Rhythmic Correspondence between Music and Speech in English Vocal Music

Xi Chen, Peggy Pik Ki Mok

Department of Linguistics and Modern Languages, The Chinese University of Hong Kong
chenxi1099@gmail.com, peggymok@cuhk.edu.hk

Abstract

This study investigates the rhythmic structures of music and speech, and the possible corresponding rhythmic patterns between the two domains in English vocal music. With fifteen English songs as samples, lexical stress of multi-syllabic words is compared with three musical dimensions—metrical stress, duration, and pitch respectively. It was found that in the chosen English songs, there is a good mapping between the metrical stress of music and the lexical stress of lyrics. In addition, the duration and the pitch patterns not only generally match the lexical stress patterns most of the time, but also serve to manifest the prominence of the primary lexical stress on one hand, and to reflect the weakness of the unstressed syllables on the other. In addition to a general good match in rhythm, this study also shows match differences within the three comparisons. Match degrees vary according to different meter patterns. Moreover, pitch takes priority over duration in their respective match with lexical stress of the lyrics. Finally, the primarily stressed syllables match duration and pitch patterns much better than the unstressed ones do.

Index Terms: Musical rhythm, Speech rhythm, English songs

1. Introduction

Widespread consensus exists among musicians and linguists that music and speech have strong parallels in many aspects. Among them, rhythm is widely acknowledged as a crucial prosodic feature in both domains, and rhythmic correspondence is an important locus of interest in this interdisciplinary field. Lerdahl & Jackendoff [1] adopted linguistic framework to compare prosodic structures of music and speech so as to investigate how the prosodic elements such as duration, pitch and intensity create structured rhythmic and melodic patterns in the two domains. Their theory of metrical structures was later developed by Todd's [2] wavelet model analyses to rhythm in music and speech. Recent empirical studies [3, 4] found that the rhythmic features of the language in a geographic location would leave imprint on its local musical instrument. And even for the different dialects of the same language, they have different rhythmic characteristics which are reflected in the local musical styles [5]. Palmer and Kelly [6] discussed the relations between linguistic prosody and musical meter in songs, and found that compound word and nuclear stress rules coincide with musical rules of metrical accent. A very recent study [7] has investigated the degree of stress-meter alignment in French vocal music and found that stronger metrical stress tends to fall on the final syllables of poly-syllabic words and mono-syllabic content words. Thus the two parallel stresses are more likely to be consistent. The author also estimated that this alignment in English vocal music might be better than in French because the lexical stress in English is less controversial than in French, but so far very few studies have systematically examined the rhythmic correspondence in English songs. The present study is designed to fill this gap.

This study investigates the underlying rhythmic correspondence between music and speech in English vocal music. Word stress is a manifestation of rhythm in speech, and there are three categories of word stress in English: primary stress, secondary stress, and unstress [8]. In music, rhythm can be realized in the variation of intensity, duration and pitch [9, 10]. First, musical accent can be presented by metrical structure. Different meter patterns are created by endowing a sequence of beats with different metrical stress, and the stronger one takes the accent [11]. In addition, previous studies showed that as a reliable cue of melodic contour, duration can also present the stress by changing the length of the notes [9] [12][13], and longer notes can emphasize the musical event [14]. Besides, pitch patterns also create stress distinctions by making the pitch of a note higher than its surrounding ones [1] [12] [15]. According to Jones' idea [15] of interval accent, a note higher than its surrounding notes is supposed to be stressed, and a note with lower pitch is supposed to be unstressed. In the present study, lexical stress in multi-syllabic words in lyrics is chosen, and is respectively compared with three dimension of musical rhythm—metrical stress, duration, and pitch. These three comparisons can reveal 1) the situation of stress-meter alignment in English vocal music; 2) whether the duration patterns and pitch patterns can manifest word stress patterns.

2. Method

2.1. Materials

Fifteen English songs were selected. They are famous folk songs, golden oldies, ballads and popular movie songs which are written in different time periods with diverse backgrounds. The styles of the songs are supposed to be simple, smooth, and traditional so as to ensure that they can be easily recognized and sung by most people. The fifteen songs cover the most common meter patterns: 2/2, 2/4, 4/4, 3/4, and 6/8, and there are three songs for each meter. The alignment analysis was based on score analysis, not the sound tracks. The scores with the assigned lyrics were found from the Internet and two university libraries in Hong Kong. For ease of reference, each song is numbered. The information of the selected songs is as following:

2/2: (1) ***Red River Valley*** (Traditional Canadian folk song)
(2)***Take Me Home, Country Roads*** (Lyrics and music by John Denver, Bill Danoff & Taffy Nivert)
(3) ***Tie a Yellow Ribbon round the Ole Oak Tree*** (Lyrics and Music by Russellbrown & Irwinlevine)

2/4: (4) ***Yankee Doodle Boy*** (From *Yankee Doodle Dandy*. Lyrics and music by George M. Cohan)
(5) ***Oh Susanna*** (Lyrics and Music by Stephen Foster)
(6) ***Do-Re-Mi*** (From *Sound of music*. Lyrics by Oscar Hammerstein II; music by Richard Rodgers)

4/4: (7) ***Sound of Silence*** (Lyrics and Music by Paul Simon)
(8) ***Yesterday once more*** (Lyrics and Music by Richard Carpenter and John Bettis)

(9) ***My Heart Will Go On*** (Lyrics by Will Jennings, Music by James Horner)

3/4: (10) ***Green Sleeves*** (English traditional folk song)

(11) ***Moon River*** (Lyrics by Jonny Mercer, music by Henry Mancini)

(12) ***My Favorite Things*** (From *Sound of Music*. Lyrics by Oscar Hammerstein II; music by Richard Rodgers)

6/8: (13) ***Silent Night*** (Lyrics by Joseph Mohr; music by Franz Gruber)

(14) ***We Are the Champions*** (Lyrics by Freddie Mercury, Music by Queen)

(15) ***Dulcinea*** (From *Man of La Mancha*. Lyrics by Joe Darion, Music by Mitch Leigh)

2.2. Procedure

This study includes three parts in which the lexical stress of the lyrics is respectively compared with metrical stress, duration patterns and pitch patterns of music. Since in English, distinguishable lexical stress cannot be found in mono-syllabic words [8], this study only focuses on di-syllabic and multi-syllabic words.

2.2.1. Match between metrical stress and lexical stress

On the score of each song, syllables of disyllabic and multi-syllabic words are circled with their aligned notes, and the categories of lexical stress of the syllables and the metrical stress of the aligned notes are identified and marked.

In the analysis, the three categories of metrical stress of music (major stress, minor stress, and unstress) are respectively represented by capital letters A, B and C, and the three categories of lexical stress in English (primary stress, secondary stress and unstress) are respectively represented by small letters a, b, and c. We follow the convention of western music system for the standard of metrical stress patterns in music [11]. And the English lexical stress patterns are based on the *Oxford Advanced Learner's Dictionary, 8th edition*, published by the Oxford University Press ELT.

In the analysis, each target is a combination of two types of stresses since it includes a syllable which takes a lexical stress, and a syllable-aligned note which is attached to a particular metrical stress. There are altogether nine possibilities of such combinations. Among them, in the cases of Aa, Bb, and Cc, the two types of stresses are consistent, so they match perfectly. In Ac and Ca, the two types of stresses are contradictory. In Ab, Ba, Bc and Cb, the match between the two types of stresses is not as perfect as that in the first group, but is still better than that in the mismatched group. In this way, the nine possibilities can be divided into three match groups—perfect match (Aa, Bb, Cc), secondary match (Ab, Ba, Bc, Cb) and mismatch (Ac, Ca).

After labeling the stress combination patterns of each target, the match degree of each category can be calculated by the following formula: $P\% = (N_P / N_t) * 100\%$, where P stands for any match category; N_P stands for the number of syllables that belong to a particular match category; N_t is the total number of syllables of disyllabic and multi-syllabic words.

2.2.2. Duration analysis

In this section, the target notes are represented by a short underline "_". The relative duration of the two notes adjacent to the target notes is examined as well. Compared with the target note, if the duration of the adjacent note is shorter, then it is marked as "S"; if longer, then it is marked as "L"; if identical, then it is marked as "I". If there is a rest, or if there is no adjacent note (the left position of the initiate note or the right position of the last note), a short dash "-" is used. There are fifteen possibilities as follows:

S_S, S_-, -_S, S_I, I_S, S_L, L_S, I_-, I_I, -_I, L_I, I_L, L_L, L_-, -_L

A capital letter X is used to represent an arbitrary metrical stress. Xa stands for a target in which the syllable takes the primary lexical stress. Xb is a target in which the syllable takes the secondary lexical stress, and Xc is a target in which the syllable is unstressed. In the analyses of both duration and pitch patterns, it was found that the number of Xb cases is quite small (12 out of 685 for duration patterns and 10 out of 687 for pitch patterns), so we only focus on Xa and Xc in this paper.

For the group of Xa, if the two adjacent notes are shorter than the target note, then the prominence of the target note is highlighted auditorily. In a similar way, for the group of Xc, if the duration of the two adjacent notes is longer than the target note, then the target note is heard as weakened. In these two cases, the duration patterns manifest the lexical stress patterns well, and the two types of stresses are in perfect match. The match degree decreases when the adjacent notes are identical to, or worse still, are longer than the target note. The match patterns are classified in Table 1.

Table 1. *Match patterns in duration analysis.*

Duration Patterns	S_S, -_S, S_-	S_I, I_S	L_I, I_L	L_L, -_L, L_-
Xa	Perfect match	Moderate match	Poor match	Mismatch
Xc	Mismatch	Poor match	Moderate match	Perfect match

Note: Xa: target with primarily stressed syllables; Xc: target with unstressed syllables

2.2.3. Pitch analysis

In this section, the analysis is similar to that in the duration comparison. If the pitch of the adjacent note is higher than the target note, then it is marked as "H (high)", if lower, then it is marked as "L (low)", if identical, then it is marked as "I (identical)". There are fifteen possibilities of patterns:

L_L, L_-, -_L, L_I, I_L, L_H, I_-, I_I, -_I, H_L, I_H, H_I, H_H, H_-, -_H

The pitch patterns for Xa and Xc are classified as in Table 2 below:

Table 2. *Match patterns in pitch analysis.*

Pitch Patterns	L_L, -_L, L_-	L_I, I_L	H_I, I_H	H_H, -_H, H_-
Xa	Perfect match	Moderate match	Poor match	Mismatch
Xc	Mismatch	Poor match	Moderate match	Perfect match

Note: Xa: target with primarily stressed syllables; Xc: target with unstressed syllables

It should be noted that in both duration and pitch analyses, for cases with the patterns of I_-, I_I, and -_I, the target note has the same duration/pitch as its adjacent notes. Thus the duration/pitch pattern has no influence on the manifestation of the lexical stress. In the patterns of S_L, and L_S in the duration analysis, and the patterns of L_H and H_L in the pitch analysis, each case is an ascending scale or descending scale. Thus, the influence of the duration/pitch pattern on the lexical stress is not clear. The number of such cases is 253 out of 683 in the duration analysis and 269 out of 687 in the pitch analysis. The present study will not focus on these patterns.

3. Results

3.1.1. *Stress match*

Table 3 shows the degrees of stress match. Regarding the average match degrees of these fifteen songs, the percentage of perfect match is 67.71%. It is 9.05% for secondary match, and 23.39% for mismatch. The degree of perfect match for each song is the highest among the three match categories, and also the percentage is higher than 50% (except 7 and 8).

Concerning the specific meter patterns, songs with 3/4 meter have the highest degree of perfect match whereas songs with 4/4 meter have the lowest degree. However, songs with 4/4 meter do not have a high mismatch degree since they still have a considerable degree in secondary match. It is found that songs with 2/4 and 2/2 meter have the highest mismatch degree (and is much higher than that of the songs with the other three meter patterns), showing that their general match are not as good as other songs. By contrast, songs with 6/8 meter have the lowest mismatch degree, showing that they have very good general match.

Table 3. *Degrees of stress match.*

Meter Patterns	Songs	Perfect Match	Average for each meter	Secondary Match	Average for each meter	Mismatch	Average for each meter
2/2	1)	82.61%	65.84%	0	0.43%	17.39%	33.73%
	2)	62.82%		1.28%		35.90%	
	3)	52.08%		0		47.91%	
2/4	4)	53.25%	57.31%	0	4.00%	46.75%	38.69%
	5)	52.00%		12.00%		36.00%	
	6)	66.67%		0		33.33%	
4/4	7)	45.37%	54.78%	28.70%	26.10%	25.93%	19.12%
	8)	47.54%		24.59%		27.87%	
	9)	71.43%		25.00%		3.57%	
3/4	10)	85.71%	85.14%	0	1.04%	14.29%	13.82%
	11)	81.25%		3.13%		15.63%	
	12)	88.46%		0		11.54%	
6/8	13)	100%	76.81%	0	13.89%	0	9.30%
	14)	51.11%		24.44%		24.44%	
	15)	79.31%		17.24%		3.45%	
Average		67.97%		9.09%		22.93%	

3.1.2. *Duration patterns*

In this section, the cases in perfect match plus moderate match in Table 1 are regarded as general match, and the cases in poor match plus mismatch are regarded as general mismatch. The data is reported in Table 4 which shows the specific match cases of each song in the duration analysis.

It is clear in Table 4 that in the group of Xa, 9 songs have more cases in general match than those in general mismatch. 6 songs have more cases in general mismatch than those in general match. If we take all fifteen songs as a whole, for Xa, there are 64 cases in perfect match, 46 cases in moderate match, 58 cases in poor match and 32 cases in mismatch. Thus there are 110 cases in general match, and 90 in general mismatch. As a result, the general match cases outnumber the general mismatch cases.

In Xc, 7 songs have more cases in general match than those in general mismatch, 5 songs have more cases in general mismatch than those in general match, and 3 songs have the same number of cases in general match and in general mismatch. If we view the fifteen songs as a whole, for Xc, there are altogether 36 cases in perfect match, 80 cases in moderate match, 26 cases in poor match and 82 cases in mismatch. Thus, there are 116 cases in general match and 108 cases in general mismatch. As a result, the general match cases again outnumber the general mismatch cases.

Table 4. *Matching cases in duration analysis.*

Duration Analysis	Xa		Xc	
	Song number	Total	Song number	Total
GM > GMi	1), 4), 5), 6), 9), 10), 12), 13), 14)	9	1), 4), 5), 6), 7), 9), 12)	7
GM = GMi	--	0	2), 10), 14)	3
GM < GMi	2), 3), 7), 8), 11), 15)	6	3), 8), 11), 13), 15)	5

Note: GM: general match; GMi: general mismatch.

3.1.3. *Pitch patterns*

Table 5 shows the data concerning the specific matching cases of each song in the pitch analysis.

Table 5. *Matching cases in pitch analysis.*

Duration Analysis	Xa		Xc	
	Song number	Total	Song number	Total
GM > GMi	1), 2), 3), 5), 6), 7), 9), 10), 11), 12), 13), 14), 15)	13	3), 4), 6), 7), 10), 11), 12), 14), 15)	9
GM = GMi	--	0	2), 9), 13)	3
GM < GMi	4), 8)	2	1), 5), 8)	3

Note: GM: general match; GMi: general mismatch

According to Table 5, for Xa, 13 songs have more cases in general match than those in general mismatch. 2 songs have more cases in general mismatch than those in general match. With regard to the whole picture of the match cases of the 15 songs, there are 82 cases in perfect match, 59 cases in moderate match, 43 cases in poor match and 49 cases in mismatch. Thus for Xa, there are 141 cases in general match and 92 cases in general mismatch. So there are more general match cases than general mismatch cases.

As for Xc, 9 songs have more cases in general match than those in general mismatch, 3 songs have more cases in general mismatch than those in general match, and 3 songs have the same number of cases in both general match and general mismatch. For a more general view, in Xc, there are 74 cases in perfect match, 26 cases in moderate match, 43 cases in poor match and 35 cases in mismatch. As a result, there are 100 cases in general match and 78 cases in general mismatch. Again, the general match cases outnumber the general mismatch cases.

4. Discussion

From the results of the degrees of stress match, it is clear that for all the songs except 7) and 8), the percentage of perfect match significantly outnumbers the other two match categories. Besides, all the songs have fewer than half of the cases with mismatched stress patterns. Therefore, it can be concluded that, the general good match situations between metrical stress in music and lexical stress in lyrics are dominating in English songs. Concerning the five specific meter patterns, songs with 3/4 and 6/8 meters have the best stress mapping, whereas songs with 2/2 and 2/4 meter patterns have the most unsatisfactory stress mapping. The reason for this may be due to the difference of the characteristics of the meters. 3/4 and 6/8 meters are in relatively free and relaxing style whereas 2/2 2/4 and 4/4 meters are more compact and strict. Consider that many rhymes and marching songs are written in 2/2, 2/4 and 4/4 meters whereas waltz and barcarolles which are commonly used to dance to are written in 3/4 and 6/8 meters. 6/8 meter is often considered as the double sets of 3/4 meter. Since it has more beats (6 beats) as well as one minor stress per measure, its meter space allows more flexibility in accommodating lexical stress.

Songs with 2/2 and 2/4 meters have the highest percentages of mismatch degree. Since in most cases, songs with 2/2 and 2/4 meters can be naturally written into 4/4 meters due to their very similar auditory features, our selected 2/2 and 2/4 meter songs can be analyzed as 4/4 meter songs. Under such condition, the match situation is totally different as Table 6 shows. The mismatch degrees for these songs sharply decline. This is because when compared with 2/2 and 2/4 meters, 4/4 meter features the strictness like them on the one hand, and has a minor stress which mediates the major stress and the unstress on the other, thus adding the possibilities of secondary match and providing this meter with more flexibility. This is also the reason why the stress match degree for 4/4 meter ranks between that of 2/2 and 2/4 meters.

It should be noted that a song can be written in different meter signatures. It is common that songs of a binary meter can be written in any other binary meters but resulting little auditory difference for the audience (same situation for songs of triple meters). Besides, composers may develop the original piece to a special version by adding syncopations and anacruses to arouse feelings of out of expectation. These factors may cause a different matching result. In this study, however, the selected songs are original versions or the accepted earliest English versions to avoid the above factors.

Table 6. *2/2 and 2/4 meter songs analyzed as 4/4 meter songs in stress match.*

Song number	Meter pattern	Perfect match	Secondary match	Mismatch	Mismatch average
1)	2/2	47.92%	35.42%	16.66%	6.84%
2)		83.33%	12.82%	3.85%	
3)		69.57%	30.48%	0.00%	
4)	2/4	70.13%	24.68%	5.19%	2.81%
5)		83.87%	12.90%	3.23%	
6)		58.33%	41.67%	0.00%	

As for the duration analysis and pitch analysis, it is found that most of the selected songs have good mapping between duration/pitch patterns and lexical stress distributions for both the primarily stressed syllables and the unstressed syllables. In these songs, the duration/pitch patterns of the musical notes correspond with lexical stress patterns very well to highlight the primary lexical stress and weaken the unstressed syllables auditorily. However, there exist variations of match degrees between the two analyses. Table 7 incorporates the general match situations in duration analysis and pitch analysis. The ratio between general match and general mismatch shows the match degrees. The higher the ratio is, the higher the match degree is. The result shows that pitch patterns have higher match ratios than those of duration patterns for both primarily stressed syllables (1.53 vs. 1.22) and unstressed syllables (1.28 vs. 1.05). In this sense, it can be inferred that in English songs, the pitch match is more prominent than the duration match. Phonetically, pitch is a more important cue than duration in the manifestation of English lexical stress [16] [17]. Our data show the same patterns in English songs. Thus, it can be concluded that vocal music shares common rhythmic features with speech.

In addition, it can be seen from Table 7 that cases of primarily stressed syllables have higher match degrees than those of mismatched syllables in both duration analysis (1.22 vs.1.05) and pitch analysis (1.53 vs. 1.28). So the primarily stressed syllables have better mapping between lexical stress and duration patterns as well as pitch patterns than that of unstressed syllables. It should be noticed that duration analysis and pitch analysis have unequal total numbers of general match cases plus general mismatch cases. This is because this study has excluded the cases in which the patterns have no explicit influence on the lexical stress (see the end of 2.2.3).

Table 7. *Ratios between general match and general mismatch.*

Analysis	Patterns	GM	GMi	GM/GMi
Duration analysis	Xa	110	90	1.22
	Xc	116	110	1.05
Pitch analysis	Xa	141	92	1.53
	Xc	100	78	1.28

Note: GM: general match; GMi: general mismatch; Xa: target with primarily stressed syllables; Xc: target with unstressed syllables

It is also noticed from Table 3, Table 4 and Table 5 that, the song *Yesterday Once More* (No. 8) is found to have very poor rhythmic match in all three aspects (the perfect match degree is low in stress match, and the general mismatch cases outnumber the general match cases in both duration and pitch analyses). However, this song is classic and popular. *Yesterday Once More* used to be the highest-debuting single in 1973 (according to *Cash Box*, a music trade magazine). And until now, it is still among the most popular golden oldies in many places. What makes this song universally loved by so many people even though it has such a poor rhythmic mapping? It is likely that there are some other songs in similar situation. Are there any special rhythmic patterns or rhythmic interactions which make them auditorily harmonious despite their poor rhythmic match? These questions await investigation in further studies.

In summary, the current study shows that there is a good correspondence between metrical stress in music and English lexical stress. Besides, most of the selected songs have good mappings between duration/pitch and English word stress, and the duration patterns/pitch patterns in music can manifest English word stress in most cases. The pitch match is more prominent than duration match in the analysis. In addition, primarily stressed syllables have better mapping between lexical stress and duration patterns/pitch patterns than unstressed syllables. The good rhythmic correspondence between lyrics and music may be because that the composers and lyricists take special care to match the rhythms in melody with the stress in lyrics when they compose melody for pre-existing lyrics or write lyrics for a pre-existing tune. However, there might be variations of match degrees for the two types of work. It remains unclear which one has better rhythmic mapping in English vocal music. This is also an interesting question for further research.

5. References

[1] Lerdahl, F., and Jackendoff, R., "A generative theory of tonal music", Cambridge, Mass.: MIT Press, 1983.

[2] Todd, N. P. M., "Segmentation and stress in the rhythmic structure of music and speech: A wavelet model", Journal of the Acoustical Society of America, 93, (2363), 1993.

[3] Patel, A. D., and Daniele, J. R., "An empirical comparison of rhythm in language and music", Cognition, 87, B35–B45, 2003.

[4] Patel, A. D., Iversen, J. R., and Rosenberg, J. C., "Comparing the rhythm and melody of speech and music: The case of British English and French", Journal of the Acoustical Society of America, 119, 3034–3047, 2006.

[5] Mcgowan, R. W., and Levitt, A. G., "A comparison of rhythm in English dialects and music", Music Perception, 28(3), 307–313, (2011).

[6] Palmer, C., and Kelly, M. H., "Linguistic prosody and musical meter in song", Journal of memory and language, 31, 525–5 42, 1992.

[7] Temperley, N., "Stress-meter alignment in French vocal music", Journal of the Acoustic Society of America, 2013.

[8] Ashby, P., "Understanding phonetics", London: Hodder Education, 2011.

[9] Antley, B. R., "The rhythm of medieval music: a study in the relationship of stress and quantity and a theory of reconstruction with a translation of John of Garland's de Mensurabili Musica", Ann Arbor, Mich.: University Microfilms International, 1983.

[10] Orbach, J., "Sound and Music: for the pleasure of the brain", Lanham: University Press of America, 1999.

[11] Henry, E., "Music Theory", Englewood Cliffs, New Jersey: Prentice-Hall, INc, 1985.

[12] Huron, D., and Royal, M., "What is melodic accent? Converging evidence from musical practice", Music Perception, 13, 489–516, 1996.

[13] Hannon, E. E., and Snyder, J. S., "The role of melodic and temporal cues in perceiving musical meter", Journal of Experimental Psychology: Human Perception and Performance, 30 (5), 956–974, 2004.

[14] Drake, C., and Palmer, C., "Accent structures in music performance", Music Perception, 10, 343–378, 1993.

[15] Jones, M. R., "Dynamics of musical patterns: How do melody and rhythm fit together?" In Tighe, T. J. & Dowling, W. J. (Eds.), Psychology and music: The understanding of melody and rhythm. Hillsdale, N.J.: L. Erlbaum. 67–92, 1993.

[16] Lehiste, I., "Suprasegmentals", Cambridge, Mass.: M.I.T. Press, 1970.

[17] Clark, J., Yallop, C., and Fletcher, J., "An introduction to phonetics and phonology", Oxford: Blackwell Pub., 2007.