

THE CLOSE LINK BETWEEN LANGUAGE AND MUSIC

PATRICE LARROQUE
JEAN-JAURÈS UNIVERSITY

Abstract – It is a well-known fact that the relationship between music and the sound of speech must involve the music of the language which, like language, functions on two axes: one is horizontal and concerned with sequences, and the other is vertical and permits the production, selection, and availability of sounds and phrases. There is, indeed, a close link between the phonic system of a language and the structure of its music, and this can be shown by comparing both systems. The present chapter explores music, and songs in particular, to demonstrate the placement of word stress, syllable structure, metrical structure, and various prosodic features which are components shared by the two domains. For example, introducing the rhythm of English and suprasegmental features in both the linguistic and musical descriptions of a sentence/utterance or lyrics line may bring to light relevant aspects of both areas. Also, this text seeks to broaden our understanding of the phonology of English and expand upon the correlation between the musical and linguistic systems, and may contribute to supplying foreign learners of English with specific and perceptual reference points.

Keywords: phonology; music-speech relationship; English rhythm; word stress; metrical structure.

1. Introduction

It is quite understood that the relationship between music and language involves complex and meaningful sound sequences which invite comparison between the two domains. According to Tomatis (1977), for example, language is music and is learned like music. Rousseau, the Enlightenment philosopher, in *Essai sur l'origine des langues* ([1753] 1993) writes: “Dire ou chanter étaient autrefois la même chose” (“Speaking or saying used to be the same thing”). The recognition that there is an obvious link between the phonic system of a language and the structure of pieces of music can also be seen in musical compositions (instrumental music or songs). Indeed, Patel (2007) conjectures that composers (or song writers) have internalized the rhythmic and intonation patterns of their native languages from an early age, and that shows in their compositions. When they write their music, they have these patterns in their ears, and draw on them in composing their work (cf. also Grout and Palisca 2000). This implies that there is a tenuous connection between linguistic and musical features, mainly, but not only, phonological.

This link is pregnant in folk music: for example, Scotland's folk song collections derive partly from the Highlands, where Gaelic speech and song contribute a characteristic rhythmic and sound pattern, and partly from the Lowlands where Scottish folklore and customs have produced popular tunes and poetry (Deas 1955, p. 210), as it may account for the reflection of speech in music. Likewise, the staccato of Japanese speech is reflected in its traditional music, and flamenco artists and *aficionados* know that the language (i.e. the pronunciation and rhythm of Andalusian Spanish) has played a fundamental part in the constitution of the musical language for the traditional flamenco chant or *cante* (cf. for instance *Flamenco Magazine*, 2013). Here must also be mentioned the commonly held views of language pleasantness-unpleasantness, which suggest a close relationship between a language and its musicality (cf. Giles and Niedzielski 1998, p. 85-93).

In English, rhythm and intonation are defined as important prosodic features (cf. Carr 2012; Hasty 1997; Larroque 2021; Roach 2009; Schlüter 2005). They contribute to the intelligibility of utterances.¹ Having recourse to music, moreover, may help foreign learners of English in their acquisition of the language (cf. Brown 2024; Graham 2002, 2006; Parkinson and Vyner 2024), or people with a stammer to overcome their speech disorder (cf. Tom Hooper's movie *The King's Speech*², 2010). In truth, the main difficulty for learners, especially French-speaking ones (cf. Carr 2012; Huart 2010), stems from the fact that they have no rhythmic references on which they can rely to organize their oral performance and stress the words or group of words to form utterances (in French, the stress always falls on the final syllable of the word as compared to the stress pattern of English words); as for music therapy, singing without a stutter, for example, may help with the fluidity of the speech production and reinforce the patients' confidence and motivation. Fluidity involves a 'normal' (that is, neither too fast nor too slow), regular cadence, which may fluctuate according to environment and influences such as emotions or feelings. Singing, for instance, may prove valuable for treating stuttering. Songs are indeed concerned with words and music, and text setting requires a constant interplay between the two: on the one hand, musical and linguistic groups must match, and, on the other hand, stressed syllables must fall on strong beats (Dell and Halle 2005, pp. 1-2), which may be one of the reasons why translating songs, for instance, is such a difficult exercise.

Giving emphasis to strong beats, that is, producing particular words or syllables with greater force than other words or syllables in utterances, can be defined as stressing, which is a prosodic feature of English. This could be illustrated, for example, by referring to the rhythm of blues (or of rock 'n' roll

¹ Although this may apply to all languages.

² The movie was inspired by a series of announcements by King George VI who overcame his stammer as he took throne when his brother, Edward VIII, abdicated in 1936.

which is the speeded-up version of it), insisting on the alternating pattern of strong and weak beats which characterizes the way blues words are sung. This view has the double advantage of showing the natural relationship between language and music and of illustrating the iambic/trochaic patterns of ordinary English speech.³ Although there are influences between preference rules in music and the phonological principles of the language, which may affect the natural conversation flow, the parallelism is likely to make users sensitive to the music of English. Moreover, it is not uncommon for linguists and musicologists to describe language in musical terms since it corroborates the tentative suggestion that the prosody of a culture (cf. Scotland's folk song collections) is reflected in the rhythm of its music (cf. Larroque 2012, 2021; Lerdhal and Jackendoff 1983; Liberman 1975; Selkirk 1984; Temperley 1999).

These observations form the core of the discussion. The transmission of the speaker's message mainly depends on the prosody of the sentence/utterance which may sound flat and monotonous in, for instance, French speakers' expressivity. In oral exchanges between French and English speakers, accentual mismatches, when said by a French learner, may sound confusing.⁴ For example, a difference in meaning may be perceived between *an 'English teacher* (a teacher of English) and *an ,English teacher* (a teacher who is English). The signs ' and , placed immediately before the appropriate syllables indicate that the syllables have respectively primary and secondary stress, with a pitch movement dependent upon the intonation of the phrase (cf. Jones 1977, introduction xxi-xxii). In sum, suggesting that music can serve as basic rhythmic patterns and intonation may provide balanced phonological structures allowing for work on relevant speech units.

2. The rhythm of English

It is often said that English is a stress-timed language, as opposed to syllable-timed languages like French and other romance tongues where each syllable takes roughly the same amount of time to be pronounced (cf. Carr 2012, p. 107; Pike 1945, pp. 34-36; Schlüter 2005, pp. 24-25). This means that English has a rhythm, that is, a combination of strong and weak beats, in the same way as there is a rhythm in a piece of music or a song.⁵ This particular rhythm relies on the stressed syllables in the sentence/utterance (mainly

³ Iamb: a pattern consisting of one short or unstressed syllable followed by one long or stressed syllable; trochee: a pattern consisting of one long or stressed syllable followed by one short or unstressed syllable.

⁴ In French, sentences are stressed on the final syllables.

⁵ The beat is the basic rhythmic unit in music.

lexical words: verbs, nouns, adverbs, and adjectives are stressed); the less salient ones (grammatical, including auxiliaries, and link words) are unstressed: they are, as it were, ‘squeezed’ between the stressed parts of the words and phrases to keep the rhythm.

The rhythm of English rests on what is commonly labeled the Principle of Rhythmic Alternation (cf. Schlüter 2005). This principle, which has the advantage of facilitating language use both for the speaker and the hearer, refers to an idealized structure which alternates stressed and unstressed syllables regularly, as in the sequence illustrated below:

(1) My 'uncle 'bought a 'red and 'yellow 'hat

(The superscript diacritic ' placed immediately before the beginning of the appropriate syllable represents primary stress.).

A sentence like this is rhythmically well balanced, stressed syllables alternating with unstressed syllables, and complies with the basic rhythm of English. Note that this type of perceptible distribution is not only to be found in poetry or in popular songs, but is also a feature of connected (or ordinary) speech.⁶ It may also be noted that such optimal rhythmic combinations are often referred to as eurhythmic (or harmonious). It follows from this that the most eurhythmic construction is a simple stressed-unstressed structure, with only one unstressed syllable following the stressed syllable. This phonological constituent, which appears to be central to the rhythm of English, is the foot (Abercrombie [1976] 1990, p. 131; Carr 2012, pp. 99-100; Giegerich 1985, p. 268; Larroque 2023, pp. 19-20; Schlüter 2005, pp. 20-21). Although the iamb is widely used in English verse, linguists usually hold that the rhythm of English is basically trochaic.

A metrical foot, according to Carr (2012, p.100), for instance, consists of a stressed syllable (primary or secondary stress) followed by any unstressed syllables intervening between it and the next stressed syllable (see also Abercrombie [1976] 1990, p. 131; Giegerich 1985, p. 268; Larroque 2023, p. 19; Schlüter 2005, p. 20). Thus, conforming to the Principle of Rhythmic Alternation the words *uncle* and *yellow* in example (1) contain a simple foot, a stressed syllable (marked) followed by an unstressed syllable. The monosyllabic word *hat* consists only of a stressed syllable. Foot structures with more than one unstressed syllable are therefore less eurhythmic. It may be noted that this preference for eurhythmy is characteristic of English and extends to sequences of metrical feet. For

⁶ It must be remembered that English prosody is based on stress. The earliest measures to develop were trochaic soon followed by iambic couplets. These sound patterns were used by dramatists because of their closeness to common speech (cf. Drabble 1985, p. 644; see also Giegerich 1985, p. 268).

instance, example (1) contains two metrical feet, 'bought a' and 'red and' which show that metrical feet do not necessarily map onto single words. Thus, the whole sequence is eurhythmic; both at the level of syllables and at the level of feet (cf. Carr 2012; Roach 2009; Schlüter 2005). A sequence containing *'yellow and 'red 'hat* would be less eurhythmic, because it exhibits a stress lapse, that is, two adjacent unstressed syllables following a stressed syllable: 'yellow and' and a stress clash, two adjacent stressed syllables: 'red 'hat', and consequently two feet consisting of single words. These constitute violations of the Principle of Rhythmic Alternation. The sequence *'red and 'yellow 'hat* is more eurhythmic and therefore preferred.

Infractions of the Principle of Rhythmic Alternation will naturally entail compensation strategies. While stress lapses are in many occurrences tolerated, stress clashes are most widely perceived as objectionable. Avoidance of stress clashes between two monosyllabic words is likely to motivate the stressed-unstressed structure. There are indeed compensatory measures to repair these violations of the Principle of Rhythmic Alternation: introducing a pause, elongating the first syllable, or adding an unstressed (grammatical or 'stray') syllable between two clashing syllables (as in *Don't you call me a dude!* or *Still tryin' to get her belt a-loose*, Chuck Berry in *No Particular Place to Go*) are some of the strategies which may restore eurhythmy. Note that the latter example taken from a rock 'n' roll song illustrates the close relationship which exists between speech and popular music: recall that binary feet are the most favored ones in speech (just as in poetry and songs), followed at a distance by ternary feet (anapest and dactyl).⁷ A more interesting compensatory process regarding stress clashes is the manipulation of the placement of stresses known as the stress shift rule. This rule of rhythm reversal applies to polysyllabic words which have final stress when a stressed-stressed sequence occurs, as in *,after'noon 'tea* or *,six'teen 'tons* (in a song by the American country and western singer and songwriter Merle Travis⁸). When the word is followed by the stressed syllable of the next foot, the main stress shifts leftward, thus:

- (2) a. *,after'noon 'tea* > *'after,noon 'tea*
 b. *,six'teen 'tons* > *'six,teen 'tons*

In examples (2a-b), the secondary and primary stresses appear to switch around by virtue of the stress shift rule, thus repairing the eurhythmic strong-

⁷ Anapest: a pattern consisting of two short or unstressed syllables followed by a long or stressed syllable; dactyl: a pattern consisting of one long or stressed syllable followed by two short or unstressed syllables.

⁸ Merle Travis (1917-1983) was an American country and western singer, songwriter and guitarist.

weak structures. In actual fact, the stress shift on *sixteen tons* (cf. 2b) is reflected in Merle Travis' singing, another example of the close relationship between the phonology of the language and its music. Notice, however, that rhythm reversal does not operate on a sequence consisting of an unstressed syllable and the first (therefore stressed) syllable of a foot, as in *alarm clock*. The word *alarm* contains one foot which consists of a stressed syllable and is not followed by a less stressed syllable; that foot comes after a 'stray' unstressed syllable (as shown in the phonetic transcription of the word: /ə'lɑ:m/). Although *alarm clock* contains a stressed-stressed sequence of syllables, it does not display a stressed-stressed sequence of feet (as in *afternoon*, for instance, which consists of two feet) and as a consequence does not undergo a stress shift (cf. furthermore Carr 2012, pp. 107-115). In a case like this, the stress clash between *alarm* and *clock* can be repaired by lengthening the stressed syllable of *alarm* or by introducing a short pause between the two clashing syllables.

As mentioned before, stress lapses are more acceptable, though only to a certain extent, than stress clashes. A stress lapse can contain a sequence of two, three, or four unstressed syllables following a stressed one. A very long series of unstressed syllables, however, is likely to disrupt the regularity of the rhythm: the greater the number of unstressed syllables intervening in the interval between two stressed syllables, the faster the tempo, the shorter the syllable length, and the less eurhythmic the utterance. So, as for stress clashes there are compensatory processes. For instance, the avoidance of more than one unstressed syllable can consist in compressing series of unstressed syllables between two stressed ones. Thus in such sequences as *working day*, *working on Sunday*, *working as a waitress*, the unstressed syllables will be shortened in length and 'squeezed' into the interval, so that the sequence will have roughly the same time duration. But an unstressed syllable can also be perceived as a rhythmic beat when it occurs in a stress lapse containing more than two unstressed syllables as in *working as a waitress*. In order to interrupt a series of three unstressed syllables, the addition of a beat on the preposition *as*, typically unstressed, will restore the trochaic rhythm (cf. Attridge 1982, p. 73; Larroque 2021, p. 92; Nespor and Vogel 1989, pp. 76, 85, 100; Schlüter 2005, p. 29). In a sequence like *the animals of the Amazon* which displays an interval of four unstressed syllables between two stressed syllables, the naturally unstressed preposition *of* will be perceived as stressed in connected speech, thus interrupting a long series of unstressed syllables. Similar evidence comes in the form of beat addition in blues songs. For example, in many interpretations a typically stressed syllable is added to mark or reinforce a beat.

Compare the following lines taken from Robert Johnson's famous song *Ramblin' on my Mind*:⁹

- (3) a. I got ramblin' on my mind
 b. I got ramblin' all on my mind

In example (3a) which contains a series of three unstressed syllables following a stressed syllable, the proposition *on*, naturally unstressed, will be perceived as salient to compensate for the stress lapse and repair the trochaic rhythm of the phrase. In (3b), which is the same line repeated, a typically stressed word, here the monosyllabic adverb *all*, is introduced to reinforce the accent. When the line is sung *all* and *on* form a single monosyllabic unit: *all'n*. This type of coalescence entails the deletion of the reduced vowel of the preposition.¹⁰ Compensation for stress lapses will, of course, have an influence on the rhythm and cadence of the sequence, and more generally on the music of the language.

It may be noted, by the way, that both realizations occur in the song and that the use of the adverb *all* as beat addition is a common compensatory process in blues lyrics:

- (4) The police run me from Cairo,
 All through Arkansas.
 The police run me all from Cairo,
 All through Arkansas.

Again, in this blues song, *Broken Levee Blues* sung by another renowned blues musician Lonnie Johnson,¹¹ the typically unstressed prepositions *through* and *from* are reinforced by the stressed (even contrastive) monosyllabic adverb *all*. Note that the first line displays a stress lapse (two adjacent unstressed syllables) between *run* and the first syllable of *Cairo*, which is compensated for in the repeated line. Also consider that in southern African American English the disyllabic word *police* (in the first and third lines) is strongly stressed on the first syllable (cf. Larroque 2021, p. 34),¹² yielding a eurhythmic strong-weak pattern between *police* and *run*.

⁹ Robert Johnson (1911-1938) was an American blues musician and songwriter.

¹⁰ In English, unaccented vowels are naturally reduced.

¹¹ Lonnie Johnson (1899-1970) was an American blues and jazz singer, guitarist, and songwriter.

¹² In southern African American English many disyllabic words which would be stressed on the final syllable in other English varieties, such as *police*, *behind*, *Detroit*, *display*, etc., are stressed on the first syllable,

In general, blues songs have a recitative form, in the sense that they use the rhythm and intonation of spoken English. These features are somewhat reflected in the music.

3. Music and language

Music functions on two axes: the axis of simultaneity and the axis of sequences. This is comparable with the paradigmatic and syntagmatic axes of language. However, the syntagmatic axis of music tends to contradict the paradigmatic axis of language which is concerned with selection and excludes intrasegmental simultaneity, which means that two sounds cannot be produced at the same moment (Beneveniste 1974, p. 56). Neither does the axis of sequences correspond exactly to the syntagmatic axis of language since the musical sequence is compatible with the simultaneity of sounds. Therefore, music can be regarded as a language which has syntax but no semiotics, because it comprises no semantics. This contrast reveals that the prosodic features, that is, features which appear when sounds are put together in both connected speech and melody, are to be emphasized when studying music and language, especially when the linguistic rhythm occurs at the level of the melody, which represents the axis of sequences.

For instance, as Patel (2007, p. 165) notes, “there may be a direct route from language to music” based on the idea that prosodic patterns are acquired in a very early age. It has, indeed, been established that young children are arguably very sensitive to the music of their native language (cf. Nazzi *et al.* 1998; Ramus 2002). They incorporate statistically, consciously or subconsciously, the prosodic patterns of their mother tongue. This is called ‘statistical learning’. It may be worth remembering, at this point, that every language has a melody and, that users have long claimed that languages have their own music – which, of course, implies that it is mainly the prosody of the language that constitutes the character of its music (cf. Rousseau [1753] 1993, pp. 145-146; see also Giles and Niedzielski 1998, pp. 85-93).

Adults as well can be sensitive to the music of a language and acquire its intonation and rhythm, especially when they learn a new idiom. The example of African slaves brought to North America is quite significant: they relied on their native languages while copying the prosody and syntax of English, the only concrete element that they could decipher (cf. Jones 1963, pp. 21-22; Larroque 2018a, p. 21). In fact, what is commonly called today ‘southern accent’, or ‘jive talk’, for instance, used to be “the accent of a foreigner trying to speak a new and unfamiliar language” (Jones 1963, p. 22), and early African American speech was words learned from slave owners, pronounced as best as possible and arranged into “aboriginal speech patterns” (Herskovits 1941, p. 80), whence Larroque’s (2012, p. 123; 2021, pp. 1-4)

initial hypothesis that early blues singers may have been influenced by the trochaic rhythm of English in their music.

But music is also a system in which the signs represent something analogous to language. Like language, it is made of sounds, the musical status of which is characterized by notes on a stave. These musical units may not be directly comparable with linguistic signs. They are organized into scales, that is, sets of musical notes ordered by increasing or decreasing pitch. Musical sounds, moreover, may be produced in single melodic lines or simultaneously, as in chords for example. There is no limitation to the multiplicity and combination of sounds. The composer or song writer is free to organize the sounds in an unconventional grammatical discourse which is only submitted to its own syntax.

As mentioned above, music cannot convey objective meaning, although within a context music can evoke emotions, for example, jazz improvisations can be regarded as conversations: one musician starts a phrase and lets the others complete it, and a kind of dialog takes place, thus creating a musical story and communicating emotions. Jones (1963, p. 30), for instance, argues that a jazz saxophonist like Charlie Parker¹³ was able to produce sounds that would “imitate the human voice with his cries, swoops, squawks, and slurs”.

Language, conversely, can be objective. Yet, the music of the language can be perceived in poetry which incorporates musical processes such as rhythm or harmony. Songs, for example, are, as it were, the manifestation of poetic texts set to music, and the rhythmic patterns of the language tend to map onto those of the music. In other words, the rhythm and intonation of the language are reflected in the music. The musical transcription of the linguistic variation exemplified above in (3a-b), repeated as (5a-b), can serve as an illustration of this:

(5)

a.



b.



¹³Charlie Parker (1920-1955) was an American jazz saxophonist, bandleader, and composer.

In example (5), the eighth note (quaver) which is aligned with the preposition *on* (cf. 5a) is subdivided into sixteenth notes (semi quavers, cf. 5b, ♪♪ = *all on my*) to make way for the stressed monosyllabic adverb *all* (the onset of the beamed sixteenth and eighth notes has greater force than the other note figures of the group). Thus, as stated previously, the set of three unstressed syllables is broken by the addition of a default stressed syllable. It may be noted, moreover, that the intonation of the musical phrase, which constitutes a foot group,¹⁴ follows a speech-like declarative falling tone.

At this point, the notion of metrical division deserves some comment. In the same way as language can be divided into regular periods, a piece of music may display equal metrical time divisions (as illustrated in 5a-b) which constitute a sensible basis for rhythm. Take the phrases *a red hat*, *a yellow hat*, and *a different hat*. When you beat time to the rhythm of the stressed syllables (by clapping your hands or tapping your foot; you may also use a metronome), you will realize that they occur at regular intervals in all three sequences. The stress pattern is as follows (the vertical lines indicate the limits of the time periods, and the stressed syllables are marked by placing the sign ' immediately before the appropriate syllable):

- (6) a | 'red 'hat |
 a | 'yellow 'hat |
 a | 'different 'hat |

Such constructions are equal in duration. As a consequence, the unstressed syllables are compressed to fit in the interval. The resulting pattern is a two-beat unit in which the beat maps onto the stressed syllables.

In musical notation, the beat will be represented by a quarter note (crotchet) value (i.e. ♩).¹⁵ We can think of a beat as the stroke of a pendulum as it swings from side to side. No matter how fast or slow the motion, it will always be even. Example (7) shows that linguistic rhythm and musical rhythm are aligned:

¹⁴ A foot group, or tone group, is the minimum unit of speech which carries intonation. It usually ends with a tonic syllable (tonic placement).

¹⁵ In a 4/4 measure bar, for example, in which a quarter note receives one count.

(7) a | 'red 'hat |

 a | 'yellow 'hat |

 a | 'different 'hat |


(It is important to note that the two eighth notes (♫) and the eighth note triplet (♫♫♫) take up the same amount of time as a quarter note; also exemplified in (8) below)

Example (7) clearly demonstrates that the phonology of the language can be translated into music, when dealing with rhythm and intonation (cf. also 5a-b). Indeed, the fact that there are stressed syllables alternating with unstressed syllables may produce pitch variation, either higher or lower, as stressed syllables are naturally elongated relative to unstressed ones (cf. Pierrehumbert 2003). Similarly, sentences containing more than two metrical feet as in example (1) fall into four-beat units depending on the number of feet, as shown in (8):

(8) my | 'uncle 'bought a 'red and 'yellow | 'hat


The beamed eighth notes represent the trochaic metrical feet. Each time is divided into two parts, two eighth notes, the first part of which having greater force than the other. As may be expected, foot structures do not match with beat units (cf. the quarter note figures below them) and divisions into metrical feet are irrespective of word boundaries, another similarity between language and music.

The following example is, moreover, an illustration of the influence of feet on the rhythm of a sentence:

(9) I'll | see you again tomorrow |(night)


This sentence falls into a three beat units or triple meter (the waltz type) with, starting from the first stressed syllable, three metrical feet 'see you a', 'gain to', and 'morrow'. If *night* is added after *tomorrow* it will constitute the first beat of the next measure bar. Additionally, in order to keep a steady beat, the pronunciation of the binary feet (containing two syllables) will be influenced by that of the first metrical foot which displays three syllables and will

imprint a ternary rhythm to the sequence. This can be obtained by elongating the first eighth note of the beamed pair (which is more stressed and naturally lengthened) into a triplet feel (♩̇̇̇ or ♩̇̇̇). It represents the triplet feel of blues or rock 'n' roll music. In actual fact, many song lyrics in English display this type of trochaic rhythm.

Consider, for instance, the sequence illustrated below, taken from Michael Jackson's song *Gone Too Soon*:

- (10) Like a comet blazing (a)cross the ev'ning sky.
 a | 'comet | 'blazing | 'cross the | 'ev'ning | 'sky

The division of the line into metrical feet shows that it corresponds to the trochaic binary rhythm in English. Each stressed syllable is followed by an unstressed syllable to form a eurhythmic stressed-unstressed series, except for the final monosyllabic foot. There are, indeed, five metrical feet in this line: 'comet', 'blazing', 'cross the', 'ev'ning', and 'sky'. This division shows that metrical structure is not sensitive to word boundaries and that the 'stray' unstressed syllable of *across* can be deleted, thus avoiding a stress lapse. This type of metrical division occurs in many pop song lyrics. One famous and much cited example (cf. Larroque 2012, p. 134; 2021, p. 116; 2023, p. 20) can be found in Pink Floyd's song *Another Brick in the Wall (The Wall 1979)*: "All in all, you're just a... nother brick in the wall". In this line, the interesting metrical feet are 'just a' and 'nother'. The pause in the singing intervenes between the two feet and the unstressed syllable of the foot 'just a' is, as it were, perceptually suspended, thus illustrating the music of the language.

Similarly, double negatives, which may address an influence of the trochaic rhythm on English grammar, have a rhythmic function both in language and in music. They may help to compensate for violations of the Rhythmic Alternation Principle. Consider, for instance, the sequence *we don't need no education* taken from the same aforementioned song. Although most people, including teachers, argue against such so-called illogical grammatical form,¹⁶ it sounds more balanced with its double negative construction than *we don't need education*, which exhibits a stress clash between *need* and the first stressed syllable (secondary stress) of *education*,

¹⁶ According to prescriptive grammar, double negatives defy logic. It claims that two negatives in a sentence cancel each other out and yield a positive. Yet, most English speakers have no difficulty interpreting this grammatical form negatively, because the basic principle that there is a direct and simple relationship between the surface syntactic structure and the logical form is spurious. In fact, the two negatives do not operate on the same stages of the predication, that is, the attitudinal and the predicative levels. While the former concerns the speaker's involvement in his or her speech and the latter the informational content, they cannot therefore collide. Besides, the duplication of marks in a sentence has an intensifying meaning (cf. Larroque 2018b).

or *we don't need any education*, which displays a stress clash between *need* and the first (stressed) syllable of *any*. The monosyllabic negative determiner *no* intervening between *need* and the first stressed syllable of *education* avoids a stress clash and is much easier to place than *any* rhythmically. Furthermore, it yields an eight-syllable line which allows a pause after *no*, marking the rhythmic point of division of the second metrical foot (*caesura*), 'need no' in the line. The same kind of analysis can be carried out with the line *I can't get no satisfaction* famously sung by Mick Jagger of the Rolling Stones. In this case, besides permitting a break after the first part of the line, the double negative intensifies the discursive negation by stressing the suspended negative *no*, thus making the chorus more effective. It is interesting to note, in passing, that the possible stress clash between the negated auxiliaries *don't* and *can't* is compensated for since the reduced negative element *n't* borrows length from the stressed auxiliary so that *don't* and *can't* can be perceived as trochaic and exhibit a ternary syncopated feel.

When the duration of the stressed objects is lengthened, it creates breaks in the rhythm. These changes are called syncopations.¹⁷ The syncopation phenomenon in the musical phrase displaces the strong beats over to the next weak beats of the measure bar. At the level of the time unit, it deletes note values. Both levels combined permits compensating for a number of infractions of the Rhythmic Alternation Principle. Syncopation, or syncope, in phonology means the loss of a sound, of a letter or a vowel, in the interior of a word, for example: *never* > *ne'er*, *over* > *o'er*, *heaven* > *heav'n*, *listen* > *lis'n*, *across* > 'cross, *because* > 'cause or simply *cos*, etc. Such forms as *ain't*, *don't*, *can't*, *won't*, *mustn't*, and periphrastic constructions as *gonna*, (for *going to*), *gotta* (for *has/have got to*), *dunno* (for *don't know*), *twern't* (for *(if) it were not*),¹⁸ which are commonly called contractions are syncopes, some of which are, as mentioned above, likely to influence the rhythm of an utterance, and help to compensate for stress lapses. They are widely used in popular music. In the phrase *I believe you gonna lose your mind* (Tommy McClennan, *Whiskey-Headed Woman*)¹⁹ which displays a regular eurhythmic stressed-unstressed sequence, the form *gonna* avoids the possible stress lapse generated by *going to*. Also note the absence of the unstressed auxiliary (*are* is reduced to 're and deleted).

There is therefore evidence that the linguistic rhythm can leave an imprint on the musical rhythm which occurs at the level of the melody, that is, at the level shared by both linguistic and musical domains. Indeed,

¹⁷ The word syncopation derives from the Greek "sygkoptô", meaning *break, disruption*.

¹⁸ This structure is taken from W. C. Handy's *Saint-Louis Blues*, and is characteristic of syllable reduction: only *were* is stressed because it carries the negation.

¹⁹ Tommy McClennan (1905-1961) was one of America's most successful blues singers.

melodic rhythms tend to reflect the cadence of the verses of a song and plausibly influence it.

4. Language and melody

In this section, further suprasegmental considerations will be taken into account; they provide extra data about stress and intonation, and their effect on the way words, phrases, and sentences/utterances are articulated, rise and fall when one speaks. As discussed earlier, syllables are not pronounced with equal force. In English, stressed syllables contrast with unstressed ones. The difference is that unstressed syllables are less salient and generally occur in reduced forms, while stressed syllables are more prominent and occur in their full form. Compare, for instance, the sounds corresponding to the realization of the letter <o> in the word *photograph*. In this word, the first syllable is stressed and contains a full vowel, the second syllable is, conversely, unstressed and contains a reduced vowel. Thus, <o> in the first (stressed) syllable will be pronounced as /əʊ/, reduced to /ə/, which is typically short and occurs in unstressed syllables. Also, the consonant /t/ in unstressed position is likely to be reduced to a flap.²⁰

Musically, the syllables, as mentioned before, are pronounced on different notes (different pitches in the scale of utterable sounds for a given performer).

The sequence of notes constitutes melodies which can be limited to two types in accordance with what is perceived at the end of the sentence: a rising melody or a falling melody. In the line taken from Ed Sheeran's song *Photograph* illustrated below, only the lexical words (*keep, love, photograph*) are stressed and alternate with function words (*we, this, in*) which typically occur in unstressed positions.

(11) We keep this love in a photograph.

Generally, the music follows this sound pattern. Listening to the song offers great help when analyzing a sentence in terms of syllables, word stress, rhythm, and intonation while practicing.

²⁰This voiced alveolar consonantal sound (also known as a 'tap'), transcribed phonetically as /r/, is produced when the blade of the tongue momentarily touches the roof of the mouth. A good example of the flap in English can be found in the way the word *better* is pronounced by Paul McCartney of the Beatles in the song *Hey Jude*. This is also the sound that most American speakers make instead of /t/ and /d/ in words like *battery, letter, ladder, bedding*, etc. (cf. Carr 2012, pp. 13, 105-106).

In a statement, for instance, in which the relation can be true or false, the intonation/melody is uttered with a falling pitch. Take the first lines of the song *Let It Be*, by the Beatles, which displays eurhythmic alternations:

- (12) a. When I 'find myl'self in 'times of 'trouble,
 b. 'Mother 'Mary 'comes to 'me.
 c. 'Whisper 'words of 'wisdom, 'let it 'be.

These lyrics have been divided into metrical trochaic feet, starting from the first stressed syllable, as indicated by the vertical lines separating the feet. They contain tone groups, the last syllable of which being perceptually more salient than the others (tonic placement), and this will be illustrated in the music, as shown in the transcription of the first two stretches, (12a-b), below:

(13)

When I find my-self in times of trouble Mother Ma - ry comes to me

In this example, the music follows the trochaic rhythm of the language: downbeats or the first part of beats correspond to stressed syllables, and two tone groups can be distinguished, as they are separated by a pause (i.e. the eighth note rest: '). The music shows that the first tone unit carries a rising pitch, conveying no impression of finality, but opening on a continuation, the 'answer', as it were, to an indirect question which is uttered with a falling pitch in the second tone group. A similar analysis can be conducted with the third line (example 12c), the first tone group of which, *Whisper words of wisdom*,²¹ being uttered with a rising pitch, and the last of which, *let it be*, being voiced with a falling tone. Furthermore, it may be noted that the tone groups in (13) correspond to the major syntactic constituents of the utterance, namely the two clauses in question. Also, as illustrated in the music (cf. example 13), the first two unstressed syllables introducing the song are transcribed as pickup notes (*anacrusis*) which precede the first strong downbeat, thus marking the beginning of the foot count. The music in this case follows the patterns of the language, that is, the rhythm, the intonation, and the English syntax.

In a complete question, namely a yes-no question, in which the validation of the subject-predicate relation is suspended, the intonation is rising (cf. Huart 2010, pp. 54-55). For example, in Elvis Presley's song *Are*

²¹ Note in passing the alliteration with /w/, which is another sound effect of the language.

You Lonesome Tonight?, there is a typically rising pitch in the singing. Conversely, in a statement such as the one exemplified in (14),

(14) What do you get when you fall in love? (Burt Bacharach, *I'll Never Fall in Love*)

in which the sequence *fall in love* is asserted, that is, in which the subject-predicate nexus [you – fall in love] is given, the pitch contour is a fall. The question applies to the object of ‘getting’. Again, this feature can be perceived in the way the sentence is sung. A scalar representation in (15) may illustrate the pitch contour of this line:

(15)



What do you get when you fall in love?

The music more or less follows the sound patterns with a drop on the word *fall*, as shown in example (16) below:

(16)



In this instance, the nucleus (tonic syllable) is *love* which carries a low rising tone and conveys a sense of mutual agreement as well as a desire to add something. The high pitch on *get* corresponds to the ‘head’, the part extending from the first stressed syllable to the tonic syllable (cf. Roach 2009). In addition, it may be noted in the musical transcription that the lexical words, *get*, *fall* and *love*, and the quarter notes (which correspond to strong beats)²² are aligned and constitute the rhythmic backbone of the phrase, both musically and linguistically. Also notice that the slur (the tied notes) in the second measure bar ends on a strong beat (cf. the quarter note on the third beat).

The description outlined above is certainly a limited one, but it nevertheless provides sufficient evidence to make significant observations

²² In a 4/4 measure bar, the first and the third beats are strong beats with predominance given to the first beat of the measure; the second and fourth beats are weak.

and trigger further investigations building upon the close relationship there is between music and language. What is important here is the fact that the prosody of the language can be reflected in its music, and maps onto its inflections. Moreover, the linguistic and musical syntaxes, as can be seen in the musical transcription of excerpt (14), display features which are shared by both areas: for instance, the melodic rhythm closely reflects the cadence of the verse. Equally, the analyzed data tend to demonstrate that the linguistic rhythm and intonation interact, as it were, with the musical rhythm, and this mainly concerns the melody, which is the defined stage shared by the two systems and the locus where the influence of one upon the other can actually operate.

5. Language is music

Beyond rhythm and intonation, prosody also involves many auditory properties of speech such as loudness, duration, pitch, and pauses. Prosodic features are mainly suprasegmental and are carried by every segment in the utterance. Modulating word stress and syllable structure has indeed the effect of influencing the rhythm of the language, and this can be perceived in its music. But in the end, all the auditory aspects of speech put together, notably compensation strategies, contribute to maintaining as much as possible its trochaic stressed-unstressed foot structure.

It can hardly escape notice, therefore, that there are similarities between language and music. All languages have their own music. In both systems there is a sense that rhythmic segmentation, or grouping, is divided into phrases of equal duration and organized in a periodical succession to create a base rhythm. This can be seen in example (16), in which the musical representation of function words, for instance, are grouped into beamed eighth notes, leaving perceptual salience to lexical words.

Rhythm is the main component of both language and music. It operates on similar systemic principles. This is one of the reasons why it is often said that language is music (cf. for example, Saraysky 2009), and introducing musical elements, such as speech rhythm, intonation, tone, and tempo, in the description of language may be of use when carrying out the study of linguistic (notably phonological) phenomena, as it is naturally related to speech.

According to this view, the present survey has been an attempt to demonstrate how the rhythmic features of English phonology are reflected in lyrics and in music, and thereby broaden our understanding of the correlation there is between the words of a song, its music, and the rhythm of English. What is more, the predominantly monosyllabic words in English make it an easy medium to align words and music, and enforce the trochaic (either

binary or ternary) patterns of the language. More thorough analyses may highlight other significant features apt to mutually clarify both systems, and the workings of the human mind as to the transmission of information. From the aforementioned observations, it may be worth delving more deeply into the description of rhythm and intonation in that special relationship between language and music, and it may be interesting to extend the study to other music forms and songs. This question is likely to be explored further.

Urban rap, which began in the 1970's, for example, sounds like a recitative rhyme scheme in which the rhythm of the language has great importance. It is as if the rhythm of American English gives the singing its tempo. Since rap music is essentially based on rhythm rather than melody, the flow of syllables in utterances enables the singers to put necessary energy into words and speech production when they rap. This music genre is liable to demonstrate that music and language share a fair amount of aspects which are firmly rooted in speech.

6. To conclude

What has been sketched herein is not intended to add to the oral grammar of English (cf. Huart 2010) by treating the phonology of the language completely, using musical forms, but rather, it focused on a number of sound and linguistic aspects such as pitch, duration, word length, and to a lesser extent loudness, in the perception of stress and intonation, thereby suggesting that these features can be connected to language use with a view to show how this relates to music.

It is not an easy task, because music and language tend to mutually influence each other, and it is sometimes difficult to decide which one serves as a reference to the other. It may, indeed, be contended that in some instances the music takes the lead owing to musical constraints,²³ but it seems plausible that speech prosody influences musical structure (cf. Patel 2007, pp. 224-225), notably at the level of the melody, that is, at the level shared by the two systems. In song lyrics, for instance, words and music are independent components, and syllables should be aligned in time, which means that stressed syllables must, as best as possible, match with strong positions, and avoid lapses and syllable compression (cf. for example, Hayes 2005).

Finally, this type of work on linguistic (mainly phonological and lexical) phenomena may contribute to providing foreign learners of English with a fun and familiar model, a musical and rhythmic reference on which

²³ Such as meter, rhythm, and tempo (speed).

they can rely and base their speech, and stress the words and phrases in order to produce more ‘swinging’ and expressive utterances.

Bionote: Patrice Larroque holds a PhD in English linguistics at the Université de Lorraine, Nancy, France. He is Professor Emeritus of English linguistics in the Department of English at the Université Jean-Jaurès, Toulouse, France, and a blues and jazz musician. His main interests are in the domains of linguistic variation (especially in nonstandard English grammar and African American English), translation studies, and the relations between language and music. He is the author of *English Rhythm and Blues, When Language and Music Come Together* (Routledge, 2021), first published in French as *English Rhythm and Blues. Les liens étroits entre le blues et l’anglais* (L’Harmattan, 2015), and *An Introduction to Linguistics through Popular Music* (Palgrave Macmillan, 2023).

Author’s address: patrice.larroque@univ-tlse2.fr

References

- Abercrombie D. [1967] 1990, *Elements of General Phonetics*, Reprint edition, Edinburgh University Press, Edinburgh.
- Attridge D. 1982, *The Rhythm of English Poetry*, Longman, London and New York.
- Benveniste E. 1974, *Problèmes de linguistique générale*, vol. 2, Gallimard, Paris.
- Brown H.D. 2024, *Principles of Language Learning and Teaching* (6th edition), Pearson Education, Philadelphia PA.
- Carr P. 2012, *English Phonetics and Phonology: An Introduction* (2nd edition), Wiley Blackwell, London.
- Deas S. 1955, *Music, In Scotland. A Description of Scotland and Scottish Life*, in Henry W. Meikles (ed.), pp. 210-213, Thomas Nelson and Sons Ltd., London, Edinburgh, Paris, Melbourne, Toronto, and New York.
- Dell F. and Halle J. 2005, *Comparing musical text setting in French and English*, in Aroin J.L. (ed.), proceedings of the conference *Typology of Poetic Forms*, Paris.
- Drabble M. 1985, *The Oxford companion to English Literature*, 5th edition, Oxford University Press, Oxford, New York, Tokyo, Melbourne.
- Flamenco Magazine*. <http://www.flamencoweb.fr>.
- Giegerich H. 1985, *Metrical Phonology and Phonological Structure: German and English*, Cambridge University Press, Cambridge, London, New York, New Rochelle, Melbourne, Sydney.
- Giles H. and Niedecki N. 1998, *Italian is Beautiful, German is Ugly*, in Baurie L. and Trudgill P. (ed.), *Language Myths*, pp. 85-93, Penguin Books, London.
- Graham C. 2002, 2006, *Jazz Chants; Let's Chant, Let's Sing*, Oxford University Press, Oxford.
- Grout D.J. and Palisca C.V. 2000, *History of Western Music* (6th edition), W. W. Norton, New York.
- Hasty C. 1997, *Meter as Rhythm*, Oxford University Press, Oxford.
- Hayes B. 2005, *Textsetting as Constraint Conflict*. https://www.researchgate.net/publication/247731408_Textsetting_as_Constraint_Conflict.
- Herskovits M. 1941, *The Myth of the Negro Past*, Boston, Beacon Press.
- Huart R. 2010, *Nouvelle grammaire de l'anglais oral*, Ophrys, Gap and Paris.
- Jones D. 1977, *English Pronouncing Dictionary*, revised by Gimson A.P. (14th edition), J. M. Dent Ltd., London and Melbourne.
- Jones L. 1963, *Blues People*, Harper Perennial, New York, London, Toronto, Sydney.
- Larroque P. 2012, *The Rhythm of English and Blues Music*, in "The American International Journal of Contemporary Research" (AIJCR) 2 [5], pp. 123-135.
- Larroque P. 2018a, *The British Drive on the Left: A contrastive Approach to Language Learning*, in Łodej M. and Strzałka A. (ed.), *Annales Universitatis Paedagogicae Cracovensia, Studia Anglica* 8.
- Larroque P. 2018b, *With Grammar you Go by Feel, E-rea*. <https://journals.openedition.org/erea/6186>.
- Larroque P. 2021, *English Rhythm and Blues. When Language and Music Come Together*, Routledge, London and New York; first published in French as *English rhythm and blues. Les liens étroits entre le blues et l'anglais* (L'Harmattan, Paris 2015).
- Larroque P. 2023, *An Introduction to Linguistics through Popular Music*, Palgrave Macmillan, London, New York and Shanghai.
- Lerdhal F. and Jackendoff R. 1983, *A Generic Theory of Tonal Music*, MIT Press, Cambridge MA.

- Liberman M. 1975, *The International System of English*, PhD, Massachusetts Institute of Technology.
- Nazzi T., Bertoncini J. and Mehler J. 1998, *Language Discrimination in Newborns. Toward an Understanding of the Role of Rhythm*, in “Journal of Experimental Psychology: Human Perception and Performance” 24, pp. 756-777.
- Nespor M. and Vogel I. 1989, *On Clashes and Lapses*, in “Phonology” 6, pp. 69-116.
- Parkinson T. and Vyner L. 2024, *More than Words: Songs for the Language Classroom*, Elements in Twenty-First Century Music Practice, Cambridge University Press, Cambridge,
- Patel A.D. 2007, *Music, Language, and the Mind*, Oxford University Press, Oxford.
- Pierrehumbert J. 2003, *Phonetic diversity, statistical learning, and acquisition of phonology*, in “Language and Speech” 46 [2-3], pp. 115-154.
- Pike K.L. 1945, *The Intonation of American English*, University of Michigan Press, Ann Arbor.
- Ramus F. 2002, *Language Discrimination in Newborns: Teasing Apart Phonostatic, Rhythmic, and Intonational Cues*, in “Annual Review of Language Acquisition” 2, pp. 85-115.
- Roach P. 2009, *English Phonetics and Phonology: A Practical Course*, Cambridge University Press, Cambridge.
- Rousseau J.-J. [1753] 1993, *Lettres sur la musique française. Essai sur l'origine des langues*, pp. 127-184, Flammarion, Paris,
- Saraysky S. 2009, *Language is Music*, Kaleidomundi.
- Schlüter J. 2005, *Rhythmic Grammar. The Influence of Rhythm in Grammatical Variation and Change in English*, Mouton de Gruyter, Berlin and New York.
- Selkirk E. 1984, *Phonology and Syntax. The Relation between Sound and Structure*, MIT Press, Cambridge MA.
- Temperley D. 1999, *Syncopation in Rock: A Perceptual Perspective*, in “Popular Music” 18 [1], Cambridge University Press, Cambridge.
- Tomatis A. 1977, *L'oreille et la vie*, Robert Laffont, Paris.