradicting phrasal stress:

(6) mismatched phrasal stress

```

                                     x
      x           x           x       x
      x   x       x   x       x   x   x   x
      x x   x       x   x xx   x   xx x x x xx x (metrical grid)
What shall we do with a drunken sailor.
                                     x x   x   x (stress grid in(5)3)
                                     x       x
                                     x

```

Based on examples like (6), HL reach the conclusion that the accentual pattern of phrasal stress is largely ignored when experienced listeners make qualitative judgments as to what constitutes acceptable text-tune composites. As should be apparent from examples such as (2), English textsetting respects local stress patterns, which is to say word level as opposed to phrasal stress tends to predominate. But even here, it will turn out to be the case that only the rough outline of word stress is significant: secondary stresses are grouped with unstressed syllables in listeners' judgments as to their appropriate assignment to metrical positions.

It follows from the above that the gradient reality of stress can, in the context of textsetting, be reduced to a binary distinction between relatively stressed and relatively unstressed syllables. To avoid confusion, these will be referred to in the following as **accented** and **unaccented** respectively within the following definitions which accomplish this necessary and sufficient simplification of the data:

³ The stress grid (6) will be seen to be inverted from its appearance in (5). This purely notational alteration has no effect on the content of the representation.

(7) Definition: accented and unaccented.

a) accented syllable (S): Syllables bearing the main stress in polysyllabic words and monosyllabic content words (nouns, verbs, adjectives).

b) unaccented syllables (s): all others (i.e. monosyllabic function words and non-primary stressed syllables of polysyllabic words).

These definitions conflate into a single category somewhat disparate linguistic sequences. Thus, “action”, “cowboy” and “go there” are taken as instantiations of the trochaic string Ss, whereas words such as “remove” will be taken as iambic sS as are the phrases “I go” and “a boy”. The phrase of monosyllables “John will go” and “John’s alive.” will be taken as the amphimacers SsS. “a Zebra” is defined as an amphibrach sSs and equivalent for the purposes of textsetting to “albino” and “release him”.⁴

⁴ We recognize that these examples ignore the possible effect of word boundaries on textsetting intuitions. See sections 3.6-3.9 for an analysis of the extent to which this empirical idealization is justified.

The definition in (9) follows DH in identifying the tactus level as a threshold above or below which notes are heard as strong or weak respectively. These categories are absolute rather than relative. Thus, while the first note in (8) is stronger than the second, both are considered, according to the absolute definition in (9), weak. There are, it will be seen, certain problems which arise with this definition some of which will be addressed immediately, others later. For present purposes we note that because the unitary definition of mismatch proposed earlier requires that we have access to an absolute rather than a relative definition of metrical accent we need to temporarily maintain (9) as our operative definition of metrical strength.

To simplify matters further, in the following sections we will concern ourselves with a song for which, unlike *The Internationale*, a binary distinction between strong and weak events is unproblematic and indisputable, namely, the first phrase of “Frère Jacques” as shown below:

(10) Frère Jacques (first phrase)⁶

x		x		x		x		x		L(0)
x	x	x	x	x	x	x	x	x	x	L(-1)
C	D	E	C	C	D	E	C			
n1	n2	n3	n4	n5	n6	n7	n8			

The string of eight notes n1-n8 marks out what we referred to in section 2.5.2 as a trochaic tune: odd numbered notes n1, n3, n5 and n7 occur on the tactus level L(0). These are designated as strong while even numbered notes occur on the subtactus level L(-1) and are weak according to the definition in (5).

⁶ By omitting levels L(1) and L(2), we recognize that this grid violates the so called “lapse” well-formedness condition. See section 3.8.6 for discussion.

(10) is, of course, a primitive rhythmic structure compared to those underlying many other well known songs. While it is arguably somewhat misleading for this reason, it will be seen that it provides a basis for examining textsetting intuitions of considerable complexity. Appealing to it will allow us to clarify certain confusions surrounding arrangements of text and tune and the results obtained will be applicable to songs making use of a wider pallet of rhythmic activity.

3.2 Mismatched Configurations

3.2.1 Unitary Mismatches

Armed with these definitions, we can now return to discussing those configurations which may potentially result in the perception of a mismatch. We begin by returning to unitary mismatches, which, as noted, are unsatisfactory in correctly predicting listeners' judgments. Unitary mismatches are of two types: a relatively stressed syllable (S) assigned to a weak (i.e. subtactus) metrical position and the relatively unstressed syllable appearing on a strong (tactus or supertactus) metrical position.

These can now be formally represented as the forms in (11) a and (11) b.

(11) unitary mismatches (definitions)

a. (accented assigned to weak)

x L (<0)
|
n
|
*S*⁷

⁷ Italics here and in the forthcoming are used to indicate that the syllable in question meets the conditions of the mismatch criterion under consideration. This will prove useful in making more conspicuous these instances when they occur within extended musical passages-most notably the entire eight note phrase in (1).

b. (unaccented assigned to strong)

x L(≥ 0)
 |
 n
 |
 s

It is now possible to conclusively demonstrate that neither of these configurations is, by itself, responsible for the perception of an unacceptable setting. To show this we return to example 2.1.1.3b which applied the text “Joe’s boy is at home on Fridays” to the first phrase of the tune “Frère Jacques.” The setting exemplifies the unitary mismatch shown in (11) a resulting from the relatively stressed syllable “boy” appearing on a subtactus level position.

(12)

x		x		x		x		L(0) (tactus)
x	x	x	x	x	x	x	x	L(-1)
C	D	E	C	C	D	E	C	

Joe's boy is at home on Friday.

Despite the unitary mismatch, indicated in the above in bold, (12) is, as pointed out earlier, an acceptable setting. Thus, the unitary definition fails in that it marks as a mismatch a form which is acceptable.⁸

An example of the unitary mismatch shown in (11) b is obtained by assigning the text “The elusive birds were robins.” to the same tune:

(13)

x		x		x		x		L(0) (tactus)
x	x	x	x	x	x	x	x	L(-1)
C	D	E	C	C	D	E	C	

The e lu sive birds were ro bins.

⁸ We will later use the term “overgenerates” to refer to definitions which incorrectly mark as unacceptable forms which are judged acceptable by listeners.

If we take the configuration in (15) as fatal, we succeed in ruling out the deviant setting in (14). However, it will be noticed that this condition is insufficiently restrictive¹¹ in that settings can be easily devised which are heard as unacceptable but where the offending syllables are not in the configuration of the composite mismatch.

One of these can be derived by replacing the unstressed syllable “is” with an inherently stressed monosyllabic content word such as the verb “works”:

(16)

x		x		x		x		x		L(0) (tactus)
x	x	x	x	x	x	x	x	x	x	L(-1)
C	D	E	C	C	D	E	C			
Si mone works at home on Fri day*										

The text is no longer an amphibrach-sSs-but a so-called bacchic foot¹² sSS with the stressed syllable in the final position failing to meet the conditions for a composite mismatch. Since the setting is experienced as unacceptable, and the composite mismatch criterion fails to reject it, it is clear this criterion does not predict the relevant intuitions.

¹¹ We will later use the term “undergenerates” when, as in this case, a given criterion fails to predict a class of unacceptably matched forms.

¹² For a definition of the bacchic foot within quantitative classical meter, see e.g. West (1982). The sense in which we are using the term is, of course, qualitative, namely with respect not to length but to metrical accent. The failure to make this distinct has created considerable confusion. See Attridge (1980) for discussion.

3.2.3 Binary Mismatches

(16), it should be noticed, also contains two unitary mismatches—the unstressed “Si-“ assigned to a strong position (represented in 11b) or the stressed final syllable “-mone” assigned to a weak position (11a). Since, as we have shown above, these cannot be uniquely responsible for the impression of a mismatch, the fatal mismatch appears to be induced by what we will refer to as a **binary mismatch**, namely the two syllable sequence “Simone” assigned to adjacent strong weak metrical positions. We will assume that it is this configuration which renders the setting in (16) unacceptable.

In the following sections we will broaden this assumption, demonstrating that all other settings can be shown to violate the binary mismatch criterion. To do so, it will be necessary to explore the properties of this configuration in a variety of contexts, noting the cases which appear to offer refutations of the criterion. These will be seen, in some instances, to require revising the basic elements we have taken as fundamental in the text setting computation.

Before we embark this discussion, we must provide a sufficiently formal definition of the configurations we are taking to be decisive. These are of two types: what we will refer to as the **left mismatch** shown in (17):

(17) left mismatch

x	
x	x
n1	n2
s	S

(e.g. “Simone”)

Also potentially fatal is the mirror image of (17), the **right mismatch**:

(18) right mismatch

```
      x
     x x
    | |
   n1 n2
    | |
   S  s
```

which appears to account for the unacceptability of settings such as (19).

(19) fatal right mismatch

```
      x          x          x          x          L(0) (tactus)
     x  x  x  x  x  x  x  x  L(-1)
     C  D  E  C  C  D  E  C
John  of ten has worked on Fri day*
```

3.3 Metrically Conditioned Focus

We say that (18) is potentially fatal because the right mismatched configuration in (19) can also appear in well-formed textsettings. For example, (12) above contains the Ss pair “boy is” assigned to a weak-strong sequence:

(20) non-fatal right mismatch

```
      x          x          x          x          L(0) (tactus)
     x  x  x  x  x  x  x  x  L(-1)
     C  D  E  C  C  D  E  C
Joe's boy is at home on Friday.
```

Furthermore, non-fatal left mismatches occur in settings such as (21):

(21) non-fatal left mismatch

```
      x          x          x          x          L(0) (tactus)
     x  x  x  x  x  x  x  x  L(-1)
     C  D  E  C  C  D  E  C
If John works, he's home on Fri day.
```

These counterexamples to the prohibition against left mismatches share a common feature: both contain a monosyllabic function word (“is” and “if” respectively) occurring in the strong position.

The notion of mismatch which we are appealing to relies on the understanding incorporated in the definition in (7) b that function words tend to be relatively unstressed compared to content words. While this is generally the case, it is also apparent that in what are called focus contexts (see e.g. Rooth 1994), some function words can be promoted¹³ to receive primary stress as a means for indicating a marked semantic or pragmatic interpretation. For example, in a normative, unfocussed reading of “He’s not working,” “working” will be receive greater stress compared to “he’s” and “not”. If the intended interpretation is meant to indicate a masculine rather than feminine subject, i.e. that John rather than Mary is working, “he’s” will be stressed. Similarly, if the intention of the speaker is to correct the mistaken impression that someone is working who is not, it will be pronounced “He’s NOT working.”

Textsettings which would contain mismatches under normative pronunciations can be construed as acceptable if focus is assumed. We have already encountered what will turn out to be a relatively pronounced instance of one such setting in 2.1.1.3 repeated below:

(22)

x		x		x		x		L(0) (tactus)
x	x	x	x	x	x	x	x	L(-1)
C	D	E	C	C	D	E	C	

The boy works at home on Friday.

This was incorrectly-albeit understandably-referred to as containing an unacceptable mismatch. In fact, it will be noticed that the setting can be taken as acceptable if prompted by the question “Does A boy work at home on Friday?”. The response- “(No.)

¹³ The sense in which we are using the term is probably equivalent to that assumed in formal prosody, e.g. Malof (1966). Also, see Attridge (1995) for discussion of promotion and demotion within a broad theory of poetic meter.

THE boy works at home on Friday.”-could be properly assigned to the trochaic tune in question. Admittedly, the utterance is somewhat odd for reasons suggested in, for example Büring (2006) and which are well outside the scope of our discussion here. For our purposes, we observe that what we will refer to as **focus conditional mismatches** such as (20), (21) and (22) are categorically distinct from the binary mismatch exemplified by (16) (shown below as (23) a).

(23)

- | | | | | |
|---|---|---|---|---------------|
| x | x | x | x | L(0) (tactus) |
| x | x | x | x | L(-1) |
| C | D | E | C | |
- a. **Si mone** works at home on Fri day.*
- b. *The*_{+f} boy works at home on Fri day.

The bold faced type in (23) a indicates what we will refer to as a **universal mismatch** in that there is no possible pronunciation or semantic or pragmatic context such that the text can be taken as forming an acceptable match. This is distinct from the focus conditional mismatch in (23) b which can be construed as acceptable according to the reading discussed above, with focus indicated (following Selkirk 1996) by the subscript +f. Specifically, the focus interpretation makes possible stress assigned to what would otherwise be an unstressed monosyllabic¹⁴ function word and, consequently, an acceptable setting.

As just mentioned, most focus conditional mismatches will be experienced as (to a greater or lesser degree) more natural than (23) b, as can be seen below.

¹⁴ Constraints on the location of monosyllabic words within verse form are central to generative metrics since Kiparsky (1977). To my knowledge, the potential limiting influence of focus stress on monosyllables has not been accommodated within these accounts.

(24)

- | | | | | | | | | |
|----|-------------------------|---------------|--------------------------|---------------|------|------|------------------------|-----------------|
| | x | | x | | x | | x | |
| | x | x | x | x | x | x | x | x |
| a. | John | is | laz | y | when | he's | <i>in_{+f}</i> | <i>France</i> . |
| b. | <i>Her_{+f}</i> | <i>home</i> , | John | said, | is | in | Kan | sas. |
| c. | Late | ly, | <i>he_{+f}</i> | <i>slept</i> | in | the | arm | chair. |
| d. | See | him | <i>when_{+f}</i> | <i>John's</i> | home | on | Fri | day. |

The acceptability of these settings appears to vary according the plausibility of the situation implied by the particular focus interpretation denoted by the utterance. Thus, “in France” in (24) a would normally be heard as mismatched but can be accepted if understood in response to the relatively unlikely question: “Is John lazy when he’s NEAR France?” The focus in (24) b can be elicited more easily, in response to the comparatively unproblematic question “Is HIS home in Kansas?” though, unlike a, it does not appear to require a special context to be experienced as an acceptable setting. Similarly, the mismatches in (24) c and d are heard as more or less natural and unproblematic and do not require the imposition of an unusual pragmatic frame to be accepted as such.

On the other extreme of this continuum¹⁵ are expletives “it” and “there” (Carnie 2004) which cannot under any circumstances receive focus stress (Seppänen 2002). Sentences such as (25) a and b are categorically distinct from those in (24) in that they are not just pragmatically odd but semantically deviant.

(25)

- a. IT seems John is home on Friday.*
- b. THERE lived once a man named David.*

When these texts are assigned to a tune which is initiated on the beat, the result is a fatal mismatch.

¹⁵ Thanks to David Pesetsky for alerting me to this fact.

(26)

- | | | | | | | | | |
|--|---|---|---|---|---|---|-----|---------------|
| | x | | x | | x | | x | L(0) (tactus) |
| | x | x | x | x | x | x | x x | L(-1) |
- a. **It seems** John is home on Friday.*
b. **There lived** once a man named David.*

Focus stress on locations containing expletives appearing within mismatched configurations cannot be justified as focus conditional and are experienced as categorically unacceptable, in contrast to the configurations in (24).

It should be noted, however, that the assignment of expletives to metrically accented positions is, it appears, not sufficient for a setting to be judged unacceptable:

(27)

- | | | | | | | | | |
|--|---|---|---|---|---|---|-----|---------------|
| | x | | x | | x | | x | L(0) (tactus) |
| | x | x | x | x | x | x | x x | L(-1) |
- a. *It is* John who's home on Friday. (?)
b. *There was* once a man named David. (?)

Since the words following the expletives are the unstressed predicates “is” and “was” the conditions for a binary mismatch are not met and the setting is (correctly-according to an informal poll of informants) predicted to be acceptable.

3.4 Stress Retraction and Stress-Beat Matching

As pointed out previously, “Simone” assigned to a strong-weak metrical sequence is a universal mismatch in that there is no possible linguistic context such that the associated rhythmic setting can be heard as acceptable. However, that not all iambic words make for universal mismatches is clearly exemplified by the settings below:

(28)

	x	x	x	x	L(0) (tactus)
	x x	x	x	x	x x L(-1)
a.	<i>Invite</i>	John	to	tea	on Fri day.
b.	<i>E lect</i>	Jill	to	high	er office. (?)
c.	<i>Rou tine</i>	stops	are	of	ten dangerous
d.	<i>Impure</i>	thoughts	dis	turbed	our pastor
e.	<i>Inside</i>	France	they	tra	velled often.

The explanation for the forms in (28) requires recognizing that the italicized configurations are not actually mismatches due to the specific sequence of stressed syllables appearing in each, namely one in which the iambic words are followed by stressed syllables. As has been discussed frequently in the phonological literature (see e.g. SPE, Liberman and Prince 1977, Hayes 1992, Kager 1996, van der Hulst in press), in these instances the rhythm rule shifts main stress leftward to a previously unstressed syllable to alleviate the “clash” created by two adjacent stresses. The actual stress patterns assigned to the tune are not the iambs of (28) but the trochaic words on the right side of (29):

(29)

		rhythm rule	
a.	<i>invite</i> Jóhn	>>	invite Jóhn
b.	<i>elèct</i> Jíll	>>	èlect Jíll
c.	<i>routine</i> stóps	>>	rouìtine stóps
d.	<i>impùre</i> thóughts	>>	ìmpure thóughts
e.	<i>inside</i> Fránce	>>	ìinside Fránce

When the conditions for the rhythm rule do not apply, the initial bisyllables revert to their original unretracted form. Thus, for example, the rhythm rule does not apply when the stressed content words in (29) are replaced by unstressed pronouns or conjunctions in normal speech. The initial words of (30) would be normally pronounced as iambs:

- (30)
- a. *Invíte* him to tea on Friday.
 - b. *Eléct* her to higher office.
 - c. *Rou tíne* but demanding duties.
 - d. *Impúre* and polluted water.
 - e. *Insíde* him there was a demon.

Given that the rhythm rule does not apply, the iambs in (30) should result in mismatches when the texts are assigned to the trochaic tune. It is somewhat surprising that the forms in (31) will not be experienced as containing fatal mismatches at these locations:

- (31)
- | | | | | | | | |
|----|-----------------|-------------------|-------|------|------|---------|-------|
| | x | | x | | x | | x |
| | x | x | | x | x | x | x |
| a. | <i>Invíte</i> | him _{+f} | to | tea | on | Fri | day. |
| b. | <i>E lect</i> | her _{+f} | to | high | er | office. | |
| c. | <i>Rou tine</i> | but _{+f} | de | man | ding | du | ties. |
| d. | <i>Impure</i> | and _{+f} | pol | lu | ted | wa | ter. |
| e. | <i>Insíde</i> | him _{+f} | there | was | a | de | mon. |

A provisional explanation¹⁶ for the acceptability of the settings in (31) implicates the interaction of metrically conditioned focus and the rhythm rule. The unstressed monosyllabic content words are promoted as stressed syllables via metrically conditioned focus (MCF). This creates a stress clash which is then subject to the rhythm rule, resulting in a final setting without mismatches. This process is represented in (32) below.

¹⁶ I am grateful to Anthony Kroch for suggesting this approach.

(32) Metrically conditioned focus (MCF) applying prior to rhythm rule (RR)

“Invite him” == sSs

x x		x x		x x
xxx		xxx		xxx
CDE	>>	CDE	>>	CDE
sSs*	(MCF)	sSs_{+f}*	(RR)	SsS _{+f}

While (32) produces the correct output it should be recognized that in at least four respects it offers a problematic account. First, the rhythm rule (its name notwithstanding) is a purely phonological principle, one which applies to normal speech in the absence of any rhythmic-musicalist context.¹⁷ It follows that the stress pattern which functions as input to the rhythm rule should not contain any pre-existing rhythmic specifications since these are not available when the rhythm rule applies in the overwhelming majority of cases, namely within normal as opposed to rhythmicized speech. Secondly, (32) incorrectly indicates the application of the rhythm rule to the focused marked sequence “Invite HIM to tea on Friday.” In fact, when the sentence is understood as normal speech, the rhythm rule does not seem to apply: “invite” appears to be pronounced in its normal unretracted form and the same appears to apply to all of the texts in (31). Thirdly, if the melody of the tune is changed very slightly by omitting the third note (n3 in (10) above) and the indicated text is assigned to this altered version, it will be noticed that “invite” can remain assigned to the strong-weak initial positions.

¹⁷ The absence of objective periodicity in linguistic performance (of the sort argued in, for example, Lehiste (1977)) now appears to be an established fact (Patel 2008, chapter 3).

(33)

x		x		x		x	
x	x	x	x	x	x	x	x
C	D	∅	C	C	D	E	C
<i>Invite</i>	∅	<i>his</i> _{-f}	<i>son</i>	<i>on</i>	<i>Fri</i>	<i>day</i>	.

However, the unaccented monosyllable “his” is now assigned to a weak metrical position and on these grounds cannot be construed as receiving metrically conditioned focus (as indicated by the subscript *-f*).

Finally, a fourth more general concern involves recognizing that the rhythm rule is a computation effected in the domain of phrasal stress not word stress. As discussed in 3.1.1, there is substantial evidence that textsetting largely ignores phrasal stress, while responding mainly to word level stress patterns. To incorporate the rhythm rule within a textsetting model would require revising this assessment, implicitly incorporated in the operative definition of stress stated in (7).

For these reasons it is preferable to reject the explanation in (32) accepting as a cost that that the initial bisyllables in (31) and also (28) remain iambs, and qualify as left mismatches under (17), rather than the trochees that would emerge had the rhythm rule applied. It will be seen in the next section that it is a relatively easy matter to designate the criteria according to which they may be viewed as exceptions.

3.5 Heavy Iambs, Universal Left Mismatches

The explanation for the acceptability of the mismatched iambs above appears to be phonological: the first syllables of the initial iambs in (28) – (31) are categorized by phonologists as heavy in that each contains either a long vowel, or the syllable is “closed”, i.e. concluded by a final consonant. The decomposition of the syllable into a

vowel (V) surrounded by a final and initial consonants (C) is shown in (34). All of the initial syllables are heavy:

(34)

	C	V	C	
a.	∅	i	n	vite (final consonant present)
b.	∅	e (ee)	∅	lect (long vowel)
c.	r	ou (uu)	∅	tine (long vowel)
d.	∅	i	m	pure (final consonant present)
e.	∅	i	n	side (final consonant present)

If heavy syllables license the forms in (31) as exceptions to the prohibition against left mismatches, it follows that iambs which contain an initial light vowel will result in unacceptable settings when they appear in the same position.¹⁸ That this is the case is indicated by the constructs below:

(35)

	x		x		x		x		x
	x	x	x	x	x	x	x	x	x
a.	Ar	rest	John	for	steal	ing	app	les.*	
b.	At	tack	dogs	can	live	in	junk	yards.*	
c.	Ra	vine	walls	sur	round	the	homestead.*		
d.	A	round	town	he	hid	the	pa	pers.*?	

The contrasting facts in (28) and (35) indicate that when the initial strong position is inhabited by a heavy syllable, a mismatch will not be perceived, even if it is unaccented and the following syllable is accented. This suggests a simple alteration to our definition of a left mismatch to require that the initial syllable be light for the mismatch to inhere:

(36) left mismatch-revised

x	
x	x
n1	n2
l	s

¹⁸ See section 3.8.3 below for discussion of exceptions to this principle.

(1 = light and unstressed.)

It will be seen that while (36) provides a resolution for the problem we are presently negotiating, there are certain exceptions which require further discussion. Before considering these, it is worth briefly entertaining the theoretically attractive possibility that the left mismatched form in (36) is symmetrical. If so, the right mismatch in (18) should be revised to require that the weak position implicated in it be occupied by a light syllable as follows:

(37) left-right mismatch: asymmetry

a. left mismatch

x		x	
x	x	x	x
n1	n2	n1	n2
l	s	<i>h</i>	<i>S</i>

b. right mismatch

	x		x
x	x	x	x
n1	n2	n1	n2
s	l	s	h

In fact, as indicated by the absence of bold text in the upper right of (37) and as (38) shows, there is an asymmetry in the behavior of left and right mismatches: the presence of a heavy unaccented syllable assigned to the strong beat in the right mismatch does not render the configuration acceptable as it does in the case of the left mismatches shown in (28).

(38)

- x x x x
x x x x x x x
a. John **often** has worked on Fri day*
 S h
- b. Now **Ethan** is home on Fri days.*
 S h
- c. John **painted** his por trait quickly*
 S h
- d. Joe's **crazy** for A lo y sius.*
 S h
- e. John's **modest** ap proach succeed ed.*
 S h

At the moment, we have no explanation for why the asymmetry in the forms shown in (37) should obtain. We mention it here as indicative of the kind of facts which emerge from a careful analysis of textsetting practice and which a theory of meter should offer an account of.

3.5.1 Light Monosyllables

A second observation to be made along these lines is based on the recognition that the left mismatches in (28), (31) and (35) all contain mismatched bisyllabic words. When the configuration implicates an unaccented monosyllable assigned to the initial strong position, as shown in 3.3, the mismatch is pre-empted by interpreting the offending syllable as functioning within a focus construction. Potentially conflicting with this is our conclusion in 3.4 that light syllables cannot occupy strong positions. The question now comes up, in the event that both of these conditions are met, what is the fact of the matter? That is, what is the behavior of light unstressed monosyllables-which are necessarily, according to the definition in (7), function words such as modals,

prepositions, determiners etc. assigned to a strong position. If the left mismatch prohibition against light syllables assigned to strong positions were enforced, texts containing the common syntactic configurations in (39) would not be able to be formed when the initial function words are phonologically light:

(39)

- | | | |
|----|---------|------|
| | x | |
| | x | x |
| | l | s |
| a. | prep. | noun |
| b. | modal | verb |
| c. | pronoun | verb |
| d. | det. | noun |

The problem with investigating this question is that there are very little empirical grounds on which to do so. This is due to the surprising fact that almost all monosyllabic function words in English are heavy. Thus, of the most common English prepositions listed in (40):

(40) most common prepositions (need reference)

at, by, down, for, from, in, like, near, of,
off, on, past, since, through, to¹⁹, up, with

all are heavy. Similarly, all English modals and auxiliaries (shall, will, do, can, may, might, etc.) are heavy as are all pronouns (I, we, you, he, she, it, they).

The two exceptions of potentially light monosyllabic function words are the determiners “the” and “a” both of which are normally pronounced with reduced vowels.²⁰

(41)

The/A barista made three lattés.

¹⁹ Morris Halle (p.c.) notes that “to” may contain a reduced vowel and is therefore light. An account of its behavior is offered in the following discussion pertaining to the determiners “the” and “a”.

²⁰ Some of these facts are discussed in Giegerich (1992).

In contrast, when the texts in (41) are assigned to a trochaic tetrameter tune, as in (42) a and b below, the vowels are more likely to be pronounced long, particularly compared to c and d.

(42)

- | | | | | | | | | |
|---|---|---|---|---|--|---|---|---|
| x | | x | | x | | x | | x |
| x | x | x | x | x | | x | x | x |
- a. *The* *ba* *rista* made three lattés.
 b. *A* *ba* *rista* made three lattés.
 c. Joe the plumber went to Ga za.
 d. Call a plumber in the morning.

It is worth pointing out that when (41) is rendered with focus applied to the initial articles in normal, unrhythmicized speech as shown in (43) the reduced short vowels also tend to become long for most speakers.

(43)

THE/A barista made three lattés.

Given that they implicate two distinct cognitive domains, it is natural to accord the facts in (42) a and b and (43) separate, independent explanations, one musical the other linguistic. As such, (42) is interpreted as mandating vowel lengthening when a syllable is assigned to a (musical) metrically accented position in a tune while in (43) vowels tend to be lengthened when they are assigned (linguistic) focus stress in normal speech.

A more theoretically attractive possibility is a unified explanation for the facts in both (42) and (43). Accordingly, just as vowel lengthening is inherited from its having been assigned to strong metrical position in (42), the imposition of focus stress results in the f-marked syllable being assigned to a strong position within a tune. According to this view, both vowel lengthening and focus stress result from the same underlying

phenomenon as in textsetting-namely assignment to strong. Of course, entertaining this possibility begs the question of the nature of the underlying “tune” for (43). For (42), the tune (i.e. Frère Jacques or some other from the trochaic tetrameter family) is explicit. For (43), we need to assume that a speaker has access to a pre-existing melody which dictates the temporal form of utterances (perhaps as rhythmic analogs to the pitch templates within the autosegmental-metrical intonation model advanced by Pierrehumbert (1980)). Or, looking at the matter somewhat differently, the speaker can be seen as deriving (i.e. composing) a tune based on the awareness of the form of the verbal utterance to which it will be assigned, possibly along the line described in the empirical study of Hayes and McEachern (1998).

We will not explore this speculative possibility here except to note that facts such as (42) and (43) seem to suggest that something of the sort must be the case. In appendix 1, we discuss other facts which seem to require a similar explanation.

3.6 Evaluating Mismatch Criteria

3.6.0 General Considerations

We introduce the following sections by restating the conclusion arrived at in the preceding section, namely that the binary mismatch criterion in (36), given again as (44) below, offers an explanation for the range of textsettings which have been presented above.

(44) Binary mismatch criterion (restated)

a) left mismatch (revised)

x	
x	x
n1	n2
l	s

b) right mismatch

	x	
x	x	
n1	n2	
S	s	

The provisional acceptance of the binary mismatch criterion, it will be recalled, was driven by the recognition that other possible candidate configurations, namely the unitary and composite mismatch criteria failed to account for certain obvious relationships of text and tune as they are experienced by listeners. In the course of discussing these, certain exceptions to (44) (the monosyllabic focus words and heavy syllables in left mismatches) emerged and were shown to be more or less naturally incorporated within it, though doing so sometimes requiring certain revisions to previously offered definitions and principles. Given that these exceptions arose somewhat haphazardly, it is reasonable to suspect that additional cases can be devised which offer exceptions and which may force revisions to (44), or possibly require abandoning it altogether. Our objective in the following will be to provide a more systematic and comprehensive empirical evaluation of (44). This data will allow us to evaluate the binary mismatch criterion, now not with reference to other possible configurations, but on its own terms.

Before we do this it is worth taking a step back to define the standards according to which it is reasonable to subject (44) or any other competing proposal to account for text setting intuitions. Without getting entangled in broad and complex meta-theoretical questions, we will assume uncontroversially that at minimum it is required of (44) that it meet two distinct sets of conditions: first, in what we will refer to as the necessity

condition, a setting manifesting either one of the configurations is required to be experienced it as mismatched. Secondly, the sufficiency condition requires that any setting experienced as mismatched, it will be found to contain an instance of one of these configurations.

These

45) Empirical Requirements of Mismatch criteria

a) Necessity Condition: Any setting meeting the conditions specified in (44) a or b is judged unacceptable

b) Sufficiency Condition: Any setting judged unacceptable must contain an instance of (44) a or b.

It follows that (44) a and b can be falsified by the appearance of two sorts of exceptions: those settings containing instances of the configurations in (44) a or b but which are experienced as acceptable, and those settings which are experienced as unacceptable but which contain no instances of (44) a or b. We refer to these as **overgeneration errors** and **undergeneration errors** respectively, or occasionally as **false positives** or **false negatives** as defined as follows:

(45) Definition of exceptions

a) overgeneration error/false positive: A setting which meets the specifications of (44) a or b but which is heard as acceptably matching stress and beat.

b) undergeneration error/false negative: A setting which is heard as containing an unacceptable match of stress and beat but which does not meet the specifications of either (44) a or b.

Like any other minimally reasonable theoretical claim, the likelihood that a particular form will offer an exception to (44) is relatively low. Thus, an arbitrary group of candidate settings drawn, for example, from the external corpus of composed vocal

works will likely not offer a sufficiently rigorous evaluation of (44). What should inform those candidate settings is not an arbitrary selection but rather a priori insight as to the type of forms which are most likely to present problems for (44). The following sections will attempt a targeted empirical evaluation of the binary mismatch criteria based on those cases which are most likely to reveal its shortcomings.

3.6.1 Revised Mismatch Criteria: Overgeneration Errors/False Positives

A hint to the class of exceptions we just alluded to involves recognizing that (44), as it currently stands, only incorporates information pertaining to the stress and metrical hierarchies relevant to language and music respectively. It contains no information indicating the constituent units of music and language, prosodic boundaries (Nespor and Vogel 1989) (hereafter **p-boundaries**). This lacuna may well be significant because it is well known that so-called external Sandhi phenomena exert a significant influence on phonological processes generally and stress in particular (Hayes 1995). Given that stress is clearly significant in determining a listener's judgments, it would be surprising if these did not exert an influence on the distribution of stressed syllables in text-tune composites.

In particular, it is reasonable to expect that the presence of p-boundaries will alter one of the configurations defined as a mismatch such that it is experienced as acceptable thus creating an overgeneration error. To unearth these instances, we will adopt a systematic approach which will involve, first, itemizing the possible locations of linguistic constituents within the mismatched configurations defined in (44) a and b and, second, constructing settings corresponding to each structural type evaluating the acceptability of each in turn.

For simplicity, we will initially incorporate into (44) only one type of p-boundary, namely the word²¹, leaving aside for the moment other constituents from the prosodic hierarchy. Furthermore, we will, for the moment, focus on the behavior of left mismatched configurations in (44) a leaving aside right mismatches which will be seen to be easily dealt with within the approach we will adopt towards the former. Within left mismatched configurations in (44) a, it is logically apparent that three possible structures obtain which is to say that the two syllables can be partitioned into 1) an entire bisyllabic word, 2) two monosyllabic words, or 3) parts of polysyllabic words on either side. Or, to describe the partitioning slightly differently, the two syllable sequence implies that a word boundary can occur in any of three locations: at the beginning, middle or end of the sequence allowing for the eight possible arrangements shown below with the word boundary indicated by |.

(47) itemization of potential word boundaries within left mismatch ((44) a)

- | | | |
|----|-------|--|
| | x | |
| | x x | |
| a. | 1 S | |
| b. | 1 S | |
| c. | 1 S | |
| d. | 1 S | |
| e. | 1 S | |
| f. | 1 S | |
| g. | 1 S | |
| h. | 1 S | |

²¹ Technical work on the prosodic hierarchy is careful to differentiate between the orthographic word, a somewhat arbitrary category, and the cognitively significant-and therefore theoretically relevant- phonological word. On those occasions where there is ambiguity, our use of term should be assumed to be the latter.

3.6.1.0 Word Boundaries/Left Mismatch

As mentioned, our strategy in the subsequent sections will be to examine the behavior of each of the configurations in (47) in turn by devising settings meeting the specifications of each. We can then determine whether the imposition of a word boundary in these locations does alter the status of the settings from unacceptable (which they are in all instances predicted by (44) to be) to acceptable. Of the configurations in (47) it will be recalled that our previous discussion pertains directly to cases a, b, and d and we will deal with these first.

3.6.1.1: (47) a and d

(47) a and d specify for their initial position | l |: a light syllable with word boundaries on each side, that is, a monosyllabic word which is both light and unaccented. But since no such cases exist (as shown in 3.4) they cannot provide counter-examples to the left mismatch criterion in (44) and we can safely ignore them as possible exceptions. It should also be mentioned that while represented as orthographically distinct, in many cases, the initial light syllable will be a function word incorporated rightward to form a polysyllabic phonological word.²² In these cases, (47) a and d will reduce to (47) b or g which we will deal with subsequently.

3.6.1.2: (47) b

It will be recalled that the revision of our initial binary mismatch criterion was precipitated by case b, | l s |, in which the two syllables within the left mismatch combine to form a single word. The examples in (35) provide evidence for taking this

²² While only some function words will allow for rightward attachment, it is uncontroversial that at least the determiners “a” and “the” should be considered as incorporated within the same phonological word as their associated nouns (e.g. the#boy, “the#man” etc. where # represents a syllabic-not a word boundary).

configuration as designating a fatal mismatch. There is, however, one potential objection to be raised against taking the settings in (35) as definitive evidence which is that each mismatch dealt with there occurs on the initial two positions of the tune. The configuration in question could also occur in three other strong-weak locations, and it is worth considering the possibility that the mismatch might be mitigated or even eliminated when it occurs elsewhere.

That this is not the case can be seen in the following:

(48)

- | | | | | | | | |
|---|---|---|---|---|---|---|---|
| x | x | x | x | x | x | x | x |
| x | x | x | x | x | x | x | x |
- a. Jas on's **arr est** made a scandal.*
 b. He told Bill to **at tack** Ja son.*
 c. He told Bill of Ja son's **ravine**.*

While these cases appear to demonstrate that the location of the left mismatch is not relevant to listeners' judgments with respect to the particular configuration, it is will be seen that in other instances it can be a factor, most conspicuously, when the mismatch occurs in the final positions of the tune. We will return to this question in 3.8.6 below.

3.6.1.3: (47) f

(47) f as shown below augments the specifications of (47) f with the context as it appears within actual settings:

(49)

(47) f. in context

- | | |
|--------|------------|
| x | x |
| x | x |
| 1 | S s |
| (47) f | |
- required successive syllable
 (must be unaccented—creates right mismatch)

Three explanatory points need to be made with respect to the augmentation of (47) f to produce (49). These are as follows: 1) a syllable is appended to the right of the sequence (47) f. 2) this syllable is unaccented. 3) The result of appending this unaccented syllable to create a right mismatch appears for the following reasons. 1) – 3) are justified as follows: With respect to 1), the absence of a word boundary on its right logically requires that there is at least one additional syllable on the right of (47) f. As for 2), this appended syllable must be unaccented (as represented by lower case s) since a word (as defined in (17)) contains only one primary stressed syllable with all others unaccented. Finally, 3) results from the subsequent unaccented syllable appearing on a strong position, thus defining a right mismatch (shown in bold face in (49)).

Based on the discussion in 3.5, we can assume that the right mismatch designated in (49) will be universally fatal. It follows that we can conclude that any setting including the configuration in (47) f. will always be unacceptable.

To further clarify and demonstrate this point, (50) a-g show seven settings containing instances of (47) f.

(50)

```

      x      x      x      x
      x x    x x x  x  x  x
a. Pa mel  a often has par ties.*
           1| s s
           (47)f \ /
                right mismatch

```

b. He sees Pame la **of ten** there.*
 1 | **S s**
 (47)f \ /
 right mismatch

c. Pa mel a **commen ted** of ten.*?
 1 | **S s**
 (47)f \ /
 right mismatch

d. Joe and Pa mel a **com men ted**.*
 1 | **S s**
 (47)f \ /
 right mismatch

e. Pa mel a **ra tion** a lized it.*?
 1 | **S s**
 (47)f \ /
 right mismatch

f. Pa mel a's **comments** annoyed him.*?²³
 1 | **S s**
 (47)f \ /
 right mismatch

g. That's what Pamela's **comment** was.*
 1 | **S s**
 (47)f \ /
 right mismatch

Example a contains a fatal right mismatched trochee similar to the mismatches in (38). b shows (in analogous fashion to (48) b and c) that the same mismatch remains fatal when it is moved to the other possible positions where it can potentially occur. Examples c, d and e show that the behavior of initially stressed tri- and quadrisyllabic verbs in inducing right mismatches is similar to that of a and b; f and g show that mismatches can be induced by nouns in addition to adverbs and verbs responsible for mismatches in a – e. The variety of examples provides conclusive evidence that (47) f will always result in a

²³ An explanation for the ambiguous forms in (50) is offered in section 3.8.8 below.

fatal mismatch, and that in this respect it behaves identically to (47) b where the word boundaries enclose, rather than occur within, the left mismatch.

3.6.1.4: (47) g

It will be seen that (47) g can be dispensed with on similar grounds to those of (47) f. Since the word boundary now occurs to the left of the configuration in question, again a subsequent syllable is required in the setting and, as a consequence of it being included within the same word as the syllables on its left, it must be unaccented as shown below:

(51)

```

  x      x
  x x   x
  | l S s
  47g |
      subsequent syllable
      (must be in same word as preceding S, therefore
      unaccented, s)
```

(e.g. “arrested”, “attacking”, “adeptly”,
“appearance” etc.)

Any Frère Jacques setting containing these forms (e.g. “Johnson arrested Johanna.”, “Mary was not attacking John.” “John had adeptly replaced it.” “Appearances are deceptive.”) will also be rejected by listeners. We will not provide transcriptions of these texts applied to the tune as it should be apparent that they are quite unnatural along similar lines to the settings in (50).

3.6.1.5: (47) h

In contrast to the other forms in (47), h indicates a complete absence of word boundaries. Because of the lack of a word boundary on the right, all settings containing this configuration can be ruled out by similar arguments used to exclude f and g. These are shown in (52): the accented syllable within (47) h requires that unaccented syllables

occur both to its left and its right. The unaccented syllable on the immediate right requires that a right mismatch inheres at this location indicated in bold face below:

(52) required context for (47) h

```

      x   x
     x x x x
a.  s l s s   (e.g. "Arizona", "constitution")
      \ /
      (47)h.

```

Any setting containing the configurations in (52) a or b will sound quite unnatural. As an exercise in what might be called negative textsetting, we will leave the composition of these to the reader.

3.6.1.6 (47) e

The form specified in (47) e will be seen to be the mirror image of (51) in that the word boundary occurs to the right rather than to the left of the prospective left mismatch. It follows that preceding syllable must belong to the same word as those within the mismatch and, to avoid a right mismatch, it must be unaccented as shown below:

(53) ((47) e in context)

```

      x
     x x
    s l s |
      |   47e
      |
required syllable to left
(must be unaccented)

```

In contrast to (51) whose deviance can be explained by the presence of a right mismatch formed by the addition of a syllable to the right, the unacceptability of (53) cannot be explained on a priori grounds. Rather, that settings containing (53) are deviant appears to be an empirical fact as shown below.

(54) Instances of (47) e

- | | | | | | | | | | |
|----|------|------|-----------|--------------|-----------|---------------|--------------------|---------------|---|
| | x | | x | | x | | x | | x |
| | x | x | x | x | x | x | x | x | x |
| a. | John | re | ar | rests | Bill | for | stea | ling.* | |
| b. | Eng | land | had | re | at | tacked | Belgium.* | | |
| c. | Mar | y | sank | E | laine's | sub | ma | rine.* | |
| d. | Cae | sar | left | the | town | un | destroyed.* | | |

Again, the forms in (47) e result in fatal mismatches and they do so wherever they appear within the phrase.

Given that all of the settings we have constructed so far based on the forms in (47) are unacceptable, we are close to having shown, contrary to our expectation, that linguistic constituency exerts no significant influence on the mismatched forms predicted by (44). That is, we have found no instance where a listener's alteration in judgment from unacceptable to acceptable is induced by the presence of a word boundary. Textsetting appears to be sensitive only to the matching of stress and beat, regardless of the location of musical and/or linguistic constituent boundaries.

3.6.1.7 (47) c

The forms in (47) do, however, offer one clear exception to this principle, namely (47) c, which contains a light syllable followed by an accented syllable enclosed within a word boundary as shown in the following:

(55) (47) c

x	
x	x
l	s

As before, we can logically infer the structure of the surrounding syllables. 1), given that there is no word boundary on the left, at least one syllable must precede the appearance of (47) c. 2), this syllable must be unaccented to avoid a right mismatch. 3), given that

its second syllable is enclosed in a word boundary, there may or may not be a syllable following the occurrence of (47) c. 4) If this syllable does occur, avoiding a fatal right mismatch requires that the syllable in question must be accented. (56) shows the sequence which meets these specifications of 1) – 4) with the optional final syllable mandated by 3) indicated by parentheses:

(56) Required context of (47) c.

```

      x      x
     x  x  x  x
    s  l  |S| (S)
        \  /
      (47)c

```

(57) shows three instantiations of the configurations specified by (56).

(57)²⁴

```

      x      x      x      x      x
     x  x  x  x  x  x  x  x
a. Pa  me  la  bought John's Toy  o  ta.
        l  | S |

b. John said Pa  me  la  bought rai sins.
        l  | S |

c. John said Bill thought Pa  me  la left.*
        l  | S |

```

²⁴ While it is not required to be, in most instances, the syllable to the left of the configuration specified in (56) will be accented resulting in the following sequence:

```

     x      x
    x  x  x  x
   S  s  l  |S|
        \  /
      47 c

```

This is an instance of a **choriambic foot**, or **choriamb** as defined in, for example, Saintsbury's (1926) glossary. We will employ this term when we return to these forms in 3.9.

Unlike all of the preceding cases, (57) a and b are heard as acceptable, contradicting (44) a which incorrectly predicts them to be unacceptable. However, as can be seen in (57) c, when the mismatch appears on the right edge of the phrase, it is judged by most listeners as unacceptable; as such, (44) a makes the correct prediction in this instance.

(57) a and b would seem to indicate that the absence of a boundary on the right of the mismatched pair pre-empts the imposition of the left mismatch criterion. This observation, however, begs the question of whether the boundary implicated is musical or linguistic. Since both a (musical) grouping and (linguistic) sentential boundary are present at the right edge²⁵, it is unclear which is responsible for the behavior manifested in (57).

²⁵ It should be noted that the musical grouping boundary is an inherent characteristic of the tune to which the text is assigned and will always occur following the eighth note- (note n8 in (10)). In contrast, the sentential p-boundary is not an inherent characteristic of the text-tune composite but rather a likely consequence of the requirement that prosodic constituents and musical groups are sufficiently well aligned. See the discussion of constituency matching in 1.3 and references therein.

3.7 P-Boundaries and Metrical Form: General Observations

Insight on this question is provided by the examples in (58).

(58)

	x		x		x		x	
	x	x	x	x	x	x	x	x
a.	Pam	e	la	died _s ,	John	re	por	ted.*
b.	"Tues	day,	Pa	me	la	died _s ,	" Jay	said.*
c.	Tues	day,	Pa	me	la	left _w	Will	iam.
d.	"Pam	e	la	left _w	Bill."	said	Am	y.*
e.	"Pam	e	la	left _s ."	Bill	told	Am	y.*

(58) shows the two non-final positions where the configuration (47) c can potentially appear.²⁶ Since a major musical grouping boundary does not appear on the right of these locations²⁷, we can conclude that it is the presence of a linguistic and not a musical boundary which is responsible for (44) a applying.

As for what type of linguistic (i.e. prosodic) boundary, the answer so far seems clear: when there is a sentential boundary to the right of the mismatched configuration (indicated by |_s) (44) a. applies normally. When the boundary is absent due to the presence of a direct object of the verb, as in (57) a and b and (58) c and d, only a word boundary (|_w) is present and this is insufficient for (44) a to apply.

Insofar as it is clear that a sentential boundary is sufficient and a word boundary is not sufficient to pre-empt the operation of (44) a, we need to investigate the behavior of an intermediate level of linguistic constituency. Again, we reference the prosodic hierarchy in defining the appropriate constituent unit, namely the clitic group (CG) which defines constituency by attaching prosodic words to adjacent clitics on the left and/or right. It is beyond our scope to evaluate proposals in the literature relating the prosodic

²⁶ Because of the absence of a word boundary on the left of (47) c, it follows that it may not appear at the left edge of the phrase.

²⁷ A minor grouping boundary does occur after the fourth note (note n4 in (10)). That no boundary occurs after note n6, relevant to cases b and c, is indisputable.

and syntactic hierarchy. For now we will assume that “panama” and “hat” in (59) group together to form (syntactic) object and subject determiner phrases (Carnie 2004) and on this basis the same boundary delimits (prosodic) CGs.

(59)

	x		x		x		x
	x	x	x	x	x	x	x
a.	John	wore	Pa	na	ma	hats	_{CG} glad ly.*
b.	Pa	na	ma	hats	_{CG} wear	out	quickly.*

That both a and b are unacceptable indicates that CGs group with sentential constituents in (58) in licensing the application of the left mismatch criterion. In contrast, less prominent constituents, namely the word boundary, will prevent (44) a from applying thus rendering the sentence acceptable.

3.7.1 P-Boundary Induced Demotion

We are now in a position to formally represent the two situations just described.

(60)

a.	
	x
	x x
	l _w S _{≥CG}
b.	
	x
	x x
	l _w S _{<CG}

Two explanatory comments need to be made on (60). First, the relative prominence of linguistic/prosodic boundaries is indicated by the arithmetic symbols \geq and $<$.

Accordingly, the former states, in reference to (60) a that a prosodic constituent equal to or higher than that of the clitic group occurs on the right edge of the mismatched unit,

while in b the constituent boundary is drawn from a lower level of the prosodic hierarchy, effectively, a prosodic word.

Secondly, the bold faced type, it should be clear, indicates not a negative prediction of the mismatch criteria with respect to the two forms, but rather the empirical fact that listeners judge (60) a to be mismatched. The problem we are confronting is that both (60) a and b meet the specifications of (44) a. Thus, what is needed is to prevent the left mismatch criterion from applying in the latter instance and not the former.

The simplest means to accomplish this is to mark (60) as a specific exception to the application of (44) a. The cost of adopting this ad hoc solution, however, as will be seen momentarily, is that additional ad hoc exceptions will need to be advanced for other forms not correctly predicted by the mismatch criteria specified in (44).

A preferable solution involves not a constraint on the application of (44) but an alteration of the underlying forms such that they do not meet the existing specifications of (44) when its application would result in an incorrect prediction. Specifically, what is required is to alter the textual sequence such that the accented monosyllable is demoted to unaccented status when it appears within a phrase while maintaining its original degree of accent when no phrasal boundary intercedes. This alteration is accomplished withing the Monosyllabic Demotion Rule (hereafter MDR) defined as follows:

(61) Monosyllabic Demotion Rule (MDR)

Given the text tune configuration

$$\begin{array}{c} \mathbf{x} \\ \mathbf{x} \quad \mathbf{x} \\ \mathbf{l} \quad |_{\mathbf{w}} \mathbf{S} \quad |_{\mathbf{x}} \end{array}$$

where $x < CG$, $S \rightarrow s$.

In informal terms, the process described in (61) involves the listener scanning the text under consideration for a prosodic boundary on the right of a left mismatch. If none is present, then stress on the right syllable is reduced and the setting is judged acceptable. If, on the other hand, a relatively prominent prosodic boundary occurs at this location, then stress is maintained, with the result that the mismatch is judged fatal and the setting is heard as unacceptable.

3.7.2 Application of Monosyllabic Stress Demotion

(62) shows the functioning of (61) with respect to (58) d and e.

(62)

a.

```

      x   x           x           x
      x x x   x     x           x x x
  "Pamela|_w left|_w Bill." said Amy.
      S s l   S

```

[(61) applies: S demoted to s.]²⁸

```

  "Pamela| left|_w
      S s l   s

```

[Neither (44) a nor b apply. Setting acceptable.]

```

      x   x           x           x
      x x x   x     x           x x x
  "Pamela|_w left|_w Bill." said Amy.
      S s l   s

```

²⁸ Justification for the sequence of steps involved in the derivation of the predicted forms will be shown in square brackets here and in the forthcoming examples.

b.

```

  x   x           x           x
  x x x   x       x   x   x x
  "Pamela|wleft.|s"Bill told Amy.
  S s l   S       s
```

[Fails to meet condition (61) a. Initial setting unaltered:]

```

  x   x
  x x x   x
  "Pamela|left|w
  S s l   S
```

[(44) a. applies; setting unacceptable.]

```

  x   x           x           x
  x x x   x       x   x   x x
  "Pamela|wleft.|s"Bill told Amy.*
  S s l   S
```

It will be found that the remaining texts in (58) are also assigned the correct prediction by applying the MDR prior to the mismatch criteria.

3.7.2 Mismatched Heavy Syllables and Prosodic Boundaries

Before concluding our discussion of overgeneration errors, it should be noted here that the same pattern of mismatches shown in (62) a and b is found in the examples below:

(63)

```

  x   x           x           x
  x x x   x       x   x   x x
  a. "Jonathan|left|w Bill." said Amy.
     S s h   S
  b. "Jonathan|left."|s Bill told Amy. *
     S s h   S
```

The problem for the derivational approach outlined above is posed by (63) b. Due to the presence of the heavy syllable it is exempted from the left mismatch criteria and is therefore predicted to be acceptable but because it is experienced as mismatched, it is a

direct counterexample to (44) a. Unlike the previous instances of (47) c discussed above which were perceived as acceptable while predicted to be unacceptable (63) b constitutes an undergeneration rather than overgeneration error. We deal with this class of exceptions in the next section.

3.8.0 Left Mismatches: Undergeneration Errors

As mentioned earlier, and as just demonstrated, in addition to designating as mismatches forms which are judged by most listeners as acceptable, (44) is likely to undergenerate. By this we mean that certain strong weak pairs judged unacceptable will manifest no instances of the structures specified in (44). Our strategy for ferreting out these undergeneration errors will be the inverse of that used to unearth overgeneration errors in the preceding section: the latter called for generating a variety of settings meeting the conditions of the left mismatch structure. If acceptable, these forms would constitute exceptions to (44) a. For undergeneration errors, what is required is to enumerate the possible forms other than those defined in (44). Those which are unacceptable constitute undergeneration errors along the lines just mentioned.

Before we undertake this strategy, we enter as an introductory caveat that our approach will be limited to alternatives to the forms designated as unacceptable by the left mismatch criteria. This means two things: First, we will focus on alternate binary sequences which is to say that we are looking for two syllables other than the light and accented sequence specified by (44) a which result in a perceived mismatch. It may be the case that mismatches could be a property of some other number of mismatched positions, though based on arguments made in 3.2, we can reject the possibility that

mismatches are a property of one or three positions. It also seems unlikely that mismatches would implicate four or more syllables.

Secondly, we will restrict ourselves to enumerating the possibilities of two syllables heard as mismatched due to their assignment to a strong-weak sequence rather than a weak-strong sequence which is to say that we will be attempting to derive exceptions to the *left* mismatch (44) a and not the *right* mismatch criterion (44) b. A rigorously systematic approach would examine both types. Our justification for not doing so is based on the recognition that the right mismatch implicates a less restrictive category than the left mismatch, i.e. any type of unaccented syllable (s) rather than the light syllable (l). It is less likely that the right mismatch criterion will fail to account for these thus the class of exceptions (if it exists) will necessarily be smaller than those applicable to the left mismatch criterion. Furthermore, it will be seen that several right mismatch undergeneration errors will come up in the context of dealing with left mismatched configurations. We will have something to say about these as they arise.

3.8.1 Itemization of Potential Cases

Our approach will begin by restating (44) a in a sufficiently general form so that the possible class of exceptions can be easily itemized.

(64) Left mismatch ((44) a restated)

x	
x	x
a	b

where both a = l and b = S.

It follows that undergeneration errors can be defined as any pair of syllables (a,b) assigned to a strong-weak sequence which are not respectively light and accented. That is,

(65) undergeneration error (redefined)

x	
x	x
a	b

where either $a \neq l$ or $b \neq S$.

We can now itemize all possible forms which undergeneration errors can (in principle) be implicated in. These consist of, for the initial position a, three syllable types shown to relevant to left mismatches: accented (S), light (l) and heavy (h). The second position b can be occupied by two syllable types: S and s. This makes available six possible pairs, with one of these, (64), that already defined as a mismatch. Thus, undergeneration errors with respect to (44) a can take the five possible forms in (66):

(66) potential exceptions to (44) a: undergeneration errors

x	
x	x
a	b

where

- a. $a = h, b = S$
- b. $a = h, b = s$
- c. $a = l, b = s$
- d. $a = S, b = S$
- e. $a = S, b = s$

It should be apparent that of the five sequences in a- e, a is the most likely to induce a mismatch in that it, unlike the others, contains a stress inversion: an unaccented-

accented syllabic sequence assigned to adjacent weak and strong positions. Indeed, this was included in our initial formulation of the left mismatch in (17) before we revised it to take the form (44) a, for reasons discussed in 3.5.

3.8.2 Undergeneration Errors: Uniqueness Requirement

Before we examine the behavior of the forms in (66), it is necessary to note that to qualify as an undergeneration error, the mismatch in question must be unique which is to say that the setting must not contain instances of defined mismatches at locations than other than where the potential mismatch occurs. For example, it will be seen that “routinely” in the following cannot be definitively categorized as a left mismatch undergeneration error, despite its containing an instance of the configuration in (66) a.

(67)

x		x		x		x
x	x	x	x	x	x	x

John was **routinely** excluded.*

h	s	s
\ /	\ /	

(66) a (44) b

Rather what is more likely to be responsible for (67) being heard as unacceptable is the right mismatch defined in (44) b as all instances of these we have observed so far have been unambiguously awkward and unnatural. But since a listener only has access to the intuition that the setting in question as a whole is unnatural it is impossible to know for certain which configuration is responsible for the negative judgment. We are therefore required to be reject (67) and all settings which contain more than one potential mismatch as a definitive example of an undergeneration error.

What is required to investigate undergeneration errors is an unacceptable setting in which one of the configurations in (66) appears in the absence of other mismatches.

The following settings will reflect the recognition that these potentially ill-formed configurations in question must be unique-that is, they must appear in the absence of any other actual or potential mismatches.

While the uniqueness condition is not directly relevant to the central topic of this following sections-namely the form of the configurations constituting undergeneration errors-we mention it here as a formal limitation on the empirical data which can be appealed to in determining the class of forms which can be categorized as exceptions to (44) a.

3.8.3 Heavy and Light Syllables: A Closer Look

The first class of exceptions of the type identified in (66) a have been noted by phonologists in the context of investigating the behavior of the rhythm rule. While it is for the most part accurate that stress retraction is blocked by the presence of an initial light syllable, as pointed out earlier this is only a general tendency rather than a universal fact. This is shown by the examples in (68):

(68)

- a. succéss + trap > succéss trap (not súccess trap)
- b. extrémé + measures > extrémé measures (not éxtreme measures)
- c. idéa + man > idéa man (not ídea man)
- d. absúrd + picture > absúrd picture (probably not ábsurd picture)
- e. obése + cat > obése cat (probably not óbese cat)

As was the case for the unretractable forms in (35), the same iambic forms which resist stress retraction also resist assignment to adjacent strong weak positions.

(69)

- | | | | | | | | |
|--|---|---|---|---|---|---|---|
| | x | | x | | x | | x |
| | x | x | x | x | x | x | x |
- a. **Suc cess** made him more re clu sive.*
b. Bill said **ex treme** speed is dangerous.*
c. Ga ry pain ted **ab surd** pictures.*
d. John re jec ted Mar y's **i dea**.*
e. Mar y's child is grossly **ob ese**.*

(69) a.- e. contain instances of the undergeneration error (66) a in that the initial vowel is taken to be heavy according to the definition in (34). There is reason to suspect, based on Kager (1989), that a reformulated definition of the relevant phonological unit (i.e. the light syllable) will explain the failure of stress retraction in (68). It is reasonable to assume that the same explanation will carry over to account for the textsetting facts in (69) and if so, these will no longer be classified as exceptions to (44) but instances of left mismatches under an appropriately revised definition of syllable weight. Since the solution to this problem is squarely within the realm of phonology, it is beyond our scope here, and we will be put it aside in anticipation that future work will provide the tools to resolve it.

3.8.4 Monosyllabic Stress Demotion (revised)

We now return to the problem posed by the setting in (63) b. As we noted, this constitutes an instance of an overgeneration error (namely, that specified by (66) a) in that a heavy rather than light syllable appears in the strong position in question and on this basis, is wrongly predicted by (44) a to be acceptable. (70) below contains additional examples of the same configuration:

(70)

- | | | | | | | | | | |
|----|------|------|-------------|---------------|-------------|---------------|-------------|------------------------|----|
| | x | | x | | x | | x | | x |
| | x | x | x | x | x | x | x | x | x |
| a. | Tra | vel | ers' | checks | cost | a | dol | lar.* | |
| | | | h | | S | | CG | | |
| b. | Ma | ry's | tra | vel | ers' | checks | van | ished.*? ²⁹ | |
| | | | | | h | | S | | CG |
| c. | John | took | Mar | y's | tra | vel | ers' | checks.* | |
| | | | | | h | | S | | CG |

These cases can be handled by shifting the location of the syllable implicated by demotion to the second rather than the first accented syllable when only a word boundary intercedes between the first two syllables. The following revision to the monosyllabic demotion rule (hereafter MDR) accomplishes this:

²⁹ An explanation for the perceived ambiguity of this setting will be provided in section 3.8.8 below.

(71) Revised Monosyllabic Demotion Rule (MDR) (c.f. (61))

Given the text tune configuration

$$\begin{array}{cc} \mathbf{x} & \mathbf{x} \\ \mathbf{x} & \mathbf{x} \\ \mathbf{s} & |_{\mathbf{w}} \mathbf{S1} |_{\mathbf{x}} \mathbf{S2} \end{array}$$

a) where $x \geq \text{CG}$: $\mathbf{S2} \rightarrow \mathbf{s}$.

b) elsewhere: $\mathbf{S1} \rightarrow \mathbf{s}$.

We deal first with case a) where a sufficiently prominent p-boundary intervenes between two accented syllables. Since right mismatches, unlike left mismatches, do not depend on the weight of the unstressed syllable, (71) treats the forms in (59) and (70) as equivalent, with the result that both are rejected by virtue of the right mismatch created by the demoted accented syllable.

(72) demonstrates the application of (71) to the two settings in question.

(72)

$$\begin{array}{cc} \mathbf{x} & \mathbf{x} & & \mathbf{x} & & \mathbf{x} \\ \mathbf{x} & \mathbf{x} & \mathbf{x} & & \mathbf{x} & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ \text{a. Travelers' } & \mathbf{checks} & \mathbf{cost} & \text{a dollar.*} \\ & \mathbf{h} & | & \mathbf{S1} |_{\text{CG}} & \mathbf{S2} \end{array}$$

$$\begin{array}{cc} \text{b. Pa nama} & \mathbf{hats} & \mathbf{cost} & \text{a dollar.*} \\ & \mathbf{l} & | & \mathbf{S1} |_{\text{CG}} & \mathbf{S2} \end{array}$$

[a and b meet conditions of (71). (71) a applies $\mathbf{S2} \rightarrow \mathbf{s}$.]

$$\begin{array}{cc} \mathbf{x} & \mathbf{x} & & \mathbf{x} & & \mathbf{x} \\ \mathbf{x} & \mathbf{x} & \mathbf{x} & & \mathbf{x} & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ \text{a. Travelers' } & \mathbf{checks} & \mathbf{cost} & \text{a dollar.*} \\ & \mathbf{h} & | & \mathbf{S1} |_{\text{CG}} & \mathbf{s} \end{array}$$

$$\begin{array}{cc} \text{b. Pa nama} & \mathbf{hats} & \mathbf{cost} & \text{a dollar.*} \\ & \mathbf{l} & | & \mathbf{S1} |_{\text{CG}} & \mathbf{s} \end{array}$$

[Right mismatch 44 b applies: settings unacceptable.]

Conversely, (71) a fails to apply when the constituent boundary is insufficiently prominent, for example when the two adjacent stressed syllables occur within the same verb phrase, as in examples (58) d and (63) b. (73) shows the application of (71) to these settings:

(73)

$$\begin{array}{cccccccc} & x & & x & & & & x & & & x \\ & x & x & x & & x & & x & & x & x & x \\ \text{a. "Jonathan} & | & \textit{left} & | & \textsubscript{w} & \text{Bill.} & \text{"} & \text{said} & \text{Amy.} & & \\ & & h & | & S1 & | & S2 & & & & \end{array}$$

$$\begin{array}{cccccccc} \text{b. "Pame} & \text{la} & | & \textit{left} & | & \textsubscript{w} & \text{Bill.} & \text{"} & \text{said} & \text{Amy.} & \\ & & \mathbf{l} & | & \mathbf{S1} & | & S2 & & & & \end{array}$$

[Does not meet conditions for (71) a. (71) b applies. S1 -> s.]

$$\begin{array}{cccccccc} & x & & x & & & & x & & & x \\ & x & x & x & & x & & x & & x & x & x \\ \text{a. "Jonathan} & | & \textit{left} & | & \textsubscript{w} & \text{Bill.} & \text{"} & \text{said} & \text{Amy.} & & \\ & & s & | & s & | & S2 & & & & \end{array}$$

$$\begin{array}{cccccccc} \text{b. "Pame} & \text{la} & | & \textit{left} & | & \textsubscript{w} & \text{Bill.} & \text{"} & \text{said} & \text{Amy.} & \\ & & s & | & s & | & S2 & & & & \end{array}$$

[Left mismatch removed in b. Both settings acceptable.]

(71), in combination with the mismatch criterion in (44) a and b successfully accounts for the forms which were correctly and incorrectly predicted to be mismatched by (61). On this basis, it must be taken as preferable to (61). To briefly review the relevant cases, (61) undergenerates in that it predicts that the form (66) a will be in all cases acceptable. As shown above, it is not, and (71) provides the mechanism to exclude this and other settings of its type.

While it is apparent that (71) is required to be preferred to (61) on empirical grounds, it should be recognized that it is descriptively less than optimal in that it assigns a somewhat counter-intuitive structure in functioning to exclude the unacceptable settings

in question. Most conspicuously, this shortcoming is revealed in the forms in (72) a and b: Both are indeed unacceptable, however, it would appear that most listeners would find that the defect inheres not in the right mismatched location taken as fatal (“left Bill”), but rather at the left mismatch location (“-than left” and “-la left”) which is only fatal in b. At this juncture, it will suffice to note this misgiving a more satisfactory account of which we will offer in 3.9 below.

3.8.5 Unaccented Trisyllabic Sequences (Tribrachs)

We now turn to a second type of undergeneration error from the list in (66), namely those settings which contain three adjacent unaccented syllables. These sequences, referred to in Saintsbury’s (1926) glossary as tribrachs, are of two types: those in which the initial syllable of the pair is heavy ((66) b) or in which it is light, ((66) c). It will be seen that we may ignore the syllable weight distinction, as either heavy or light syllables can result in settings which are quite egregiously unnatural. Some of these can be seen in the following:

(74) (66) b and c (heavy = h, light = l)

- | | | | | | | | | | |
|----|-----|-----|---------|------|-----|-------------|------------|----------------------------|---|
| | x | | x | | x | | x | | x |
| | x | | x | | x | | x | | x |
| a. | Jim | was | stopped | from | pil | fer | ing | it. ³⁰ * | |
| | | | | | | h | | | |
| b. | Jim | had | picked | the | Ar | a | bic | one. * | |
| | | | | | | l | | | |
| c. | Jim | was | sure | he'd | ra | tion | al | ized. * | |
| | | | | | | h | | | |

³⁰ The final syllable appears in bold faced type based on the possibility-as yet to be determined –that a right mismatch may inhere in this setting and account for its unacceptability. For this reason, here, and in subsequent settings, we will represent this position in bold.

in the existing grids. At the same time, it should be noted that the grids we have made use of so far are, from a formal point of view, incomplete. In fact, as discussed in 2.1 above, two additional supertactus levels are required to meet the well formedness conditions applicable to the geometry of the grid.

Conformity to this requirement dictates the four possible complete grids shown below assigned to the eight locations of the first phrase of the Frère Jacques tune:

(77) Complete grids (supertactus levels in parentheses).

a.

					(x)			L(2)
		(x)			(x)			L(1)
x		x		x		x		L(0)
x	x	x	x	x	x	x	x	L(-1)
C	D	E	C	C	D	E	C	

b.

		(x)						L(2)
		(x)			(x)			L(1)
x		x		x		x		L(0)
x	x	x	x	x	x	x	x	L(-1)
C	D	E	C	C	D	E	C	

c.

(x)								L(2)
(x)			(x)					L(1)
x		x		x		x		L(0)
x	x	x	x	x	x	x	x	L(-1)
C	D	E	C	C	D	E	C	

(80)

						(x)	
		(x)				(x)	
x		x		x		x	
x	x	x	x	x	x	x	x
{C	D	E	C}	{C	D	E	C}}

It should be noted, however, that the grid in (80) is not strictly metrical in that it represents both structural and metrical accentuation. As such it might be referred to as a **composite grid** with the understanding that it is not necessarily more broadly applicable to other musical structures, including, for example the Mozart A major sonata shown in (78), (for reasons discussed in GTTM).

While it is questionable whether it is appropriate to conflate structural and metrical accent within musical form generally, in the case of certain simple songs viewing the rightmost elements of groups as inherently structurally accented allows us an explanation of textsetting facts which would otherwise go unexplained. In particular, it will be noticed that appealing to a composite accent (indicated by parentheses) corresponding to the rightmost tactus level grid positions within groups allows us to make a choice among the well formed grids in (77), namely (77) a. Furthermore, this choice allows for the representation of the deviant settings in (76) as mismatched:

(81) composite grid: structural accents parenthesized

						(x)		L(2)
		(x)				(x)		L(1)
	x	x		x		x		L(0)
	x	x	x	x	x	x	x	L(-1)
a.	Jim	was	stopped	from	pil	fer	ing	it.*
b.	Jim	had	picked	the	Ar	a	bic	one.*
c.	Jim	was	sure	he'd	ra	tion	al	ized.*
d.	Pam	ex	pects	a	va	can	cy	soon.*
e.	John	de	clared	his	can	did	a	cy.*
f.	Tim's	re	duced	to	pen	ni	less	ness.*

If we ignore the interceding grayed-out syllable, the bold faced syllables shown in (81) indicate right mismatches inhering between the stressed and unstressed syllables and levels L(1) and L(2) of the composite grid.

(82)

	(x)	L(2)
	(x)	L(1)
x	x	L(0)
x	x x	L(-1)
S	s s	

e.g. **pilfering**

In contrast, the acceptable settings in (75) do not contain mismatches.

(83)

		(x)	L(2)
	(x)	(x)	L(1)
x	x	x	L(0)
x	x	x x	L(-1)

a. Jim was *pilfering* them dai ly.

b. Jim took *Arabic* lin guistics.

c. John had *rational* ized go ing.

d. Pam said *vacancies* are fre quent.

e. Johnson's *candida* cy faltered.

f. John said *peniless* ness saved him.

3.8.7 Extrasyllabic Marking Rule (EMR)

Given the representations of the facts in (81) and (83), it should be clear what is now required of the theory, namely, a mechanism to remove from view the second syllables of the words implicated in the mismatches such that only the two syllables appearing on levels L(0) and above are “visible” to 44 b. We will refer to the locations which contain these excised syllables as **extrasyllabic**, replacing the grayed out locations in (81) and (83) with parentheses surrounding the syllable assigned to the position in question, when such an extrasyllabic location is identified. Marking these

locations as extrasyllabic will have the effect of ruling out the settings in (81) as normal right mismatches while registering as properly matched the acceptable forms in (83).

Our approach to accomplishing this involves two steps. First we define the syllables represented in gray in (83) as extrasyllabic by means of the following definition.

(84) Definition: Extrasyllabic events (s) are invisible to mismatch criteria in (44).

Next, we specify the circumstances according to which a location is to be marked as extrametrical:

(85) Extrasyllabic marking rule³¹ (EMR)

For any sequence of three or more unaccented syllables within a major prosodic group (i.e. a clitic group or above), mark as extrasyllabic the first member of the sequence.

(86) a and b shows the application of (84) and (85) to “pilfering him” in (81) and (83) a.

(86)
a.³²

(x)		(x)	L(2)
(x)		(x)	L(1)
x x		x x	L(0)
x x x	x	x x x x	L(-1)

“pilfering it” >> “pil(fer)ing it*
(EMR)

³¹ More formally, we can state the rule as
Given $s_1 s_2 \dots s_n |_x$ where $n \geq 3$ and $x \geq CG$, $s_1 \rightarrow (s_1)$ (as defined in (84)).

³² It should be mentioned that the EMR will not apply when the final element of the unaccented sequence is not followed by a major p-boundary. If correct, it follows that,

x	x	x	x	L(0)
x	x	x	x	L(-1)

Pilfering the goods was le gal.

will be judged to be acceptable compared to the mismatched settings in (81) and (83) a. This appears to be the case but should be confirmed empirically.

b.

(x)		(x)		L(1)
x	x	x	x	L(0)
x	x	x	x	L(-1)

“pilfering them” >> “pil(fer)ing them” (no mismatch)
(EMR)

3.8.8 Remaining Exceptions

Before concluding this section, two loose ends merit brief attention. First, we note here that the cases in (76) where the mismatch appears at the left boundary are correctly predicted to be unacceptable by applying the composite grid to the “Frère Jacques” tune as proposed above. However, as shown in (87) below, since the mismatch inheres between L(1) and L(0), as opposed to, in (81), the more prominent grid location L(2) and L(0), the perception of a stress beat clash is somewhat minimized. This is consistent with listeners’ judgments which, as indicated by “?” in the following tend to be somewhat less clear-cut than those in relation to (81).

(87)

				(x)		L(2)
	(x)			(x)		L(1)
x	x		x	x		L(0)
x	x	x	x	x	x	L(-1)

a. **Pilfer** ing it was il le gal.*?
b. **Ar** ab ic **ones** are ex pensive.*?
c. **Ration** al **iz** ing was ea sy. *
d. **Va** cancies **are** ver y frequent.*?
e. **Candi** da **cies** are encouraged.*
f. **Penni** less **ness** is too common.*?

Finally, scholarly objectivity requires us to mention a class of possible exceptions to the explanation proposed above:

(88)

	x		x		x		x		x
	x		x		x		x		x
a.	Tim	per	formed	the	ce	re	mon	y.	?
b.	Mar	y	used	to	ba	by	sit	him.	?
d.	Joe	had	failed	to	agg	ra	vate	him.	?
e.	E	than	could	not	casti	gate	him.	?	
f.	Who	did	John	send	Hender	son	to	?	

All of these contain instances of the adjacent pair of unstressed syllables denoted in (66) b or c and meet the conditions of extrasyllabic marking rule, but are not obviously experienced as mismatched, as they are predicted to be within our system.

Two facts are worth noting with respect to the forms in (81) and (88). First, those in (81) tending to be experienced as mismatched are morphologically compositional (Koopmans, Sportiche, Stadler, forthcoming) in that they contain definable suffixes appended to a root (e.g pilfer (+ing), Arab (+ic), rational (+ized), vacan(t) (+cy), candida(te) (+cy)). The forms in (88), in contrast, are morphologically idiomatic: “ceremony” has no apparent morphological decomposition, whereas although “babysit” can be decomposed, its connection between a verb “sit” affixed to a root noun “baby” is not apparent. This may have some effect on the underlying stress pattern resulting in the dilution of what would otherwise constitute an unproblematically perceived mismatch.

Another possible explanation implicates the well known fact that two (or sometimes more) pronunciations are available to the speaker for many English words. Among these are the words “recognize”, “castigate” and “aggravate” which while not normally stressed on the final syllable in most American English dialects, can be in certain British English dialects (Wells 2000). The availability of this alternate form may create the potential for ambiguity in the listener’s judgment with respect to the word’s assignment to metrical form.

3.9 Interaction of Extrasyllabic Marking and Metrical Demotion Rules.

In the preceding sections we have posited two distinct principles to account for the exceptions to the left mismatch criterion in 44 a. These were **demotion** which has the effect of reducing stress on the syllable following a relatively major prosodic group and **extrasyllabic marking** which results in non-adjacent syllables being made visible to the mismatch criteria.

As these two transformational rules were developed as independent solutions for specific exceptions to the mismatch criteria, it escaped notice that they have the potential to interact. In particular, a sequence which is transformed by monosyllabic demotion could result in an unaccented trisyllabic sequence triggering the application of extrasyllabic marking. The output of this sequence of two transformation on the text then would subject to the mismatch criteria.

This hypothetical situation is represented as follows:

(90)

text-tune composite >> MDR >> EMR >> mismatch criteria (44)

We say hypothetical because the revised form of the MDR as it appears in (71) demotes the syllable following the accented monosyllable and not the accented monosyllabic word itself. Thus, the conditions for the EMR are not met and the interaction of rules shown in (90) does not obtain. To see this we apply the MDR to the text in (70) (“Travelers’ checks cost a dollar.”) This results in a sequence of only two unaccented syllables as shown below:

(91)

```

  x   x           x       x
  x x x         x       x x x x
Traveler's checks cost a dollar.*
  h   |  S1 |CG S2
```

[revised MDR (71) applies]

```

  x   x           x       x
  x x x         x       x x x x
Traveler's checks cost a dollar.*
  s h   |  S  |CG s
```

Since the EMR specifies a sequence of at least three unaccented syllables, it does not apply in this instance.

For the EMR to apply to this setting, we are required to return to the previous version of the MDR in (61) which demotes the initial rather than the final syllable of the accented pair. This has the effect of demoting “checks” rather than “cost” as shown below.

(92)

```

  x   x           x       x
  x x x         x       x x x x
Travelers' checks cost a dollar.*
  s s   |  s  |CG S
```

The right mismatch is now removed with the consequence that we are temporarily unable to mark the setting as mismatched. However, with the EMR now available to us, this altered structure can be taken as input into it.

(93)

```

  x   x           x   x
  x x x       x   x x x x
Traveler's checks cost a dollar.*
  h   |   S   |CG S

```

[original MDR (61) applies]

```

  x   x           x   x
  x x x       x   x x x x
Traveler's checks cost a dollar.*
  s h   |   s   |CG S

```

[EMR applies]

```

                                     (x)
                                (x)
  x   x           x   x
  x x x       x   x x x x
Tra(ve)lers' checks cost a dollar.*

```

As indicated by the bold faced type, the setting in (93) is now correctly designated as unacceptable.

It should be noted, however, that the application of the revised MDR in (91) also results in a fatal right mismatch marking the setting as unacceptable. We are therefore confronted with the situation in which both the revised MDR and the original MDR in interaction with the EMR result in a correct prediction.

While there is no immediate empirical justification for preferring it, there are two reasons for why the interactive account resulting in (93) is preferable to that in (91). First, as we observed, although (93) results in the correct prediction, it is inconsistent with our experience of the passage in its targeting of the location of the mismatch. This is heard not at the juncture between “checks” and “cost” as indicated in (93) but within the setting of the word “travelers”. As (91) correctly indicates the locus of the mismatch, it is preferred on these grounds.

Secondly, it will be seen that the incorporation of the EMR allows us to alter the revised monosyllabic demotion rule to be more generally applicable. In its final version in (94) the MDR applies not only to accented monosyllables preceded by a light syllable but to those preceded by any unaccented syllable. Furthermore, this statement mandates a single conditional transformation, removing the inelegant and admittedly counter-intuitive “elsewhere” condition from the revised MDR as it appears in (71).

(94) Monosyllabic Demotion (final version)

Given the text tune configuration

```

x
x   x
s |w S |x

```

where $x > W$, $S \rightarrow s$.

In this form, MSD will be seen to apply to (70) resulting in the correct prediction for it as well as all cases we have dealt with above.

The following shows rederived predictions for the choriambic sequences which presented problems for both the original and revised forms of the MSR. We first deal with a choriamb having a right bounded accented monosyllable, i.e. one which is attached to a clitic group or higher prosodic constituent on its right.

(95)

```

x   x           x           x
x x x x       x           x x x
"Pame|left|w Bill." said Am y.
S s s S

```

[MDR applies.]

```

x   x           x           x
x x x x       x           x x x
"Pame|left|w Bill." said Am y.
S s s s

```

[w < CG, EMR does not apply. Setting remains intact.]

```

  x   x           x           x
  x x x x       x           x x x
  "Pame|la|left|_w Bill." said Am y.

```

[No mismatch. Setting acceptable.]

In (96), the final syllable of the choriamb is now left bounded; that is, it is enclosed within a major prosodic constituent on its left-in this case a sentence.

(96)

```

  x   x           x           x
  x x x x       x           x x x
  "Pame|la|left".|_s Bill told Am y.
  S s s S

```

[MDR applies.]

```

  x   x           x           x
  x x x x       x           x x x
  "Pame|la|left".|_s Bill told Am y.
  s s s

```

[S ≥ CG, EMR applies.]

```

                (x)
  (x)           (x)
  x   x           x           x
  x x x x       x           x x x
  "Pa(me)la|left.|_s Bill" told Am y.*
  S       s

```

[Right mismatched detected. Setting unacceptable]

The MDR and EMR combine to rule out settings in which the choriamb occurs line-finally.

(97)

```

  x           x           x x
  x   x x x       x x x       x
  John took Mary's travelers' checks.*
                S s s       |S|_s

```

[MDR applies.]

```

  x          x          x  x
  x    x    x x      x x x      x
John took Mary's travelers' checks.*
                S s s      |s|s

```

[EMR applies.]

```

                                (x)
                                (x)
          (x)                    (x)
  x          x          x  x
  x    x    x x      x x x      x
John took Mary's tra(ve)lers' checks.*
                S s s      |s|s

```

[44 b applies.]

Finally, (98) shows the derivation for a setting we categorized as ambiguously unacceptable:

(98)

```

  x          x  x          x
  x x      x x x      x  x x
Mary's travelers' checks vanished.*?
                s |  S |cg

```

[MDR applies.]

```

  x          x  x          x
  x x      x x x      x  x x
Mary's travelers' checks vanished.*?
                S s s |  s |cg

```

[EMR applies]

```

                                (x)
                                (x)
          (x)                    (x)
  x          x          x          x
  x x      x x x      x          x x
Mary's tra(ve)lers' checks vanished.*?
                S s s |  s |cg

```

[(44) a or b not applicable. Setting predicted to be acceptable.]

While the prediction does not strictly accord with the designation of the setting as somewhat unnatural, it is reasonable to consider that some textsetting principle other than stress to beat matching might be responsible for this negative judgment. Specifically, it is apparent that the setting in (98) contains a constituency mismatch. Constituency mismatches, to recall the discussion in 1.3, are those locations in a text-tune composite where (musical) grouping boundaries occur in the absence of p-boundaries, or, conversely, where prosodic boundaries occur in the absence of group boundaries. Instance of these are shown in bold face below.

(99)

$\{ \{ \quad C \ D \quad \quad E \ C \} \{ C \quad \quad D \quad \quad E \ C \quad \quad \} \}$
 $| \ | \ | \ \text{Mary's} \ |_w \ \text{trave} \ \text{lers}' \ |_w \ \text{checks} \ |_w \ |_{CG} \ \text{vanished.} \ | \ |_w \ |_{CG} \ |_s \ * ?$

The major musical and linguistic constituents are acceptably matched: as is the case with minimally acceptable texts assigned to the Frère Jacques tune (see section 1.3), the entire sentence aligns the eight note full group indicated by double (right) curly braces. In contrast, the main internal grouping boundary which partitions the sequence into two symmetrical (and parallel) subgroups (indicated by opposed left and right curly braces) is not aligned with a corresponding (linguistic) p-boundary. It is reasonable to suspect that this constituency mismatch (as opposed to the conceptually distinct stress-beat mismatch) is what accounts for listeners' minimally negative judgment in this case. If so, the prediction of (98) that the setting is acceptable from the standpoint of stress-beat mismatches is defensible and likely correct.

Partial References:

Wells, J. C. (2000). *Longman Pronunciation Dictionary*. Harlow, England: Pearson Education Ltd

Malof, Joseph (1966). Meter as Organic Form, *Modern Language Quarterly* 27(1):3-17

Kager, Rene (1989) A metrical theory of stress and destressing in English and Dutch

Piston, Walter (1941) *Harmony*. New York: W. W. Norton & Company, Inc.

Ladefoged, Peter and Ian Maddieson (1996) *The Sounds of the World's Languages* Oxford, UK and Cambridge, MA: Blackwell

Seppänen, Aimo (2002) On analysing the pronoun it. *English Studies* 5:442-462.

Liberman, Mark and Alan Prince (1977) "On Stress and Linguistic Rhythm." *Linguistic Inquiry* 8: 249-336.

Büring, Daniel (2006) "Focus Projection and Default Prominence". In Valéria Molnár and Susanne Winkler (eds) *The Architecture of Focus*. Berlin/New York: Mouton De Gruyter.

Hulst, H.G. van der (in press). Brackets and grid marks or theories of primary accent and rhythm. In: E. Raimy and C. Cairns (eds.). *Contemporary views on architecture and representations in phonological theory*. MIT press, 225-245.

West, M. L (1982) *Greek Metre*,

Attridge, Derek (1980) *Well-Weighed Syllables: Elizabethan Verse in Classical Metres*. London: Cambridge University Press.

Saintsbury, George (1926) *Historical Manual Of English Prosody*. London: MacMillan and Co.

Giegerich, Heinz (1992) *English phonology: an introduction*. London: Cambridge University Press

Lehiste, I. (1977) Isochrony reconsidered. *Journal of Phonetics*, 5 (3), 253-263.