TWO KINDS OF METRICAL SHIFTS IN ENGLISH TEXT-SETTING

Noriko Hattori

Mie University
hattori@human.mie-u.ac.jp

ABSTRACT

The way text (or lyrics) is assigned to musical notes (or tunes) partly depends on native speakers' intuition. Earlier studies on English text-setting, as it is called, used sung or chanted verses, and an algorithm or a set of constraints was proposed to explain the intuition. These attempts avoided vocal music because of its complexity and idiosyncrasy.
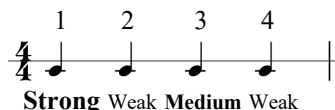
This study examined twenty pieces of English vocal music and tried to explain apparent exceptions to text-setting rules, which predict a well-formed matching of stressed syllables with musically strong beats. The results showed that most of the apparent exceptions are no longer exceptions if two kinds of metrical shifts are introduced into the framework; one is a leftward metrical shift proposed in the literature, and the other is a rightward metrical shift, which is dealt with in this work. This work also suggests the need to introduce the idea of intonation phrases into text-setting.

Keywords: English, stress, text-setting, intonation, vocal music.

1. INTRODUCTION

Text-setting is another term of text-to-tune mapping and is one aspect of speech-music interface. In Western music tradition, the following relationship holds for music with a time signature of 4/4 [8, 9]:

- Beat strength:
  1st beat > 3rd beat > 2nd and 4th beats

The Syllabic Distribution Algorithm (SDA) [3] is a rule-based formalisation of the strong tendency of 'stress-to-beat' matching observed in English, which refers to a grid notation such as that depicted in Figure 1 [1, 2, 3, 5]. The grid notation is a linguistic parallel of a musical notation.

Figure 1: The grid notation used in text-setting.

whole note	ES:			x							x									
half notes	S:			x			x				x			x						
quarter notes	M:		x	x	x		x		x		x		x		x					
eighth notes	W:	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x			
sixteenth notes	EW:																				

The rows of x depict a series of isochronic beats, and the columns indicate the strength of individual beats. 'ES', 'S', 'M', 'W', and 'EW' stand for extra strong, strong, medium, weak, and extra weak, respectively.

The SDA is summarised as follows [3]:

(a) Map stressed syllables one-to-one, left-to-right onto S (Strong) positions;

After each iteration:

(b) Map stressless syllables one-to-one, right-to-left onto the highest grid level able to accommodate them.

Once (a) and (b) have applied as many times as they can:

(c) Map any remaining stressless syllables one-to-one, left-to-right onto the highest grid level able to accommodate them.
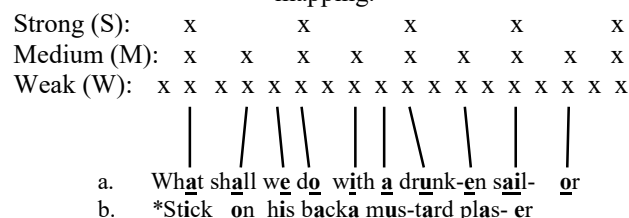
2. BACKGROUND

This section deals with earlier studies on text-setting.
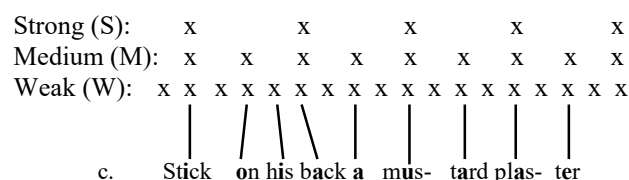
2.1. How the SDA works: Chants

Figure 2 shows a frequently cited example of English text-setting: verses from 'What Shall We Do with a Drunken Sailor.' There are ten syllables in verse (a), and nine in verse (b). The solid lines indicate how syllables are mapped onto the grid. The way of mapping in (a) does not lead to well-formed mapping in (b) because of a different arrangement of stressed/stressless syllables in (a) and (b) [1, 2, 3]:

Figure 2: Grid notation and the two ways of text-tune mapping.

The SDA rightly predicts (c) in Figure 3.

Figure 3: Grid notation and a well-formed mapping.

2.2. Constraints or 'Metrical Well-Formedness Rules'

Another attempt to explain English text-setting was carried out within the framework of Optimality Theory [3] and a similar line of research. Several constraints [3] or 'Metrical Well-Formedness and Preference Rules' [2, 4] were proposed in the work. The following are some of the constraints related to this work:

(1) *LAPSE, *NULL IN S, *STRESSLESS IN S, STRONG IS LONG

The last one summarises the characteristic that 'rhythmically strong units tend to be long as well' [3, 4].

2.3. Leftward metrical shift in a surface structure

Temperly [10] and Temperly and Temperly [11] showed that the rules of text-setting are valid if 'syncopation is viewed not simply as the accenting weak beats, but some sort of deviation from an underlying structure, where stressed syllables are shifted over to the strong beats on which they belong' [10, 11].

Their claim is illustrated in Figure 4 using the actual score examined in this work. The solid lines indicate alignments in a deep structure, whereas the dotted lines indicate alignments in a surface structure.

Figure 4: Leftward metrical shift in a surface structure
(from # 2: *Bette Davies Eyes*).

```
Extra Strong (ES):        x            x
Strong (S):        x      x      x     x
Medium (M):   x    x   x  x   x  x   x  x
Weak (W):   x x x x x x x x x x x x x x x x x
              / / \  / \  \  | \  / \  \
          she'll un- ease you all the  bet- ter
```

Temperly [10] proposed two levels of metrical structure and stated that 'in a deep structure, a syllable is in its 'correct' position, but it is shifted one eighth or sixteenth beat to the left in a surface structure'.

3. METHOD

The scores examined in this work are from *The Billboard® HOT 100: 50th Anniversary Songbook*, a compilation of 50 chart-topping songs of the last five decades [6]. There are 42 scores with a time signature of 4/4, and the first 20 of them were the objects of analysis here. The reason scores with 4/4 were chosen is that the SDA is concerned with music with four beats, the so-called *common time*, in which 'the first and third beats are the strongest and usually receive an accented word from the lyrics; the second and fourth beats are weak, usually receiving an unaccented word'[7]. In music with a time signature of 3/4, only the first beat is strong, and interesting interactions of strong/weak beats and stressed/unstressed syllables are rare to observe.

3.1. Data

The following criterion were set to examine the scores. The first verse and the first chorus were the objects of analysis. When a song had a sung part and a rap, only the former was included as data because the rap part consisted of notes of the same length (the sixteenth note) throughout (e.g. #14), or no notes were given in the score (e.g. #12).

Table 1 contains a list of titles and the number of bars examined in each score. The number of beat positions examined in each title does not necessarily coincide with the four times (i.e. four beat positions, strong and weak, in each bar) the number of bars. This is because the rests in beat positions were excluded from the counts.

Table 1: List of song titles and number of bars examined.

	Title	Number of bars	Number of beat positions
1	Aquarius/ Let the Sunshine in	63	241
2	Bette Davis Eyes	28	80
3	Flashdance … What a Feeling	40	122
4	Big Girls Don't Cry	37	124
5	Billie Jean	37	92
6	Centerfold	18	60
7	Dilemma	22	62
8	Endless Love	30	88
9	Every Breath You Take	41	121
10	I Do it for You	38	112
11	Eye of the Tiger	25	77
12	Family Affair	8	27
13	Gold Digger	15	37
14	Gangsta's Paradise	13	45
15	Hot Stuff	21	69
16	How Do I Live	20	70
17	How Deep Is Your Love	25	85
18	How You Remind Me	18	55
19	I Will Always Love You	26	74
20	I Just Want to Be Your Everything	29	77
	Total	554	1718

3.2. Beat positions examined

It has been pointed out that 'the stress-to-strong principle ignores the assignment of unaccented syllables, which may appear in either strong or weak musical positions' [2]. To give an example from our data (# 1), the preposition 'with' (boxed in the score below), which is usually unstressed, is mapped onto the strong (first) beat:

(2)

Ju‑pi‑ter_____ a‑ligns___ with___ Mars_____

The curved line linking the two notes on which the lyric *with* is mapped is called a 'slur', and it 'links two different notes' [9]. At the same time, the three notes on which *-ter* in *Jupiter* is mapped in (2) are all the same pitch; this is called a 'tie' and 'joins two notes of the same pitch to produce one extended sound' [9]. Following the music literature, tied notes were treated as one long note in this work.

The 'stress-to-strong principle' does not deny matching of a stressless syllable to musically strong beat [2]. Accordingly, cases where stressed syllables in the lyrics failed to fall in musically strong positions were examined in this study. There were four types of possible situations where this could happen:

(3) Stressed syllables fail to fall on the strong beats. That is,
  (i) Stressed syllables fail to fall on the first beat.
  (ii) Stressed syllables fail to fall on the third beat.
Stressed syllables fall on the weak beats. That is,
  (iii) Stressed syllables fall on the second beat.
  (iv) Stressed syllables fall on the fourth beat.

4. RESULTS

Each bar (total = 554 bars) was examined to determine how stressed syllables are aligned with the above-mentioned four beat positions. As an example, the results of the examination of the first score are shown here:

Table 2: Results of text-setting in #1.

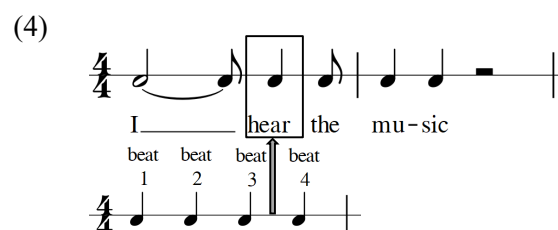|	Part of lyrics	(i)	(iii)	(ii)	(iv)
		First	Second	Third	Fourth
1	sev(enth)	✓			
2	house			✓	
3	Mars			✓	
4	stars	✓			
5	This		✓		
6	dawn(ing)	✓			
7	(under)stand(ing)	✓			
8	(a)bound(ing)	✓			
9	false(hoods)		✓		
10	(de)ri(sions)	✓			
11	dreams			✓	
12	(reve)la(tion)	✓			
13	mind's	✓			
14	true		✓		
15	(liber)a(tion)	✓			
16	sun(shine)	✓			
17	sun(shine)	✓			
18	sun(shine)	✓			
19	sun(shine)	✓			

There were 19 examples of misalignments in #1. Among them, 16 examples (Nos. 1, 2, 3, 4, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, and 19) were all preceded by the 8th note in either the first or the third beat positions (they are shadowed in Table 2). This is the leftward shift mentioned in 2.3. above. If two levels of structure, deep and surface, are postulated, they are no longer exceptions to 'the stress-to-strong principle'.

Discussion of the remaining three examples (Nos. 5, 9, and 14) and similar cases will be conducted in Section 5 below. The apparent exceptions to the SDA total 353. The breakdown is shown in Table 3:
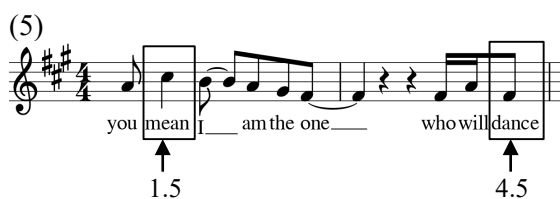
Table 3: Apparent exceptions to the SDA.

	(i)	(iii)	(ii)	(iv)	
	First	Second	Third	Fourth	Total
1	13	3	3	0	19
2	14	6	2	1	23
3	9	1	2	9	21
4	9	3	11	6	29
5	11	4	9	1	25
6	0	4	3	4	11
7	1	0	16	0	17
8	1	0	5	0	6
9	2	13	8	1	24
10	1	2	0	10	13
11	4	6	9	4	23
12	0	5	1	6	12
13	3	0	0	2	5
14	0	8	4	4	16
15	10	1	2	4	17
16	2	4	4	3	13
17	16	0	5	3	24
18	5	8	5	7	25
19	0	1	0	1	2
20	3	16	7	2	28
Total	104	85	96	68	353

There were some cases where notes fell on 0.5 after either beat position (that is, 1.5, 2.5, 3.5, or 4.5 beat). To cite an example from #3, the lyric *hear* starts on 0.5 beat after the third, that is on 3.5 beat:

(4)

I_____ hear the mu‑sic
beat beat beat beat
1 2 3 4

Such examples were not included in the total.

Another example from #5 will make the criterion of my classification clear; the lyrics *mean* and *dance* boxed in (5) start on 0.5 beat after the 1st and 0.5 beat after the 4th, respectively. Accordingly, they start on 1.5 beat and 4.5 beat. Such examples were not included in the above count.

(5)

you mean I am the one who will dance

1.5   4.5

Among the 353 examples, 230 were preceded by an 8th or 16th note, as shown in Figure 4. The breakdown of the examples is as follows:

Table 4: The distribution of beat positions where notes are preceded by 8th or 16th notes.

	(i)	(iii)	(ii)	(iv)	Total
	Position of beat				
	First	Second	Third	Fourth	
1	13	0	3	0	16
2	14	0	2	0	16
3	9	0	2	0	11
4	9	1	11	2	23
5	11	3	9	0	23
6	0	1	3	1	5
7	1	0	16	0	17
8	1	0	5	0	6
9	2	1	8	0	11
10	1	0	0	0	1
11	4	0	9	2	15
12	0	2	1	6	9
13	3	0	0	0	3
14	0	4	4	4	12
15	10	0	2	0	12
16	2	1	4	0	7
17	16	0	5	0	21
18	5	0	5	1	11
19	0	0	0	0	0
20	3	1	7	0	11
Total	104	14	96	16	230

5. DISCUSSION AND CONCLUSIONS

The number of scores examined was twenty and that of beat positions was 1,718. The number of apparent exceptions to the SDA, or misalignment, was 486. Among them, 133 examples were of cases where a stressed syllable fell on 0.5 after either beat (see (4) and (5) above), and these were excluded from the count. Therefore, 353 (= 486−133) beat positions were the object of analysis. Among them, 230, that is 65.2%, were explained as a result of leftward metrical shift in a surface structure. The breakdown is shown in Table 5.

Table 5: The ratio of left metrical shift.

	(i)	(iii)	(ii)	(iv)	
	Position of beat				Total
	First	Second	Third	Fourth	
N	104	14	96	16	230
%	100	16.5	100	23.5	65.2

It is worth noting that 100% of the first and the third beat positions were explained by postulating two levels of metrical structure, and they are no longer exceptions to the SDA. Furthermore, the notes were given an extra duration by preceding the 8th or 16 notes. The first and third beat positions were strong or medium strong, and this is a good reflection of one of the constraints in (1): STRONG IS LONG.

Another example of apparent exceptions to the SDA should be explained if the idea of intonation phrases is introduced in text-setting. For example, (6) is a prediction of the SDA. However, the actual score shows a mapping like (7). In a phrase consisting of an adjective + a noun (e.g. *silent tears*), a noun, which is the last content word in a phrase, usually receives more prominence than the preceding adjective, but this fact is not reflected in the SDA. When a noun is aligned with the first beat as in (7), the preceding adjective 'silent' is to be submitted to a weaker position than the noun.

(6) || silent tears || full of pride || (from # 3)
　　　1 2　**3** 4　**1** 2　**3** 4
　　　　　　　(where '||' means a bar in a score)
(7) silent || tears full of || pride …
　　　4　　　**1** 2 3 4　　　**1**

The word that is a nucleus (*tears* in this case) appears in the strongest position, that is, immediately after a bar. This work calls such a shift a 'rightward' metrical shift. If the idea of intonation phrases is introduced into text-setting, 45 examples will no longer be exceptions to the SDA.

Table 6: Total number of examples of shifted mapping.

leftward metrical shift	230	65.2 %
rightward metrical shift	45	12.7 %
Total	275	77.9 %

The two kinds of metrical shifts explain 77.9% of the apparent exceptions to the SDA. Thus, the real number of exceptions amounts to 78 (= 353−275). Mapping of stressed syllables with the second and the fourth beats is not distributed evenly in the 20 pieces of vocal music examined. Some songs prefer this type of mapping, whereas others never use it. Further study will be needed to decide whether this can be attributed to 'backbeat', which emphasises beats two and four in some styles of music, especially in rock music [8].

6. ACKNOWLEDGMENTS

7. REFERENCES

[1] Dell, F., Halle, J. 2009. Comparing musical textsetting in French and English songs. In: Aroui, J.-L., Arleo, A. (eds), *Towards a Typology of Poetic Forms: from Language to Metrics and beyond.* Amsterdam/ Philadelphia: John Benjamins. 63−78.

[2] Halle, J., Lerdahl, F. 1993. A generative textsetting model. *Current Musicology* 55, 3−23.

[3] Hayes, B. 2009. Textsetting as constraint conflict. In: Aroui, J.-L., Arleo, A. (eds), *Towards a Typology of Poetic Forms: from Language to Metrics and beyond.* Amsterdam/Philadelphia: John Benjamins. 43−61.

[4] Lerdahl, F, Jackendoff, R. 1983. *A Generative Theory of Tonal Music*. Cambridge, Mass.: The MIT Press.

[5] Liberman, M. 1975. *The Intonational System of English*. The Indiana University Linguistics Club, Bloomington.

[6] Nielsen Business Media, Inc. 2009. *The Billboard HOT 100: 50th Anniversary Songbook*. Van Nuys, CA.: Alfred Music Publishing Co., Inc.

[7] Petrik, J., Austin, D., Lynn, C. 2010. *Songwriting for Dummies*. 2nd edition. Indianapolis: Wisley Publishing.

[8] Powell, J. 2010. *How Music Works*. New York: Little, Brown and Company.

[9] Taylor, E. 1989. *The AB Guide to Music Theory.* Part 1. London: The Associated Board of the Royal Schools of Music.

[10] Temperly, D. 1999. Syncopation in rock: a perceptual perspective. *Popular Music* 18, 19−40.

[11] Temperly, N, Temperly, D. 2011. Music-language correlations and the "Scotch Snap". *Music Perception* 29, 51−63.