

# **The metrics of folk song: a comparative study of text-setting in Spanish and English**

**Rosalía Rodríguez Vázquez<sup>1</sup>**

**University of Edinburgh, Universidade de Vigo**

**Abstract:** This article re-examines the typological distinction between stress-timed and syllable-timed languages from an interdisciplinary perspective. A theoretical and empirical analysis is developed to determine whether the study of the setting into music of specific languages – in this case, Spanish and English – can shed light on the examination of the speech and verse prosody of those languages. The varying rules of text-setting for Spanish and English are seen to reflect the prosodic characteristics of the corresponding languages.

**Key words:** stress-timed, syllable-timed, rhythm, metre, verse, text-setting, English, Spanish.

## **Introduction: the relation between speech prosody, verse prosody and musical rhythm**

Several branches of linguistic and literary theory have dealt with the role of rhythm in speech and verse, highlighting the similarities between these two and music. The central hypothesis in this article is that when we talk about rhythm, three levels come into play, namely (i) speech rhythm, also known as *prosody*, which refers to the analysis of the rhythm of ordinary speech, (ii) verse prosody, which is the formal system used as a tool for poetic composition and scansion, and (iii) musical rhythm, which deals with the organisation in time of beats and note values. If we accept the statement that rhythm pervades at least these three levels of human cognition, it must follow that the time in which they are articulated is measured, that is to say, it has a perceptually regular pulse or metre. As Couper-Kuhlen (1993:112) observes,

[t]he fact that speech, verse, and music all have hierarchically organised metrical structure implies [...] a common cognitive origin. Not only are the principles of organisation surprisingly similar for all three faculties, but they also allow for the same play-off between abstract construct or underlying structure and actual realisation.

Although conversational speech is clearly less regular than poetry and than most musical forms, experiments have shown that, even in conversation, “speakers even attune to each other’s rhythm, so that the same pulse underlies the speech [...] of both” (van Leeuwen 1999:43).

The core research question addressed in this article has to do with how verse is set into music in Spanish and in English, that is, what is the relation between speech prosody,

---

<sup>1</sup> This paper will be published as well in *Rhythmica. Revista española de métrica comparada*.

verse prosody and musical rhythm in each of the two languages, and what generalisations can be made regarding the matching of speech rhythm with poetic metre and with musical rhythm.

The article is divided into two main sections. The first one constitutes, on the one hand, an overview of the typological classification of Spanish and English in reference to rhythm, and on the other, a brief exploration of the correspondences between the prosodic typologies of those two languages and their verse prosody. The second part is an analysis of the process of text-setting in Spanish and English. The materials which I am using in my analysis are two folk songs extracted from Miguel Manzano Alonso's compilation of Castilian folk songs *Cancionero de Burgos*, and Peter Kennedy's collection of English songs in *Folksongs of Britain and Ireland*.

With this analysis I hope to prove the existence of a correspondence between the timing typologies of language and the rhythmic typologies of music. I shall argue that there are mismatches between speech prosody, verse and music rhythm, and that these inconsistencies work in different ways for different languages. While in Spanish I hope to find a natural counterpoint between speech prosody and verse prosody/musical rhythm, in English this counterpoint will be considered arrhythmic. In other words, the rules of Spanish speech prosody and verse prosody are subservient to the rules of musical rhythm, while the rules of English speech prosody, verse prosody and music rhythm and metre are in almost total agreement with one another.

### **Speech and verse prosody in Spanish and English**

English and Spanish, like most languages in Europe, are stress languages. This means that they make use of stress, that is, the phonetic peak or climax occurring typically in every word (Allen 1973:86, Hyman 1977:39), in order to mark syllable prominence. Stress languages have been further classified into stress-timed and syllable-timed languages (Pike 1945, Abercrombie 1967). In syllable-timed languages syllables tend to be perceived as being rhythmically equal, while in stress-timed languages stresses fall at perceptually equal intervals. In the former group the syllable is considered the major unit of rhythm and in the latter this role is played by the foot – a stretch of utterance which begins with a stressed syllable and includes all unstressed syllables following it. Focusing on specific languages, Pike (1945) classifies English as the stress-timed language *par excellence*, while Spanish is considered one of the prototypically syllable-timed languages. Up to the present day, nearly all the works that deal with the study of speech rhythm in these and other languages take

Pike's theory as a theoretical point of departure, either to reassert the validity of his classifications or to contradict them. Since there have been numerous articles and reviews written on this topic – see, for instance, Bolinger (1965), Abercrombie (1967), Halliday (1967), Ladefoged (1967), Dauer (1983), on English, and Navarro Tomás (1918), Allen (1975) Gili Gaya (1950), Pointon (1980), Fant (1984) on Spanish – I will not be reviewing them here yet once again, but will summarise the characteristics that define a stress-timed language such as English in opposition to a syllable-timed language like Spanish.

Following Dauer (1983) and Bertinetto (1989), there are several characteristics that determine the classification of a language as stress-timed or syllable-timed. First, in stress-timed languages, stressed vowels have full articulation, while unstressed ones are reduced or centralised. In syllable-timed languages, vowels retain their distinct quality in both stressed and unstressed syllables.

| Spanish      |              | English            |                   |
|--------------|--------------|--------------------|-------------------|
| piso /'piso/ | pisó /pi'so/ | dígest /'daɪdʒəst/ | digést /dɑːdʒəst/ |
| póso /'poso/ | posó /po'so/ | súrvey /'sɜːveɪ/   | survéy /sə'veɪ/   |
| páso /'paso/ | pasó /pa'so/ | tórment /'tɔːmənt/ | tormént /tə'ment/ |
| péso /'peso/ | pesó /pe'so/ | ímport /'ɪmpət/    | impórt /ɪm'pɔːt/  |

Table 1: Vowel quality in stressed and unstressed syllables

Second, in stress-timed languages, segmental quality and quantity distinctions are manifest in stressed syllables. In syllable-timed languages, quality, quantity and stress are independent to a large extent, as seen in Table 1 above.

Third, in stress-timed languages, stressed syllables are usually longer than unstressed syllables and the latter usually undergo compression. In syllable-timed languages, stressed syllables are only slightly longer.

Fourth, in stress-timed languages, the stressed vowel is more or less compressed depending on the number of unstressed syllables surrounding it. In syllable-timed languages, there is no compensatory shortening, as stresses do not need to fall at equal intervals.

Fifth, in stress-timed languages, there is a wide variety of syllable types, and relatively uncertain syllable boundaries. In syllable-timed languages, CV and CVC predominate, with well-defined syllable boundaries. Table 2 contains the most frequent syllable structures in English and Spanish (adapted from Laver 1994:531, where 'C' stands for consonant and 'V' for vowel):

| Syllable type        | English % | Spanish % |
|----------------------|-----------|-----------|
| CV (light, open)     | 34        | 58        |
| V (light, open)      | 8         | 6         |
| CVC (heavy, closed)  | 30        | 22        |
| VC (heavy ?, closed) | 15        | 6         |
| CVCC (heavy, closed) | 6         | –         |

Table 2: Most common syllable structures in English and Spanish

Sixth, in stress-timed languages, heavy syllables tend to be stressed, while light syllables tend to be unstressed. In syllable-timed languages, syllable weight and stress are, in principle, independent. The independence of stress and weight in syllable-timed languages means that they are quantity-insensitive languages. However, Spanish seems to be a controversial case in this respect. Harris (1983:88) states that Spanish has inherited from Latin the principle that “antepenultimate stress is impossible if the penultimate syllable has a branching rhyme”, according to which Spanish would have to be classified as a quantity-sensitive language. Nevertheless, Roca (1988:417) states that Spanish stress is not quantity-sensitive. His arguments are reinforced from different perspectives, both by Hayes’ theory that trochaic systems such as Spanish are prototypically quantity-insensitive (in Lipsky 1997:563) and by Trubetzkoy’s idea that stress sensitivity to syllable quantity presupposes a systematic vowel length contrast (in Roca 1997:621), as well as a tense/lax contrast (Hyman 1977:48), both of which are absent in Spanish.

Seventh, in stress-timed languages, word stress is free – as opposed to fixed – and it is signalled by a combination of length, pitch change, loudness and quality. In most syllable-timed languages, there is fixed word stress. Although Spanish non-verbal stress must fall on one of the last three syllables of the word, this does not mean that Spanish is a fixed-stress language. Spanish, like English, is a free stress language, where stress can virtually occupy any position within the word. In these languages stress is unpredictable, and therefore lexical – at least partially – and phonemic. Other syllable-timed languages, like French or Polish, are characterised by the fixed position of their stress, which is predictable, and therefore grammatical and non-phonemic.

Eighth, in stress-timed languages, the intervals between stresses are relatively short; when they are long, new stresses can be introduced. In syllable-timed languages, intervals have variable durations; there can be stress clashes, but not stress shift – the opposite of what happens in stress-timed languages (see (1) below).

(1)

| English                          | Spanish                       |
|----------------------------------|-------------------------------|
| <i>Thirtéen vs. thirteen mén</i> | <i>Algún and algún hómbré</i> |

Last, in stress-timed languages, pitch change corresponds with stress. In syllable-timed languages, they are independent variables. In English, for instance, higher pitch in the accented syllable corresponds with longer duration and louder intensity, while the vowels of unstressed syllables are reduced in both loudness and duration (Allen 1975:80). This correspondence is not as obvious in Spanish, where higher pitch and louder intensity do not render a longer duration.

In conclusion, the most salient feature of stress-timed languages like English in opposition to syllable-timed languages like Spanish is the existence of vowel reduction, that is, the neutralisation of vowel quantity and quality – unstressed vowels tend to be centred and realised as a schwa (/ə/) –, which is linked to the rhythmic requirement that stresses fall at roughly equal intervals. In Spanish, given that stresses do not have to fall isochronously, vowels are not reduced, and syllables are kept untouched, thus becoming the markers of rhythm in this language.

Turning now to verse prosody, there is no surprise in saying that the speech prosody of a language correlates to its verse prosody. Spanish speech prosody, as already mentioned, is based on the isochronic occurrence of syllables. If we refer to verse prosody, Spanish verse is syllabic, that is, the syllable constitutes the main rhythmic unit in Spanish verse, which can be proved by the fact that much Spanish verse has a perfectly regular syllable count. In fact, the names of Spanish lines – ‘heptasyllabic’, ‘octosyllabic’, ‘hendecasyllabic’, and so on – are an explicit reference to the number of syllables that they contain. Let us have a look at the following example by Antonio Machado:

(2)

|                                                              |                    |
|--------------------------------------------------------------|--------------------|
| <i>Adoro la hermosura, y en la moderna <u>est</u>ética</i>   | (15 - 1 syllables) |
| <i>corté las viejas rosas del huerto de Ronsard;</i>         | (13 +1 syllables)  |
| <i>mas no amo los afeites de la actual <u>cosm</u>ética,</i> | (15 - 1 syllables) |
| <i>ni soy un ave de esas del nuevo gay-trinar.</i>           | (13 +1 syllables)  |

The evolution of Spanish metrics has been marked by the contact with other Romance traditions. Originally, Spanish had an accentually-based metrics that turned into syllabism due to French influence (Duffell 1999:45). Apart from being isosyllabic, Spanish verse also takes into account metrical accents. The lines in each kind of poetic composition carry a fixed number of rhythmic accents on specific syllables, one of which is invariably the last-but-one syllable in each line. In (3), another fragment of a poem by Machado, the underlined syllables correspond to poetic accents, which fall regularly on the 3<sup>rd</sup> and 7<sup>th</sup> syllable of each line, with an extra one on the 1<sup>st</sup> syllable of the 4<sup>th</sup> line:

(3)

*Caminante, son tus huellas  
El camino, y nada más.  
Caminante, no hay camino,  
Se hace camino al andar.*

Spanish isosyllabic poetry uses a series of metrical devices in order to regularise syllable counts. Those same metrical devices also serve the purpose of forcing rhythmic accents to fall on their corresponding syllables, thus rendering the line metrical. Two of them – synaeresis and dieresis – take place at the word level, while the other two – synalepha and hiatus – take place at any level beyond the word.<sup>2</sup> Summarising, Spanish verse prosody is syllabic in that the vast majority of Spanish verse is composed and scanned by counting the syllables in each line. In order for lines to be isosyllabic, several devices are used in order to adjust the number of syllables, namely syneresis and synalepha – which respond to a natural tendency of the language – and dieresis and hiatus, together with several other devices. Apart from isosyllabism and despite the marked syllabic character of Spanish, there is another principle at work in Spanish verse prosody, which is the recurrence of stresses at roughly regular intervals (see (3) above). This is helped by the devices quoted above. We shall emphasise, once again, that in Spanish poetry, just as in Spanish speech, the unstressed syllables which fall in between stressed ones are never reduced, so the underlying rhythm is dependent not so much on the stresses themselves, but rather on the fact that those stresses are mediated by the same number of syllables – three in (3).

---

<sup>2</sup> Compression devices such as synalepha are attested in oral speech in Spanish (see Piera 1980). For a thorough explanation of these devices as they occur in poetry and speech, see also Canellada and Madsen (1987).

English is a stress-timed language, which would make us think of its verse prosody as accentual. Nevertheless, English verse is typically accentual-syllabic. The characteristics of PdE verse prosody have to do with its hybridity, inherited from a fertile interaction between the native traditions and the Romance one. Old Germanic verse evolved from syllabic into accentual (see Lass 1994). In the Middle Ages, the Latin quantitative foot and the English accentual foot were connected through French, which was purely syllabic. Around the fourteenth century English verse gradually became what it currently is, that is, an accentual-syllabic type of verse. Present-day English verse is built upon the basic principle that there has to be a specific number of stresses per line, each of which gives rise to and governs its respective foot. The second principle states that the number of syllables per foot must be identical. The foot as a unit of scansion was inherited from the Greco-Latin tradition. English feet are said to be isochronous, although, as happened in speech prosody, isochrony is perceptual. The names of the four basic types of English feet – ‘iamb’ (x /), ‘trochee’ (/ x), ‘anapaest’ (x x /) and ‘dactyl’ (/ x x) – have also been inherited from the classical tradition. As for the devices which are used in English in order to adjust the duration of feet, synalepha is a standard metrical device which has to do with the abstract metrical representation of the poem. Independently of this, there is also to pronounce or not to pronounce the syllable in question. We can check the realisation of synalepha in the following line by E.B. Browning (example borrowed from Fabb 2002:10):

(4) “Say over again, and yet once over again”

|            |          |            |          |              |            |            |             |          |            |          |             |
|------------|----------|------------|----------|--------------|------------|------------|-------------|----------|------------|----------|-------------|
| <i>Say</i> | <i>o</i> | <i>ver</i> | <i>a</i> | <i>gain,</i> | <i>and</i> | <i>yet</i> | <i>once</i> | <i>o</i> | <i>ver</i> | <i>a</i> | <i>gain</i> |
| *          | *        | 0          | *        | *            | *          | *          | *           | *        | 0          | *        | *           |

As a sort of conclusion, we could say that English verse prosody responds to the building principles of English speech prosody, that is, reduction of vowel sounds in order to get a specific number of isochronous feet per line.

In the case studies that follow this section we are going to explore the behaviour of English and Spanish verse when it is set into music, so that the principles sketched up to now can be confirmed or contradicted.

### **Case study I: English folk song**

“The farmer’s boy” is a strophic song collected in Peter Kennedy’s (1984[1975]) *Folksongs of Britain and Ireland*, more specifically in the subsection entitled “Songs of country life”. According to Kennedy’s (1984:547) introduction, “The farmer’s boy” was a very popular song throughout England, although the version collected by Kennedy incorporates a tune originally from Galloway, in Scotland. As happens with the song chosen to exemplify text-setting in Spanish, the choice of this English song is not random. The genre of folk song comprises vocal compositions which have been passed from generation onto generation; in this respect, the matching of lyrics and music must necessarily respond to some kind of native language of text-tune alignment, where both the rules and the exceptions are clearly part of the song grammar of that specific tradition. The two songs analysed here have been only minimally edited by the collectors themselves, something which ensures the native quality of the songs. Since the genre of song is obviously intended to be sung, we can presume that the prosody of the Spanish text will relate to the rhythm of the music according to traditions of Spanish word-setting, and that the same will happen with the English text.

Going back to this specific case study, “The farmer’s boy” is a narrative divided into four four-line stanzas, each of which is followed by a refrain (in italics in Table 3) with two variable lines at the beginning – which work as a link between the narrative of the previous stanza and that of the refrain itself – and two invariable lines at the end.

As can be observed in the scansion of the lyrics given in Table 3 below, the number of syllables per line varies between six and nine. There are two instances of synalepha (underlined in Table 3), which in English is a correspondence rule (Halle and Keyser 1971:71) with non-phonological status. In this case, the realisation of the two syllables in ‘little’ as one is a poetic device with no phonological status.<sup>3</sup> The case of ‘little’, pronounced /’lɪtl/, corresponds to what Fabb (2002:10) terms “non-projection rule c”, quoted below for clarity’s sake:

Non-projection rule c: Optionally, do not project a syllable which has as its nucleus one of the following sonorant consonants: [l], [r], [m] or [n], or which has as its nucleus the weak vowel schwa followed by one of these sounds.

---

<sup>3</sup> The non-phonological status of non-projection rule is arguable, as the degree of sonority of phonological segments is taken into account in several prosodic processes. Non-projection rule c corresponds to the invisibility of certain syllables with a syllabic consonant as their nucleus in the stress assignment process. As an illustration of this, observe how a word like ‘badminton’, antepenultimate stress can only be expressed by taking the last syllable, with a syllabic [n] as nucleus, as invisible or non-projected. In other words, the syllable /ən/ does not count as such for the purposes of stress assignment – if it did, the stress would fall on the penult.



On the other hand, the synalepha between the article ‘the’ and the first syllable of the adjective ‘eldest’ corresponds to Fabb’s (2002:9) “non-projection rule b”:

Non-projection rule b: Optionally, do not project a syllable which ends on a vowel, when that syllable precedes a syllable which begins on a vowel.

Non-projection rule b is equivalent to synalepha proper in Spanish, which uses the linking of vowels in order to bring the actual number of syllables in the line to the right number in metrical terms. In this case, given that the third line of the first stanza and the third line of the second stanza have six lines, we can presume that the third line of the third stanza will have six lines, too, something which is only achievable by the realisation of the synalepha between ‘the’ and ‘eld-’ in ‘eldest’, whereby the schwa in ‘the’ is elided and the sound /ð/ becomes the onset of the first syllable in ‘the eldest’, realised as /’ðeldɪst/

The lines in the poem are structured into alternating iambic tetrameters and trimeters, that is, odd lines consist of four feet with the structure weak-strong per line, with some minor variations, while even lines are arranged into structures of three feet with the same internal organisation weak-strong. This structure corresponds to the so-called “common metre” in Church hymns.

A minor variation in the structure of iambic tetrameters is the substitution of an anapaest (x x /) for an iamb (x /). In spite of this, all the odd lines in the poem conform to the structure of iambic tetrameter, which in some cases entails the promotion of certain unstressed syllables – ‘to’ (4<sup>th</sup> line), ‘I’ and ‘you’ (5<sup>th</sup> line), ‘will’ and ‘me’ (6<sup>th</sup> line), ‘if’ and ‘me’ (9<sup>th</sup> line), ‘this’ (12<sup>th</sup> line), ‘her’ (18<sup>th</sup> line), ‘what’ (19<sup>th</sup> line), ‘for’ (28<sup>th</sup> line) – as well as the demotion of certain stressed syllables – ‘ask’ (5<sup>th</sup> line), ‘give’ (6<sup>th</sup> line). Promotion and demotion of syllables is a thoroughly standard rhythmic device in English poetry, and could be compared to the phenomenon of stress shift in English prosodic phonology. As for rhyme, even lines rhyme in each stanza, while odd ones are left without rhyme, that is, trimeters rhyme, while tetrameters are left without rhyming.

| Line                                              | Syllable count | Rhyme | Foot structure |
|---------------------------------------------------|----------------|-------|----------------|
| The sun went down, beyond yon hills               | 8              |       | x/x/x/x/       |
| Across yon dreary moor                            | 6              | a     | x/x/x/Ø        |
| When weary and lame, a boy there came             | 9              |       | x/xx/x/x/      |
| Up <b>to</b> the farmer's door.                   | 6              | a     | x/x/x/Ø        |
| <i>May <b>I</b> ask you, if any there be</i>      | 9              |       | x/x/x/xx/      |
| <i>That <b>will</b> give me employ</i>            | 6              | b     | x/x/x/Ø        |
| <i>To plough and sow, to reap and mow,</i>        | 8              |       | x/x/x/x/       |
| <i>And to be a farmer's boy?</i>                  | 7              | b     | xx/x/x/Ø       |
| And <b>if</b> that thou won't <b>me</b> employ    | 8              |       | x/x/x/x/       |
| One thing I have to ask                           | 6              | c     | x/x/x/Ø        |
| Will you shelter me, till break of day            | 9              |       | xx/x/x/x/      |
| From <b>this</b> cold wintry blast?               | 6              | c     | x/x/x/Ø        |
| <i>At break of day I'll trudge away</i>           | 8              |       | x/x/x/x/       |
| <i>Elsewhere to seek employ</i>                   | 6              | b     | x/x/x/Ø        |
| <i>To plough and sow...</i>                       |                |       | x/x/x/x/       |
| My father's dead, my mother's left                | 8              |       | x/x/x/x/       |
| With <b>her</b> five children small               | 6              | e     | x/x/x/Ø        |
| And <b>what</b> is worse for mother still         | 8              |       | x/x/x/x/       |
| I'm <u>the</u> <b>eldest</b> of them all          | 6              | e     | x/x/x/Ø        |
| <i>Though <u>little</u> I be, I fear not work</i> | 8              |       | x/x/x/x/       |
| <i>If thou wilt me employ</i>                     | 6              | b     | x/x/x/Ø        |
| <i>To plough and sow...</i>                       |                |       | x/x/x/x/       |
| In course of time, he grew a man                  | 8              |       | x/x/x/x/       |
| The good old farmer died                          | 6              | f     | x/x/x/Ø        |
| And left the boy the house now has                | 8              |       | x/x/x/x/       |
| And his daughter <b>for</b> his bride             | 7              | f     | xx/x/x/Ø       |
| <i>The boy that was, the farm now has</i>         | 8              |       | x/x/x/x/       |
| <i>He thinks and smiles with joy</i>              | 6              | b     | x/x/x/Ø        |
| <i>Of the lucky day he came that way</i>          | 9              |       | xx/x/x/x/      |
| <i>For to be a farmer's boy.</i>                  | 6              | b     | xx/x/x/Ø       |

Table 3: Scansion of “The farmer’s boy” lyrics

The fact that the number of syllables changes from line to line in English is not unimportant, as it has musical implications. If the premise for text-setting is that, as a general rule, one syllable corresponds to one musical note value, then a higher number of syllables will render a higher number of values per bar, which means that the melodic and rhythmic contour of the first stanza will not necessarily be preserved throughout the whole composition. There is, therefore, a relatively high degree of freedom for the singer to insert or delete syllables and, accordingly, note values. As an example of this, observe how one of the two synalephas realised in the scanned version of the lyrics – the one between ‘the’ and ‘eld-’ in the twentieth line – is eliminated in the musical rendition of the text, where each of the syllables ‘the’ and ‘eld-’ is assigned to a different note value; ‘little’, on the other hand, is kept as a single syllable assigned to a single note value, as seen in Figure 1 below. In this song, as in

most songs, the only element that must remain constant is the occurrence of accents at equal intervals, something which is automatically achieved in vocal music, as there is always a strong accent on the first beat of each bar. The difficulty arises when accented syllables must be made to correspond to primary and secondary counts – that is, salient beats –, as is the case in English song. The singer’s skill is challenged at that level. In this sense, the musical rendition of the text is perfectly aligned with the scansion given in Table 3, in such a way that all the cases of promotion and demotion of syllables are made to correspond to strong and weak beats respectively, while no other syllables are mismatched. We can say, thus, that the agreement between verse prosody and music rhythm and metre is virtually total in this song and, presumably, in English folk song.

Figure 1: Musical score of “The farmer’s boy”<sup>4</sup>

We could summarise the above explanation by saying that, in English text-setting, there are two groups of constraints at work, namely metrical and grouping constraints. The first metrical constraint predicts that the rises and falls of stress within the line are matched to the rises and falls of the metrical pattern (Hayes and Kaun 1996:10). Nevertheless, it is violated on several occasions in the song.

There are several instances of mismatch on primary counts:

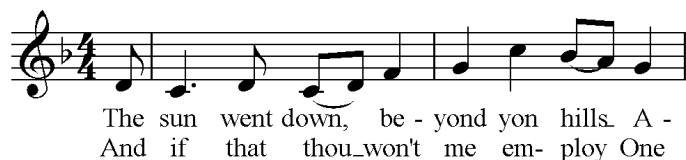
- First stanza: ‘to’ (4<sup>th</sup> line), ‘be’ (8<sup>th</sup> line).
- Second stanza: ‘if’ (1<sup>st</sup> line, as seen in the example below), ‘me’ (1<sup>st</sup> line, see (1)), ‘you’ (3<sup>rd</sup> line), ‘this’ (4<sup>th</sup> line), ‘elsewhere’ (7<sup>th</sup> line).
- Third stanza: ‘her’ (2<sup>nd</sup> line), ‘what’ (3<sup>rd</sup> line), ‘the’ (4<sup>th</sup> line).
- Fourth stanza: ‘the’ (7<sup>th</sup> line), ‘he’ (7<sup>th</sup> line), ‘to’ (8<sup>th</sup> line).

<sup>4</sup> In his collection, Kennedy underlays only the first stanza and the first refrain of the song, while the other stanzas are written out beneath the score. As a consequence of this, there are cases where the alignment of melody and text can be ambiguous. I have aligned all the stanzas following native intuitions of text-setting, which does not mean that my setting is the only possible setting for these specific stanzas.

There are also mismatches on secondary counts:

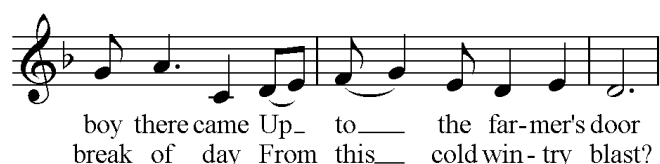
- First stanza: 'you' (5<sup>th</sup> line), 'be' (5<sup>th</sup> line), 'me' (6<sup>th</sup> line)
- Second stanza: 'thou' (1<sup>st</sup> line, see (1)), 'me' (3<sup>rd</sup> line)
- Third stanza: 'of' (4<sup>th</sup> line), 'be' (5<sup>th</sup> line), 'me' (6<sup>th</sup> line)
- Fourth stanza: 'for' (4<sup>th</sup> line), 'that' (7<sup>th</sup> line)

(5) Mismatches in the first line of the second stanza



Curiously, mismatches are not always signalled in terms of musical accent, but also in terms of note values. This is the case of what I call “duration mismatches”, where note values, not only the position of the beat within the bar, signal a mismatch. Let us observe the following example, where the function word ‘to’, realised as /tə/ in speech, not only falls on a main beat – stress mismatch –, but is made to correspond to a long note value – a quaver plus a crotchet –, something which renders it a more salient syllable than the content word ‘boy’ in the bar before, which, in spite of containing a diphthong, is aligned with a quaver. In this sense, we could say that the mismatch of the grammatical word ‘to’ is doubly salient.

(6) Duration mismatch



Duration mismatches such as the one above contradict Hayes and Kaun’s (1996:16) “Syllable Duration Rule”, quoted below:

Syllable Duration Rule: reflect the natural phonetic durations of syllables in the number of metrical beats they receive.

The second metrical constraint predicts that the four strongest positions in the line are filled with syllables (Hayes and MacEachern 1998:492). In this song, this constraint is violated only at the end of each couplet, as all the strong positions are filled with a syllable with the exception of the fourth position of every even line – as explained above, even lines are structured as trimeters, not tetrameters. Let us have a look at the grid notation for the first two lines, where we can observe that the fourth position of the second line is left empty.

(7) Non-filling of strong positions

|     |     |      |      |     |      |     |       |
|-----|-----|------|------|-----|------|-----|-------|
|     | x   |      |      |     | x    |     |       |
|     | x   |      | x    |     | x    |     | x     |
| x   | x   | x    | x    | x   | x    | x   | x     |
| The | sun | went | down | be- | yond | yon | hills |

|    |       |     |       |    |      |   |   |
|----|-------|-----|-------|----|------|---|---|
|    | x     |     |       |    | x    |   |   |
|    | x     |     | x     |    | x    |   | x |
| x  | x     | x   | x     | x  | x    | x | x |
| A- | cross | yon | drea- | ry | moor | Ø |   |

This phenomenon is linked to truncation, which, as explained above, means that the last foot in the line – in this case, even lines – is left unfilled. In terms of the correspondence between verse and musical grouping, this phenomenon gives rise to systematic non-correspondences between prosodic constituents (lines) and musical constituents (bars and phrases). In musical terms, the violation of the second metrical constraint entails that the unfilled strong position is forced to be linked to the previous one(s) – as in this song, where the end of each musical phrase is marked with a long note value –, but also that the very last beat might be left without no corresponding prosodic position in that line – in this song, the very last beat of each even bar is filled with a syllable from the next poetic line. This is what Hayes and MacEachern (1996:16-17) call “theft” of positions, which, in English, happens only to the left: “it is only in the leftward direction that the positions are there to be stolen”.

(8) Theft of positions (where { } signal a line, and [ ] signal a bar)

|     |        |     |       |    |       |   |              |
|-----|--------|-----|-------|----|-------|---|--------------|
|     | x      |     |       |    | x     |   |              |
|     | x      |     | x     |    | x     |   | x            |
| x   | x      | x   | x     | x  | x     | x | x            |
| {A- | [cross | yon | drea- | ry | moor} | Ø | <b>When]</b> |

The second metrical constraint is closely related to grouping constraints. This is the reason why it can only be violated when the grouping constraint that regulates the saliency of constituents – related to truncation – is active. In any other case, a violation of the metrical constraint that requires the filling of strong positions would render the text-setting instance unacceptable.

As for the third metrical constraint, it states that sequences in which no syllable is placed in the interval between any two of the four strongest positions in the line should be avoided. In this song, the constraint is violated whenever the previous constraint is, as weak positions between the third and fourth positions in even lines are not filled. A lapse of a single syllable is regarded as normal; it is only when the lapse comprises more than three syllables that the text-setting instance is unacceptable.

(9) Presence of lapses

|   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|
|   | x |   |   |   | x |   |   |
|   | x |   | x |   | x |   | x |
| x | x | x | x | x | x | x | x |

A- cross yon drea- ry moor Ø

The second group of constraints is that of grouping constraints, subdivided into SALIENCY and PARALLELISM. SALIENCY is directly related to the action of what Hayes and MacEachern (1998:476) call “rhythmic cadences”, which are the characteristic grid placements of the final syllable or two of the line. Hayes and MacEachern classify cadences into four main types, namely G, 4, 3 and 3f (for a complete explanation of these terms, see Hayes and MacEachern 1998). As explained when talking about the second metrical constraint, SALIENCY works at the level of the couplet, which is systematically made salient by the insertion of a 3 cadence, that is, the second line in each couplet has three strong positions filled, while the fourth one is left unfilled. The rhythmic cadence is reinforced by rhyme, which is at work in even lines. Let us have a look at how this works in the couplet of the first stanza (43):

(10) SALIENCY

|     |       |      |       |     |      |     |       |     |
|-----|-------|------|-------|-----|------|-----|-------|-----|
|     |       | x    |       |     |      | x   |       |     |
|     |       | x    |       | x   |      | x   |       | x   |
| x   | x     | x    | x     | x   | x    | x   | x     |     |
| The | sun   | went | down  | be- | yond | yon | hills | (4) |
|     |       | x    |       |     |      | x   |       |     |
|     |       | x    |       | x   |      |     |       | x   |
| x   | x     | x    | x     | x   | x    | x   | x     |     |
| A-  | cross | yon  | drea- | ry  | moor | Ø   |       | (3) |



PARALLELISM works at two levels, the quatrain level and the stanza level. At the quatrain level, it states that the cadences that end units of the maximal analysis – in this case, the couplet – are always identical (Hayes and MacEachern 1998:17) and must rhyme with each other (Hayes and MacEachern 1998:17), something which is true for the song in question, where the structure of couplets is always 43 and the two 3s in each stanza rhyme between themselves. At the strophic level, positional parallelism (Halle and Dell 2005:10) establishes that settings to the same tune have the same distribution of onsets with respect to the grid. This constraint is violated on the third line of the second, third and fourth stanzas, where there is a deletion of a syllable, which implies the deletion of a pitch.

(11) Positional PARALLELISM



May I ask you, if an-y there be That  
At break of day I'll trudge a - way Else

|     |       |   |     |   |     |      |   |               |   |          |       |     |    |   |   |   |   |      |   |
|-----|-------|---|-----|---|-----|------|---|---------------|---|----------|-------|-----|----|---|---|---|---|------|---|
|     |       | X |     |   |     | X    |   |               |   | X        |       |     |    | X |   |   |   |      |   |
| X   |       | X |     | X |     | X    |   | X             |   | X        |       | X   |    | X |   | X |   | X    |   |
| X   | X     | X |     | X | X   | X    | X | X             | X | X        | X     | X   | X  | X | X | X | X | X    | X |
| May | I     |   | ask |   | you | if   |   | <b>an-</b>    |   | <b>y</b> | there |     | be |   |   |   |   | that |   |
| At  | break |   | of  |   | day | I'll |   | <b>trudge</b> |   | a-       |       | way |    |   |   |   |   | Else |   |

In general, we observe how the melodic contour of the song “anticipates” the need of syllable insertion at certain points – where melismata happen –, in such a way that violations of PARALLELISM as such cannot be strictly observed at the musical level, although at the prosodic level there is a clear insertion of extra syllables, as we can see in the following example:

#### (12) Melismata



boy there came Up\_\_ to\_\_ the far-mer's door  
break of day From this\_\_ cold win - try blast?  
mo - ther still I'm the eld\_\_ dest of them all  
house now has And his daugh-ter for his bride

As a sort of conclusion to this section, we could say that any alteration of the musical structure is made in order not to violate the matching of stresses and beats or, at least, to minimise the number of violations of this metrical constraint, to which all other constraints are subsidiary. This is clearly seen in the violation of positional PARALLELISM, which happens in order to avoid mismatches between stressed syllables and strong beats. In this sense, metrical constraints seem to regulate text-setting in English, while grouping mismatches bear the consequences of the former.

#### Case study II: Spanish folk song

“Señor regidor que duerme” is a strophic song collected in Manzano-Alonso’s (2003) *Cancionero popular de Burgos*. More specifically, this song belongs in the subsection entitled “Cantos de aguinaldo en la fiesta de Santa Águeda”, which comprises five songs devoted to Saint Águeda’s festivity in the Castilian province of Burgos.

From the analysis of the lyrics we can conclude that the song was originally intended to persuade the audience – the most socially salient members of which were named in the lyrics – into giving a relatively small amount of money (*aguinaldo*) for the celebration of Santa Águeda’s party. In this respect, the singers themselves were entitled by traditional custom to change the lyrics and adapt them to the circumstances and the type of audience to whom they were singing. In terms of the general relation between lyrics and tune in this

specific song, it is worth noticing that, in spite of the singers' compositional freedom, the lyrics always remain constrained by the tune, that is, neither the melodic contour nor the rhythmic structure of the song can be changed in order to fit the lyrics, in such a way that, if a song has four note values per bar, they will have to be kept untouched no matter which word is intended or needed to fall on that bar. We can observe how this works in the notated score, where the two stanzas that configure the song are set to exactly the same tune, with no added or eliminated note values:

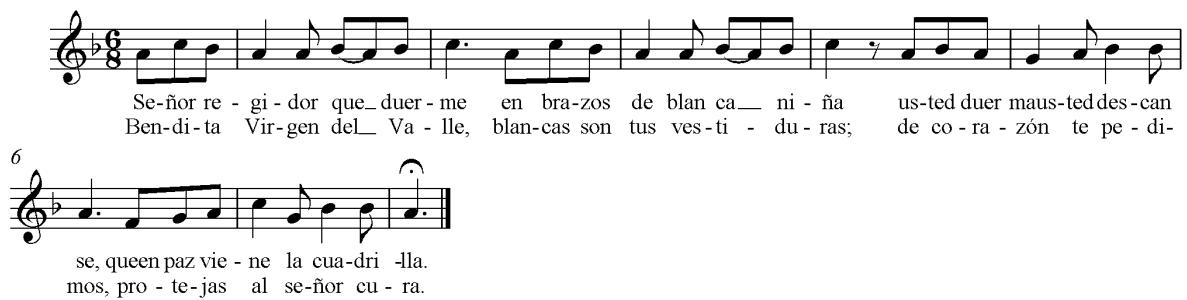


Figure 2: Musical score of “Señor regidor que duerme”

The fact that the melodic and rhythmic contours have to remain untouched has several prosodic consequences – one could also say that the prosody determines the melodic and rhythmic contour, but I will leave the matter of directionality aside for the moment –, one of which is the preservation of isosyllabism. In the above score we can observe how each note value corresponds to a prosodic syllable, with the exception of the melismata in bars one and three, in such a way that the preservation of the number of values in each musical phrase naturally leads to the preservation of the number of syllables in each line. Before going on, it is necessary to make clear that, in spite of the confusing and often inaccurate use of the term ‘beat’ in the literature, this term is not a synonym for ‘note value’. The term beat is normally equated with ‘tactus’, defined as the level at which a listener would clap.<sup>5</sup> In an instance of 6/8 metre, like that in “Señor regidor que duerme”, the tactus, that is, the beat, is located at the dotted crotchet level. On the other hand, a 6/8 bar can comprise various note values, of which the most common ones are the crotchet and the quaver combined, or as series of three quavers. Taking this into account, we can observe how, in this song, some syllables fall on the beat – like, for example, the syllable ‘se-’ – but have a smaller value than the beat itself,

### (13) Beat

<sup>5</sup> Although ‘beat’ and ‘tactus’ are nowadays used as synonyms, the later is better applied to music written before the bar era, while the former is applied to music organised into bars.





while others, while falling on the beat, take a bigger value, which may be realised on the same musical pitch – like ‘-gi-’ in ‘regidor’ – or on two pitches (what is known as a melisma), like ‘que’ or ‘del’:

(14) Beat and melismata

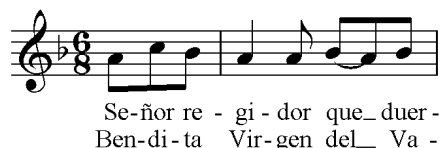



Table 4 presents us with the scansion of the lyrics in this song. We can observe that the eight lines which configure the recorded version of the song have eight syllables each. The only compression device used in this song is synalepha, both between two different vowels, in the third line – *duerma usted* –, and between different vowels, in the fourth line – *que en*. These two instances of synalepha are kept in the song (see (15)). It is interesting to note that, in the scanned version, the accented syllables in the poem fall roughly on the same syllables – with certain minor variations –, thus strictly following the rules of Spanish verse prosody. However, when the lyrics are set to music, the placement of prosodic accents is disregarded in favour of the musical accentuation, which, in 6/8 metre, naturally falls on the first beat of each bar – primary count –, and on the second beat of the bar, corresponding to the second dotted crotchet or the fourth quaver – secondary count. This has an effect of total disagreement between verse prosody and music rhythm in the song, with the latter winning over the former. Rhyme is systematically kept on even lines, while odd lines are left without rhyme. The resulting poetic form, the *copla*, is a very widely used stanza in popular poetry and folk song.

### (15) Musical realisation of synalepha



us-ted duer maus-ted des-can se, queen paz vie - ne la cua-dri -lla.  
de co - ra - zón te pe - di-mos, pro - te-jas al se-ñor cu - ra.

(16) Mismachthes

|     |     |     |     |   |     |     |   |       |    |   |   |
|-----|-----|-----|-----|---|-----|-----|---|-------|----|---|---|
|     |     |     | x   |   |     |     |   |       | x  |   |   |
| x   |     |     | x   |   |     | x   |   |       | x  |   |   |
| x   | x   | x   | x   | x | x   | x   | x | x     | x  | x | x |
| Se- | ñór | re- | gi- |   | dór | que |   | duér- | me |   |   |

Second, the constraint that requires the filling of strong positions in the line is never violated, as all the strong positions are filled with syllables. The violation of this constraint would render the setting of the lyrics unacceptable for a native listener of Spanish folk song. The

following example gives an alternative setting for the same lyrics, where the second and fourth strong positions are left empty.

(17) Filling of strong positions

|      |     |     |     |   |     |     |     |       |    |    |   |
|------|-----|-----|-----|---|-----|-----|-----|-------|----|----|---|
|      |     |     | x   |   |     |     |     |       | x  |    |   |
| x    |     |     | x   |   |     | x   |     |       | x  |    |   |
| x    | x   | x   | x   | x | x   | x   | x   | x     | x  | x  | x |
| Se-  | ñór | re- | gi- |   | dór | que |     | duér- | me |    |   |
| *Se- | ñór | re- | Ø   |   | gi- | dór | que | duér- | Ø  | me |   |

Third, the constraint that disallows lapses is not violated within stanzas, as there are no empty positions between any two strong positions. Nevertheless, violating this constraint would not render the composition ill-formed. Indeed, it is the case that the prohibition of lapses is violated in many other songs of the same tradition.

It is essential to observe that mismatches happen in a quantitatively and qualitatively different way in English and Spanish. While in the former all the violations of this constraint are “translations” of the verse scansion into the musical score – which implies that those mismatches do not really count as such –, in Spanish the text-setting process separates itself from the scansion of the lyrics in isolation and re-creates the rhythm of the poem by assigning syllables to beats in an apparently anarchic manner, so that neither linguistic stresses nor poetic accents necessarily fall on main beats. In this respect, the link between verse prosody and musical setting is much weaker in Spanish than in English.

The second big group of constraints is that of Grouping, which is subdivided into SALIENCY and PARALLELISM. As mentioned in the introduction to this chapter, SALIENCY is directly related to the concept of rhythmic cadence, which in turn has to do with how the last or two last positions in the line are fulfilled. In “Señor regidor que duerme”, the first and second stanzas share their rhythmic cadence structure, 4444, which refers to the fact that the four strong positions in each line, including the last one, are filled with syllables – in metrical terms, this would be classified as the second metrical constraint. This means that neither couplets nor stanzas themselves are salient in cadential terms. Nevertheless, there is a structural principle that signals constituency levels in this song, namely rhyme. The change in rhyme scheme (from –a–a to –b–b and so on) signals a change in stanza.

PARALLELISM works both at the quatrain level and at the song level. First, at the quatrain level, we have already mentioned how the cadences ending the units of the maximal analysis of the quatrains are identical (4 in every case), and how the cadences of the maximal analysis rhyme with each other (in this case, the cadence marking the couplet is realised in terms of rhyme, as pointed out before). Second, at the strophic level we observe positional parallelism, which has to do with the fact that settings to the same tune have the same

distribution of onsets with respect to the grid. This is directly related to the observation that the number of beats per bar cannot change in Spanish, which results in the preservation of isosyllabism throughout the song. Synalepha happens twice in the first stanza in order to achieve parallelism, which again responds to isosyllabism.

Summarising, in Spanish, musical structure dominates over prosodic structure, in such a way that text-setting constraints are applied much more leniently in Spanish than in English. The only underlying principle in Spanish text-setting is the parallel alignment of syllables and beats, which must remain constant throughout the whole song. PARALLELISM is more important than any metrical constraint in Spanish.

## **Conclusions**

The analyses carried out in this article point to a correspondence between the timing typologies of language and rhythmic typologies of music. As we have seen, both English and Spanish show inconsistencies or mismatches between speech prosody, on the one hand, and verse and music rhythm, on the other. However, these inconsistencies work differently in a syllable-timed language like Spanish than in a stress-timed language like English. While in the first type of languages there is a natural counterpoint or dialogue between speech prosody and musical rhythm, in the second type this counterpoint is considered arrhythmic and, therefore, unacceptable. Spanish word setting would be impossible in English. The small mismatches in English could be regarded as nearly insignificant instances of compositional clumsiness that signal the folk song style. On the other hand, the continuous radical stress mismatches observed in Spanish folk song would make nonsense of an English text because English word stress is vital to the rhythmic arrangement of the language at all levels and, therefore, to meaning. In other words, there exists a difference in kind in relation to the dialogue between prosody and music for each of the two types of languages. In English, the level of agreement between the two rhythmic patterns is really high, while in Spanish the counterpoint between the two is actually used as an expressive device. What we have described is, therefore, not only a difference in the ranking of metrical and grouping constraints for English and Spanish, but a radical disparity in the degree of violation into which each constraint is allowed to incur.

## **References**

- Abercrombie, David (1967) *Elements of general phonetics*. Edinburgh: Edinburgh University Press.
- Allen, George D. (1975) "Speech rhythm: its relation to performance universals and articulatory timing". *Journal of Phonetics* 3:75-86.
- Allen, William S. (1973) *Accent and rhythm*. Cambridge: Cambridge University Press.
- Bertinetto, Pier Marco (1989) "Reflections on the dichotomy 'stress- vs. syllable-timing'". *Revue de Phonétique Appliquée* 91/93:99-130.
- Bolinger, Dwight (1965) *Forms of English: accent, morpheme, order*. Tokyo: Hokuou.
- Canellada, Josefa and John Kuhlmann Madsen (1987) *Pronunciación del español*. Madrid: Castalia.
- Couper-Kuhlen, Elizabeth (1993) *English speech rhythm: form and function in everyday verbal interaction*. Amsterdam: Benjamins.
- Dauer, R.M. (1983) "Stress-timing and syllable-timing reanalyzed". *Journal of Phonetics* 11:51-62.
- Duffell, Martin (1999) *Modern metrical theory and the 'verso de arte mayor'*. *Papers of the Medieval Hispanic Research Seminar* 10. London: Queen Mary and Westfield College.
- Fabb, Nigel (2002) *Language and literary structure*. Cambridge: Cambridge University Press.
- Fant, Lars (1984) *Estructura informativa en español. Estudio sintáctico e informativo*. *Studia Romanica Upsaliensia* 34. Stockholm: Almqvist & Wiksell International.
- Gili Gaya, Samuel (1950) *Elementos de fonética general*. Madrid: Gredos.
- Halle, Norris and Samuel L. Keyser (1971) *English stress: its form, its growth, and its role in verse*. New York: Harper & Row.
- Halliday, Michael (1967) *Intonation and grammar in British English*. The Hague: Mouton.
- Harris, James W. (1983) *Syllable structure and stress in Spanish*. Cambridge, Mass.: The Massachusetts Institute of Technology Press.
- Hayes, Bruce and Abigail Kaun (1996) "The role of phonological phrasing in sung and chanted verse". *The Linguistic Review* 13:243-303.
- Hayes, Bruce and Margaret MacEachern (1996) "Are there lines in folk poetry?". *UCLA Working Papers in Phonology* 1:125-142.
- Hayes, Bruce and Margaret MacEachern (1998) "Quatrain form in English folk verse". *Language* 74(3):473-507.
- Hyman, Larry M. (ed.) (1977) *Studies in stress and accent*. Los Angeles, Ca.: University of Southern California.
- Kennedy, Peter (1984[1975]) *Folk songs of Britain and Ireland*. London: Oak Publications.

- Ladefoged, Peter (1967) *Three areas of experimental phonetics: stress and respiratory activity, the nature of vowel quality, units in the perception and production of speech*. London: Oxford University Press.
- Lass, Roger (1994) *Old English*. Cambridge: Cambridge University Press.
- Laver, John (1994) *Principles of phonetics*. Cambridge: Cambridge University Press.
- Lieberman Mark and Alan Prince (1977) "On stress and linguistic rhythm". *Linguistic Inquiry* 8:249-336.
- Lipsky, John M. (1997) "Spanish word stress: the interaction of moras and minimality". Eds. Fernando Martínez-Gil and Alfonso Morales-Front. *Issues in the phonology and morphology of the major Iberian languages*. Washington DC: Georgetown University Press, 559-594.
- Manzano Alonso, Miguel (ed.) (2003) *Cancionero popular de Burgos: V canciones del ciclo anual y vital*. Burgos: Diputación Provincial de Burgos.
- Navarro Tomás, T. (1965[1918]) *Manual de pronunciación española*. Madrid: Gredos.
- Piera, Carlos (1980) *Spanish verse and the theory of meter*. PhD Dissertation. Los Angeles, Ca.: University of California Los Angeles.
- Pike, Kenneth L. (1945) *The intonation of American English*. Ann Arbor, Mich.: University of Michigan Press.
- Pointon, Graham E. (1980) "Is Spanish really syllable-timed?". *Journal of Phonetics* 8:293-304.
- Roca, Iggy (1988) "Theoretical implications of Spanish word stress". *Linguistic Inquiry* 19:393-423.
- Trubetzkoy, Nikolai S. (1969) *Principles of phonology*. Berkeley: University of California Press.
- van Leeuwen, Theo (1999) *Speech, music, sound*. London: Macmillan.