

FROM SPEECH TO SONG:

An interdisciplinary investigation of rhythm in English and Spanish

Rosalía Rodríguez-Vázquez

Doctor of Philosophy in Linguistics and English Language
The University of Edinburgh. 2007



Table of Contents

Declaration	I
Abstract	II
Dedication	III
Acknowledgements	IV
Introduction	1
1. Rhythm as the connecting element between language and music	5
1.1. What is rhythm?	5
1.2. Rhythmic hierarchies	8
1.3. The study of rhythm: prosody	11
1.4. Linguistic stress and its relation to rhythm	12
1.4.1. Introduction: stress and accent	13
1.4.2. Types of languages in relation to accent	13
1.4.3. Stress	15
1.4.3.1. Definitions	16
1.4.3.2. Phonetic realisation of stress	17
1.4.3.3. Stress placement rules: a preliminary account	18
1.4.3.4. The relationship between stress, weight, length and tenseness	20
1.5. Summary	25
2. Speech prosody	27
2.1. Prosodic typologies: the dichotomy between stress-timed and syllable-timed languages. English and Spanish as typological models	27
2.2. The prosody of Spanish	35
2.2.1. From Latin to Spanish	35
2.2.1.1. Stress and the evolution of the quantitative system	35
2.2.1.2. Vowel systems and syllable structure	43
2.2.2. Stress assignment in Present-day Spanish	46
2.2.2.1. Harris's analysis amended by Roca	46
2.2.2.2. Navarro Tomás's analysis amended by D'Introno <i>et al.</i>	50
2.2.3. Syllable-compression phenomena in Spanish: synaeresis and synalepha	55
2.3. The prosody of English	66

2.3.1. From Old English to Present-day English.....	66
2.3.2. Stress assignment in Present-day English.....	73
2.3.3. Stress and related features.....	83
2.3.3.1. Vowel length and reduction.....	83
2.3.3.2. Vowel quality and tenseness.....	89
2.3.4. Syllable structure and syllable weight.....	91
2.4. Summary.....	92
3. Verse prosody.....	98
3.1. Introduction: poetic rhythm and metre: their relation to speech prosody. Analysis of verse prosody in English and Spanish.....	98
3.2. Verse prosody in Spanish.....	111
3.2.1. Introduction: from Latin quantitative verse to Spanish syllabic verse...	112
3.2.2. The principle of isosyllabism in Spanish verse.....	115
3.2.3. Syllable counting and its related devices.....	119
3.2.4. The displacement of stress in Spanish poetry.....	126
3.3. Verse prosody in English.....	131
3.3.1. From Old English alliterative verse to Present-day English accentual- syllabic verse.....	131
3.3.2. Stresses and syllables in contemporary English verse.....	139
3.3.3. Synalepha in English: a comparison with Spanish.....	141
3.4. Summary.....	149
4. The relationship between speech, verse and music.....	153
4.1. The interaction between verse prosody and musical rhythm.....	154
4.2. The text-setting problem in the existing literature.....	157
4.2.1. Janda and Morgan (1988) “El acentó dislocadó – pues cantadó – castellanó” and Morgan and Janda (1989) “Musically conditioned stress shift in Spanish revisited”.....	160
4.2.2. Hayes and Kaun (1996) “The role of phonological phrasing in sung and chanted verse”.....	162
4.2.3. Hayes and MacEachern (1996) “Are there lines in folk poetry?”.....	172
4.2.4. Hayes and MacEachern (1998) “Quatrain form in English folk verse”.	178
4.2.5. Hayes (in press) “Text-setting as constraint conflict”.....	186

4.2.6. Dell and Halle (in press) “Comparing musical text-setting in French and English songs”	189
4.2.7. Halle (forthcoming) “Text, tune and metrical form”	197
4.2.8. Kiparsky (2006) “A modular metrics for folk verse”	202
4.2.9. Summary	207
4.3. Object of study: Folk song	209
4.3.1. Definition and characteristics of folk song	210
4.3.2. Transcription and classification criteria	215
4.3.3. The importance of lyrics in the study of folk song	217
4.3.4. Text-tune relationships in folk song	218
4.3.5. The choice of a specific folk song corpus	220
4.3.5.1. Manzano Alonso’s <i>Cancionero de Burgos</i>	220
4.3.5.2. Kennedy’s <i>Folk songs of Britain and Ireland</i>	221
4.3.6. Analysis of the folk song corpus: overview and general observations ..	221
4.3.7. Case studies	226
4.3.7.1. Case study I: “The farmer’s boy”	226
4.3.7.2. Case study II: “Alegría, caballeros”	235
4.3.7.3. Conclusions	243
4.4. Stretching out the theory: Art song	244
4.4.1. Definition and characteristics of art song	244
4.4.2. Art song in Britain	245
4.4.3. Art song in Spain	246
4.4.4. Case studies	247
4.4.4.1. Case study III: “Loveliest of trees”	248
4.4.4.2. Case study IV: “Olas gigantes”	259
4.4.4.3. Conclusions	269
5. Summary, conclusions and further implications	271
References	279
Appendices	292
Chapters 1-4:	292
Appendix I: Glossary of terms and symbols	292
a) Terms:	292

b) Symbols:	297
Appendix II: Instances of synalepha and synaeresis in a corpus of 444 Spanish folk songs (Manzano Alonso 2003)	298
Chapter 4: The relationship between speech, verse and music	299
Appendix III: Index of OT text-setting terms and constraints with examples.	299
Appendix IV: Mismatches between stresses and beats in Manzano Alonso (2003)	305
Appendix V: Mismatches between stresses and beats in Kennedy (1984)	359

I declare that this thesis is my own work, and that it has not been submitted for any other degree or professional qualification.

Abstract

The general theoretical frame of this dissertation has to do with the study, from an interdisciplinary and interlinguistic point of view, of the typological dichotomy between stress-timed and syllable-timed languages, inasmuch as this distinction is valid at all. As a preliminary step, I carry out a comparative examination of the basic prosodic characteristics of English and Spanish, in order to then analyse the standard versification systems of these two languages. In the central part of my dissertation, I explore the most important text-setting Optimality Theory constraints as applied to a corpus of English and Spanish folk and art songs.

My main objective in carrying out these three-level analyses is to check whether the actual setting of verse to music responds to some kind of underlying rhythmic constraints common to language prosody, verse prosody and music, and whether those constraints are ranked differently from language to language.

The conclusions have to do with a correspondence between the timing typologies of language and the rhythmic typologies of music. I find clear inconsistencies or mismatches between speech prosody, on the one hand, and verse and music rhythm, on the other. 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 I find a natural counterpoint or dialogue between speech prosody and musical rhythm, in the second type this counterpoint tends to be considered arhythmic. In other words, I establish 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 stress-patterns is really high, while in Spanish the counterpoint between the two patterns is used as an expressive device.

Ós meus pais, Filomena e Virxilio.
Ó meu irmán, Virxilio.
Sempre.

Acknowledgements

First and foremost, this dissertation is indebted to the work, support and encouragement of my two supervisors, Professor Heinz Giegerich and Professor Raymond Monelle, whose contribution to the final result goes far beyond the academic. They believed in this project from the very beginning, asserting the value of such an interdisciplinary enterprise. They made enormous efforts to always facilitate my work by offering me both human and material resources whenever I needed them. They devoted long hours to the thinking and discussion of each and every page contained in the final manuscript. They came up with ideas and thoughts which enriched the writing-up process substantially. Last, but not least, they supported me in every possible way.

I am very grateful to the numerous people who have contributed to this work by generously sharing their invaluable knowledge, experience and ideas with me: Dr. Javier Pérez-Guerra, Dr. François Dell, Dr. Carlos Piera Gil, Prof. Nigel Fabb, Dr. Ricardo Bermúdez-Otero, Prof. Donna-Jo Napoli, Dr. Patrick Honeybone, Dr. Katie Overy, Prof. April McMahon, Dr. Chris McCully, Dr. John Halle, Dr. Marta Falces Sierra, Dr. Jean-Louis Aroui, Prof. José Domínguez Caparrós, Prof. Peter Nelson, Rafa Martín Castilla.

The accomplishment of this dissertation would not have been possible without the financial support of the *Fundación Caixa Galicia*, the *Arts and Humanities Research Council*, the Xunta de Galicia and the School of Philosophy, Psychology and Language Sciences (University of Edinburgh).

Special thanks to all the staff and fellow postgraduate students at the department of Linguistics and English Language and the faculty of Music (University of Edinburgh), and the department of English, French and German (University of Vigo, Galicia, Spain). Also, to my English Language students at both institutions, who turned my working weeks into truly enjoyable learning adventures.

A very big thanks to the Edinburgh University Music Society, the Edinburgh University Chamber Orchestra, the *Noteworthy Trio*, the *Rosalba Duo*, the Orquesta Clásica de Vigo and my group, *Nao d'Ire*, for giving me the chance to cultivate one of my biggest passions.

To my friends Marta Pérez Quiroga and Helena Miguélez Carballeira, with whom I have shared long hours of fruitful conversation, lots of laughs and a few tears for some years now. ¡E que sexan moitos máis! To Nerina Sunara, for always saying the right words at the right time.

Colin, your presence in my life has shown me that beauty exists even beyond good poetry and symphonic music. I will never find enough words to thank you for this happiness. It's been a long time waiting...

To my brother, Virxilio Rodríguez Vázquez, my best life-long companion. Thanks for your true friendship, dialectic capacity, optimism, generosity, eagerness to learn, sensitivity, captivating joy and, most of all, your unstoppable idealism.

Last and forever first, to my parents, Filomena Vázquez Paz and Virxilio Rodríguez Silva, for giving me the language of my dreams, for your always on-time words, your welcoming smiles, your openness, your silent and constant longing to have us close, your support, your immense love, your inspiration, your sense of freedom. Thanks for this boundless love for a small corner on the Iberian Finisterre.

Vós sodes a miña aldea, a miña identidade, o meu tesouro.
Sodes todo o que me dá a alegría.
Quérovos.

“Agora teño saudade
do futuro que hei de andar
lonxe da miña cidade.

Cúpulas brancas ó ar:
E sinto na eternidade
o corazón a soñar”

(A. Avilés de Taramancos)

Introduction

What art is, you know as well as I do: it is nothing more than rhythm. And if that's true, I don't have to burden myself with imitation or with soul, but can modestly and simply give you rhythm, in any material whatsoever: bus tickets, oil paints, building blocks [...], or words in poetry, or sounds in music, or you just name it (Schwitters 1993[1926]:229).

In 1926, the avant-garde artist Kurt Schwitters used these words to refer to the essentially rhythmic nature of any artistic work that aimed at being described as such. Strange as it might seem, Schwitters's categorical assertion about the ontology of the artistic object could be transferred to the linguistic field and thus become a partial definition for language.

In this piece of work, I assume that the status of natural language is the same as that of artistic language and, more specifically, music, in two respects, namely that both are essentially semiotic systems and both are produced and perceived in real time.¹ While a painting or a sculpture can be visually apprehended at once and then analysed thoroughly, music, like speech, can only be experienced over a period of time, the length of which is chosen by the speaker.² This fact imposes both on the speaker and on the listener the necessity to devise, consciously or unconsciously, some kind of systematic organisation of the sound stream. Rhythm is one of the multiple means – perhaps the most obvious one – to which people resort in order to impose a regular structure on any flow of sounds occurring in time. As Liberman (1979:313) observed, “[a]ll temporally ordered behaviour is metrically organised”, an assertion which points to a deep relation between what Schwitters referred to as ‘art’, that is, poetry and music, on the one hand, and everyday speech, on the other. Such a relation would be mediated by rhythm:

The fact that speech, verse, and music all have hierarchically organised metrical structure implies [...] a common cognitive origin. Not only are

¹ I am using the term ‘language’ in a broad sense, referring to any human manifestation that articulates thoughts by means of an arbitrary group of signs, whose main goal is to serve as a vehicle for communication (see Halliday 1973). For further discussion about the temporal nature of music, see Cureton's (1993:68-72) classification of arts into ‘spatial arts’ and ‘temporal arts’.

² For clarity's sake, I use ‘speaker’ to refer not only to the role of the individual that utters words or sentences, but also to the role of the performer of a musical composition.

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 (Couper-Kuhlen 1993:112).

Several branches of linguistic theory have dealt with the role of rhythm in speech and language, highlighting the similarities between music and speech (Lieberman 1979). In fact, Metrical Phonology draws on formal tools developed for rhythm in music. From a more literary perspective, there have been studies done on poetic rhythm (Attridge 1982) and on the technical analysis of text-tune relations (Cureton 1992).

The initial theoretical assumption for the purposes of this dissertation is that when we talk about rhythm, three levels come into play: (i) speech rhythm, also known as 'speech prosody', which refers to the rhythm of ordinary speech, (ii) verse prosody, which is the formal system used as a tool for poetic scansion and (iii) musical rhythm and metre. The main focus of this work is the relation between these three levels. The core research question that this work addresses has to do with the rhythmic typologies of language and how they are manifested in verse and vocal music. What are the differences in the nature of stress between an allegedly syllable-timed language, such as Spanish, and an allegedly stress-timed language, such as English? This question entails a series of others, which need to be answered before reaching a final conclusion: What are the prosodic characteristics of Spanish and English? What are the rules of stress assignment in each of these languages? What is the status of the syllable in each of them? How are those prosodic characteristics adapted to verse in Spanish and English? What role do stress and the syllable play in the versification systems of the two languages? How is verse set into music in Spanish and English? What are the generalisations to be made about the matching of speech rhythm (i) with poetic metre, and (ii) with musical rhythm/structure?

In order to find appropriate answers for all the above questions, I have divided my dissertation into four main parts. In chapter one, I carry out a preliminary review of the current theories on language rhythmic typologies. In chapter two, I focus on the main rhythmic element in the so-called 'stress languages', namely stress, and I analyse the evolution and present status of stress assignment rules in Spanish and English. In the third chapter, I explore some of the methods developed

for the scanning of verse in stress-timed and syllable-timed languages, and explore a theory that relates linguistic prosody and poetic prosody in Spanish and English. In the fourth chapter, I analyse the process of text-tune alignment in English and Spanish in order to establish the differences in the musical setting of stress-timed verse and syllable-timed verse. The methodology applied to the analysis of text-setting stems from a series of articles published in recent years, mainly – although not exclusively – within the Optimality Theory (henceforth OT) framework. I review each of those papers – Janda and Morgan (1988), Morgan and Janda (1989), Hayes and Kaun (1996), Hayes and MacEachern (1996), Hayes and MacEachern (1998), Kiparsky (2006), Hayes (in press), Dell and Halle (in press), Halle (forthcoming) – in order to be able to establish an analytical apparatus which extracts the most relevant elements from them. In this respect, I must point out that the method applied to my analysis does not strictly belong in the OT school.

I carry out a series of analyses of specific case studies, which have been carefully chosen among a vast corpus of songs. This corpus comprises two subgroups, corresponding to the subgenres of folk and art song. These subgenres show a number of specific characteristics which are carefully explained before introducing the case studies. The first subgroup is constituted by 444 songs extracted from Manzano Alonso's (2003) compilation of Castilian folk songs *Cancionero de Burgos*, on the one hand, and 239 songs extracted from Kennedy's (1984) collection of English songs in *Folksongs of Britain and Ireland*, on the other.³ There are several reasons why this corpus of songs constitutes the analytical focus of this dissertation. First, folk songs have been passed from generation to generation by people – most often non-musicians – who inherited them as popular 'composers' had created them; 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 collections with which I am dealing have been only minimally edited by the collectors themselves, something which ensures the 'native' quality of the songs. On

³ Chapter 4 contains an in-depth analysis of two folk songs (4.3.7) and two art songs (4.4.4.). I have also analysed the whole corpus of 444 Spanish folk songs and 239 English folk songs in terms of syllable/stress/beat mismatches (see 4.3.6.). Details of such an extensive analysis are shown in appendices IV and V.

the other hand, the genre to which they belong – song – is obviously intended to be sung, which means that any instrumental parts are composed bearing in mind vocal parts, and vice versa. We can therefore presume that the prosody of the Spanish texts will relate to the rhythm of the music according to traditions of Spanish word-setting, and that the same will happen with the English texts. The second subgroup within the corpus is constituted by a sample of two art songs extracted from two collections, namely Butterworth's (1911) *Six songs from 'A Shropshire lad'*, and Mompou's (1971) *Becquerianas*. The analysis of art songs is intended as a test to find out whether the 'native' rhythmic constraints applied to folk songs also work for songs which have been written by trained composers, who intentionally step away from any device considered musically 'incorrect' or aesthetically 'unacceptable'.

My main objective in carrying out these three-level analyses is to check whether the musical setting of verse – either written before or after the instrumental score – responds to some kind of underlying rhythmic constraints common to language prosody, verse prosody and music, and whether those constraints are ranked differently from language to language.

My conclusions show a correspondence between the timing typologies of language and the rhythmic typologies of music. The analysis of the folk and art songs brings to light clear inconsistencies or mismatches between speech prosody, on the one hand, and verse and music rhythm, on the other. These mismatches work differently in a syllable-timed language like Spanish than in a stress-timed language like English. While in the first type of languages I find a natural counterpoint or dialogue between speech prosody and musical rhythm, in the second type this counterpoint will tend to be considered arhythmic. In other words, I establish 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 stress patterns is really high, while in Spanish the counterpoint between the two patterns is used as an expressive device. This confirms the hypothesis that there is a difference in what we could call 'stressability' of syllables in English and Spanish. The originality of this dissertation is that it shows that such a difference is manifested not only at the speech level, but at all three domains – speech, verse and, as discovered in this study, music.

1. Rhythm as the connecting element between language and music

Given that both natural language and music are sound systems – that is, they are constituted by a small, very productive number of sound units arranged according to certain theoretical parameters –, it seems plausible to assert that there are certain aspects of music theory which bear a strong resemblance to some of the work done in phonology. As Jackendoff (1989:15) points out, “musical analysis might be brought to bear on some current problems in phonological theory”. At a more practical level, there is no doubt that some of the methodological tools used in music analysis parallel certain aspects of phonological structure. Actually, scholars such as Halliday (1967, 1994) and Liberman (1979) introduced their theories by presenting rhythm as the underlying building block in phonology and invoked chants and poetry as evidence for the alignment of rhythmic beats with speech. Moreover, Liberman’s use of music notation to represent metrical structure clearly emerged from an interfertilisation of work on music and on language.

Suprasegmental phonology, that is, the branch of phonology that deals with intonation and rhythm, is the field where the most fruitful parallelisms arise. The next sections concentrate on an explanation of the rhythmic phenomenon. In section 1.1. I review some of the works that deal with rhythm as a general cognitive capacity, in order to establish a comprehensive definition for it. In section 1.2. I expand on an important characteristic of rhythm, namely its hierarchical nature. In section 1.3. I concentrate on the discipline that deals with rhythm in language, that is, prosody. Section 1.4. is devoted to the analysis of stress and its role in the rhythm of the so-called ‘stress languages’. Section 1.5. presents a summary of the chapter.

1.1. What is rhythm?

As mentioned at the beginning of this chapter, rhythmic perception and production is one of our basic mental capacities and in the evolution of the human species it almost certainly preceded our ability to use language (see Blacking 1973; Downing 1995). It is widely assumed that the comprehensive and stabilising factor of our temporal consciousness is rhythm. Despite this acknowledged fact, there is still much confusion with regard to the definition of rhythm itself.

One of the first contemporary treatises on rhythm defined it as “that property of a *sequence of events in time* which produces *on the mind of the observer* the *impression of proportion* between the duration of the several events or groups of events of which the sequence is composed” (Sonnenschein 1925:16, my emphasis).

Underlying this definition, there is a series of principles commonly associated with the nature of rhythm. First of all, rhythm is a mental construction.⁴ The ‘sequence of events’ to which Sonnenschein refers can only be perceived through the observer’s senses; they are the objects of the observer’s experience and, as such, they are subject to time and space. Time and space are previous intuitions – ‘*a priori* forms’ in Kantian terminology (see Kant 1884) – in relation to perception; in other words, they are the conditions that make experience possible. In this sense, time is even more transcendental than space because, while the latter is the organising form of outer experience, the former organises both inner and outer experience. We could say that perception does not constitute experience unless it is real for one’s own consciousness. An event in the physical realm must be ordered with respect to the ‘here and there’ and also to the ‘before and after’.

Secondly, given that rhythm is dependent on inner time, it is a serial phenomenon. The events that constitute the basis for rhythm to be constructed take place in sequence, one after another.

Thirdly, the conception of rhythm is based on quantitative measurements, either conscious or unconscious, delimited by the action of stress, which produces a sense of proportion between the duration of different segments.⁵ As Barry (1990:262) observes,

The organisation of time is always the search for points of reference and articulation, and durational spans are created by information shaped between those points. It is through points of articulation that time can be structured into durational spans, and through defined characteristics design can be understood and remembered.

⁴ Although rhythm can effectively be conceived as a mental construction (see Pike 1945; Abercrombie 1967), it is important to note that it has physiological correlates (see Luce 1971; Glass and Mackey 1988).

⁵ Stress is the basis of rhythm: “Quantity alone cannot have been decisive for rhythm. Therefore, with it we have – not only in a regulating but in a shaping capacity – stress” (Weil 1960:675, in Allen 1973:100).

Cooper and Meyer (1963:16) define rhythm as “the way in which one or more unaccented beats are grouped in relation to an accented one”.⁶ According to this, an event is a group composed by at least one main beat, and one or more weak beats arranged around it. Malof (1970:2) complements these definitions by pointing to the fact that the observer’s sense of rhythm has to do not only with the beat, but also with other linguistic features such as pitch, loudness and duration, which are markers of prominence.

Rhythm is based on iteration, which is, in turn, manifested in repetition and periodicity. These two are the most important perceptual hints used by listeners to determine whether a sound stream is rhythmical or not. In Attridge’s (1982:77) words,

To perceive a regular rhythm is to comprehend a sequence of events as a pattern in time, with two mutually reinforcing features, repetition and periodicity. That is to say, a series of stimuli is understood as the same stimulus occurring again and again, and these repetitions are felt to be occurring at equal, or at least equivalent, temporal intervals. It is not an objective, measurable phenomenon, but a perceptual one.

A certain amount of repetition is an essential condition for rhythm to be perceived as such by the listener. In music theory, it is said that repetition, both small- and large-scale, can be considered the prime form-building agent and the most fundamental construction principle of musical form (Barry 1990:71). Similarly, in literary theory, readers are reminded that “in all levels of language, the essence of the poetic artifice has to do with periodic repetitions” (Jakobson 1966:23). Repetition makes it possible for listeners to order objects into a specific pattern, so that their recurrences create a coherent design. This has processing consequences: it reduces complex information to intelligible shape and coherent structuring; the more coherent the structure of the perceived object, the more easily the information will tend to be learned and remembered. Periodicity has to do with the fact that repetitions are perceived to occur at equal intervals in time. It contributes to the

⁶ In music, as in speech, there are two kinds of accents, namely ‘metrical accent’ – analogous to lexical stress –, which refers to the strongest beat in a bar, and ‘phrasal accent’ – analogous to sentence stress –, which refers to the most prominent beat of a musical phrase. These two concepts are directly related to the notions of ‘metrical hierarchy’ and ‘grouping hierarchy’ that I develop in the next section.

establishment of the so-called ‘invariants’ inside the listener’s mind (see Dowling and Harwood 1986): once structural constants have been established, any surface change in local pattern features would be processed as such, never as a change in the underlying structure. Periodicity complements repetition and reinforces the rhythmic impression. Once repetition and periodicity have been fixed in the mind of the listener, rhythm can be projected into the future, so that each event continually satisfies an expectation at the same time as that event itself generates a new one.

1.2. Rhythmic hierarchies

If we accept the statement that all human actions, artistic or non-artistic, have rhythm, we are implicitly asserting that the time in which any action takes place is measured in one way or another, that is to say, it has a perceptually regular pulse or metre. Conversational speech is clearly less regular than poetry and than most musical forms; nevertheless, experiments have shown that, in conversation, “speakers even attune to each other’s rhythm, so that the same pulse underlies the speech [...] of both” (van Leeuwen 1999:43).⁷ This is the basic idea underlying Liberman’s original thesis work, based on the so-called ‘Metrical Organisation Hypothesis’. This hypothesis, mentioned in the introduction, states that “all temporally ordered behaviour is metrically organised” (Liberman 1979:313).

According to this broad conception of metre, we can say that rhythm consists of two separate hierarchies, namely the metre hierarchy, which has to do with the relative strength of beats, and the grouping hierarchy, which has to do with the organisation of beats into phrases.⁸

In order to decode a metrical sequence, listeners try to find a more or less regular pulse or beat pattern, that is, a beat rate in a comfortable range.⁹ Metre imposes a salience structure on beats, and that structure is instinctively inferred by the listener, who terms ‘strong’ those beats which are salient and ‘weak’ those which are not salient. Studies in rhythm perception point to the fact that, in order to be

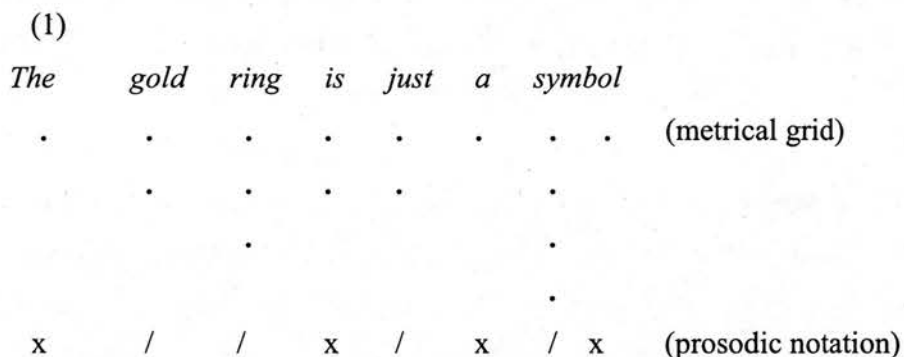
⁷ This conception is not new. Saint Augustine had already observed that “the role that metre plays in all movements is seen more readily in speech” (Epistola CI, in Dogil 1979:45).

⁸ One of the metatheoretical problems of dealing with prosody is that the technical terminology is used in many different ways. ‘Metre’ seems to be one of the most controversial terms. I am using it here in a broad sense, referring to any kind of regular pattern, and not just to the study of poetic patterns.

⁹ At the metrical level, the strong impulses in the rhythmic sequence are called ‘beats’, and the weak ones are called ‘offbeats’ (Attridge 1982:77).

perceived as such, beat intervals cannot be much longer than 1.5 seconds or tempos much slower than 40 beats per minute, where ‘tempo’ refers to the rate at which beats occur. Another perceptual restriction is that beat intervals are always divided, preferably into patterns of equal subdivisions in 2:1 ratios. More complex subdivisions are difficult to encode by listeners (Dowling and Harwood 1986:196).

The conception of a periodic alternation between strong and weak beats, that is, the metrical hierarchy, can be represented graphically. Traditionally, it has been represented either in the so-called ‘prosodic notation’ or in ‘metrical notation’ (Lieberman 1979; Selkirk 1984). The latter derives from Liberman’s hypotheses, which suggested the existence of two abstract structures, one of them imposed on any form of complex behaviour and the other one imposed on the time line. Those structures were called ‘metrical pattern’ and ‘metrical grid’, respectively, and the relation between the two was defined as their ‘alignment’.



In prosodic notation, there is no hierarchy among the different accented syllables. In the grid notation, the hierarchy is clear: the more dots, the stronger the beat. At each hierarchical level, the distance between beats must be either two or three times that of the immediately lower level. Given that the grid can actually represent a hierarchy of prominences, both between strong and weak beats, and also among strong and weak beats themselves, this latter type of metric representation is more transparent than the former. As we can see, Metrical Phonology – initiated by Liberman and Prince (1977) and Liberman (1979) – is the most relevant model for rhythm and intonation in language, and the fact that it uses some of the tools employed in music puts these two in connection.

In music, “the metrical organisation of a composition is designated in the time signature and measured by the bar lines” (Cooper and Meyer 1963:4-5). The concept of ‘bar-line’ connects metre with the second hierarchy, that is, grouping. Departing from the comprehension of metre, listeners organise sounds into groups. The sounds of natural speech are divided into syllables, words, phonological phrases and so on (Fabb 1997:25). Poetic sounds are organised into feet, phonological words, phonological phrases, utterances and so on. In music, these groups are called bars, motives, themes, phrases, periods, theme-groups, sections and the piece (Lerdahl and Jackendoff 1983:12).

The hierarchy of groups emerges from the listener’s rhythmic response to points of structural culmination within delimited structural spans; each group contains a grouping peak around which the other elements are organised. The delimitation of groups has to do with *Gestalt* principles such as distance and similarity. In musical terms, the listener tends to group events that are coherent in terms of register, texture, dynamics, timbre and tempo; also, events that are delimited by clearly-defined pauses; symmetrical groups are preferred to non-symmetrical ones. Apart from all the perceptual variables that determine grouping, the latter is also subject to the listener’s own imagination and creativity: “the grouping of sounds is not dictated by their real relations, but is completed by us, in an act that is subject to the will” (Scruton 1997:29). This is related to the fact that grouping is one of the most important variables which the performer can manipulate in projecting a particular conception of a piece (Lerdahl and Jackendoff 1983:63).

Summarising, rhythmic organisation consists of two hierarchies, namely metre and grouping, which interact so that the listener perceives any sound stream as being structured. Metrical structure consists of beats arranged hierarchically while grouping consists of more or less extended units arranged hierarchically. Once metre has been established and groups roughly delimited, listeners expect the given parameters to remain constant until the end of the poem or composition. The kind of rhythmic response to anticipated points of structural resolution gives rise to a hierarchy of prolongational regions defined by those points of arrival, as well as by points of departure (Cureton 1992:124).

1.3. The study of rhythm: prosody

The term 'prosody' comes from Greek *prosôidia*, which means 'parallel to song'. The Latin word *accentus* is a morph-to-morph translation of the former – *pros=ad*, and *ôidê=cantus*. The earliest recorded use of *prosôidia* appears in Plato's (1998:99) *Republic*, in a discussion of the social and emotional correlates of the different musical modes and the different kinds of speech they imitate:

- Well at any rate you can agree easily enough that song consists of three elements, words, mode, and rhythm.
- Yes, I agree to that.
- As for as the words are concerned, then, the same principles will apply as those we have just laid down for words not set to music, both for their content and form.
- True.
- And surely the mode and rhythm should suit the words.
- Certainly.

The previous text asserts the long-acknowledged relationship between words, on the one hand, and melody and rhythm – the so-called 'suprasegmental aspects of speech' – on the other. In its origins, prosody was a part of grammar which dealt with the study of accent. Accent, in turn, referred to the phonetic properties of syllables and words in relation to the measure of rhythm, especially in verse and, more generally, those of the different metres and forms of verse.

Currently, the term 'prosody' is used to refer to two different, though closely related, fields. When it refers to the study of the elements and structures related to the rhythmic and intonational dimensions of speech and language in general, it is known as 'linguistic prosody' or 'suprasegmental phonology'. Suprasegmental phonology is broadly concerned with the analysis of the phonological features that occur within a time span greater than that of a single significant sound or phoneme. As opposed to segmental features, which have to do with the sounds of individual vowels and consonants, suprasegmental or prosodic features – intensity, duration and pitch – may occur on the level of the syllable, the word, the phrase or even the complete sentence, and may thus discriminate between different meanings for the same stretch of speech. One of the most important theories within suprasegmental phonology is

Metrical Phonology (Lieberman and Prince 1977; Lieberman 1979), which deals with the study of the alignment of speech sounds into a temporal grid, as well as with the conventions that allow a linguistic community to discriminate between prose and verse.

The second field to which prosody refers is literary prosody, which deals with the analysis of rhythm in verse, that is, with the laws of versification. This second meaning of prosody has always been the most commonly used among scholars, probably because it is the most directly connected with the original meaning of the word. Thus, it is not surprising that this term was originally applied by Halle and Keyser (1971) exclusively to the distribution of stress in poetry. According to them (1971:139-140), “an adequate theory of prosody [...] should be expected to [...] help us to understand the nature of metrical verse and illuminate the relationship between a speaker’s everyday linguistic competence and his ability to judge verses as metrical or unmetrical and as complex or simple”.

Nevertheless, as explained in the previous section, in the last twenty five years – mostly since the publication of Lieberman and Prince (1977), and Selkirk (1984) – linguists’ understanding of speech rhythm has deepened significantly, to the point that the statement made by Metrical Phonology that the distribution of stress in the words and phrases of speech is metrically conditioned has come to be widely accepted. As a result, the phenomenon of patterning in normal speech has been taken to form part of the general theory of prosody.

1.4. Linguistic stress and its relation to rhythm

As already pointed out, all languages articulate the flow of sound in a rhythmical way, where weaker segments are grouped around a salient one. What makes languages differ from one another is the mechanism used to assign prominence to certain units over the others. Section 1.4.1. introduces the topic and argues that all prominence devices can be included in the category ‘accent’. Section 1.4.2. analyses the kind of accent used by different groups of languages, which gives rise to the typological distinction between tone languages, pitch-accent languages and stress languages. Section 1.4.3. concentrates on stress languages, since both Spanish and English belong in this group.

1.4.1. Introduction: stress and accent

The issues addressed in this section have to do with the nature of the phenomenon which articulates an otherwise unstructured flow of sound in language. I am referring to linguistic stress. There are several questions which constitute the core of my dissertation and prompt a comprehensive investigation of this often marginal subject within the phonological field: What exactly is stress? Do all the languages in the world have stress? If not, which other articulating devices may languages have? How many types of stress are there? Are there any rules for the placement of stress? How is stress manifested in actual speech?

As several scholars (Allen 1973:86; Hyman 1977:37) have pointed out, most of the existing treatises on stress use this term as a synonym for ‘accent’, especially when referring to emphasis. Nevertheless, it is necessary to point out that accent is a more general concept which refers to the mechanisms that all languages display in order to signal the prominence of certain syllables or words within utterances: “[s]tress-accent [...] is a property found in many if not most languages of the world. What is more, it is often observed to play a key role in the phonological and/or morphological structure of languages” (Hyman 1977:37). Accent can in turn be subdivided into three main classes, one of which is stress, the other two being tone and pitch. Depending on which of the previous accentual mechanisms they show, languages can be classified into stress languages, tone languages and pitch-accent languages – for a full discussion of this topic, see Beckman (1986). Even though this dissertation focuses on stress languages and their characteristics, in what follows it seems important to explain the basic differences between stress languages, tone languages and pitch-accent languages. This I undertake in section 1.4.2.

1.4.2. Types of languages in relation to accent

As already pointed out, not all the languages of the world articulate the flow of sound by means of stress. There are another two devices present in many languages, namely tone and pitch-accent. Languages are classified according to the kind of accent which they use.

Most languages in Europe, like, for instance, English and Spanish, are stress languages. This means that they make use of stress – a phonetic peak or climax occurring typically in every word (Allen 1973:86; Hyman 1977:39); for an overview

of the current definitions of 'stress', go to section 1.4.3.1. – in order to mark syllable prominence and, therefore, highlight boundaries between words. Stress is, therefore, a tool used to parse a string of sound which would otherwise constitute a shapeless phonemic sequence. As a consequence, the basic principle in stress languages is that only one syllable per word can receive primary stress (Hyman 1977:38), a fact which can be summarised by stating that the most general property of stress is its culminative function (Allen 1973:87). This does not mean that languages cannot show degrees of stress. Actually, the well-known occurrence of secondary stresses in languages such as English does not affect the general principle that each accentual unit – normally the word – has one and only one main or primary stress. The occurrence of secondary stresses is in fact directly related to the occurrence of the main one; the former can be conceived of as 'echoes of the latter' (Allen 1973:90).

The majority of languages belong in the category of tone languages (Laver 1994:465). Tone languages have been defined as those languages which use pitch movements in order to signal prominence. The different levels or movements of pitch are commonly known as 'tones'. Given that tones operate in each syllable, it is obvious that they are not culminative. In stress languages, the fact that it is not permitted to have more than a primary stress per word reduces the possible stress patterns in disyllabic words to unstressed-stressed or stressed-unstressed. By contrast, given that tones can be combined freely and placed on any syllable, tone languages show a wider range of patterns for disyllabic words, from high-high, high-low, to low-high, low-low. Despite this richness of combinations, the placement of tones shows some restrictions, which can be either phonological or grammatical.

Attending to the kind of unit on which the tones operate, tone languages can be classified into word-based tone languages and syllable-based tone languages. In word-based tone languages a contour-pattern is associated with the entire word, over a variable number of syllables (Laver 1994:162). Within this group, there is a subgroup of languages, like Slovenian, the Lewis dialect of Scots Gaelic, Swedish and Norwegian, which exploit two contrasting patterns on many words of two or more syllables, of which the first syllable is stressed, while monosyllabic words

show no tonal contrast (Laver 1994:463).¹⁰ In Pike's (1948:3) words, syllable-based tone languages are those "having significant, contrastive, but relative pitch on each syllable". This pattern of a series of such tones on the syllables of the word serves to identify the word. Most of these languages are found in Central America, Africa and Southeast Asia – Chinese is, probably, the best-known example of this type of language. Depending on what feature of the tone behaviour is more salient, syllable-based tone languages can be subdivided into register tone languages, where the relative height of pitches – level, rise or fall, or combinations of them – used by the speaker is the relevant feature for word identification, and contour languages, where the relevant feature is the shape of the pitch curve, like in Vietnamese, Mandarin Chinese and Thai. There are also some languages that show a combination of both register and contour tone systems (Laver 1994:466).

There is a third group of languages that could be placed in between stress languages and tone languages: pitch-accent languages. In pitch-accent languages, a tone is assigned to the most prominent syllable of the word. According to this, pitch-accent languages are tonal in the sense that the assigned feature is a tone which can be combined and render a contour; on the other hand, they are stress languages in the sense that only one syllable per word can receive the tonal prominence, that is, in pitch-accent languages tone is culminative. Examples of pitch-accent languages can be found in Africa and Asia – Japanese is one of these languages.

1.4.3. Stress

Since this dissertation deals with two stress languages, it seems in order to explore in depth the notion of 'stress' and its characteristics. Section 1.4.3.1. introduces a series of definitions for stress. Section 1.4.3.2. explores the phonetic – that is, physical – correlates of stress. Section 1.4.3.3. constitutes an introduction to stress-placement rules. Finally, section 1.4.3.4. analyses the potential interactions between stress, on the one hand, and weight, length and tenseness, on the other.

¹⁰ The status of Scandinavian languages seems to be controversial. While Laver (1994) classifies them as tone languages, Roca and Johnson (1999) group them, together with Lithuanian and Serbo-Croatian, as pitch-accent languages, while Gärding (1977:7) clearly states that "the Scandinavian languages are not tone languages".

1.4.3.1. Definitions

When dealing with stress languages, one needs to take into consideration the various conceptions to which the term 'stress' has given rise in the linguistic field. Its definitions vary depending on the point of view from which stress is examined.

From a physiological perspective, stress has been defined as

- “the comparative force with which the separate syllables of a sound-group are pronounced” (Sweet 1878:203).
- “force of breath-impulse” (Abercrombie 1923:19).
- “greater amplitude of sound-waves, [...] produced by means of more energetic movement, such as pumping more breath, bringing the vocal chords [...] closer together for voicing, and using the muscles more vigorously for oral articulation” (Bloomfield 1933:110-111).
- “kinaesthetic sensations of muscle and pressure changes” (Heffner 1949:224-225).

From a psychological point of view, stress has been said to consist of

- “an impulse (primarily of a psychological nature) which expresses itself in the first place by an increase of pressure in the speech-canal and approximately coincides with the point of greatest pressure” (Classe 1939:37).
- “relative strength or loudness” (Trager & Smith 1951:36).
- “the degree of force with which a sound or syllable is uttered. It is essentially a subjective action” (Jones 1956:243).
- perceived prominence imposed within utterances (Bolinger 1958).
- a mental phenomenon picked up by the hearer (Lehiste 1970; Weinreich 1954).

All the above definitions highlight the fact that the realisation of stress involves an extra degree of energy, an impulse, both physically perceptible and psychologically apprehensible. The first group of those definitions makes constant reference to the phonetic cues that indicate stress, namely pitch, loudness and duration. In fact, although stress can be analysed as a phonological concept, it is necessary to refer to the phonetic cues which both speaker and listener use in order to recognise stress as such.

1.4.3.2. Phonetic realisation of stress

With regard to phonetic realisation, there are three main strategies used to signal stress, namely pitch change, greater duration and greater intensity (Bolinger 1958, in Hyman 1977:40). In what follows, I shall examine these three phonetic cues in an attempt to determine their real importance both in the realisation and in the perception of linguistic stress.

First, 'pitch' is normally defined as "that attribute of auditory sensation in terms of which sounds may be ordered on a scale from low to high. It depends primarily upon the frequency of the sound stimulus, but it also depends upon the sound pressure and waveform of the stimulus" (Crystal 1969:108). Second, 'loudness' is "that aspect of auditory sensation in terms of which sounds may be ordered on a scale running from 'soft' to 'loud'" (Crystal 1969:113). Last, 'duration' is "that aspect of auditory sensation in terms of which sounds may be ordered on a scale from 'short' to 'long'" (Crystal 1969:121). In many stress languages there is a clear correlation between these three concepts, in such a way that a stressed syllable is frequently characterised by a pitch change, by greater duration and by a greater intensity, in that order, as well as by an "increased articulatory precision" (Allen 1973:93) which has to do with the occurrence of richer vowel systems in the accented syllable.

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 length (Allen 1975:80). This correspondence is not as obvious in Spanish, where higher pitch and louder intensity do not render more length. Stressed syllables in English are on average 1.5 times longer than unstressed syllables; in Spanish, they are on an average 1.3 times longer. As Yavas and Goldstein (1998:49-60) point out, this is even more noticeable in non-final open syllables – the concepts 'open' and 'closed' syllables are explained in section 1.4.3.4. – where in Spanish stressed syllables are only 1.1 longer than unstressed syllables compared to 1.6 in English and 1.5 in German. French seems to be an in-between case, where successive syllables within a sense group are very similar in duration, pitch and loudness, with the salient exception of the last, which is, on average,

higher in pitch, slightly softer and much longer in duration (Delattre 1966, in Allen 1975:78).

In sum, not all stress languages manifest prominence using the same devices in the same way. In English, one syllable is more prominent than another to the extent that its constituent segments display higher pitch, greater loudness, longer duration or greater articulatory excursion from the neutral disposition of the vocal tract (Laver 1994:450), something which is not the case in Spanish, where higher pitch and loudness do not necessarily entail longer duration.

1.4.3.3. Stress placement rules: a preliminary account

As mentioned in the previous section, stress languages are diverse with regard to the phonetic realisation of stress. They also differ in the types of rules that govern stress assignment. While some languages do not have specific stress placement rules, others show a complex system of rules.

With regard to the presence or absence of rules of placement, there are two main types of stress: on the one hand, lexical stress, which does not respond to any rules, and whose position is thus unpredictable phonologically; on the other hand, grammatical stress, whose position is morphologically conditioned (Hyman 1977:39). Grammatical stress is in turn subdivided into morphological stress and demarcative stress. In the case of morphological stress, the position of stress is free but largely determined by the morphological structure of the word. Morphological stress serves a distinctive function, as its presence disambiguates the morphological status of a word – for instance, Spanish *termino* ('I finish') vs. *terminó* ('he finished'), English *import* (noun) vs. *import* (verb). Generally speaking, there are few pairs of words distinguished exclusively by their stress – most of which could not occur in comparable contexts (Allen 1973:88-89). Demarcative stress is fixed by rule and always occurs on the same syllable, indicating either that a word commences or ends with the syllable in question, or ends with the next syllable – this type of stress is found in Czech, Icelandic and Hungarian, falling on the initial syllable; in Armenian, on the final syllable; in Polish, on the penult (Allen 1973:87). Summarising, the primary difference between morphological and demarcative accentuation is that the former can, in certain cases, help to indicate the morphological structure of a word.

Independently of the presence or absence of rules, each stress language shows an idiosyncratic tendency to place stress on a specific syllable, a fact that is related to the phonological and morphological characteristics of that language.

The placement of stress in each language conditions the perception of native speakers. As Jakobson, Fant and Halle (1951:10-11, in Allen 1975:78) observed,

Knocks produced at even intervals, with every third louder, are perceived as groups of three separated by a pause. The pause is usually claimed by a Czech to fall before the louder knock, by a Frenchman to fall after the louder; while a Pole hears the pause one knock after the louder. The different perceptions correspond exactly to the position of the word stress in the languages involved: in Czech the stress is on the initial syllable, in French, on the final and in Polish, on the penult. When the knocks are produced with equal loudness but with a longer interval after every third, the Czech attributes greater loudness to the first knock, the Pole, to the second, and the Frenchman, to the third.

Despite this perceptual idiosyncrasy, it would be possible to state that there are certain universals with regard to linguistic stress placement: no language has general initial stress and all languages avoid both second syllable and also final syllable stress. As shown in Table 1 – adapted from Hyman (1977) – the most common placements are on the penultimate and on the initial syllable:

Initial	Medial	Final	No dominant placement
Finnish	a) On the 2 nd syllable: 12	French	English
Total number of languages: 114	b) On the penult: 77 Spanish and Portuguese	Total: 97	Total: 113

Table 1: Stress placement in different languages.

There have been some hypotheses that point to a close relationship between stress assignment and intonation. As Hyman (1977:44) states, these hypotheses conceive stress as a grammaticalisation of intonation at the level of the word, in such a way that pitch, duration and intensity are encoded as components of a word in the mind of the speaker-hearer:

Stress comes from intonation [...] Intonation becomes grammaticalised as word-stress when the suprasegmental features of pitch, duration, and intensity that would have characterised a word in isolation are encoded with the word, and thus come to function in words not in isolation.

In Hyman's (1977:43) words, "languages will tend to assign stress to those positions where a pitch-change will be most perceptible". In clausal intonation, a falling pitch is always more common than a rising one. This also happens in word stress. This seems to be the reason why penultimate stress is so common: it corresponds to a falling pitch. It is perfect, both phonetically and conceptually – there is an alternation of unstressed-stressed-unstressed syllables. According to this theory, the reason why stress tends not to be placed word-finally is that final stress is less natural for the realisation of a falling pitch – this fall has to be identified within an only syllable: "[p]lacing stress non-finally will enhance the two functions of stress: i) phonetically, the fall from H to L is enhanced and ii) conceptually, the culmination of prominence is enhanced by virtue of the fact that a syllable lacking stress follows" (Hyman 1977:46).

1.4.3.4. The relationship between stress, weight, length and tenseness

There are three variables which may condition the placement of stress in different languages, depending on whether they act as phonological parameters or not. Those variables are length, weight and tenseness. I shall focus on English and Spanish in order to show how those three variables interact in order to delimit the possible placements of stress.

It seems in order here to specify the characteristics of the domain on which all these variables have an effect, that is, the syllable. A syllable can be defined as the "basic unit in the prosody of all languages [which can be] subdivided into an onset and a rhyme. The rhyme is further subdivided into peak/nucleus and margin/coda" (Hyman 1985:1). The peak of a syllable is usually a vocalic element. What makes syllables complex in terms of weight is the presence of one or more consonants in the coda. Syllables vary depending on the combinations of vocalic and consonantal elements in their rhyme.

Table 2 shows the most frequent syllable structures in English and Spanish (adapted from Laver 1994:531 ['C' stands for consonant and 'V' for vowel]), while Table 3 indicates the percentages of open and closed syllables in both languages:

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 frequent syllable structures in English and Spanish.

Open	44%	70%
Closed	56%	30%

Table 3: Open and closed syllables in English and Spanish.

As evinced by Tables 2 and 3 above, English enjoys a greater variety in permissible syllable types. Syllables may contain a short vowel, a long vowel or a diphthong as their nucleus and may be closed by as many as four consonants. In Spanish, the syllable nucleus contains a single vowel or diphthong, and syllables may be closed by at most one consonant finally and two consonants medially. In the absence of a vowel contrast, in Spanish we simply have a contrast between open and closed syllables. Over half of the syllables in Spanish have a simple CV structure, that is, open syllables clearly predominate in Spanish (Dauer 1983:55).

Unlike Spanish, English is a weight-sensitive language, that is, it belongs to the group of languages which display a heavy versus light syllable distinction. In a weight-sensitive system, the segments within the rhyme that bear the weight of the syllable are called 'morae'. According to this, a heavy syllable is said to consist of two morae, and a light syllable of one (Hyman 1985:9-10). The phonetic difference between heavy and light syllables is that the former have "full articulations" and the latter, "reduced timing" (Allen 1975:83).

In general, weight-sensitive languages follow one of the following two patterns. A first group of languages treats a syllable whose rhyme consists solely of a short vowel as light, whereas a syllable whose rhyme has a long vowel and/or a final consonant – or more – is treated as heavy. Syllable weight is thus tied to the existence of vowel length or vowel tenseness (Hyman 1977:6); in other words, in order for a -VC rhyme to be counted as heavy, the language must also have -V: rhymes. This is the case in Latin, most dialects of Arabic, or English (Hyman 1985:5). The second group of languages treats a syllable whose rhyme has a short vowel as light and a syllable whose rhyme has a long vowel as heavy – but, independently of whether the syllable is closed by a consonant or not, this group treats a closed syllable with a short vowel as light. This is the case of some South-American and Asian languages (Hyman 1985:5-6).

In English and other stress languages with the dichotomy heavy-light, stress is placed on heavy syllables, that is, syllables formed by a long vowel or by a vowel and at least one consonant (Dauer 1983:55). As previously mentioned, stress systems that require accenting of heavy syllables, as English does, are said to be quantity-sensitive (Roca and Johnson 1999:357). The general case in quantity-sensitive languages is that a short vowel in open syllable position shows reluctance to accept stress, and may therefore pass it on to a neighbouring syllable when this is feasible (e.g. *A-mé-ri-ca*). The result is that syllables of the shape V or CV, which are said to be light, sometimes function differently from syllables of the shape CVC, CV:, VC, C:, which are said to be heavy (Hyman 1977:47). According to this, we could say that, in languages like English, the weight of a syllable determines whether it can carry the stress of the word or not.

There is a second variable which relates to weight and stress in quantity-sensitive languages, namely length. The relationship between stress, weight and length is quite complex. The most widely accepted hypothesis is that syllable weight owes its existence directly to the nature of stress itself and the consequent lengthening of vowels: “the tendency towards lengthening is the most obvious feature observed as a physiological correlate to stress” (Hyman 1977:48). In accordance with this hypothesis, the reason why stress may shy away from a light (CV) syllable in a language with a vowel length contrast, like English, is that the

vowel of that syllable would tend to lengthen and therefore might merge with the corresponding long vowel. In a language without a vowel length contrast, the stressed short vowel would be free to lengthen, and therefore no syllable weight distinction between CV and CVC would be observed (Hyman 1977:48). In sum, length owes its existence to the assignment of stress, because this entails a greater duration on the stressed syllable. As a consequence, stress will be rebuffed by a syllable which has articulatory or perceptual reasons to stay short (Allen 1964). A very complex example of the relationship between syllable weight and stress placement is Latin: if a light syllable is accented, the correspondent word ends with the next syllable but one – for instance *fácilis* ('easy'); if a heavy syllable is accented, the word ends with the next syllable if that is also heavy – *compóno* ('I compose') – or with either the next or the following one if the next syllable is light – *compóne*, *compónere* ('she composes', 'to compose').

What is clear is that not only the structure but the segmental composition of syllables reinforces the difference between stressed and unstressed syllables in English. Spanish syllables are much more equal than English syllables with respect to their length and their retaining vowel quality (Pointon 1980:300-301). In Spanish, quality tends to obscure rather than accentuate the difference between stressed and unstressed syllables. This might be the reason why the perceived durational contrast between stressed and unstressed vowels is likely to be less striking in Spanish than in English (Dauer 1983:57). Assuming that segments have minimum and maximum durations, then based on structure alone syllable length is bound to vary more in English than in Spanish.

The third variable which relates to stress is tenseness. According to Jakobson and Halle (1964:96), a tense vowel compared to its lax counterpart is "produced with a greater deviation from the neutral position of the vocal tract [...]; consequently a tense vowel displays a greater deviation from the neutral formant pattern".¹¹ Tenseness is directly related to length and weight. The heightened subglottal air

¹¹ The dichotomy tense-lax has a specific physical basis. While Bell (1867), who first drew attention to the tense-lax distinction, ascribed the decisive role to differences in the behaviour of the pharynx, Sweet (1878) put the chief emphasis on the 'shape of the tongue.' Later investigations, however, as summed up in Heffner's *General Phonetics* (1949), shifted the reference "from tongue elevations and tongue muscle tensions to laryngeal positions and air pressures" (Jakobson and Halle 1964:97).

pressure in the production of tense vowels is indissolubly paired with longer duration. As has been repeatedly stated by different observers, tense vowels are necessarily lengthened in comparison with the corresponding lax phonemes. Tense vowels have the duration needed for the production of the most clear-cut optimal vowels and, in comparison with them, lax vowels appear as quantitatively and qualitatively reduced, obscured and deflected from their tense counterpart toward the neutral formant pattern (Jakobson and Halle 1964:97).

In stress languages such as Spanish, the dichotomies between tense vs. lax vowels and long vs. short vowels do not exist, which means that vowel reduction does not operate in unstressed position or, if it does, it operates differently than in English. A sequence of adjacent vowels in Spanish can indeed be reduced and pronounced as a single syllable – a phenomenon called ‘synaeresis’ – but this reduction is conditioned by phonological environment rather than by stress. Actually, it seems to be the case that in Spanish greater speed and ‘ease of articulation’ – the reasons given for vowel reduction in English – are achieved at the expense of the consonants rather than the vowels, and this does not affect subjective syllable length. In English not all unstressed vowels are reduced, but in ordinary speech a very high proportion of unstressed syllables have schwa (/ə/), or syllabic /r/, /l/, or /n/ as their nucleus. In addition, these syllables often occur in function words and morphological endings so that they do not seem to carry as much semantic information and thus seem subjectively shorter than full vowels in stressed syllables (Dauer 1983:57-58).

All this has important consequences for the rhythm of the languages under investigation. Although English and Spanish are considered stress languages, in the former the structure of the syllable, the word, and the utterance, is determined by stress to a much greater extent than in the latter, where those units retain their characteristics whether stressed or not. That may be why languages like English are called ‘stress-timed’, while languages like Spanish are referred to as ‘syllable-timed’. As already mentioned, stress-timed languages have lexical or word stress – usually lexical free stress –, realised by a complex set of changes in length, pitch contour, loudness and quality, which clearly make stressed syllables more prominent than unstressed syllables, that is, there is a clearly discernible ‘beat’. Languages which have been called ‘syllable-timed’ either have no lexical stress – the case of French –

or stress, if present, is more likely to be realised by variations in the pitch contour and in the loudness scale –, as happens in Spanish. In sum, it could be said that the greater the effects of stress on the linguistic system in both its phonetic realisation and in the number of stressed-related rules that are necessary in a grammar of the language, the more likely a language is to be stress-timed.

1.5. Summary

Chapter 1 focuses on the study of rhythm, a phenomenon which is articulated at multiple levels. Three of those levels, namely speech, verse and music, share some general structural characteristics. The hypothesis argued for in this dissertation is that the prosodic characteristics of a specific language both determine and are determined by the other two levels, namely verse and music. The questions raised have to do with the kind of mechanisms that rule this three-fold interaction.

Section 1.1. defines rhythm as the impression of proportion between the duration of the various segments or groups of segments in a sequence of events. The quantitative measurements involved in the process of apprehending rhythm are not necessarily conscious and work at the perceptual level. In the case of the languages analysed in this dissertation – English and Spanish –, the articulating device that delimits those segments or groups of segments is stress. The markers of stress are, in turn, pitch, loudness and duration, three features which work at the level of the phoneme and the syllable.

In section 1.2., the hierarchical nature of rhythm is explored. Rhythm consists of two hierarchies. The first one is the metre hierarchy, which has to do with the relative strength of beats, and the second one is the grouping hierarchy, which has to do with the organisation of beats into phrases. In phonological terms, the metre hierarchy has to do with syllables and stress levels, while the grouping hierarchy relates to the foot, the phonological word, and so.

Section 1.3. concentrates on prosody or suprasegmental phonology, the discipline that deals with rhythm in language. Suprasegmental phonology analyses the phonological features that occur within a time span greater than that of a single significant sound or phoneme. The suprasegmental or prosodic features by means of which stress is manifested – loudness, duration and pitch – may indeed occur at the

level of the syllable, the word, the phrase or even the complete sentence, and may thus discriminate between different meanings for the same stretch of speech.

Section 1.4. is devoted to the analysis of stress and its role in the rhythm of the so-called 'stress languages'. Regarding the type of accentual device that they use in order to articulate an otherwise arhythmic sequence of events, languages can be classified into stress languages, tone languages, and pitch-accent languages. Tone languages use pitch movements in order to signal prominence. Pitch-accent languages assign tone only to the most prominent syllable of the word; in this respect, they are a mixed type of languages. Both Spanish and English are stress languages, but stress does not play the same role in each of these languages. The placement of stress is conditioned by the presence or absence of three phonological parameters, namely length, weight and tenseness. The syllable structure of English, where weight, length and tenseness play an important role, determines the placement of stress, and vice versa. This is not the case with Spanish, where syllable structure and stress are, in principle, independent. All this has clear consequences for the rhythm of the two languages. While in English stress can function as a rhythmic device, in Spanish it is only a prominence device, that is, it does not alter the rhythmicity of the utterance. All this has an ultimate consequence on the classification of English as a stress-timed language and Spanish as a syllable-timed language – analysed in chapter 2. This controversial classification seems to work in the sense that the greater the effects of stress on a linguistic system – in both its phonetic realisation and in the number of stress-related rules that are necessary in a grammar of the language – the more likely a language is to be stress-timed, which is indeed the case of a language such as English.

2. Speech prosody

In the next sections I shall show that English and Spanish differ in terms of rhythm and, as a consequence, in the constraints on the stressability of their syllables: Section 2.1. constitutes a summary of the different studies, analyses and experiments carried out in relation to the study of prosodic typologies, paying especial attention to those theories devoted to the statuses of Spanish and English as syllable- and stress-timed languages, respectively. This classification will be questioned throughout this section. In sections 2.2. and 2.3. I shall analyse the speech prosody of Spanish and English, exploring their historical evolution with regard to prosody in order to then concentrate on a specific theory of stress assignment for each of the two languages. In my analysis, I will exclusively pay attention to the phonological behaviour of these languages, rather than to phonetic evidence. Section 2.4. presents the reader with a summary.

2.1. Prosodic typologies: the dichotomy between stress-timed and syllable-timed languages. English and Spanish as typological models

In 1938, Trubetzkoy described languages from a rhythmic point of view, although he never got to develop a complex theory. Some years later, in 1945, the anthropologist Pike proposed a series of criteria for the typological classification of languages. Pike's hypotheses were completely based on a temporal view of speech rhythm, according to which similar events recur at regular intervals in time, that is to say, they are isochronous. The nuclear idea in Pike's theory is that languages will be different depending on which elements constitute those isochronic events.

The two types of speech rhythm considered by Pike are, on the one hand, syllable-timed rhythm, characterised by the fact that syllables come at equal intervals in time and, on the other hand, stress-timed rhythm, where stresses occur at equal intervals. According to the above differentiation, there are two major types of languages, namely syllable-timed languages – where syllables tend to be perceived as being rhythmically equal – and stress-timed languages – where the stresses fall at perceptually equal intervals. In the former group the syllable is obviously considered the major unit of rhythm, while 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, up to (but not including) the next stressed syllable.

I shall now discuss the scientific basis of isochrony, the pillar on which Pike's theory is supported. Measurements have shown that strict isochrony does not really exist and that deviations from foot isochrony depend on the number of syllables in a foot – just like deviations from syllable isochrony are related to the number of segments in a syllable. There is only limited evidence for stress-timing, but none of any value for the physiology of stress-timing. Nevertheless, it has also been shown that, under certain conditions, such deviations tend to be psychologically disregarded in speech perception and, as a result, speech tends to be heard as more rhythmical than it really is (Barnwell 1971; Huggins 1972 a, b; Coleman 1974).¹² In relation to this issue, some scholars (Lehiste 1977:255) have pointed to the necessity to determine the limits of the speaker to produce isochrony, as well as on the limits of the listener to perceive it, which emphasises the fact that it has never been clear how much deviation from strict isochrony might be permissible, and perhaps required, for sentences to be perceived as regular (Cummins and Port 1998:146). In any case, I would like to underline that, from a phonological point of view, talking about isochrony does not presuppose the existence of acoustic exactness (Giegerich 1985:185).

Going back to Pike's rhythmic typologies, many linguists have contributed to their further development. Abercrombie (1967) makes an important claim when he says that Pike's two rhythm categories of stress-timed and syllable-timed languages are mutually exclusive – nowadays, Pike's classification is seen as a continuum rather than a dichotomy. Dauer (1983:51) emphasises the characteristics of purely stress-timed languages, highlighting the fact that interstress intervals often contain different number of syllables, which forces those syllables to be either compressed or spread out by the speaker in order to maintain a perceivable similar spacing of

¹² In order to determine whether perceptual isochrony actually existed, Woodrow (1951) and Fraise (1964) carried out several psychological experiments. First of all, they showed that, in order for sounds to be perceived in terms of a rhythmic pattern, their rate of succession must fall within a limited temporal range: if the rate is much greater than five repetitions per second, the sounds tend to be heard as a continuous stream; if, on the contrary, the rate falls below one repetition every two seconds, the sounds tend to be perceived individually (Hyman 1977:2). Secondly, they proved that, due to psychological factors, humans always tend to perceive temporal patterned events as rhythmically structured whether they are or not (Cummins and Port 1998:146).

stresses in stress-timed languages. On the other hand, in syllable-timed languages, syllables are never compressed or spread out and, consequently, the distance between two stresses will vary according to the number of intervening syllables.

Focusing on specific languages, Pike (1945) classifies English as the prototype of a stress-timed language, while Spanish is considered a very good example of a syllable-timed language. 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.

Abercrombie (1967), Ladefoged (1967, 1975) and Allen (1975), among many others, have supported Pike's views, contributing to the spreading and acceptance of the latter. These linguists follow Pike and enlarge the inventory of both stress- and syllable-timed languages, including Arabic, all Germanic languages, Russian and the already mentioned English among the languages of accentual isochrony (see Pike 1945; Bolinger 1965; Abercrombie 1967; Halliday 1967; Ladefoged 1967, among others), and French (see Delattre 1951, 1965; Abercrombie 1967; Ladefoged 1975; Catford 1977), Spanish (see Pike 1945; Gili Gaya 1950) and Japanese (Ladefoged 1975) among those of syllabic isochrony. Nevertheless, in the last years there have been some scholars (Pointon 1980; Fant 1984) who have cast doubt on the classification of certain languages – among which we can find Spanish – as strictly stress-timed or syllable-timed, thus pointing to the necessity to come up with new criteria for the rhythmic typological classification of the languages of the world.

In relation to English, there seems to be little doubt concerning its status. Most theorists agree on its classification as a stress-timed language. Its rhythm is what we can call an "accentually measured rhythm" (Sosa 1999:53), which means that the primary stresses in a sentence are situated at approximately equal intervals, without taking into account the unstressed syllables between them. As mentioned before, speakers tend to adapt their rate of speech to the number of interstress syllables: if there is none, the speaker talks slower, and if there are several ones, the speaker talks faster, so that the beat established by the stressed syllables is maintained throughout the utterance.

The application of the notion of isochrony to the prosody of English is not new at all. In fact, before Pike or Abercrombie wrote about the rhythm of English, Jones (1950:237-238) had already defended the isochronic nature of the interstress spaces as one of the most salient characteristics of the rhythm of English, stating that “[t]here is a strong tendency in connected speech to make stressed syllables follow each other as nearly as possible at equal distances”. Jones illustrated his theory with the help of musical transcriptions, already pointing to the connection between speech rhythm and musical rhythm, an idea that Liberman (1979) would apply some years later for the development of the grid theory. The scientific basis of isochrony was also championed by Allen (1973:77), who argued that our feeling that English is stress-timed arises from the preponderance of intervals whose durations lie in the narrow range from 0.4 to 0.7 seconds, which is in agreement with the findings made by Woodrow (1951) and Fraisse (1964). Later research works replicated this idea. Ladefoged (1975:103) insisted on the various factors that, to his knowledge, are responsible for the isochronic phenomenon. According to him, the combination of those factors operates as though there were a conspiracy to make English maintain a regular rhythm, a statement which Roach (1982) and Dauer (1983) would use years later. Bolinger (1981) added a nuance to the analysis of isochrony in English when he mentioned that isochrony is a phenomenon forced in production. In Bolinger’s (1981:44) own words, “[s]peakers do adjust accents, where it is practical to do so (that is, where the stress and vowel structure permits), in order to get roughly equal intervals”. Just like the former, Halliday (1985:272) emphasised the idea that “[t]here is a strong tendency in English for the salient syllables to occur at regular intervals; speakers of English like their feet to be all roughly the same length”. The notion of isochrony as the main regulating rhythmic device gives rise to the idea of the foot as the basis of speech rhythm in English. Catford (1977:182) states that “English speech tends to be delivered in a series of feet, and that feet tend to be isochronous and isodynamic”.

In contrast to all the previous authors, who, on the one hand, assert the existence of isochrony and, on the other, classify English as stress-timed, there are other scholars – Classe (1939), Dauer (1983) – whose opinion is more sceptical with regard to these two issues. For Classe (1939:85), isochrony may exist, but only under

favourable circumstances: “the groups concerned must not contain very different numbers of syllables; the phonetic structure of the component syllables must not differ too widely; the grammatical connection between the groups must be similar”. Classe’s main contribution to the topic has been proving that the number of syllables in a word or utterance is always directly proportional to the length between contiguous stresses, so, strictly speaking, isochrony is just a perceptual illusion. With regard to the classification of English as a stress-timed language, the most sceptical author is Dauer (1983:52), who states that no one has been able to prove that there is a tendency towards isochrony in English.

To this point, I have tried to summarise the view of several significant authors in reference to two issues, namely the existence of isochrony and the classification of English as a stress-timed language. The authors who defend isochrony do not have doubts about the status of English, whereas those who see isochrony as a non-existent phenomenon are not convinced about the straightforward classification of any language according to the isochronic parameter. Since the notion of isochrony seems a very useful tool for the analysis of the rhythms of different languages, for the purposes of this dissertation, I shall assume that English is an isochronic language.

When one analyses the status of Romance languages, the isochronic view turns out to be even more controversial. The most uncontroversial language seems to be French, often quoted as being the perfect prototype of syllabic isochrony: “[s]uccessive syllables within a sense group are very similar in duration, pitch, and loudness, with the exception of the last, which is, on the average, higher in pitch, slightly softer, and much longer in duration” (Delattre 1966:78). French has been claimed to belong to the accentual group (Grammont 1933; Wenk and Wioland 1982). Italian, on the other hand, is claimed to be clearly stress-timed, while Portuguese – especially Brazilian Portuguese – shows a tendency towards stress-timing (Major 1981:350).

Spanish has been described in either group. Pike’s view of Spanish as the prototype of syllabic rhythm is shared by Gili Gaya (1950), Delattre (1965, 1966), Abercrombie (1967) and Olsen (1972), and rejected by Navarro Tomás (1918) Allen (1975) and Pointon (1980). Contrary to Pike’s view is the Spanish linguist Navarro

Tomás (1918), who, on the basis of the results of a series of acoustic experiments, affirmed that Spanish is uneven in the duration of syllables, while accentuated feet remain stable. Allen (1975) reasserts Navarro Tomás's idea, emphasising the idea that there are occasions when there seem to be accentual alternations in Spanish (Allen 1975:80-81). Other authors underline the irregularity of accentual groups compared to the relative uniformity of syllabic duration in Spanish. Delattre (1965, 1966) rejects the validity of Navarro Tomás's experiments and claims that rhythm in Spanish is syllable-timed in contrast to English, a view with which Olsen agrees (1972). On the other hand, going a step further from Gili Gaya and Olsen's analyses, Pointon (1980:302) boldly suggests that Spanish is neither a stress-timed language nor a syllable-timed language:

The most probable solution seems to be that, on the evidence, Spanish has no regular rhythm in the sense of an isochronous sequence of similar events, be they syllables or stresses, but there is some form of segment-timing, in which the number and type of segments in each syllable, together with the presence or absence of stress, determine the duration of a syllable.

However, Pointon (1980:302) admits that "to the ears of native English speakers, Spanish sounds syllable-timed, and it is not surprising that English speaking investigators should want to find syllable-timing, even when it does not exist", a view shared by Bond and Fokes (1985), who say that non-native speakers of English tend to assume psychologically the differences between their own languages – when they are syllable-timed – and stress-timed languages like English.

Having in mind the controversies that Pike's typologies have generated, there have been some attempts to expand the dichotomy by adding to it new rhythm categories – for instance, the category of mora-timing, to which Japanese is said to belong –, to replace the traditional classification by some other opposition – 'leader-' vs. 'trailer-timed' –, or to relax the requirement for isochrony of inter-stress and inter-syllable intervals – stress- vs. syllable-based languages. Most significantly, many phonologists accept the hypothesis that there are languages which show features of both stress-timing and syllable-timing. As a result, and in view of the lack of empirical support for the stress-timing/syllable-timing classification, scholars such

as Dauer (1983) prefer to view speech rhythm not as a dichotomy but rather as a scale or continuum, with two hypothetical languages, one perfectly stress-timed and the other one syllable-timed, marking the ends of the scale (Dauer 1983; Bertinetto 1989):

—————→

Japanese French Spanish Greek Portuguese English

In this vein, after carrying out a series of perceptual experiments, Miller (1984) rejects the idea of a dichotomy. He defends the scalar hypothesis by saying that languages share features of both syllable and stress isochrony. For him, while some languages belong clearly to one or the other group, others fluctuate between the two (Miller 1984:75-76).

Theorists who support the scalar hypothesis claim that the rhythmic differences between languages result from the interaction of a variety of components. Depending on this interaction, each language can be given an overall rhythm score which will correspond to the position of that language on the continuum seen above. According to Dauer (1983) and Bertinetto (1989), these are some of the components that determine the score of a language:

- a) Vowel quality: 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.
- b) Consonant quality and quantity: In stress-timed languages, segmental quality and quantity distinctions are manifest in stressed syllables. In syllable-timed languages, quality, quantity and stress are independent.
- c) Syllable duration: In stress-timed languages, stressed syllables are usually longer than unstressed ones and unstressed syllables usually undergo compression. In syllable-timed languages, stressed syllables are only slightly longer.
- d) Compensatory shortening: 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.
- e) Syllable structure: 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.

- f) Relationship between syllable structure and stress: 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 independent.
- g) Position of stress: 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, not stress shift.¹³
- h) Intonation: In stress-timed languages, pitch change corresponds with stress. In syllable-timed languages, they are independent variables.
- i) Tone: In stress-timed languages, it falls only on stressed syllables. In syllable-timed languages, it can fall on both stressed and unstressed syllables.

After examining the different theories on rhythm typologies, we can conclude that the scalar hypothesis seems to be the most plausible one. The fact that it is actually based on Pike's dichotomy means that the two alternatives are not as at odds as they could seem. What these controversies seem to prove is that languages have rhythm, even if this rhythm is not based on a perfect isochrony. The perception of a relatively constant periodicity between prominent elements, be they syllables or stresses, is what makes them rhythmical.

¹³ Someone might object that the so-called 'emphatic constructions', which can take place both in stress-timed languages and also in syllable-timed languages, are cases of stress shift. In this respect, I must say that, according to linguistic theory, there are two main types of stress shift (see Dogil 1979). The first one is phonologically driven: for eurhythmic purposes – in metrical phonology, the term 'eurhythmy' refers to the modification of lexical stress patterns to achieve strong/weak alternation at the surface level (Selkirk, 1984; Hayes, 1984) –, the main stress of a word is shifted to another syllable, so that the metrical pattern of the corresponding utterance is better formed rhythmically speaking – in English, the phenomenon known as 'iambic reversal' is an instance of this (Giegerich 1985:206-218; Hogg and McCully 1987:132-148). The second type of stress shift is semantically driven; it is subdivided into two further types: in the first subtype, known as 'contrastive stress', when speakers want to highlight the contrast between two prefixes or suffixes which are present in different words of the same utterance, they can put stress on those otherwise unstressed syllables – this corresponds to the so-called 'parallel constructions' by Chomsky (1971). In the second subtype, called 'emphatic stress', speakers can put an extra degree of stress upon the focus word, in which case we cannot speak about stress shift, or they can put this extra degree of stress upon another syllable of the focus word, in which case there is stress shift. For the purpose of this dissertation, where I am dealing exclusively with phonological issues, I will only consider the first type of stress shift as a proper instance of this phenomenon. Once this point has been made, we can say that syllable-timed languages, like French or Spanish, do not experience phonologically-driven stress shift.

2.2. The prosody of Spanish

This section analyses the nature of Spanish prosody in order to define the most salient prosodic characteristics of this language. Section 2.2.1. explores the emergence and evolution of the Spanish prosodic system departing from Classical Latin. In section 2.2.2. I deal with stress assignment in Spanish, focusing on two specific theories, namely Harris (1983) and Navarro Tomás (1965), both of them amended in subsequent works.

2.2.1. From Latin to Spanish

In order to determine the prosodic characteristics of Present-day Spanish, it is necessary to overview the historical evolution of the Romance prosodic system, from Classical Latin to the present day. Section 2.2.1.1. explores the Latin quantitative system and its gradual disappearance, as well as the consequences of this process on the stress system of Present-day Spanish. Section 2.2.1.2. analyses the interaction between the vowel system, the structure of syllables and the assignment of stress in Classical Latin, and how the evolution of the first has had noticeable consequences on the other two.

2.2.1.1. Stress and the evolution of the quantitative system

As already pointed out in section 1.4.2., the languages of the world can be classified according to the kind of prosodic features they show, that is, according to the way in which they articulate the flow of sound. Languages such as English or Spanish are said to be stress languages, as opposed to tone languages and to pitch-accent languages. This means that English and Spanish make use of stress in order to mark syllable prominence. As explained by Hyman (1975:38), the most important characteristic of stress is its culminative function, which entails that in a stress language only one syllable per word can receive primary stress.

Although in all stress languages stress is culminative in nature, they can be further classified regarding the degree of freedom in the assignment of stress (Hyman 1975:204, 1977:39; Quilis 1993:388). In some languages, such as English or Spanish, stress can virtually occupy any position within the word; these are 'free stress' languages, where stress is unpredictable, and therefore lexical – at least partially – and phonemic. Other languages, like French or Polish, are characterised by the fixed position of their stress; these are called 'fixed stress' languages, where

stress is predictable, and therefore grammatical and non-phonemic. In languages with fixed or grammatical stress, stress can have a demarcative function, that is, it can be used to signal a word boundary, or a morphological function, whereby the position of stress marks the morphological category of a specific word – see section 1.4.3.3.

It is important to note that stress languages do not always belong to one of the two groups – that is, free stress and fixed stress languages – exclusively. In fact, the most common group is that of intermediate languages, which show lexical stress for most of their lexicon but also some instances of words with fixed or grammatical stress; moreover, the occurrence of one type of stress may be overridden by the occurrence of another (Hyman 1975:230, 1977:39-40). This is indeed the case of Spanish and English, although the amount of instances which show grammatical stress in both languages is really reduced – that is why these languages are classified as free stress languages.

Let us now have a look at Table 4 below, where all the possible types of stress in stress languages – lexical or unpredictable stress, and grammatical or predictable stress – are applied to examples in Spanish and English:

Lexical	Grammatical		
	Demarcative	Morphological	
Spanish	Spanish	Spanish	English
<i>término</i> (noun)	<i>termín-o</i> (stem marker)	<i>termin-ó</i> (tense and person marker)	<i>cónvert</i> vs. <i>convért</i> (noun) (verb)

Table 4: Different types of stress in stress languages.

Table 4 shows, on the one hand, an instance of lexical stress on the Spanish noun *término* ('term; boundary'), which could only be explained by means of a reference to the etymology of that word since there is no prosodic justification for the assignment of stress to the antepenult. On the other hand, grammatical stress falls on two verbal items, namely *termino* ('I finish'), where the penult is stressed in order to demarcate the end of the stem, and *terminó* ('she finished'), where the morpheme *-ó*, corresponding to the third person singular in the simple past, is marked prosodically. In English, grammatical – morphological – stress is applied to those instances such

as *có*vert (noun) and *convért* (verb), in order to determine the morphological status of the words in question.

According to the classification of Spanish as a predominantly free stress language, the stress of a hypothetical Spanish noun is unpredictable (see *término* in Table 4), and the same applies to adjectives and adverbs. This means that, for the vast majority of cases, prosodic information must be provided in the lexicon with reference to the stress of individual items. Nevertheless, Spanish nonverbal stress retains a number of important residual effects of the Latin Stress Rule, something which can shed light on the stress assignment system of Present-day Spanish.¹⁴ Before going on, I need to point out that, contrary to the view held by Latin grammarians that Latin inherited the tonal system from Ancient Greek, it is now widely accepted that Latin stress was always dynamic (Lloyd 1993:150).¹⁵ Penny (1991:42) explains the evolution of Latin stress and says that “early Latin had a type of accent in which pitch was the predominant element [pitch-accent] [...] However, for reasons that remain obscure, spoken Latin underwent a change in accent-type and came to have an accent in which energy-deployment dominated [stress-accent]”. As Allen (1964:4) explains, the rules for the location of stress in Classical Latin were simple:

It falls on the penultimate syllable if this is heavy; or on the antepenultimate if the penultimate is light. Obvious exceptions are provided in the case of monosyllables, of syncopated and apocopated forms (e.g. *illíc* < *illíce*), and of disyllables with light penultimate.

Classical Latin showed a clear preference for stress to fall on heavy syllables in words of three or more syllables, while light ones qualified preferentially as non-stressable syllables – although the latter could be stressed if a polysyllabic word did not contain any heavy syllable, or if the first syllable of a disyllabic word was light. Allen’s quotation does not specify which syllables could be classified as heavy and which ones as light in Latin, something which turns out to be essential if we take into

¹⁴ Although Spanish nonverbal stress still shows some residual effects of the Latin Stress Rule, it is impossible to conclude that Spanish inherited the Latin Stress Rule directly (see Lipski 1997:549).

¹⁵ We must take into account that Latin grammarians translated the prosodic terminology directly from Greek, so their testimony is not reliable (Lloyd 1993:150). Nevertheless, some scholars still hold the view that Latin was a pitch-accent language before it became a stress-accent language (see Beare 1957:210).

account that, in polysyllables, the position of stress was determined by the length – and, therefore, by the weight – of the penultimate syllable. Latin polysyllables can be classified into three groups: (i) the first one is constituted by those words which have a long penultimate vowel – and therefore, a heavy penultimate syllable; (ii) the second group includes a penultimate with a short vowel followed by two or more consonants, except clusters formed by consonant + /r, l/ – this would be, again, a heavy syllable; (iii) the third and last group comprises words with a penultimate formed by a single short vowel or by a short vowel followed by one consonant – in this case, we are talking about a light syllable (Penny 1991:41):

(i) Penult = long vowel	E.g.: <i>moli:nu</i> , <i>de:be:re</i> ('mill', 'to have to')
(ii) Penult = short vowel + \geq two consonants (except cons. + /l, r/)	E.g.: <i>sagitta</i> , <i>quadraginta</i> ('arrow', 'forty')
(iii) Penult = short vowel (consonant)	E.g.: <i>filius</i> , <i>vetulus</i> , <i>corrigia</i> ('son', 'old', 'strap')

Table 5: Classification of Latin polysyllables according to the weight of their penult.

Thus, length in Latin had the function of determining which syllable in a word could be stressed, though it was always subject to two rules, namely the 'three-syllable' rule, which did not allow stress to recede further than the antepenult – not even in the cases where there was a heavy syllable beyond that position –, and the 'iambic shortening' rule, which would apply to a disyllabic word, so that its final syllable was lightened and the stress could fall on the penult. According to this, Latin stress appeared to depend directly on syllabic weight, and therefore, on vowel length. In Allen's (1964:5) words, in Latin "the rules of quantity are rules of stressability". The term 'quantity' is used as a synonym for 'length', and more specifically for 'vowel length'.

Hyman (1975:206) agrees with Allen in that Latin is the best example of the connection between syllable weight and stress: "Stress is normally assigned to the penultimate syllable in Latin, except when that syllable is light. In this case, the stress is assigned to the antepenultimate position, as in *réficit* ('she fixes/remakes'),

where the penultimate syllable *fi* is light”. Hyman’s view coincides with that of Penny in that they both argue that Latin stress was always penultimate except when the penult was a light syllable; in that case, stress fell on the antepenultimate, no matter whether it was light or heavy. Nevertheless, this clear-cut view of the relationship between weight and stress in Latin is contradicted by Hyman (1977) himself, who states the exceptionality of Latin with regard to other languages with syllable weight phenomena. What Hyman (1977:53) observes is that, in Latin, the search for a stressable – heavy – syllable “cannot go back further than antepenultimate position [note *réficit* (antepenultimate stress because the penult is light) vs. *conspício* (‘I contemplate’) (antepenultimate stress on a light syllable in a word with a heavy pre-antepenult)]”, something which makes him question the assumption that Latin words have an underlyingly penultimate stress. Hyman’s (1977:53) revised theory is that Latin had a dominant antepenultimate stress, although a heavy penult would always attract stress. According to such a theory, stress changed from an underlying initial position – that is, on the first syllable of the word – to an underlying antepenultimate position, so that the number of unstressed syllables in a row could be reduced.

(2) *fácilius* > *facílius* (‘easier’)

Hyman’s theory proves the controversial state of the discussion on the assignment of stress in Latin. For the purpose of our discussion, we will adopt Penny’s widely-accepted view (see Table 5).

As regards poetic prosody, what was special about Classical Latin is that some syllables, being phonemically heavy but not accented due to positional rules – note that there could also be light accented syllables –, qualified as potentially stressable. This allowed, first, for the fact that syllables which were not stressed when a specific word was taken in isolation could be stressed in a specific context and, second, for the phenomenon that certain primarily stressed syllables could see their stress ‘devaluated’ due to positional factors – that is, in order to avoid stress lapses or clashes. We can develop a clearer idea of these two concepts by looking at a poetic fragment in Latin (Horace, in Gasparov 1996:86), where the stressed

syllables are signalled with a graphic accent, while heavy – and, in Latin, stressable – syllables are in bold:

- (3) *Ínteger* *ví-tae* / *scelerísque* *púrus*
 Honest life-of crime-of pure
 ‘The person who is honest and a good citizen/
Non éget *Máu-ris* / *iáculis* *neque árcu*
 Not need-3rd p. Moors-of darts-of nor bow-of
 does not need darts or bows/
Nec vene-ná-tis / *grávida* *sagít-tis*
 Nor venomous-of full arrows-of
 or a quiver full of venomous arrows/
Fúsce, pharét-ra...
 Fuscio quiver
 oh, Fuscio’.¹⁶

As one can observe in the above fragment, the poetic ictus in each foot – represented in bold – does not necessarily coincide with the linguistic stress of every word. On the one hand, we observe that there are words with non-stressed, yet stressable syllables – they are heavy and stressable because they contain a long vowel – to which an ictus is assigned. On the other hand, there are words without any stressable syllable, so the stress – and the ictus – has to fall on a light syllable. The most common case is the first one, while the instances of light syllable on which a metrical stress falls are rare:

Stressable (heavy) syllables with poetic ictus (´) = word stress X = stressable syllable	Non-stressable syllables with poetic Ictus (´) = word stress X = light syllable
E.g.: <i>Ínteger, vítae, Máuris, iáculis, árcu, venenátis</i>	E.g.: <i>Éget, grávida</i>

Table 6: Stress and ictus in Latin.

¹⁶ Unless otherwise stated, all translations Latin-English and Spanish-English are mine.

In sum, Latin stress assignment rules relied, on the one hand, on syllabic weight, which was a synonym for quantity, and, on the other, on the existence of three types of syllables with regards to stress, namely stressed syllables, unstressed syllables and stressable syllables. Stressable syllables were heavy, but they did not necessarily carry stress. This translated into an idiosyncratic freedom to shift stress around (as seen in (3)) without giving rise to any conflict between stress and ictus in poetic texts (Allen 1964:11).

Latin, as most languages which have a vowel length contrast, had also vowel quality differences in the corresponding series of long and short vowels (Hyman 1977:48). There were two series of vowels, namely tense and lax, which roughly corresponded to the categories of stressable and non-stressable vowels. If a lax vowel became stressable in the special circumstance that there were no tense – stressable – vowels in a word, it would not only be its length that changed, but also its quality – from lax to tense; similarly, a tense vowel which was not accented would become lax and, eventually, reduced, so it would no longer be stressable.

The relationship between stress, syllable weight, vowel length and quality could be summarised by claiming that the Latin vowel system had a direct impact on its stress system. As already noted, in Classical Latin the position of stress, which was fixed on the antepenult, depended essentially on the phonological distinction between long and short vowels, which in turn were mapped onto the distinction between tense and lax vowels and, therefore, on the heaviness of syllables. Indeed, as soon as vowel length stopped being phonemic and became exclusively phonetic, stress itself became phonemic, as it is in contemporary Spanish (Lloyd 1993:191).

The reasons for the gradual disappearance of the quantitative system in Latin are not exclusive to that language; rather, they seem to be shared by most languages. Scholars have long pointed out that a well-known characteristic of Indo-European languages in general is the tendency to eliminate phonological quantity as a feature of the vocalic system (Lloyd 1993:180-181). This phenomenon has to do with the fact that quantity distinctions in language are universally reduced to two degrees: long and short. Qualitative distinctions based on the height of the tongue and on the presence or absence of tension are not so limited. Quantity is, thus, less important than the other factors in the vocalic system of a language, that being the

reason why the former can easily be left behind when the conditions are not favourable. In Latin, quantity was not an independent distinctive feature any longer and became a non-distinctive phonetic feature which depended on the position of stress within the words, as well as on some other factors (Lloyd 1993:179). In Beare's (1957:215) words, "by the fourth century A.D. Latin vowels were losing their old quantitative difference, though they retained their qualitative difference [...]. These qualitative differences are retained by the Romance languages".

In sum, the vocalic system in Latin evolved from a quantity-sensitive system to a quantity-neutral system, affected by several circumstances, amongst which we observe (i) the relative inefficacy of quantitative distinctions in comparison with qualitative ones; and (ii) the action of stress, which made quantity depend partially on the position of the former, and so it reduced quantity distinctions to stressed syllables exclusively. The concurrence of these factors brought about a collapse of quantity as an independent phonological feature in the vocalic system (Lloyd 1993:184).

Spanish did not inherit the connection between stress, syllable weight and vowel length. As Quilis (1993:377) observes, after the complete collapse of the quantitative system, "quantity is not pertinent in Spanish, although it is true that in certain positions, and more often in a high register, a quantitative difference can be observed".¹⁷ Thus, length in Spanish is just a phonetic correlate of stress, and no longer an independent feature of vowels. That is, vowels are not inherently long or short in a phonological sense; however, when they are stressed, they are also lengthened from a phonetic point of view. In D'Introno *et al.*'s (1995:129) words, "quantity is, together with tone, the most important index of stress in Spanish. Tonic syllables are regularly longer than short ones, so the total energy tends to be bigger".

¹⁷ Certain Hispanists (García-Bellido 1993) do not agree with the view that phonological quantity is no longer present in contemporary Spanish. Indeed, scholars such as Harris (1983) relate the assignment of stress in Spanish nominals to the quantity of the penult, thus building up a system parallel to that of Latin. On the other hand, Aske (1990), Eddington (2000) and Bárkányi (2002), among others, argue that "the apparent quantity sensitivity of Spanish nominals is a historical heritage, not an active constraint", and that "stress is rather the lexical property of words" (Bárkányi 2002:375); in other words, stress is lexical.

2.2.1.2. Vowel systems and syllable structure

The vowel system in Classical Latin consisted of ten vocalic phonemes, which could be classified according to three features, namely degree of openness, place of articulation and length. In spelling, grammarians usually represented the differences in length by adding a macron or a micron on top of the corresponding vowel – in Table 7, length is represented with the diacritic <:>. The phonemic status of those ten vowels is shown in Table 8, which presents some Latin minimal pairs:

	Front	Centre	Back
High	/i:/i/		/u:/u/
Mid	/e:/e/		/o:/o/
Low		/a:/a/	

Table 7: Latin vocalic system (Penny 1991:45).

Long vowel		Short vowel	
<i>HI:C</i>	‘here’	<i>HIC</i>	‘this’
<i>VE:NIT</i>	‘he came’	<i>VENIT</i>	‘he comes’
<i>MA:LUM</i>	‘apple’	<i>MALUM</i>	‘misfortune’
<i>PO:PULUS</i>	‘white poplar’	<i>POPULUS</i>	‘people’

Table 8: Minimal pairs in Latin (adapted from Penny 1991:45).

As in most languages, Latin vowels constituted the syllabic nucleus. Syllables were mostly open in Classical Latin, a tendency which has been inherited by Romance languages. As Lloyd (1993:86) highlights, “Romance has preferred open syllables to closed ones from the time of Old Latin”. The fact that the prototypical syllable was open in Latin had a consequence on the consonants that followed the syllabic nucleus, namely that they were “especially subject to erosion” (Lloyd 1993:148). This fact was not exclusive to Latin; rather, it has always had an important effect on syllable-final consonants in Romance languages. In Spanish, for instance, codas have historically experienced some kind of erosion due to “the predominance of open syllables together with the tendency to make syllables as open as possible” (Lloyd 1993:552). Quilis (1993:368-371) argues that Spanish evinces a

clear tendency towards open syllables. The syllables that end with a vowel (V, CV, CCV) are the most common; among those, the syllable CV represents more than half of the Spanish syllabic structures; in contrast, the percentage of syllables with two postnuclear consonants is very low.

Present-day Spanish has five vocalic phonemes (D’Introno *et al.* 1995:104):

	Front	Back
High	i	u
Mid	e	o
Low	a	

Table 9: Spanish vocalic phonemes.

The corresponding open and closed realisations happen randomly, and they cannot be associated with specific phonetic contexts. Quilis chooses not to include them in the description of the allophones corresponding to the vowel phonemes in Spanish (D’Introno *et al.* 1995:93-94). In articulatory terms, there is not a great variation between stressed and unstressed vowels. In fact, the phonological inventory of the two positions is not very different, although the oppositions in tonic position are more fruitful than those in unstressed position, as exemplified in Table 10:

<i>pí</i> so ‘apartment’	<i>pé</i> so ‘weight’
<i>pó</i> so ‘sediment’	<i>pú</i> so ‘he put’
<i>pá</i> so ‘step’	

Table 10: Spanish vocalic phonemes in tonic position.

In any case, the smaller amount of energy which is generally applied to unstressed syllables makes articulatory tension lower, and so the quality of vowels is affected. In their phonetic realisation, all the vowels tend to get closer to the acoustic structure of the schwa vowel (D’Introno *et al.* 1995:97-98).

When vowels are in context, their quality may change. The contexts where a vowel tends to lose its quality will be those where several vowels or vowels plus vocalised consonants such as liquids and voiced approximants go one after the other:

(4) *Nunca he oído hablar de ellos* ('I have never heard about them').

In (4), there is a succession of six vowels, only interrupted by the approximant realisation of /d/. D'Introno *et al.* (1995:99-100) explain that, in such contexts, there are sounds which do not belong in the Spanish allophone inventory – for instance, schwa. In general terms, the realisations which are closest to schwa are found when there are two contiguous vowels, one of which disappears. In the case of diphthongs, particularly the falling ones, the quality of the second vowel normally gets altered towards the configuration of schwa – indeed, it is often pronounced as such – or even disappears (D'Introno *et al.* 1995:101-102).

As mentioned above, in Spanish the most common syllable structure is CV. Núñez-Cerdeño (1999:211) has carried out a study based on an electronic 91000-word corpus, in which he observes a clear relation between the segmental structure of the last syllable in a word, and the placement of stress in that word:

Last syllable ends in a vowel			Last syllable ends in a consonant		
Penult	57,911	88.00%	Ultima	26,642	97.80%
Antepenult	7,327	11.10%	Penult	512	2.03%
Ultima	573	0.87%	Antepenult	35	0.05%
TOTAL	65,811	100%	TOTAL	25,189	100%

Table 11: Syllable structure and stress placement in Spanish.

Table 11 illustrates three main facts about the relationship between syllable structure and stress placement in Spanish: (i) that the vast majority of Spanish words end in a vowel, (ii) that the unmarked stress position for words that end in a vowel is the penult and (iii) that the unmarked stress position for words that end in a consonant is the ultima.

All this could lead us to think that Spanish has preserved the Latin stress position; whether it has also preserved the Latin quantity-sensitive system itself or not is, as we have noted, less clear.

2.2.2. Stress assignment in Present-day Spanish

After that succinct analysis of the evolution of the Romance prosodic system, it seems in order to concentrate on the characteristics of stress assignment in Present-day Spanish. In sections 2.2.2.1 and 2.2.2.2. two classical theories are presented and amended.

2.2.2.1. Harris's analysis amended by Roca

In his (1983) book, Harris offers one of the most widely accepted accounts of the Spanish stress system, and more specifically, of the rules of stress assignment for nouns, relying on the theoretical tools provided by the school of Generative Phonology. Framing his study on the same theory, Roca (1997) offers a revised version of Harris's analysis, adding some significant new ideas.

Harris (1983:83) bases his analytical apparatus on the observation that the characteristics of the Spanish stress system follow from the peculiar morphological structures of words, where etymology plays an essential role, plus the contrast marked vs. unmarked stress placement. In his view, Spanish nonverbal stress is subject to a number of restrictions from the right edge of the lexical word (Harris 1983:85; Roca 1988:398).

As a general observation, Spanish stress must fall on one of the last three syllables of the word, never removed further to the left. The least-marked type of stress – especially if the word ends in a vowel – falls on the penult (see Table 11).¹⁸

(5) *cá*sa, *armá*rio, *encué*sta, *latifú*ndio ('house', 'wardrobe', 'poll', 'estate')

However, penultimate stress is unacceptable if the last syllable contains a falling diphthong – [a_ɨ, o_ɨ, e_ɨ, a_ɯ, o_ɯ, e_ɯ].

(6) **cón*voy ('convoy')

As for antepenultimate stress, the least marked instances occur when a word ends with a vowel:

¹⁸ Penultimate stress is four times more frequent than oxytone stress and antepenultimate stress, which is the least frequent one. See Quilis (1993:400-403) for a thorough account of the frequencies of the different accentual patterns in Spanish.

(7) *teléfono, auténtica, tránsito* ('telephone', 'authentic', 'transit')

Unless the singular form of a noun is antepenultimate, which implies that the plural will be so – the plural basically adds -s to the last vowel of the singular, leaving the stress pattern unchanged –, antepenultimate stress in consonant-final words is severely restricted:

(8) *teléfono, teléfonos*

Antepenultimate stress occurs in only two sets of forms: (i) a closed group of words of Greek origin:

(9) *hipérbaton, metátesis* ('hyperbaton', 'metathesis')

and (ii) a small group of 'random idiosyncratic words':

(10) *régimen, Júpiter, Álvarez* ('regime', 'Jupiter', 'Álvarez')

On the other hand, if the penultimate syllable is heavy, which in Spanish means that it is closed by a consonant or containing a semivowel, antepenultimate stress is highly marked or, for Harris (1983:88), even impossible:

(11) **masájista, *catápulta, *énigma* ('masseur', 'catapult', 'enigma')

Antepenultimate stress is also impossible if the last syllable contains a diphthong:

(12) **expériencia, *négocio, *cálvicie* ('experience', 'business', 'baldness')

Last, oxytone stress is unmarked if the last syllable is heavy, that is, closed:

(13) *albañíl, miradór, cajón* ('bricklayer', 'viewpoint', 'drawer')

It is possible to have consonant-final words where the stress falls on a non-final syllable; such instances are seen as marked yet acceptable:

(14) *exámen, dúctil, azúcar* ('exam', 'ductile', 'sugar')

Nevertheless, if the final 'consonant' of a word is a glide, non-final stress is unacceptable, as we have seen above, while glide-final oxytones are well formed:

(15) *convóy, caráy* ('convoy', 'good gracious!')

Vowel-final oxytones are widely attested in Spanish. In contrast with vowel-final penultimates, these oxytones seem to be marked in that they feel less natural than the former. In sum, Spanish nonverbal items can end either in a vowel or in a consonant. If they end in a vowel, stress is prototypically placed on the penult; if they end in a consonant, stress falls on the last syllable. Although exceptions to these rules are possible, and indeed numerous, stress will invariably fall on one of the last three syllables of the word (Roca 1997:631).

Table 12 summarises the acceptable and unacceptable occurrences of stress in Spanish:

Type of stress	Consonant-final words	Vowel-final (glide-final) words
Oxytone	<i>albañil, miradór, cajón</i>	<i>convóy, caráy</i>
Penultimate	<i>exámen, dúctil, azúcar</i>	<i>cása, armáριο, encuésta, latifúndio,</i> <i>*cónvoy</i>
Antepenultimate	<i>hipérbaton, metátesis</i> <i>régimen, Júpiter,</i> <i>Álvarez</i>	<i>teléfono, auténtica, tránsito</i> <i>*masájista, *catápulta, *énigma ,</i> <i>*expériencia, *négocio, *cálvicie</i>

Table 12: Stress in Spanish.

By observing the previous descriptive account of Spanish nonverbal stress, the question which arises is why antepenultimate stress with penultimate or final heavy syllables is restricted in Spanish. In order to account for that problem, 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”, which is the case in (11) and (12) above. According to this observation, Spanish would have to be classified as a quantity-sensitive language.

However clear Harris’s theory might seem, the exceptions to a potential quantity-sensitivity rule are quite numerous and varied in nature (Roca 1988:417). For once, referring to penultimates, although the majority of superheavy syllables, with three or more positions in their rhyme, must – and indeed do – bear final stress, in a noun like *tóra*[ks] and a few other Latinisms “a superheavy ultima is manifestly weak” (Roca 1988:404). Second, we have already analysed two exceptional word groups – see above, Greek words and Spanish idiosyncratic words – with antepenultimate stress, something which contradicts the theory. But the case that most definitely overrides Harris’s theory is that of a stressless penultimate heavy syllable in a non-idiosyncratic, perfectly Spanish antepenultimate place-name like *Frómista*, mirrored by some foreign words which seem to be freely exempted from the above constraint, like *Róbinson*, *Wáshington* or *Mánchester* (Roca 1988:416). According to Roca (1988:418), the occurrence of such foreign words effectively turns the *régimen* group into “an open class, thus increasing the pressure for a principled solution”. As noted before, although most members of the class do indeed have a foreign or erudite flavour – see *hipérbaton* – at least *régimen* is genuinely Spanish. The reason for the existence of these exceptions is indeed historical: the stress locus inherited from Latin has been almost completely preserved in Spanish. Given that all the exceptions previously explored climb to an extraordinarily high amount, Roca (1988:417) concludes that Spanish stress is not quantity-sensitive. His arguments are reinforced from different perspectives, both by Hayes’s 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.

The contradictory evidence given by different authors to defend the quantity sensitivity or insensitivity of Spanish leaves a theoretical gap with regard to Spanish stress assignment, which Roca (1997:627) tries to fill in by introducing the term

'accent', used to encode deviations from the normal pattern. To start with, Roca adapts the following rule from Halle (1991:145): "Accent [...] the rightmost (metrical) syllable". Subsequently, Roca formulates the principle that in Spanish stem-final stress is unmarked, especially if the corresponding rhyme is heavy or followed by a palatal consonant. Finally, Roca (1997:644) concludes that any deviation from this principle must be lexically marked.

Once we have dealt with the assignment of primary stress in Spanish, it is important to note that some Spanish words can show weaker stresses, too. These may appear either on the initial syllable of the word – *gènerativo* ('generativist'), *gràmaticàlidád* ('grammaticality') – or on even-numbered syllables counting leftward from the primary stress – *genèrativo*, *gramàticàlidád*. The only condition attached to non-primary stresses is that they cannot occur adjacent to each other or to the primary (Harris 1983:86) – that is, stress clashes are avoided.

2.2.2.2. Navarro Tomás's analysis amended by D'Introno *et al.*

In his (1965) book, Navarro Tomás established a series of rules of stress assignment which have long been regarded as an authority for any study of Spanish phonology. Nevertheless, Navarro Tomás's principles have recently been considered insufficient, and have thus been amended in a comprehensive volume by D'Introno *et al.* (1995). Taking a different path from Quilis (1993), these authors start by stating that Spanish stress shows only two main functions, namely a distinctive and a culminative function (D'Introno *et al.* 1995:124-225). The latter refers to the fact that Spanish words have just one tonic syllable, while the former points to the phonemic value of stress in Spanish, that is, the position of stress distinguishes between meanings of words formed by the same sequence of phonemes.

Always relying on Navarro Tomás's observations, D'Introno *et al.* (1995:161-173) make a first very important observation when they assert the existence of three different degrees of stress in Spanish, namely primary, secondary and tertiary stress. As a second step, they proceed to explain the rules of primary stress assignment in nouns and adjectives.

In Spanish, the domain of stress assignment is what D'Introno *et al.* (1995:162) call "the minimal word", which is constituted by what they call "a lexeme" and a gender suffix – prototypically *-a*, *-o*, although there are others. Their

notion of 'lexeme' includes the lexical morpheme alone or, if there is any derivational suffix, the lexical morpheme together with this one. Stress is assigned according to two basic rules, which will be quoted hereafter as 'Rule A' and 'Rule B' (D'Introno *et al.* 1995:164-165):

Rule A - *Closed Syllable Stress Rule*: The first closed syllable – counting from right to left – of the minimal word must be stressed.

Rules B - *Default Rule*: Stress the second syllable.

There are several observations to be made on the behaviour of Spanish words with reference to the above rules, which I will illustrate with some examples. Rule A works for oxytones ending in consonant, penultimates with a closed penult and closed antepenultimates with the two or three following syllables open. Nevertheless, words such as *hermano* ('brother'), *persona* ('person') or *cartera* ('purse') pose a problem to this rule, given that their structure would require them to be antepenultimates (see D'Introno *et al.* 1995:165) – the solution to this apparent problem will be explained later. Similarly, Rule B does not seem to apply to antepenultimates like *término/sábana* ('sheet') or *sílaba* ('syllable'). The solution in this case is to establish an opposition marked/unmarked stress, where oxytones which end with a vowel – like *sofá* ('sofa') –, paroxytones which end with a consonant – like *cadáver* ('corpse') – and proparoxytones with the last three syllables open – like *sílaba* – must be assumed to be lexically marked if there is no phono-morphological reason for them to be stressed. The rest of the words get their stress according to Rule A or, this one lacking, to Rule B. In any case, as D'Introno *et al.* (1995:165) highlight, this theory points to an interesting fact: stress never falls on inflectional morphemes.

There are other groups of words which do not seem to respond to any of the above rules, and require a further explanation (D'Introno *et al.* 1995:165ff.). First of all, the analysis under examination leaves words such as *metamorfosis* ('metamorphosis'), *hipnosis* ('hypnosis'), *hipótesis* ('hypothesis'), *hepatitis* ('hepatitis') or *sífilis* ('syphilis') without an explanation. D'Introno *et al.* assume that the endings *-sis*, *-tis*, and *-is* do not count as derivational suffixes, and that final *-s* does not count at all – note that they do not change to form the plural. According to this solution, words like *hepatitis* or *artritis* would be obtained applying Rule B,



while the others would be labelled as lexically marked. The second case includes words ending with *-es*, like *viernes* ('Friday') and *Hércules*. In these words, the final *-s* final is not a suffix, so Rule B applies again. A word like *Júpiter* is said to be lexically marked. Last, words such as *baúl* ('trunk') or *maíz* ('corn'), with hiatus, contrast with words like *jaula* ('cage') or *Raiza*, with a diphthong where we find a the glides [ɥ] in the first case and [j] in the second, followed by a vowel; if we admit that glides are derived from high vowels, we can sustain the view that before the application of Rule B there is a process by means of which a high vowel in contact with another vowel becomes a glide – sliding. At this point, Rule B can no longer be applied to the glide, but rather to the vowel.

The explanation of the apparent exceptions seen above forces D'Introno *et al.* to establish a sequence in the application of the rules, which goes as follows: (i) apply Rule A – except when there is a derivational suffix –, (ii) apply sliding – except when there are several derivational suffixes – and (iii) apply Rule B – except when there are several derivational morphemes.

A more complicated apparatus must come into play when we have a look at words with derivational morphemes (D'Introno *et al.* 1995:167ff.). The first group includes words with the derivational suffixes *-on*, *-cion*, *-es*, *-or*. They end with a consonant, so they must be assigned stress according to Rule A. The second group includes minimal words with a derivational morpheme ending in a vowel, which means that Rule A applies:

(16) *zapat-ero*, *ca-sita*, *cas-ona* ('shoe maker', 'little house', 'big house')

There are words with several derivational morphemes – class and diminutive/augmentative – ending in a vowel. In this case, the above rules are applied to the last derivational morpheme, placing the stress on that morpheme:

(17) *cancioncita*, *poesía*, *zapatería* ('little song', 'poetry', 'shoe shop')

In the last two words in (17), the process of sliding gets blocked because there is a condition that requires the application of stress rules. In the case of words with an

only suffix such as *-acia*, *-ancia*, *-encia*, *-ario*, *-oria*, sliding applies. As a last observation, some words with an added derivational suffix – *-ito*, *-ítico*, *-cito* – are lexically marked, in the sense that the suffix gets the stress. The exception to this is the suffix *-tico*, which does not admit stress, so it moves back to the syllable before the suffix. In the case of *-ólogo*, stress is lexically marked on the suffix, too (D’Introno *et al.* 1995:168-169).

Once the rules for words with derivational suffixes have been explored, D’Introno *et al.* (1995:168) find a good solution for words like *hermano*, seen above, which apparently override Rule A. He takes *hermano*, *cartera/-o* or *calvicie* to have a derivational suffix, over which stress is applied, even though it is applied on the closed syllable.

We could summarise the rules of primary stress assignment for Spanish nouns and adjectives as follows: as a general rule, stress can never be assigned to a fourth syllable from the end of the word; put differently, the three-syllable rule inherited from Latin is still at work. Primary stress is assigned to the minimal word; if that word contains any derivational morphemes, stress is assigned to the last one of those morphemes. Last, sliding is inapplicable if the minimal word has a derivational suffix which contains only one vowel.

As for secondary and tertiary stress, D’Introno *et al.* (1995:172) establish a series of rules, summarised as follows:

- a) Primary stress is assigned, as in *em-pe-ra-dó-res* (‘emperors’)
- b) Secondary stress is assigned to the first unstressed syllable in the word – in case there is one –, as in *è-m-pe-ra-dó-res*.
- c) Stress is assigned to other syllables in accordance with the rules of secondary accent: tertiary stress next to primary stress, secondary stress next to tertiary stress, tertiary stress next to secondary stress, etc, as in *è-m-pè-ra-dó-res*.
- d) The sequence secondary stress + secondary stress changes into secondary stress + tertiary stress, as in *è-m-pe-ra-dó-res*.

In compounds and adverbs ending in *-mente*, stress gets assigned to the first lexical element, then to the second one; the stress of the first element is reduced, becoming secondary (D'Introno *et al.* 1995:173).¹⁹

D'Introno *et al.*'s (1995:412-414) implementation of Navarro Tomás's model has the added value of rendering a brief, yet comprehensive, account of stress assignment rules in verbs, organised with reference to the different tenses in Spanish. First, in the present tense of regular verbs, both in the indicative and in the subjunctive, stress falls on the penult (except for the second person in the plural) independently of the type of syllable and of the preceding syllable. As for second person of the plural, stress falls on the last syllable, which means that Rule B is applied, then sliding in the first and second conjugations, and elision/synaeresis in the third one, where the thematic vowel is in the same syllable as the morpheme of person. Second, in the imperfect (indicative and subjunctive), the thematic vowel is stressed (note that it gets elided in the second and third conjugations). Third, the past tense carries stress on the thematic vowel, which is then elided, while the accent remains on the original syllable. The exception is the third person of the singular, lexically marked on the morpheme *-ó*, especially in the second and third conjugations. Last, the future and conditional are marked with the stress on the morphemes *-ré* and *-rí*, respectively. In non-finite forms Rule B is applied on the participle and gerund, while the infinitive is lexically marked on the last syllable – morphemes *-ar, -er, -ir*.

¹⁹ Núñez-Cerdeño (1999:225) postulates that secondary stress is rhythmic in Spanish. Secondary stress does not stem from the assignment of primary stress, as it does in other languages. It is, on the contrary, a manifestation of the metrical rhythm in the phrase. So, while primary stress is assigned according to lexical rules applied at the level of the phonological word, secondary stress responds exclusively to postlexical rules, applied at the level of the phonological phrase.

Finite forms				Non-finite forms		
Indicative and Subjunctive				Infinitive	Gerund	Part.
Present	Imperfect	Past	Future/ Conditional	Marked Suffix	Rule B	Rule B
Penult (except 2 nd plural)	Stress on thematic vowel	Stress on thematic vowel + elision	Marked stress on morphemes -ré, -rí			

Table 13: Rules of stress assignment for Spanish verbal forms.

2.2.3. Syllable-compression phenomena in Spanish: synaeresis and synalepha

As mentioned in sections 1.4.3.4. and 2.1., the so-called ‘stress-timed languages’ like English show a tripartite connection between stress, vowel length and quality, and syllable weight (Dauer 1983). In general terms, heavy syllables, that is, syllables whose rhyme is formed by a long vowel or a vowel plus a consonant, are more likely to attract stress than light syllables – “prevocalic segments in the syllable (i.e. onset segments) are prosodically inert” (Hayes 1995:51). Stressed vowels are, in turn, longer than unstressed ones, and unstressed vowels usually undergo compression. This means that, while stressed vowels have full articulation, unstressed ones are reduced or centralised. A vowel in an unstressed position tends to be made shorter, laxer, and with a neutralised quality; this is to say, it tends to become a schwa or, in some cases, /ɪ/ in stress accent languages (Giegerich 1992:287).

Vowel reduction in English is dependent on word stress. This has, in turn, direct implications for the general rhythm of the language: in English speech, only the stressed syllables in an utterance tend to be evenly spaced in time. This means that unstressed syllables are forced to hurry into short time-limits, a fact which is responsible for the reduction and/or loss of quality of English unstressed vowels (MacPherson 1975:35).

Syllable-timed languages like Spanish do not show the connection stated for stress-timed languages. The vowel system in Spanish, with just five length-neutral vowels – that is, neither phonologically short nor long (see section 2.2.1.2.) – is

responsible for the fact that weight depends on a different variable, which is, in this case, the open or checked status of syllables: in Spanish, a syllable is heavy exclusively when it is checked by a consonant. According to this, it is impossible to state that heavy syllables attract stress in Spanish. On the other hand, given that vowels are neither long nor short, unstressed syllables do not necessarily count as non-stressable syllables.

The syllable-timed rhythm of Spanish does not explain why this language requires that stresses fall at specific points in the utterance; in other words, stresses at the level of the utterance in Spanish are not completely free, but rather dependent on the overall rhythm of utterances, which require a certain number of rhythmical 'leaning-points' at certain intervals. This entails that, at points where unstressed syllables are very numerous, there will be a need to increase the speed of discourse, which means that certain processes of 'compression' must come into play. There are two compression devices in Spanish, which have two corresponding 'decompression' devices. The former are synaeresis and synalepha; the latter are dieresis and hiatus.

As observed in 2.2.1.2., five vowel phonemes can be identified in Spanish on the basis of a five-way minimal contrast (see Tables 9 and 10). Vowels constitute syllabic nuclei, and all the nuclei contain a vowel – note that there are no syllabic consonants in Spanish. Vowels may be adjacent in a word, as well as in different words, in which case either a hiatus or a diphthong may arise.

The phonemes /i/ and /u/ can be realised as glides – [ɨ], [ɨ̃] (also called 'semivowels' in Spanish) and [j], [w] (also called 'semiconsonants'). A vowel may be preceded by an *-i* or a *-u* in spelling, realised as an onglide – [wa, we, wi, wo, ja, je, jo, ju] –, which leads to an increase in sonority over the length of the combination; in such a case, the outcome will be a rising diphthong. Second, a vowel is followed by an *-i* or a *-u* realised as an offglide – [aụ, eụ, iụ, oụ, ại, ẹi, ọi, ụi] –, which entails a decrease in sonority throughout the combination; under these circumstances a falling diphthong will arise. The combinations [wu, iu, ji] are very rare in Spanish, while [uɨ] is completely unattested. Any other combination will normally make the two segments involved belong to different syllables, in an instance of hiatus.

In sum, the glides can be seen as the forms taken by /i/ or /u/ when they occur in a diphthong with another vowel. Each glide can occur immediately before or immediately after the vowel with which it forms a diphthong, as seen in Table 14.

	<i>Before vowel</i>	<i>After vowel</i>
/i/	[ˈtjera] <i>tierra</i> ‘land’	[ˈbojna] <i>boina</i> ‘beret’
/u/	[ˈfwego] <i>fuego</i> ‘fire’	[euˈropa] <i>Europa</i> ‘Europe’

Table 14: The behaviour of Spanish glides.

It is important to note that Spanish diphthongs, unlike English ones, are not single phonemes which correspond to a single syllable, but rather two independent vocalic phonemes which are combined to form one syllable. Every diphthong has a clear peak of prominence, which is the most sonorous vowel. As observed above, the most normal situation involves one of the vowels being a high vowel, so that the other vowel – high, medial or low – can occupy the syllabic nucleus. Therefore, at the lexical level, the sequences that can form a diphthong are those which involve /i/ or /u/ and any other vowel. Under certain morphological circumstances, like the addition of a plural morpheme, the two parts of the diphthong can be redistributed between two different syllables – *ley* [lej] vs. *leyes* [ˈlejes]; ‘law’, ‘laws’ (MacPherson 1975:52-54).

As I have just mentioned, when two vowels which appear in succession and would naturally form a diphthong are pronounced in separate syllables, they are in hiatus. This is normally the case when none of the vowels involved is high. Nevertheless, even when a high vowel is involved, there are some exceptional cases where hiatus is preferred over diphthong, something which does not have any morphological explanation – for example, in cases like *cli-en-te* (‘customer’), *di-a-blo* (‘devil’), *Ru-an-da* (Núñez-Cerdeño 1999:183). According to Canellada (1987:51-53), hiatus has seventy-five theoretical combinative possibilities in Spanish, twenty-five for each combination – two unstressed syllables, unstressed plus stressed syllable, and stressed plus unstressed syllable –, as observed in actual speech. Table 15 summarises the possible combinations of Spanish vowels.

/-V.V-/
< (from a more closed to a more open segment): /i/ and /u/ + /e/, /o/, /a/
> (from a more open to a more closed segment): /e/, /o/, /a/ + /i/, /u/
< : /e/, /o/ + /a/
> : /a/ + /e/, /o/
= (segments of equal openness): /i/, /u/ + /i/, /u/
= : /e/, /o/ + /e/, /o/
= : /a/ + /a/
/-V.V'-/
< : /i/, /u/ + /'e/, /'o/, /'a/
> : /'e/, /'o/, /'a/ + /'i/, /'u/
< : /e/, /o/ + /'a/
> : /a/ + /'e/, /'o/
= : /i/, /u/, /'i/, /'u/
= : /e/, /o/ + /'e/, /'o/
= : /a/ + /'a/
/-V'.V-/
< : /'i/, /'u/ + /e/, /o/, /a/
> : /'e/, /'o/, /'a/ + /i/, /u/
< : /'e/, /'o/ + /a/
/'a/ + /e/, /o/
= : /'i/, /'u/ + /i/, /u/
= : /'e/, /'o/ + /e/, /o/
= : /'a/ + /a/

Table 15: Hiatus in Spanish.

It must be observed that, in popular speech and in rapid educated speech, hiatus is not always preserved, so that vowels which would otherwise belong in separate syllables of the same word – typically, two medial vowels, like /a/ and /o/, or /e/ and /o/ – are compressed to form a single syllable, a fact that reinforces the assertion that Spanish diphthongs are not independent entities, but two vowels made to belong in the same syllable. This entails that, at the postlexical level, we have to accept the existence of both high and medial glides.

If two adjacent vowels within the same word are made to correspond to a single syllable, the resulting phenomenon is called ‘synaeresis’. The most frequent instances of synaeresis occur when two unstressed alike vowels come on after the other, something which gives rise to the fusion of those two vowels into one of the same quality:

(18) *neerlandés* → *ner-lan-dés* (‘Dutch’)

In contrast, it is not very frequent when one of the adjacent vowels bears a stress:

(19) *leer* → *lér* (‘read’)

Synaeresis between unlike vowels is particularly common in sequences of rising sonority – /ia, ie, io, ua, ue, uo/. Nevertheless, there are some exceptional cases where hiatus is preferred:

(20) *riamos* → *ri-a-mos* (‘let us laugh’); *diente* → *di-en-te* (‘tooth’); *dueto* → *du-e-to* (‘duet’)

There are several possible causes for this unexpected choice of the marked option (Hualde 2002:217-218). One of them is that the hiatus is preserved because the root word also has hiatus:

(21) *riada* → *ri-a-da* (‘flood’) < *rí-o* (‘river’) ← *río*

Another explanation is that there can be a perceived word boundary between the two vowels, and the hiatus is preserved to signal it:

(22) *bienio* → *bi-e-nio* (meaning ‘two years’, from Latin).

Last, the phonological context can cause the preference for hiatus over diphthongisation. This is the case when a word starts with a trill, or when the word-initial sequence is #(C)ia, #(C)io.

The least common of the possibilities arises when two unlike vowels come into contact to form a single syllable, particularly if one of the two is lexically stressed:

(23) *teoría* → *teo-rí-a* (‘theory’)

This fact forces us to observe a lexical contrast between sequences in hiatus and diphthongs, where the position of stress determines the realisation of one or the other (Hualde 2002:217):

(24) *pié* → *pi-é* (‘I cheeped’) vs. *pie* → *pjé* (‘foot’)

The realisation of synaeresis depends, on the one hand, upon linguistic factors, such as the position of the word in the rhythm group – when the word in question is completely unemphatic, synaeresis is very likely to be found. On the other hand, the realisation of groups of vowels as hiatus or synaeresis has a sociolinguistic component, a fact that makes it difficult to reduce these features to a set of rules: uneducated speakers make a larger use of synaeresis; educated speakers in fast conversation are also likely to use synaeresis (MacPherson 1975:37). Beyond normal speech, synaeresis is a regular and completely accepted feature of Spanish verse, as will be observed in section 3.2.3.

Quilis (1993:370) explains synaeresis in terms of the position of the tongue in the articulation of the vowels involved (Table 16). The contact between two non-high vowels gives rise to two different syllables; the contact between a low or medial

vowel, and a high vowel, or vice versa, gives rise to a single syllable; when the contact is between a non-closed, unstressed vowel and a closed, stressed vowel, two different syllables arise.

Articulation of the vowels	Resulting syllables: examples
Two non-high vowels → Two syllables	a-é-re-o <i>aéreo</i> ('aerial'), pe-le-ár <i>pelear</i> ('to fight'), lé-a <i>lea</i> ('read!')
Low/medial vowels + high vowel (or vice versa) → One syllable	ái-re <i>aire</i> ('air'), eu-ró-pa <i>Europa</i> , á-sia <i>Asia</i> , bué-no <i>bueno</i> ('good')
Non-closed, unstressed vowel + closed, stress vowel	a-bí-a <i>había</i> ('there was'), pa-ís <i>país</i> ('country'), re-ú-no <i>reúno</i> (‘I gather’), ba-úl <i>baúl</i> ('trunk')

Table 16: Synaeresis in Spanish.

D’Introno *et al.* (1995:209) equate synaeresis with elision and state that, “[w]hen there are two equal vowels next to each other, there is a tendency to elide the unstressed one”. There are four main elision rules according to this model:

Rule	Formula
<i>Vowel elision</i> : elide a vowel if it is followed by another vowel with the same degree of aperture and same point of articulation.	$V \rightarrow \emptyset / \text{_____} [a \text{ aperture, } b \text{ point}]$ $V [a \text{ aperture, } b \text{ point}]$
<i>Elision of syllabic boundary</i> : elide \$ between a consonant and a vowel, or between a displaced vowel and a vowel, and vice versa.	$\$ \rightarrow \emptyset / C, D \text{ _____ } V; V \text{ _____}$ D
<i>Sliding</i> : turn a vowel [+high] into a displaced one if it is preceded or followed by another vowel.	$V \rightarrow D / \text{_____} [\text{high}] V; V \text{ _____} [\text{high}]$
<i>Vowel rising</i> : turn a [+medial, +lax] vowel into a [+high, +lax] vowel if it is in contact with a stressed vowel within the same word.	$[\text{Medial}] \rightarrow [\text{High}] / \#X 'V \text{ _____} [\text{Vowel, Lax}] Y \#; \# X \text{ _____} [V, R] 'V Y \#$

Table 17: Rules of elision (D'Introno *et al.* 1995:223-224).

A similar type of process takes place when two or more vowels which belong to different words that come into contact in a rhythm group are combined to form a single syllable: $V\#V \rightarrow (V) (V)$. This is known as 'synalepha', and often occurs when the final vowel or vowels of one word is joined together with the initial vowel or vowels of the following word. A curious phenomenon is observed when synalepha brings unlike vowels together to form a single syllable. In that case, the main stress in that syllable will fall upon the more or most perceptible of the vowels, not on the one that belongs to the most lexically significant word.

(25) non-tonic pronoun *se* + verb *únen* \rightarrow *séunen* ('they unite')

This entails that the stressed vowel in a group of vowels brought together in synalepha may be a vowel which is not normally stressed when the word to which it belongs is pronounced in isolation (MacPherson 1975:57).

The maximum of vowels joined together under synalepha is five:

(26) *Subió a Eulalia al árbol* → *su-bióaeu-la-liaal-ár-bol* ('She lifted Eulalia onto the tree')

The most common synalepha groups two vowels. It normally covers two words, but sometimes it extends over three:

(27) *Venció a un enemigo* → *ven-cióaun-e-ne-mi-go* ('he defeated his enemy'); *rompió en aplausos* → *rom-pióen-a-plau-sos* ('he burst out clapping')

There are as many combinatorial possibilities for synalepha as we have seen there are for hiatus (see Table 15).

Generally speaking, there are three conditions for synalepha to take place, namely that the tempo of speech must be fast, the register must be colloquial and, when more than two unlike vowels come together, they may be combined to form a single syllable provided that the syllable contains a continuous rise or a continuous fall in prominence, or a peak of prominence. If a combination of vowels contains a trough of prominence, the vowels are spread over two syllables (MacPherson 1975:56). In other words, when there are three or more vowels involved, there must not be a closed one between two open ones; if this happens, hiatus will take place:

(28) *estoy aquí* → *es-toi/a-kí* ('I am here')

According to D'Introno *et al.* (1995:224-225), the processes of vowel elision, syllable boundary elision and sliding (see Table 17) also apply to synalepha. Elision is more frequent when one of the vowels belongs in a clitic, that is, an unstressed and unstressable monosyllable, like determiners, prepositions, non-tonic pronouns, relatives, conjunctions, etc. It is also possible to elide the vowel in a clitic followed by a different vowel, something that does not occur in synaeresis. Any [+high] vowel

in contact with a different vowel can slide and be syllabified with the latter. Sliding of a [+medial] vowel does not happen across words.

Summarising what we have explored so far, the union of two or more vowels may be of four types in Spanish (Canellada 1987:54):

/V.V-/→

Hiatus		[-V.V-]
Synalepha ²⁰	Reduction	[-VV-]
	Diphthongisation	[-SV-]; [-VS-]
	Elision	[-V-]

Table 18: Compression and decompression devices.

As can be observed in Table 18, synalepha can be divided into three further subtypes, namely reduction, diphthongisation and elision. Diphthongisation is the prototypical type of synalepha, and occurs when a vowel comes together with a high vowel, in such a way that the latter becomes glide. This is why this process is sometimes referred to as ‘gliding’, represented as /VV/ → [Vj], [Vɥ], [jV], [wV].

(29) *mi amigo* → *mja-mi-go* (‘my friend’)

In the case of reduction, there is a vowel – typically, the first in a sequence of two medial vowels – which becomes non-syllabic postlexically, that is, it is not ‘placed’ on the nucleus of the resulting syllable. Although it preserves its quality, it loses a part of its quantity and intensity, and it becomes a lax vowel. The vowel resulting from this union is perceptually longer than in elision: /VV/ → [(V)V].

(30) *se acabó* → *s(e)a-ka-bó* (‘it is over’)

²⁰ Studies such as Cabré and Prieto (2003:1687-1690) only use the term ‘synalepha’ in the following sense: two vowels in a V#V sequence occupy a single metrical position – in written verse – or a single note of music, as if melismata did not exist – a melisma is a passage of several notes sung to one syllable of text. In this view, reduction, diphthongisation and elision would be three related phonological phenomena which could sometimes give rise to poetic and/or musical synalepha.

Elision is the least frequent type of synalepha and, as the name indicates, it occurs when a vocalic sound is completely deleted in favour of the following one, which comes to adopt the onset of the previous vowel and preserve its own rhyme. This is particularly normal when the two vowels joined together are the same – e.g. /e/ + /e/: /V1V2/ → [V1]; [V2].

(31) *se acabó* → *sa-ka-bó*

The realisation of synalepha as diphthongisation, reduction or elision follows four general rules (Canellada 1987:54-55; MacPherson 1975:55-56), namely:

- a) If high vowels /i/ or /u/ appear, there may be diphthongisation. That is to say, when one of the vowels is /u/ or /i/, it is pronounced, according to its position in the syllable, as an onglide or offglide.
- b) The more open vowel prevails over the more closed one.
- c) The tonic vowel prevails over the non-tonic one.
- d) The element which goes in the second position prevails over the one in first position.

These rules can, in turn, be reduced to two main ones and one exception which regulate diphthongisation in Spanish (Piera, personal communication), namely:

- a) The first type of diphthongisation is lexical, and it involves the glides [j], [w], [i̯], [u̯].
- b) The second type is postlexical and involves any sequence of two adjacent vowels. In this case, the only inviolable principle is that, if /a/ is involved, it must always be the syllabic nucleus of the resulting syllable.
- c) As an exception, there are a limited number of cases of medial vocoids realised as diphthongs already at the lexical level, namely:
 - c.1) -AO: *Bilbao*, *bacalao* ('codfish'), *cacao* ('cocoa'), *Curaçao*, and the participles in -ao (<-ado) and their nominalisations – e.g. *pescao* ('fish'), *empleao* ('employee'), *pesao* ('heavy')
 - c.2) -OA: *coalición* ('coalition') (vs. *ko-a-la*), *coartada* ('alibi')
 - c.3) -EA: *real* ('royal') (vs. *re-al*, 'real')
 - c.4) -AE: *trae* ('bring that here')

Although the phonological dichotomy long vs. short for vowels does not exist in Spanish, it seems obvious that, whenever two vowels are joined together under synalepha – particularly under diphthongisation and reduction –, the outcome will be a longer vowel. Actually, the length of a vowel made long under synalepha can be up to double the length of a simple vowel. Nevertheless, the fact that we can always undo synalepha proves that vowel length does not have phonemic value (Canellada 1987:54): -V.V-/→[-V:-]/[-V-].

Although primarily considered a linguistic phenomenon attested in normal speech, synalepha is also an enormously productive device in the poetry of syllable-timed languages like Spanish. As a general principle, syllable count in the poetry of these languages is based on the principles of phonetics and euphony which govern the speech of the individual languages. In Spanish, as well as in Provençal, Italian and Portuguese, the adjustment of the verse line to the required number of syllables is achieved by synalepha, by elision or by hiatus (Preminger 1974:713-714). Synalepha and its realisations in verse are covered in the next chapter.

2.3. The prosody of English

This section analyses the nature and characteristics of the English prosodic system. Section 2.3.1. constitutes an examination of the emergence and evolution of the English prosodic system starting with Old English. Section 2.3.2. deals with stress assignment in English, focusing on the theories developed by Halle and Vergnaud (1987), Giegerich (1992) and Gasiorowski (1998). In sections 2.3.3. and 2.3.4. I examine the interaction between stress, vowel length and quality and syllable structure in English, as well as the consequences of this interaction for the assignment of stress in this language.

2.3.1. From Old English to Present-day English

The evolution of the stress assignment system in English is a controversial, as well as intricate, topic in English historical linguistics. This chapter amounts to an overview of that evolution, which will help understand the basic principles of Present-day English (henceforth PdE) stress assignment.

Following the tendency of most Germanic languages, in Old English (henceforth OE) the primary stress of native words and, according to Halle and

Keyser (1971:88), also borrowed ones, was placed on the first syllable of the stem. A stem is the combination of the basic form of a word – the root – plus any derivational morphemes, which excludes inflectional elements; put differently, the stem is the form of the word to which inflectional morphemes can be added. In Lass’s (1994:91) words, OE stress falls “on the first syllable of a simplex word, but on the first syllable of the lexical root in morphologically complex words whose first element is a prefix”.

(32) OE stem stress:

scíp ('ship')
word ('word')
cýning ('king')
drincan ('to drink')
círice ('church')
súþerne ('southern')

Two observations seem in order here with regard to the assignment of primary stress in OE. On the one hand, in compounds the first element was always the strong one, just as in PdE (see Lass 1994:90; Halle and Keyser 1971:89).

(33) OE compounds:

méodo-hèall ('mead-hall')

On the other hand, OE prefixation was quite complex, with a considerable number of prefix doublets, where an old Germanic prefix could appear both stressed and unstressed (Lass 1994:92).

(34) OE prefixes:

Stressed prefix	Unstressed prefix
'æf-,þanca ('offence')	of- 'þyncan ('displease')
'an-,gin ('beginning')	on- 'ginnan ('begin')
'bi-,genga ('inhabitant')	be- 'gan ('occupy')

At a first glance, the principle at work seems to be that noun prefixes were stressed, while verb prefixes were unstressed. Nevertheless, Lass (1994:92) lists a number of significant exceptions to this conclusion, as we can observe in (35).

(35) Exceptions to (34):

Verbs with stressed prefixes	Nouns with unstressed prefixes
<i>'inn-,gangan</i> ('enter')	<i>be-'bod</i> ('command')
<i>'æfter-,swyrian</i> ('inquire after')	<i>for-'gifness</i> ('forgiveness')
<i>'and-,swarian</i> ('answer')	

As Lass (1994:92) explains, there were two types of prefixation in OE, namely (i) 'non-genuine' prefixation, resulting from (ia) an independent and, as such, stressed adverb (*inn, æfter*) joined to a noun or verb, or (ib) an initially-stressed noun which then becomes an initially-stressed verb (*'and-,swarian* < *'and-,swaru*), and (ii) 'genuine' prefixation, resulting from a verb with an unstressed prefix which then becomes a noun with non-initial stress (*for-'gifness* < *for-'gifan*).

If we look at the cases in (ia) above, where prefixed nouns were stressed initially, we will observe that certain noun prefixes behaved as though they were independent roots, so that the resulting nouns had the stress contour of compounds – that is, the first element was the strong one:

(36) *ánd-gièt* 'sense'

In spite of prefix-stress not being placed in accordance to the morphological category of the corresponding word in OE (see (35)), the Middle English (henceforth ME) period was characterised, among many other things, by the emergence of a group of prefixed verb-noun doublets which were indeed differentiated by the placement of the stress (37). As happened in the vast majority of the OE cases, in ME the noun carried the stress on the prefix, whereas the verb carried it on the root. To make matters even more complicated, this applied not only to Germanic words,

but also to French loans, and even to Germanic roots with Latinate prefixes – in most cases, the prefixes were no longer transparent in ME, and the prefix plus the root was taken as a stem. This situation was inherited in PdE.

(37) Verb-noun doublets:

Verb	Noun
<i>re-sít</i>	<i>ré-sit</i>
<i>im-pórt</i>	<i>ím-port</i>
<i>con-vért</i>	<i>cón-vert</i>

As one can observe in (33), (34), (35) and (36) above, OE showed several degrees of stress, crucially primary and secondary. As already mentioned, “primary stress was obligatorily lined up with the initial boundary of the lexical root” (Gasiorowski 1998:141).

Summarising, as Halle and Keyser (1971:97) point out, OE stress assignment responded to three main rules, namely the Initial Stress Rule – whereby the first syllable of the stem carried the main stress –, the Stress Retraction Rule – which retracted the main stress from that first syllable to the prefix in some derived nouns and adjectives – and the Compound Rule – by means of which the first element of a compound carried the main stress.

I will now highlight the most important characteristics of the change in stress assignment system from OE to PdE. The key difference between OE and PdE stress assignment is that OE stress rules refer to the left edge of the root – this means that stress fell on the first syllable of the stem –, while PdE stress rules work from the right edge of the word. That is to say, PdE stress is assigned starting from the last-but-one syllable in the word.²¹ Moreover, in OE the location of primary stress in polysyllabic words did not depend on syllable weight – OE primary stress was quantity-insensitive –, while it does in PdE. Lass (1994:93) summarises the OE Stress Rule as follows:

²¹ In PdE, like in Latin, final syllables are extrametrical unless they are the only syllable in the word, in which case they are accented by default.

- a) Starting at the left-hand edge of the word, look for a syllable bounded on the left by a major category label.
- b) Construct a binary foot s [strong] w [weak].
- c) Add at most one additional w.

The reason of the change in the stress assignment system of English – from left- to right-handed – is not clear. It could have been due to language contact; nearly all the PdE words that do not bear stress on the first syllable of the root were borrowed from Latin or Greek, either directly or via French.

In spite of these differences, OE and ME share several characteristics with PdE, something which favours the deep connection between the two systems. OE, just like PdE, builds bimoraic feet, that is, rhythmic units consisting of two weight units or morae.²² In order to support this idea, we can draw on three pieces of evidence. First, OE shows High Vowel Deletion (HVD) (see Lass 1994:98-99), a rule that deletes the ending -u (/u/) of the nominative and accusative plural of strong neuter nouns after a stressed heavy syllable or after a stressed light syllable plus another light – after a bimoraic foot.

(38) High Vowel Deletion:

	SINGULAR	PLURAL
	<i>scip</i>	<i>scip-u</i>
BUT	<i>word</i>	<i>word-∅</i>
	<i>werod</i>	<i>werod-∅</i>

HVD also affected the vowel /i/ in the same contexts, that is, after a bimoraic foot. After a light syllable, /i/ would remain, but would usually lower to /e/.

(39) preOE **win-ī* > OE *win-e* ('friend') (Lass 1994:98-99)

²² As pointed out in 1.4.3.4., a mora is a unit of sound that determines syllable weight. The syllable onset does never represent any mora. The syllable nucleus represents one mora – if it is a short vowel – or two morae – if it is a long vowel or a diphthong. The coda of a stressed syllable represents a mora, but it is not clear whether the coda of an unstressed syllable represents a mora or not.

As Lass (1994:100) explains, HVD reflected a weight-based constraint on foot structure; a heavy syllable + /i, u/ would be considered ‘overheavy’, while a light syllable + /i, u/ would be legal.

Second, like PdE, OE has a Minimal Word Constraint which implies that stressed light monosyllables are not tolerated in OE – this principle is applied to PdE in the next section.²³

Third, as explained in section 2.3.1., in OE verse a stressed heavy syllable is equivalent to a stressed light syllable followed by an unstressed syllable, which is called ‘resolution’;²⁴ both of those syllables together constitute a bimoraic foot.

The existence of bimoraic feet in both OE and PdE points to a historical relationship between the two systems. Before the change of the stress assignment system, there was an overlap between left-hand and right-hand stress. In the native vocabulary of OE, which stemmed from Old Germanic, no root was ever more than three syllables long. In ME, native words underwent several processes of vowel loss, in such a way that, by the end of the Early Modern period, most of the native Germanic words were either mono- or disyllabic. What this meant was that stress would fall in the identically whether the rules worked from left to right or from right to left. On the other hand, when Latin and Greek loans with right-hand stress started to flood the language, speakers were able to reanalyse the native vocabulary as a special case of right-hand stress.

The Latin Stress Rule (Lass 1994:88) was subsequently applied to a vast amount of Latinate lexicon:

Final syllables in Latin, unless they are the only syllable in the word, in which case they are accented by default, are extrametrical; that is, they are outside the domain of accent-assignment, ‘invisible’ to the accentuation rule. Therefore any disyllabic word, regardless of the weight of the first syllable, will be initial-accented. In polysyllables,

²³ OE is quantity-insensitive in that, in native polysyllabic words, stress is assigned to the first syllable of the stem regardless of its weight. Nevertheless, lexical monosyllables cannot be light, which means that, to a certain extent, stress and weight are correlated. Foot structure, on the other hand, seems to be weight-sensitive, too.

²⁴ In Lass’s (1994:101) words, “a sequence of two light syllables “resolves to” or is metrically/quantitatively equivalent to, one heavy one”.

accent the penult if it is heavy. Otherwise accent the antepenult, regardless of weight.

The Latin system was thus right-handed and quantity-sensitive, two characteristics which were diametrically opposed to the Germanic system.

There is controversy as to when the change from a left- to a right-hand system really happened. Minkova (1997) has shown that in ME verse French loans are predominantly stressed on the initial syllable – except line-finally in rhymed verse. In fact, those loans could be stressed either initially or finally, as required by the metre:

(40) *citée – cite; comfórt – cómfort; divérs – díverse; présént – préésent*

During the ME period, originally end-stressed loans from Norman French became nativised and received stress on their initial syllable, in accordance with the Germanic stress rules, something which indicates that etymologies were not always taken into account for prosodic matters.

(41) *ábbot, bárren, fórtune, hónour, mércy, témpet, vírtue*

As Halle and Keyser (1971:102-103) observe, in later stages stress came to be assigned according to the origins of the word – either Germanic or Latin; nevertheless, sometimes they were taken to have Germanic origin when they were Latinate, and vice versa, so the stress was assigned contrary to the rule.

To sum up, the stress assignment system of PdE is a hybrid between a native Germanic system, which was made complex during the OE period, and a borrowed Latinate system, which started to gain influence in ME and ended up becoming the major stress assignment system for PdE vocabulary. The Germanic heritage can be seen in monosyllables, where stress and weight correlate, in native disyllabic words, which preserve the stress as it was assigned in OE, even when the structure of the word changes, and in prefixed nouns and verbs which contrast in the placement of stress. The Latin heritage, on the other hand, can be seen in the English vocabulary

with Latin or French origin, where syllable-weight determines the placement of stress, and where stress itself is assigned from right to left.

We must be aware, however, that the evolution from the OE stress system to the PdE system was far from simple, and that there were numerous intermediate stages that contributed to shape the system as it is nowadays. The sixteenth, eighteenth and nineteenth centuries stand as important critical points as regard the interaction between the Initial Stress Rule and the Latin Stress Rule, on the one hand, and the Stress Retraction Rule, on the other.²⁵

2.3.2. Stress assignment in Present-day English

As already mentioned, stress has to do with the arrangement of syllables in rhythmic structures according to their relative perceptual prominence. Those structures are, in turn, organised hierarchically. Syllables are arranged into feet, where the stressed syllable is one that heads a foot; feet are gathered into higher-order metrical constituents such as prosodic words. Within a word, primary stress falls on the syllable that heads the foot that heads the prosodic word. For instance:

(42) *music*:

WORD	
FOOT	
SYLLABLE 1 - HEAD	SYLLABLE 2
STRESSED	UNSTRESSED
<i>mú</i>	<i>sic</i>

PdE is a stress language, that is, every lexical word – noun, verb, adjective or adverb – must carry a stressed syllable; in the case that more than one syllable carries stress, one of those stresses will be the main or primary stress, and the others will be subordinated to the former (Giegerich 1992:179) – primary stresses are signalled with an acute graphic accent on the nuclear vowel of the corresponding syllable, while secondary stresses are signalled with a grave accent.

²⁵ Given that the main topic of this dissertation is not stress in English, I am forced to leave the discussion here. Halle and Keyser (1971) give a thorough account of the intermediate stages in the evolution of English stress.

(43) *lésson, àbracadábra*

To a small extent, stress in English is phonemic. This is proved by the fact that there are some pairs of words that are segmentally identical but distinct in terms of stress placement. This correlates, in turn, to the morphology of each word in the pair; typically, one of them is a noun while the other one is a verb.²⁶

(44)

NOUN	VERB
<i>import</i>	<i>impórt</i>
<i>récord</i>	<i>recórd</i>

A syllable must satisfy certain structural requirements in order to be eligible to carry stress in English. First, stressed syllables must be heavy while unstressed syllables may be light or heavy.

(45) Weight and stress:

STRESSED	UNSTRESSED	
	HEAVY	LIGHT
<i>accórdion</i> (V: or VC)	<i>hónest</i> (VC)	<i>eróde</i> (V)

English is, therefore, a quantity-sensitive language, that is, stress is assigned according to syllable weight; weight is, in turn, closely related to vowel length and vowel tenseness. Those features are arranged into binary oppositions, in such a way that each of the terms in the oppositions would prototypically have a corresponding term in each of the other oppositions. Table 19 shows those oppositions, as well as the potential correspondences between each of them:

²⁶ The existence of these doublets goes back to ME, as we saw in the previous section (see (37)).

STRESS	Stressed	Unstressed
SYLLABLE WEIGHT	Heavy	Light
VOWEL LENGTH	Long (Short)	Short (Reduced)
VOWEL TENSENESS	Tense (Lax)	Lax

Table 19: Stress and its related features in English.

There is a set of phonological rules governing this relationship between stress, syllable structure and vowel quality, to which I will return in the next sections.

In English, stresses, syllables and feet are arranged according to a set of characteristics and rules. As pointed out in the last section, these rules were inherited directly from OE, and work for PdE as follows:

- a) The (bi)Moraic Trochee: As signalled for OE, PdE feet comprise two morae. It follows from this that an English foot can consist of either a heavy syllable, that is, a syllable with two rhymal X-positions, or two light syllables, that is, two syllables with one rhymal X-position. On the other hand, those feet are trochaic, that is, the first element in the foot is stronger than the second element, that is, the first element is the head of the foot (see (42)).
- b) Word Minimality: As observed for OE, PdE stressed monosyllables must be heavy. This follows from the fact that a syllable can only be stressed if it heads a foot, and an English foot requires two morae, as we have already seen, in order to be built.

When trying to summarise the rules of primary stress assignment in English, we come across quite an extensive number of different theories. One of the most important treatises is Halle and Vergnaud (1987), which devotes a chapter to the exploration of English word stress (1987:227-276). This volume inherits its theoretical and empirical apparatus from a previous treatise about English stress, co-authored by Halle and Keyser (1971) which, in turn, draws some of its theoretical principles from Chomsky and Halle's (1968) *The Sound Pattern of English* (SPE). As Halle and Vergnaud (1987:227) acknowledge, "[a] major empirical result of SPE was the discovery of the central role played in stress assignment by the contrast between 'strong' and 'weak' clusters – that is, between syllables with branching and non-branching rhymes". As mentioned on several occasions throughout this

dissertation, such contrast gives rise to the main rule of stress assignment in nouns, which bears a strong resemblance to the Latin Stress Rule (see section 2.2.1.1.). This resemblance was discovered by Chomsky and Halle (1968) and set a point of departure for contemporary treatises of English stress, which left behind the view that English stress was not predictable in order to assert the exact opposite.

The rule states that, for nouns, “main stress falls on the antepenult when the penult contains a non-branching rhyme – that is, when the penultimate syllable is light – and on the penult otherwise” (Halle and Vergnaud 1987:227).

(46) Cánada, agénda, marína, túna, hénna, alúminum, conúndrum, cerébrum,
póssum, vénom

Main stress in nouns does not normally fall on the last syllable because, as Halle and Vergnaud (1987:227) indicate, this syllable is extrametrical, that is, it is ‘invisible’ to the rules of stress assignment. However, there are several groups of exceptions to the rule of extrametricality, the first of which is formed by some nouns whose last syllable has a long vowel as its nucleus.

(47) políce, brocáde, bazáar, attaché, kangaróo

A further group of exceptions to the extrametricality rule – mostly words with secondary stress on the last syllable – proves more difficult to be accounted for.

(48) Berlín, ínsèct, decáthlòn vs. Lóndon, témpèst, súbject²⁷

There is a further rule at work in English, the so-called ‘Rhythm Rule’ (Halle and Vergnaud 1987:234-235), which retracts the original main stress of a noun from a final syllable with a branching rhyme to a previous syllable which originally carried a secondary stress. If there is no such syllable, stress is not retracted. The words to which this rule applies are lexically marked.

²⁷ Some varieties of English realise *súbject* as *sùbjèct*.

(49) désigné, anecdote, téléphone, jeopardise, elongate

Of the numerous studies of English stress that came after Chomsky and Halle (1968), Halle and Keyser (1971) and Halle and Vergnaud (1987), I focus on two specific theories, namely Giegerich (1992) and Gasciorowski (1998). These two theories are helpful for the purposes of this study because they do not resort to a complex system of theoretical tools in order to explain the basics of stress assignment.

The rules of primary stress in English nouns are explained by Gasciorowski (1998:143) as follows. Following Halle and Vergnaud (1987), primary stress in English falls on the rightmost non-final stressed syllable. The rightmost non-final stressed syllable must be heavy in order to be assigned primary stress; if it is light, the stress is retraced to the antepenult, regardless of its weight. Primary stress never goes beyond the antepenult.

(50) *Novémber, disáster, exámple, agénda*

There are, nevertheless, many exceptions to the above rule. The first exception is a group of trisyllabic nouns – not of Anglo-Saxon provenance – which carry stress on the first syllable – that is, on the antepenult, where the Latin rule would predict penultimate stress.

(51) *áncestor, cáelendar, sínister, mínister, cháracter, áadjective, órchestra, tálisman, prótestant, méssenger, lávender, cýlinder*

Stress on the first syllable is particularly regular on trisyllabic nouns ending in [i:] ~ [ɪ].

(52) *Nórmandy, mónarchy, énergy, tápestry, gálaxy*

These irregularities can be grouped under the following categories:

- In words with a light initial syllable, the second syllable becomes heavy due to the contact between [st] or [+son] and a homorganic consonant.
- The ‘energy’ class in RP (Received Pronunciation) shows a light second, because the /r/ is not pronounced. However, in rhotic varieties, where the second syllable is heavy, the stress falls on the first syllable, too.
- The rest render irregularly stressed derivatives.

(53) *ancestral, cylindrical vs. sinistral, ministrant, Protestantism*

- For words ending in *-ic*, stress falls on the penult, even when this is light.

(54) *periódic, melódic, phonétic, Mathemátics*

Exceptions to the rule are *rhétoric, cátholic, lúnatic*.

- The words underlined in (51) and (52) are truly irregular, so they must be lexically listed.

For most of the exceptions seen above, the oddity of the stress assignment to apparently light syllables could be solved by resorting to the concept of ‘ambisyllabicity’, whereby a consonantal segment working as the onset of an unstressed syllable can also function as the coda of the preceding syllable if this is stressed, thus rendering the latter heavy.

The second exception is a set of long words – four syllables – with the primary stress on the first syllable; these cases are truly exceptional.

(55) *álligator, cáterpillar, hélicopter, cáricature, dándelion*

There are various subclasses of these words:

- a) Greek loans in *-lepsy, -plexy, -doxy, -gogy, -archy, -thermy, -mancy, -urgy*.
- b) Latin loans in *-mony, -versy*.
- c) Idiosyncratically stressed words: *-ary, -ory, -ery*.

(56) *inventory, répertory, désultory, dýsenterý*

Following Gasiorowski (1998:145), the explanation for this is that the second syllable in these words is closed with [+son]; it can work as a light syllable and, therefore, it can be unstressed. The first syllable, on the other hand, is light in most cases, although, as happened with trisyllabic exceptions, it can be made heavy by means of ambisyllabicity.

The third exception is a mixed group adjectives and nouns with final primary stress.

(57)

NOUNS: *degré*, *police*, *duréss*, ADJECTIVES: *seréne*,
efféct, *boutique*, *ballóon*, *Carlisle*, *extréme*, *exáct*, *bizárre*, *innáte*,
Dundée, *Madrid*, *Berlín*, *grotésque*
enginéer, *kangaróo*

There are several explanations for the cases listed in (57) (Gasiorowski 1998:146):

- They end in ‘autostressed suffixes’ of French origin: *-ade*, *-aire*, *-eer*, *-oon*, *-esque*, *-esse*, *-ette*, *-ee*.
- Many are geographical names in which there is a tendency to keep the native pronunciation – French, Spanish, etc.

Summarising what we have seen up to now, stress in PdE nouns is always assigned to a heavy syllable. In the case of monosyllabic nouns, they must be heavy; in the case of polysyllabic nouns, they must contain at least one heavy syllable, or one of their light syllables must become heavy through ambisyllabicity. The rules of stress assignment for verbs are simpler. The major generalisation is that, for two-syllable verbs, the final syllable will receive main stress only if it has a long vowel; otherwise, the initial syllable will have the main stress.

(58) *achieve*, *emplóy*, *guffáw*, *ignite*, *maintáin* vs. *árson*, *finish*, *gíggle*,
lánguish, *mánage*

Giegerich (1992) summarises the rules of stress assignment in English nouns in accordance to the position of the stressed syllable, as well as the morphological category of the word in question:

I. Nouns bearing final stress:

- As a principle, there are no final-stressed words in English that end in a light syllable. On the other hand, no English word with final stress ends in a lax vowel. Among the disyllabic examples, secondary stress occurs on the first syllable only where that syllable is also heavy.

(59) *bàmbóo* vs. *ballóon*

- While nouns with final stress are comparatively rare in English, verbs and adjectives with final stress are quite common. The final syllable in all such cases is heavy; but unlike in nouns, heavy initial syllables do not necessarily have secondary stress.

(60) *obscúre*, **òbscúre*

- Sometimes, secondary and primary stresses invert their positions, but only when both syllables are heavy.
- As mentioned above, there are a number of noun-verb pairs in English that are distinguished by their stress patterns, the nouns having non-final and the corresponding verbs final stress.

(61)

NOUNS	VERBS
<i>dígest</i>	<i>digést</i>
<i>súrvey</i>	<i>survéy</i>
<i>tórmént</i>	<i>tormént</i>

II. Nouns bearing non-final stress:

- The penult is stressed if it is heavy; otherwise, stress falls on the antepenultimate syllable.

(62) *aróma, horízon, potáto*, with two rhymal X-positions occupied by vowels

(63) *enígma, agénda, sínópsis*, with two rhymal X-positions occupied by a vowel and a consonant. Exceptions (lexically listed): *bádminton, cáalendar*

(64) *cámera, cápital, vértebra*, with one rhymal X-position, so stress is retraced to the antepenult. Exceptions: *vanilla* or *madónna*, where the penult is light, so ambisyllabicity occurs to make it heavy

Giegerich (1992:189) observes that stress is not assigned only in accordance with phonological conditions, but also following an interaction of phonology and morphology. More specifically, he analyses the role of suffixation in the assignment of stress. In phonological terms, suffixes can be stress-neutral, in which case they never make any difference to the stress pattern of their base and are never stressed – inflectional *-s, -ing, -ed*; derivational *-less, -hood, -ly, -able, -ness* – or stress-shifting, in which case the stress pattern of the base may differ radically from that of the base with a suffix – derivational *-ic, -ity, -ian* – and can bear the main stress of the word.

(65) *-ee, -ette, -ese, -esque*, which are always heavy syllables

A suffix that shifts stress in some instances does not necessarily do so in all bases that it may attach to.

(66) *solémnity* vs. *divínity*

There are some cases where foot-level rules, which have to do with syllabic structure, interact and override word-level rules, which deal with morphological information. As an obvious example, heavy penults cannot be stressed if the final

syllable carries stress. In terms of foot-level rules, a foot aligned with the penultimate syllable has to be disyllabic, also including the final syllable. It cannot do that if the final syllable is itself a foot, like in (67).

(67) *níghtingàle, càvalcáde, chímpanzée*

Monosyllabic nouns always constitute heavy syllables; as we saw when quoting the Minimal Word Constraint, a monosyllabic lexical word cannot be light.

(68) *bit, camp*

At the level of the foot, feet are formed from right to left, and they can be bi- or trisyllabic; as a principle, every lexical word must have a foot. Monosyllabic feet do nevertheless occur on heavy penultimate syllables, but only where that syllable also happens to be the first one of the word.

(69) *bàmbóo, chàmpágne, Julý, ballóon*

Giegerich (1992:201) concludes by theorising about Foot Assignment in nouns, which works according to the following rules:

- a) Assign a foot to the final syllable if it contains a long vowel, or exceptionally, if it is otherwise heavy.
- b) Assign a disyllabic foot to the penultimate syllable if it is heavy.
- c) Assign a foot to the penultimate syllable if it is heavy and initial.
- d) Assign a maximal di- or trisyllabic foot to any remaining string of syllables from right to left, and ensure that the word has at least one foot.

These rules apply from right to left: first the final syllable is checked by rule a); then the penult is checked by b); and the default rule d) also assigns feet from right to left.

As seen in this section, English stress is assigned by following two series of rules, namely phonological rules and morphological rules – in the case of suffixes. Last, there is a group of words which must be listed lexically.

2.3.3. Stress and related features

In the following sections I analyse the interaction between stress and vowel length (2.3.3.1.), on the one hand, and between stress, quality and tenseness (2.3.3.2.), on the other. These two levels of interaction have important consequences on the assignment of stress in English and its classification as a stress-timed language.

2.3.3.1. Vowel length and reduction

It is an acknowledged fact that, in English, there is a correlation between stress and vowel length. Long vowels display a statistical tendency to be stressed more often than short vowels; on the other hand, most unstressed vowels tend to undergo reduction. According to Hayes (1995:15) and Gasiorowski (1998:134), stressed vowels are always full in English; whether the converse is true remains to be demonstrated. In fact, as Burzio (1994:112) observes, “stress is a necessary but not sufficient condition for vowel reduction, and [...] there exist a class of unstressed but unreduced vowels”.

(70) *Halicarnássus, condensátion*

In (70), we observe that the unstressed vowels in the syllables preceding the main stress are full, and do not reduce. This means that there are other factors, apart from stress itself, which contribute to reduction.

Before going on with this discussion, it is necessary to have a look at the vowel system of RP:

a. short	ɪ	e	æ	ʌ	ɒ	ʊ	ə
b. long	i:	ei	ai		ɔi		
	iu		au	ou		u:	
	iə	eə	ɑ:	ɜ:	ɔ:	uə	

Table 20: Vowel system of RP (Gasiorowski 1998:130).

The vowel quality that shows up quite commonly in vowel reduction is /ə/. Reduced /ə/ is not specified for any vowel features – except for minimal opening. As Giegerich (1992:68) emphasises, schwa is, in terms of its articulation, “neither high

nor low, neither front nor back. It is a vowel produced with a neutral setting of the articulators and is in this respect a ‘minimal’ vowel, involving, as it does, no displacement of the articulators from the neutral positions”. In terms of its occurrence, schwa can occur in all positions – word-initially, word-medially and word-finally – and all kinds of syllables – open or closed (71). If it appears in a monosyllabic, non-lexical word, it can be the only vowel in that word.

(71) /ə/ *táin*, *inf*/ə/(r)*mátion*, *ús*/ə/(r)

In spite of the predominance of schwa in unstressed positions, although all schwas are stressless vowels, not all stressless vowels are schwa. There are some other segments that can appear in unstressed position, as we observe in the list below (Hayes 1995:13):

- Never stressed: [ə], syllabic [n], syllabic [m]
- Variable: syllabic [r], syllabic [l]

[i, ou] / _____ #, / _____ V, and in prefixes.

[ɪ] / _____ ŋ

[yu] ~ [yə]

In contrast to the segments above, certain vocalic segments can never be unstressed:

- Always stressed: [eɪ, ε, æ, a, ɔ, ʌ, u, ʊ]

[i:, ou, ɪ] when not in the contexts above.

Going back to the segments in unstressed positions, where schwa can occur, no other vowel can, and vice versa. In other words, there are no minimal pairs contrasting schwa with any vowel other than /ɪ/. This is the reason why Giegerich (1992:69) points to the fact that, strictly speaking, schwa should not be called a ‘phoneme’ of English.

(72) *family* > *fám*/ə/ɹy, *fám*/ɪ/ɹy.

Let us go back now to the observation quoted above, made by Burzio (1994:112): “There exists a class of unstressed but unreduced vowels; stress is a

necessary but not sufficient condition to vowel reduction". This observation points to the fact that the correlation between stress and vowel length is not as transparent as one would predict. Indeed, there are several factors to take into account apart from stress, and most of those other have to do with the segmental characteristics of syllables.

First of all, there is a distinction between open and closed syllables. According to this, unstressed open syllables reduce quite generally, while closed ones do not.

(73)

OPEN (reduced)

América

párasi:te

ecónomy

CLOSED (unreduced)

próduct

dérelict

mánifest

The exceptions to those cases seem to be those syllables closed by sonorant or /s/, which frequently permit reduction.

(74) *information, carpenter, orchestrate*

The position of a syllable within a word also determines whether the nucleus in that syllable can be reduced or not. Below (following Burzio 1994:117) we compare final and internal positions:

I. Word-finally:

- a) If the unstressed vowel is checked by obstruents – except /s/ –, it remains unreduced.

(75) *pársnip, kídnap, lílac, Shérlock, nútmeg, kétchup*

- b) If, in turn, it is checked by sonorants and/or /s/, it gets reduced.

(76) *bálsam, útensil, ápron, sýllabus*

- c) There are some cases where the vowel is checked by sonorants and/or /s/, but the vowel remains unreduced.

(77) *cháos, méteor, wigwam*

II. Word-internally:

- a) If the unstressed vowel is checked by obstruents – except /s/ –, it remains unreduced.

(78) *áutopsy, architetónic, conductivity, enigmatic, hyperactivity*

- b) If the vowel is checked by sonorants and /s/, it gets reduced:

(79) *contemplation, affirmation, information, usurpation*

- c) Like in word-final syllables, sometimes a vowel checked by sonorants and /s/ remains unreduced.

(80) *incantation, detestation, exorcise, inculpate*

The distribution in both cases exhibits what Burzio refers to as ‘Arab-rule’ effects (1994:119). That is, both in word-final and in word-internal position, “if the syllable bearing stress is light, then the unstressed syllable is affected by reduction, regardless of the quality of the postvocalic consonant, thus precisely as in ar[ə]b, versus carn[æ]p” (Burzio 1994:119). Let us have a look of some examples where this happens.

- a) Word-finally:

(81) *bárrack, góssip, hávoc*

- b) Word-internally:

(82) *recognisable, resignátion, adaptátion*

There are, nevertheless, cases that violate this generalisation:

(83) *ad[æ]ptation vs. att[e]station, macr[v]n*

Finally, it is worth pointing that certain segments in clusters are more likely to license reduction than others. Ross (1972, in Burzio 1994:120) noted that coronal stops /t, d/ are unlike other stops in not inhibiting vowel reduction,²⁸

(84) *mýriad, inválid, périod, lilliput*

where unstressed vowels are reduced, contrasts with unreduced unstressed vowels in:

(85) *nóm[æ]d, nímr[v]d*

In any case, there is a difference between coronal stops and other stops which, aside from Arab-rule effects, generally block reduction. Final coronal stops also fail to inhibit reduction when they occur as the second member of a cluster that has a sonorant or s as its first. Words such as

(86) *élephant, élement, catalyst, cómfort*

with reduced unstressed vowels, contrast with

(87) *áardvark, ásterisk, ábelmosk (tropical plant), pódunk (small, isolated town) (Burzio 1994:121)*

On the other hand, stops always inhibit reduction when combined with /t, d/ (Burzio 1994:122).

²⁸ A stop is a consonant articulated by closing the airflow in the vocal tract. A coronal consonant is defined by being articulated with the blade of the tongue.

(88) *cátaract, insect, product, edict, concept, district*

All the cases of non stress-dependent reduction discussed here can be summarised in the rules that follow (Burzio 1994:124-125):

1) In the structure VCX, where X is not a vowel:

- a) Reduction of V is generally inhibited, because C requires vocalic support – not provided by X.

(89) *adirond[æ]ck, adj[e]ctival*

b) Reduction of V is permitted if either (i) or (ii):

- i) C has high intrinsic sonority, being either a sonorant or /s/.

(90) *ser[ə]ndipity, apr[ə]n*

ii) C requires a low sonority downstep – being /t, d/.

(91) *connectic[ə]t*

c) Reduction of V is forced – plus or minus idiosyncrasies – by ‘constant transition’, when the preceding stressed syllable is light – Arab rule.

(92) *hamm[ə]ck, rec[ə]gnition*

2) In the structure VC1C2X, where X is not a vowel:

- a) If C1 is relatively sonority-transparent, being a sonorant or s, then C2 is maximally licensed if it requires a low sonority downstep (/t/).

(93) *elephant*

and only minimally licensed if it requires a high sonority downstep (/p, k/).

(94) *podunk*.

Minimal licensing – as with clusters Cp, Ck – results in both of (i), (ii):

i) Relative infrequency.

(95) *podunk*

ii) Non-reduction of V – whose sonority is critical.

(96) *pod[ʌ]nk*

b) If C1 is relatively sonority-opaque, being a stop, then C2 is maximally licensed only if it either requires a low sonority downstep /t/.

(97) *abrup[t]*

or if it is intrinsically sonorous – sonorant, /s/.

(98) *particip[l]*

Clusters stop-p, stop-k are thus excluded.

The conclusion of this brief overview on vowel length and its determining factors is that stress and vowel length are intimately related, but this relation is mediated by a series of other variables, the most important one being the structure of the syllable where the unstressed vowel lies.

2.3.3.2. Vowel quality and tenseness

As seen in the previous section, not only does stress interact with vowel length, but these two also correlate with vowel quality or tenseness. In Halle and Keyser's (1971:4) words, "[t]he location of primary stress in a word is closely correlated with the distribution of tense vowels". In other words, stressed syllables in polysyllabic words have a tense vowel as their nucleus, while unstressed syllables can have either a tense or a lax vowel as nucleus.

Table 21 shows the inventory of English vowels and diphthongs with reference to quality:

a. lax	ɪ	e	æ	ʌ	ɒ	ʊ	ə
b. tense	i:	ei	ai		ɔi		
	iu		au	ou		u:	
	iə	eə	ɑ:	ɜ:	ɔ:	uə	

Table 21: Tense and lax vowels in English.

The correspondence between the parameters of length and quality is striking. All short vowels are lax, while long vowels are always tense. As Giegerich (1992:101) observes, quality differences normally have a quantity correlate in the vowels of English.

(99) /i:/ is [+tense, +long] and /ɪ/ is [-tense, -long]

In RP and GA (General American), [+tense] gives rise to [+long], while [-tense] gives rise to [-long], which means that the feature [long] is redundant.

If it is true that long vowels have to be stressed and that the corresponding syllables will be heavy, then one could say that there is a triple correlation between syllable structure, vowel length, and vowel quality.

SYLLABLE STRUCTURE	VOWEL LENGTH	VOWEL QUALITY
Heavy	Long	Tense
Light	Short	Lax

Table 22: Correlation between parameters.

As was the case with length, this correlation is not so clear-cut, and exceptions are indeed numerous. For instance, as Halle and Keyser (1971:13) observe, some dialects of English have the characteristic feature that most vowels in word-final position appear tense rather than lax:

(100) *pit[i]* (not [ɪ]), *válʃu* (not [jʊ])

This means that “phonetically tense vowels are not subject to reduction even when totally without stress” (Halle and Keyser 1971:29). In (100), we observe that the vowels in question lose a degree of length, but their quality remains the same, so they become ‘neutralised’.

As a sort of summary, Giegerich (1992:142) establishes a rule by means of which we can relate stress, length, and quality or tenseness in English.

Vowel-Length Rule (for RP and GA):

- a) Associate a [-tense] vowel with one mora. This means that the syllable in which such vowel can be a light syllable – if it is not followed by a consonant. Being so, it is not eligible to carry the stress of the corresponding word.
- b) Associate each element of a diphthong with one mora. So, a diphthong occupies two morae, and therefore, the syllable is heavy and eligible for stress.
- c) Associate a [+tense] vowel with two morae. The corresponding syllable is, again, heavy. It is eligible for stress.

2.3.4. Syllable structure and syllable weight

Syllable weight depends primarily on the length and quality of the syllabic nucleus, that is, the vowel. In turn, stress is assigned according to syllable weight. Hyman (1985:5) notes the importance of the syllable weight concept in determining the placement of stress in different languages.

As explored in the previous sections, English treats a syllable whose rhyme consists solely of a short/lax vowel as light, whereas a syllable whose rhyme has a long/tense vowel, a diphthong or a short vowel plus one or more final consonants is treated as heavy. Schematically, a -V rhyme defines a light syllable, and either a -V: or a -VC rhyme defines a heavy syllable. Giegerich (1992:146) explains syllable weight in terms of rhymal X-positions. According to this, stressed syllables have minimally two X-positions – that is, they are heavy –, while unstressed syllables may have a rhyme with a single X-position – that is, they may be heavy or light.

(101) *alliance*:

/ə	'lai	əns/
V-	-VV = V:-	-VCC = VC
light	heavy	heavy

English, like Latin, assigns stress in such a way that a heavy penult will always be stressed, while a light penult will be skipped to assign antepenult stress – regardless of the weight of the antepenult.

(102) *allocación* vs. *mérriment*

It follows from what we have seen up to now that lexical polysyllabic words must have, at least, one heavy syllable, while lexical monosyllables must be heavy; there are no English lexical words of the form CV, where V is a short vowel. Function words, on the other hand, may be light – with a heavy variant.

(103) *sea* (/si:/), *bit* (/bit/) vs. *the* (/ðə/, heavy variant /ði/); **bi* (/bi/)

It must be observed that, in disyllabic words, the afore-mentioned Arab rule operates: if the first syllable is light, the second must have a full vowel as its nucleus.

Summarising, syllable weight is “necessarily tied to the existence of a vowel length – or vowel tenseness – opposition” (Hyman 1985:6). As we have seen, heavy syllables characteristically attract stress, whereas light syllables receive stress only in the absence of an eligible heavy syllable.

2.4. Summary

Chapter 2 focuses on the analysis of the prosodic characteristics of Spanish and English speech prosody, in order to find out whether there are any rhythmic differences between the two.

Section 2.1. summarises previous research work carried out in relation to the study of prosodic typologies, paying especial attention to those theories devoted to the statuses of Spanish and English. The concept on which most prosodic

classifications are based is that of 'isochrony'. One of the first researchers to use the term was Pike (1945). The nuclear idea in Pike's theory is that languages are different depending on which elements are perceived as isochronic. The two main groups of languages are syllable-timed languages, characterised by the fact that syllables come at equal intervals in time, and stress-timed languages, where stresses occur at equal intervals. Other linguists who used isochrony as the basis for their studies were Jones (1950), who illustrated his statements with musical transcriptions, already pointing to the connection between speech rhythm and musical rhythm and Liberman (1979), who used musical notation for the development of the grid theory, among others.

In any case, the isochrony hypothesis has been contradicted by experiments which prove that physical isochrony does not really exist and that, rather than a measurable phenomenon, isochrony is a perceptual illusion. Despite the fact that the existence of isochrony is a controversial issue, even sceptical scholars such as Pointon (1980) admit that to the ears of the speakers of a language like English, Spanish sounds syllable-timed.

Given the lack of empirical support for the stress-timing vs. syllable-timing classification, scholars such as Dauer (1983) have developed the theory that speech rhythm is not a dichotomy but rather a scale or continuum, with two hypothetical languages, one perfectly stress-timed and the other one syllable-timed, marking the ends of the scale. Theorists who support the scalar view (Dauer 1983; Bertinetto 1989) claim that the rhythmic differences between languages result from a variety of components which interact in such a way that a language can be given a position in the continuum. Some of the components that determine the score of a language are vowel quality, consonant quality and quantity, syllable duration, compensatory shortening, syllable structure, relationship between syllable structure and stress, nature of stress, position of stress, intonation and tone.

In sections 2.2. and 2.3., a thorough account of the speech prosody of Spanish and English is given. First of all, I examine their historical evolution with regard to prosody, and then I concentrate on a specific theory of stress assignment for each of the two languages.

The evolution of the Spanish prosodic system stems from the Latin Stress Rule. Classical Latin showed a preference for stress to fall on heavy syllables in words of three or more syllables, while light ones qualified preferentially as non-stressable syllables. Latin is, thus, the best example of the connection between stress, syllable weight, and vowel length and quality, which could be summarised by claiming that the Latin vowel system had a direct impact on its stress system. In Classical Latin, the position of stress, which was fixed in principle on the antepenult, depended on the phonological distinction between long and short vowels, which in turn were mapped onto the distinction between tense and lax vowels and, therefore, on the heaviness of syllables. As soon as vowel length stopped being phonemic and became exclusively phonetic, stress itself became phonemic, as it is in contemporary Spanish.

In Spanish, the relationship between syllable structure and stress placement can be summarised into three main characteristics, namely that the vast majority of Spanish words end in a vowel, that the unmarked stress position for words that end in a vowel is on the penult and that the unmarked stress position for words that end in a consonant is on the ultima.

According to Harris (1983), the rules of Spanish stress placement could be summarised as follows. Nonverbal items can end either in a vowel or in a consonant. If they end in a vowel, stress is prototypically placed on the penult; if they end in a consonant, stress falls on the last syllable. Although exceptions to these rules are possible, and indeed numerous, stress will invariably fall on one of the last three syllables of the word. Spanish stress is not quantity-sensitive, an idea reinforced by Hayes's 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).

In D'Introno *et al.* (1995), stress is assigned according to two basic rules:

Rule A - *Closed Syllable Stress Rule*: The first closed syllable (counting from right to left) of the minimal word must be stressed.

Rule B - *Default Rule*: Stress the second syllable.

We could summarise the rules of primary stress assignment for Spanish nouns and adjectives as follows: as a general rule, stress can never be assigned to a fourth syllable from the end of the word; put differently, the three-syllable rule inherited from Latin is still at work. Primary stress is assigned to the minimal word; if that word contains any derivational morphemes, stress is assigned to the last one of those morphemes. Last, sliding is inapplicable if the minimal word has a derivational suffix which contains only one vowel.

Stresses at the level of the utterance in Spanish are not completely free, but rather dependent on the overall rhythm of utterances, which require a certain number of rhythmical 'leaning-points' at certain intervals. This entails that, at points where unstressed syllables are very numerous, there will be a need to increase the speed of discourse, which means that, in a way or another, certain processes of 'compression' must come into play. There are two compression devices in Spanish, synaeresis and synalepha, which have two corresponding 'decompression' devices, dieresis and hiatus.

In English, following the tendency of most Germanic languages, the primary stress of native words in OE was placed on the first syllable of the stem, with two observations: (i) in compounds the first element was always the strong one, just like in PdE and (ii) OE prefixation was quite complex, with a considerable number of prefix doublets, where an old Germanic prefix could appear both stressed and unstressed.

The key difference between OE and PdE stress assignment is that OE stress rules refer to the left edge of the root – this means that stress fell on the first syllable of the stem – while PdE stress rules work from the right edge of the word – that is to say, PdE stress is assigned starting from the last-but-one syllable in the word. Moreover, in OE the location of primary stress in polysyllabic words did not depend on syllable weight – OE primary stress is quantity-insensitive –, while it does in PdE.

The reason of the change in the stress assignment system of English – from left- to right-handed – is not clear. It could have been due to language contact; nearly all the PdE words that do not bear stress on the first syllable of the root were borrowed from Latin or Greek, either directly or via French.

OE and ME share several characteristics with PdE, something which favours the deep connection between the two systems. OE, just like PdE, builds bimoraic feet, that is, rhythmic units consisting of two weight units or morae. Second, like PdE, OE has a Minimal Word Constraint which implies that stressed light monosyllables are not tolerated in OE. Third, as explained in section 2.3.1., in OE verse a stressed heavy syllable is equivalent to a stressed light syllable followed by an unstressed light syllable, which is called 'resolution'.

During the ME period, originally end-stressed loans from Norman French became nativised and received stress on their initial syllable, in accordance with the Germanic stress rules. The Latin Stress Rule (Lass 1994:88) was subsequently applied to a vast amount of Latinate lexicon.

To sum up, the stress assignment system of PdE is a hybrid between a native Germanic system, which was made complex during the OE period, and a borrowed Latinate system, which started to gain influence in ME and ended up becoming the major stress assignment system for PdE vocabulary. The Germanic heritage can be seen in monosyllables, where stress and weight correlate, in native disyllabic words, which preserve the stress as it was assigned in OE, even when the structure of the word changes, and in prefixed nouns and verbs which contrast in the placement of stress. The Latin heritage, on the other hand, can be seen in the English vocabulary with Latin or French origin, where syllable-weight determines the placement of stress, and where stress itself is assigned from right to left.

To a small extent, stress in English is phonemic. This is proved by the fact that there are some pairs of words that are segmentally identical but distinct in terms of stress placement. On the other hand, stressed syllables must be heavy while unstressed syllables may be light or heavy.

Gasiorowski (1998) articulates a theory which can be summarised saying that stress in PdE nouns is always assigned to a heavy syllable. In the case of monosyllabic nouns, they must be heavy; in the case of polysyllabic nouns, they must contain at least one heavy syllable, or one of their light syllables must become heavy through ambisyllabicity. The rules of stress assignment for verbs are simpler. The major generalisation is that, for two-syllable verbs, the final syllable will receive

main stress only if it has a long vowel; otherwise, the initial syllable will have the main stress.

Giegerich (1992) summarises the rules of stress assignment in English nouns in accordance with the position of the stressed syllable, as well as the morphological category of the word in question. He also analyses the role of suffixation in the assignment of stress. In phonological terms, suffixes can be stress-neutral, in which case they never make any difference to the stress pattern of their base, and they are never stressed or stress-shifting, in which case the stress pattern of the base may differ radically from that of the base word with a suffix and can bear the main stress of the word.

According to Hayes (1995, in Gasciorowski 1998:134), full vowels are always stressed in English; whether the converse is true remains to be demonstrated. In fact, as Burzio (1994:112) observes, “stress is a necessary but not sufficient condition for vowel reduction, and [...] there exist a class of unstressed but unreduced vowels”. Burzio goes on to explain how, in spite of the predominance of schwa in unstressed positions, and although all schwas are stressless vowels, not all stressless vowels are schwa as well as how the position of a syllable within a word also determines whether the nucleus in that syllable can be reduced or not.

The correspondence between the parameters of length and quality is striking. All short vowels are lax, while long vowels are always tense. As was the case with length, this correlation is not so clear-cut, and exceptions are indeed numerous.

Syllable weight depends primarily on the length and quality of the syllabic nucleus, that is, the vowel. In turn, stress is assigned according to syllable weight. Hyman (1985:5) notes the importance of the syllable weight concept in determining the placement of stress in different languages. English, like Latin, assigns stress in such a way that a heavy penult will always be stressed, while a light penult will be skipped to assign antepenult stress – regardless of the weight of the antepenult.

3. Verse prosody

In the introduction, I outlined the view that rhythm is manifested on multiple levels – three of which are speech, verse and music – and that the characteristics of each of those levels affected the other two. In chapter 2, I explored the differences between English and Spanish at the level of speech rhythm. Chapter 3 analyses the verse prosody of Spanish and English, emphasising those elements which hold a direct relationship with speech prosody. Section 3.1. examines the relationship between speech prosody and poetic rhythm and metre before introducing some general characteristics of verse prosody in English and Spanish in connection to the typological classification examined in chapter 2. In sections 3.2. and 3.3. I analyse the verse prosody of Spanish and English, exploring their historical evolution and specific characteristics. Section 2.4. is devoted to summarising the chapter.

3.1. Introduction: poetic rhythm and metre: their relation to speech prosody. Analysis of verse prosody in English and Spanish

In section 1.2. I pointed out that rhythm is constituted by two hierarchies, namely metre and grouping. I also mentioned that, for reasons of terminology, I would be using the word ‘metre’ in its broadest sense, with the meaning of ‘any kind of patterning’, either consciously used and delimited by rules, as in the case of verse, or unconsciously used and loosely regulated, as in the case of normal speech. The two types of patterning are intimately related. In fact, as already mentioned, many theorists have suggested that the origins of some of the metrical patterns used in verse may be closely linked with the spoken language on which they are based, deriving not so much from abstract principles as from the prosodic patterns of actual phrases uttered in normal speech. As Abercrombie (1967:98) observes, “[t]he rhythm of everyday speech is the foundation of verse, in most languages”. We could go further and consider the idea that the relationship between language and metre is two-fold, that is, not only are metrical patterns ultimately founded on phenomena of ordinary speech, but also the manifestation of such patterns is in terms of speech. This close relationship between verse and normal spoken language has been commented upon by numerous writers, particularly in recent years. In T.S. Eliot’s

(1942:17) words, “[t]he music of poetry must be a music latent in the common speech of its time”.

The difference between ordinary speech and verse is that, when spoken language passes into poetry, it is no longer used for simple communication; rather, it is employed to create aesthetic effects. The aesthetic effects that lie at the core of poetic endeavour are never intended to be apprehended straightforwardly by the observer’s mind. On the contrary, poetic form – more specifically, poetic rhythm – challenges people’s minds with elements that interfere with the regular course of natural speech:

The more prosaic a discourse, the more it loses its singing stress pattern, and is simply articulated dryly. Poetry has the opposite tendency. Thus, it must create its own time sequence to make clear that it is a discourse with its own end [...] In this way the listener will be drawn out of reality and put into an imaginary time sequence; he will perceive a regular subdivision of sequences, a measure inherent in the discourse itself. Hence this marvellous phenomenon, the fact that in its deepest expression, when it is used as a game, language spontaneously loses its arbitrary character, which otherwise rules it firmly, and it now follows a law apparently alien to its contents. This law is measure, cadence, rhythm (A. W. Schlegel 1963:103-104, in Todorov 1988:17-18).

In order for poetry to achieve those intended aesthetic effects, its rhythms are put on exhibition, skilfully organised (Malof 1970:1). In sum, poetic language may be considered as a stylisation of ordinary spoken language.

As pointed out in chapter 1, the discipline that deals with poetic rhythm is known as ‘verse prosody’. Verse prosody is also called ‘metrics’, a term which is, in turn, used in two senses, meaning either the art of versification, or the study of the art of versification (Halle and Keyser 1971:140; Duffell 1999:8).

At this point it is necessary to make an important terminological clarification. In the field of verse prosody, the word ‘metre’ is used with a narrower sense than the one explained in the preceding chapters. In this second, narrower sense, the term ‘metre’ has two related meanings. On the one hand, it refers to the basic pattern

contained in a particular piece of verse, a notion linked to the idea that verse is often defined as the result of the matching between an abstract metrical pattern and language (Nespor and Vogel 1986:295). Modern metrics usually refers to this pattern as the 'template' of a piece. Both traditional and modern metrists use the technical term 'tension' for the phenomenon of variation from a template in verse (Allen 1973:110), pointing to the fact that, the more tension a pattern shows, the more complex it is (Halle and Keyser 1971:142). It is important to note that the abstract pattern of any poetic metre, that is, the template, can be violated to a certain extent before the reader perceives a breakdown in the system; in fact, the tension between the abstract pattern and the actual stress of the words provides variety and interest in the verse. The second meaning of 'metre' makes reference to the template plus a specified amount of tension. Jakobson (1960:365) introduced the term which best fits this definition of metre, namely 'verse design', which differs from the lines of verse, called 'verse instances', as well as from realisations of those, called 'delivery instances'.

In accordance with the narrow sense in which metre is used in the field of verse prosody, we can say that this discipline is primarily concerned with the identification of the different metres and forms of verse. The backbone of metrics – at least from a generative point of view – is to distinguish lines that are metrical from those that are not, and determine when verse can be called so. Thus, a grammar of metrics would be a set of rules that generated only structural alternatives that could be actually attested in verse texts and precluded those that are not (Duffell 1999:7).

Let us concentrate on the notion of 'metre' which refers to the template plus an added amount of tension. According to this notion, the metre of a line can be of three types, depending on which unit is taken to be the periodic element. The first type is 'syllabic metre', where the number of syllables in the line is the most important factor in order to consider it well-formed rhythmically. This metrical type corresponds to the versification of the so-called 'syllable-timed languages' in speech prosody. As explained in section 2.1., in these languages syllables are isochronous, so their distinct identity is never lost; equal timing of syllables naturally induces

syllable-counting in verse, which, in turn, becomes its timing.²⁹ The second type is ‘accentual metre’, which is normally used in the versification of stress-timed languages. These languages present variable syllable duration. As a consequence, versification does not rely on syllable-counting, but rather on the number of accented syllables in a line and the contrast between strong and weak stress as forming feet – a definition of ‘poetic foot’ is given in the next paragraphs. Verse metrists distinguish a third type, called ‘accentual-syllabic metre’. As happened with speech rhythmic typologies, accentual-syllabic metre is said to be a combination of accentual and syllabic metres, and pays attention to the more or less regular alternation of stressed and unstressed syllables in a line. This kind of metre would be used, although not necessarily, with the so-called ‘mixed-type languages’ in speech rhythm typologies, that is, those which are neither syllable, nor stress-timed, which would hypothetically occupy a middle position in the timing scale introduced in section 2.1.³⁰

As was the case with language typologies in reference to speech, I shall focus on English and Spanish as hypothetical models of accentual-syllabic metre and syllabic metre, respectively. I will concentrate on the methods of scanning poetry in these languages, in order to analyse the differences between the verse prosody of both, and then determine whether there is any correspondence between the rhythmic typologies of speech and those of poetry.³¹ This analysis will also serve to determine whether Spanish can be considered a syllable-timed language or a mixed-type language, in which case it would show features both of syllable- and stress-timed languages.

Historically, accentual metres were amongst the most common in Europe. OE verse is a clear example of accentual verse, but it also shows some peculiar characteristics. As Halle and Keyser (1971:147) explain, “the Old English alliterative verse had a specified number of vowels with primary stress, but [...] required, in

²⁹ As was highlighted in chapter 2, it is a widely accepted view that perfect isochrony can never be reached in production; in other words, isochrony is a perceptual phenomenon.

³⁰ The correspondence between speech timing and metre is not always perfectly univocal. In stress-timed languages, the timing of speech can correspond either to an accentual metre or to an accentual-syllabic metre. In the case of English, while its speech rhythm became more and more stress-timed, its metre evolved from accentual – in OE poetry – to accentual-syllabic – Chaucer, influenced by Italian and French poetry, was the first to use accentual-syllabic metre, in the fourteenth century (Adams 1997:5).

³¹ To scan a line or passage is to determine – by noting the pattern of stressed and unstressed syllables – its normative, or typical, metrical pattern and deviations from that pattern.

addition, that the consonant clusters preceding certain stressed vowels in the line alliterate".³² Apart from alliteration, another characteristic of OE verse was that each line was composed of two-half lines, separated by a pause or 'caesura'. From Chaucer on, English became more accentual-syllabic, while the speech gradually turned more stress-timed (see Bertinetto 1989).

The classical approach to English metre inherits from the Greco-Latin tradition the notion of the 'foot' as the fundamental rhythmical unit, and consequently, the minimal unit of scansion.³³ Being a measurable, patterned unit of poetic rhythm, the English foot is customarily defined as a measure of rhythm consisting of one accented or stressed syllable and one or more unaccented or unstressed syllables, matching – however uncomfortably, but not equating, the 'long' and 'short' of the classical originals (Attridge 1982:6-7).³⁴ Although stress and length are not synonyms in English, some scholars, like Abercrombie (1967), have called attention to the fact that syllable quantity in English is dependent on three factors, namely the presence of stress, the phonetic structure of the syllable, and the position of word boundaries. English is 'quantitative' in the sense that some syllables are perceived to be longer than others.³⁵

Abercrombie (1973:11[1964], in Couper-Kuhlen 1986:55) defines 'foot' as "the space in time from the incidence of one stress-pulse up to, but not including, the next stress-pulse". This definition points to the fact that, in addition to being conditioned by the principles of stress-timing and isochrony, English rhythm is characterised by the principle of rhythmic alternation.

One of the most important characteristics of feet is that, within the same poem, they are assumed to be isochronous; nevertheless, as happened in normal speech, this assumption has been showed to be, rather than the norm, the exception. The reason for this lack of isochrony is that the lines of every poem consist of feet, which in turn consist of words and, within them, syllables, which in their turn contain

³² The native OE line was an accentual line which consisted of four beats with a caesura, bound together by alliteration (Adams 1997:37).

³³ Some metrists do not agree with the conception of the foot as the basic unit of scansion in English (see Giegerich 1985; Hogg and McCully 1987). This topic would deserve a deep discussion on its own, which I am forced to leave for future research.

³⁴ We must take into account that, in English, a strong syllable must be stressed, but a non-strong syllable may be stressed or unstressed.

³⁵ It is necessary to emphasise that long and/or heavy syllables are not always perceived as (primary) stressed, which leads to huge fuzziness in all the commonly-accepted definitions.

phonemes. As explained in chapter 1, measurements have shown that deviations from strict foot isochrony depend precisely on the number of syllables in a foot, and that deviations from syllable isochrony are related to the number of segments in a syllable. However, and this is what makes the above assumption work, I emphasised that, in perception, such deviations often tend to be disregarded and, as a result, spoken verse, like ordinary speech, tends to be heard as more rhythmical than it really is.

The names of the four basic types of English feet are inherited from the classical tradition. They are called 'iamb' (x /), 'trochee' (/ x), 'anapaest' (x x /) and 'dactyl' (/ x x) (Hamer 1966:9; Malof 1970:29).³⁶ Apart from those four basic metrical feet, there are two further ones which can serve as rhythmic variations within the line, but never as bases; they are called spondee (/ /) and pyrrhic (x x) (Adams 1997:10-11)³⁷.

According to English 'classical' verse prosody, metrical patterns are named according to the predominant type of foot and the number of feet per line. Thus, a line of 'iambic pentameter' consists of five iambic feet. Similarly, we may speak of 'trochaic tetrameter', 'anapaestic trimeter', 'iambic hexameter' and so on. The iambic pentameter has been, since the time of Chaucer, the favourite of English poets. Its pattern allows for a great deal of freedom, although it also presents some important constraints (Halle and Keyser 1971:171). Halle and Keyser consider it the only basic type of foot in English verse, while the other feet are taken to be allowed deviations from the iambic pattern.

The stanza, in turn, is a sequence of lines arranged in a definite pattern of rhyme and metre, which makes up a structural unit repeated throughout the work. To mention some of the most characteristic stanzas in English poetry, two successive rhymed lines are a 'couplet', and if they are in iambic pentameter and the sense of the pair is closed, that is, it does not carry over into the following line, a 'heroic couplet'. Four lines rhymed *abba* or *abab* make up a 'quatrain'; three quatrains followed by a couplet, all in iambic pentameter, make up the conventional

³⁶ I must highlight the fact that in Greek poetry, these names corresponded to quantitative patterns. English verse prosody has adopted Greek names, but they do no longer refer to quantity; rather, they refer to stress patterns, which makes them somewhat misleading (see Hamer 1966:10).

³⁷ Malof (1970:37) argues that these variant feet, also called 'cadences', are "perhaps unnecessary", because they tend to obscure, rather than illuminate, metrical patterns.

‘Shakespearean sonnet’. A ‘Petrarchan sonnet’ comprises an octave of two quatrains rhymed *abba* and a sextet with various arrangements of additional rhyme sounds – for instance, *cdcdcd* or *cdecde*. For instance, the ‘ballad stanza’ contains four lines, the first and third being unrhymed in iambic tetrameter and the second and fourth rhymed in iambic trimeter.³⁸

The best way of illustrating all the concepts explained up to here is presenting the scansion of a poem written in contemporary English, and then discussing its metrical characteristics. For this purpose, (104) marshals a scansion of Rossetti’s *Song* – adapted from Adams (1997:21). This analysis is aimed at serving as an illustration for the concepts explained above; obviously, it is impossible to draw general principles from a single case study:

(104)	
When I am dead, my dearest,	oó, oó, oó(o)
Sing no sad songs for me;	oó, oó, oó
Plant thou no roses at my head,	óo, oó, oò, oó
Nor shady cypress tree:	oó, oó, oó
Be the green grass above me	óo, oó, oó(o)
With showers and dewdrops wet;	oó(o), oó, oó
And if thou wilt, remember,	oó, oó, oó(o)
And if thou wilt, forget.	oó, oó, oó
I shall not see the shadows,	oó, oó, oó(o)
I shall not feel the rain;	oó, oó, oó
I shall not hear the nightingale	oó, oó, oó, oó
Sing on, as if in pain:	oó, oó, oó
And dreaming through the twilight	oó, oó, oó(o)
That doth not rise nor set,	oó, oó, oó
Haply I may remember,	óo, oó, oó(o)
And haply may forget.	oó, oó, oó

³⁸ An alternative analysis of the ballad stanza is to regard lines two and four as instances of silent beats.

The poem in (104) is composed using a song-like stanza which is an extension of the ballad stanza, constituted by a quatrain of variable metre – in the first stanza, followed by a quatrain of iambic trimeter, rhyming – *a* – *a*. Both stanzas conclude with a sort of refrain, playing with the antithesis between ‘remember’ and ‘forget’. As we can observe, there are some lines (1, 5, 7, 9, 13, 15) which contain an added unstressed syllable.³⁹ The first quatrains in each stanza are a good illustration of the different types of feet used in English poetry. The basic pattern of the poem, that is, the template, corresponds to the iambic foot (oó, oò) combined in groups of three feet, thus yielding an iambic trimeter; nevertheless, there are also some instances of trochaic (óo), and spondaic feet (óó), which work as variations of the basic iambic design. Furthermore, while the first, second, and third lines of the first quatrains in both stanzas are trimeters, the third one is a tetrameter in both cases. Apart from this, the arrangement of the stresses within the lines does not always respond to the stress of the words in the language – see, for instance, the stress on ‘when’ in line 1, ‘at’ in line 3, or ‘if’ in lines 7, 8. This confers a certain degree of tension and complexity to the poem, as well as variety and interest. In any case, what seems to be clear is that, taking the foot as the basic metrical unit of scansion, the poem is metrically arranged in such a way that a sense of isochrony is perceived. Each foot contains a main stress and a non-stress, and the number of syllables in each foot is two in most of the cases.

From the analysis of this case study, we can conclude that the stress-timed rhythm of PdE is the basis of the structure of contemporary English verse.⁴⁰

Turning now towards Spanish verse prosody, most modern reference works in metrics for the general reader state categorically that Spanish verse is syllabic. There are three main reasons for believing that the essential unit of scansion in most Spanish verse is the syllable and not the foot, namely that much Spanish verse has a perfectly regular syllable count, that there is a tendency in modern Spanish for syllable timing to predominate over stress timing and that vowel reduction does not occur in Spanish (Duffell 1999:23-24). In spite of all that, Duffell (1999:45) argues that, from a diachronic point of view, Spanish has an accentually-based metrics that

³⁹ This phenomenon is also known as ‘hypercatalexis’ (Malof 1970:41).

⁴⁰ Some late twentieth century poets, like Marianne Moore, W.H. Auden, Dylan Thomas, or Sylvia Plath, carried out experiments using syllabic verse in English. Nevertheless, in general terms it does not work for English poetry (Adams 1997:64).

turned into syllabism due to the influence exerted by French literature and literary theory on the poets of other Romance languages. According to Duffell (2000:113), verse became rigidly syllabic only in the middle of the fifteenth century, when a Spanish poet, Marqués de Santillana (1398-1458), first articulated the norm that skilled poets, writing for an educated public, had to count syllables. Santillana is said to be the first poet who brought to perfection in Spanish the two types of syllabic verse for which he had French models: short-line metres – those with lines shorter than nine syllables – and composite metres – those with lines of ten or more, divided by a fixed caesura. As a consequence, over the last six hundred years it has become firmly established that Spanish versification is syllabic, because “generations of cultivated poets and audiences have observed the aesthetic code urged and exemplified by the medieval poet” (Duffell 2000:125). So French syllabism has prevailed in Spanish for social reasons.⁴¹

Traditionally, in order to scan a verse written in Spanish, three factors must be considered, namely the final word of the line, synaeresis and synalepha, and hiatus and dieresis. In what follows I shall explain each of those factors. The first factor has to do with the final word of the verse, which determines whether the verse is *verso agudo* (‘oxytone line’), *verso llano* (‘paroxytone line’) or *verso esdrújulo* (‘proparoxytone line’).⁴² An oxytone line ends with an oxytone word, that is, one which is accentuated on its final syllable. Spanish verse prosody prescribes that this syllable counts as if it consisted of two syllables, as happens in (105).

⁴¹ I am aware that this assertion deserves further research and discussion. Nevertheless, for the purpose of this dissertation, suffice it to say that contemporary Spanish verse is syllable-timed.

⁴² As explained in section 2.2.2., depending on the syllable that carries stress, Spanish words can be oxytone, paroxytone or proparoxytone. This means that the stress falls on the last syllable, on the last but one or on the last but two or three syllable, respectively. With the sole exception of the adverbs finished in *-mente* (the equivalent to the English suffix *-ly*), Spanish words can have only one primarily stressed syllable.

(105)

Caminante, son tus huellas (8 syllables)

‘Walker are your tracks

el camino, y nada más, (7+1 syllables)

the path and nothing else

caminante, no hay camino, (8 syllables)

walker not there is path

se hace camino al andar (7+1 syllables)

it is done path while walking’

(Antonio Machado)

A paroxytone line ends with a paroxytone word, that is, one which is accentuated on its penultimate (106). In this case, which is the most common one, syllables are neither added nor subtracted from the ‘llano’ verse.

(106)

No te quiero sino porque te quiero (11 syllables)

‘Not I love you but because I love you

y de quererte a no quererte llego (11 syllables)

and from living you to not living you I get

y de esperarte cuando no te espero (11 syllables)

and from waiting for you when I not wait for you

pasa mi corazón del frío al fuego. (11 syllables)

goes my heart from cold to fire’

(Pablo Neruda)

A proparoxytone line ends with a proparoxytone word, that is, one which is accentuated on its antepenultimate syllable (107). On counting its syllables, the last word loses a syllable. This is the least frequent of the three, something which is in accordance with the fact that, in Spanish, proparoxytone words are much less frequent than paroxytones and oxytones, in that order.

(107)

Adoro la hermosura, y en la moderna estética (15 - 1 syllables)

'I adore beauty and in modern aesthetics

corté las viejas rosas del huerto de Ronsard; (13 +1 syllables)

I cut the old roses of the orchard of Ronsard

mas no amo los afeites de la actual cosmética, (15 - 1 syllables)

but I not love the makeups of current cosmetics

ni soy un ave de esas del nuevo gay-trinar. (13 +1 syllables)

nor I am a bird of those of the new happy-birping'

(Antonio Machado)

The second factor to be taken into account when scanning Spanish verse has to do with the phenomena called 'synalepha' and 'synaeresis'. As explained in section 2.2.3., synalepha, a very normal and usual device in Spanish verse, is the union of two or more contiguous vowel sounds between two, sometimes three, different words. Synaeresis is the union of two vowels in the same word which are normally pronounced as two syllables.

(108)

Ella dio al desmemoriado (8 syllables due to the signalled synalepha)

'She gave the forgetful one

una almohadilla de olor. (8 syllables due to the signalled synaeresis

A cushion of odour' and synalephas)

(José Martí)

The third factor has to do with hiatus and dieresis, which are the opposites of synalepha and synaeresis, respectively. Hiatus is the separation of contiguous vowels of different words (109). Dieresis is the separation of vowels within a single word, which divides an existing diphthong into two syllables. It is usually signalled with two dots placed above the weaker vowel of the affected diphthong (110).

(109)

el que | a mi canta va la missa cada día (14 syllables due
'the one who to me sings goes to church everyday' to hiatus)

(Gonzalo de Berceo)

(110)

La catedral es un gran relicario. (11 syllables)

'The catedral is a big shrine

La bahía unifica sus cristales (11 syllables)

the bay unifies its glass pieces

en un azul de arca_̄icas mayúsculas (11 syllables due to dieresis)

in a blue of archaic capitals

de los antifonarios y misales. (11 syllables)

of the antiphonies and missals'

(Rubén Darío)

Summarising, in order to scan the number of syllables in a verse-form used in a particular poem in Spanish, the syllables of the first line must be counted without taking into account any of the above phenomena. Second, the number of syllables must be adjusted by considering whether the line is oxytone, paroxytone or proparoxytone, and whether synalepha, synaeresis, hiatus or dieresis (or several of those) affect the syllable count. When the syllable-count of the 'model' line has been determined, the other lines can be scanned to yield the same number of syllables, also taking into account the above phenomena. If, after making use of these poetic licenses, a correct syllable-count is still not achieved in any of the lines, that line is said to be 'irregular' or 'defective', depending on whether the irregularity is or is not intentional.⁴³

Spanish names of lines are determined by the number of syllables they contain. The most commonly used names are *heptasilabo* ('heptasyllabic') – seven syllable lines –, *octosilabo* ('octosyllabic') – eight syllable verses –, *endecasilabo*

⁴³ In general, much poetry establishes a normative pattern, but does not adhere to it mechanically; some lines will deviate from the norm as an aesthetic or rhetoric device.

(‘hendecasyllabic’) – eleven syllable verses – and *alejandrino* (‘Alexandrine’) – fourteen syllable verses. Verses from two to eight syllables are classified as *versos de arte menor* (‘minor art lines’), whereas verses of more than eight syllables are classified as *versos de arte mayor* (‘major art lines’).

Although the rhythm in Spanish poetry is basically syllabic, when a verse is pronounced accents occur at certain intervals. Let us have another look at (105), reproduced here as (111), in order to find out whether this can be proved:

(111)

Caminante, son tus huellas

El camino, y nada más.

Caminante, no hay camino,

Se hace camino al andar.

In the above example, the underlined syllables are stressed, whereas the rest are only weakly stressed or not at all. As we can see, stresses fall at equal distance in all the lines except for the last one, where there is one extra interstress syllable. Obviously, this recurrence of equally-spaced stresses confers a highly rhythmical character to the poem when it is read aloud. Nevertheless, in Spanish poetry, 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 – in this case, three.

Generally speaking, all lines require a rhythmic accent on the penultimate syllable. Moreover, lines of five or more syllables have one or more auxiliary accents. For instance, a heptasyllabic line (*heptasilabos*) has an accent on the sixth syllable and at least one more on any syllable but the fifth one; an octosyllabic line (*octosilabo*) has an accent on the seventh syllable and at least one more on any syllable but the sixth one; with a hendecasyllabic line (*endecasilabo*), apart from the obligatory accent on the tenth syllable, various distributions are possible: 1,4,7; 2,6; 1,6; 4,6; 4,8; 4,6,8; an Alexandrine (*alejandrino*) is a line of fourteen syllables divided into two hemistichs of seven syllables each, and each hemistich

(*hemistiquio*) follows the rule for heptasyllabic lines (112). For this reason, some metrists argue that it is plausible to assume the existence of a fragmentary foot structure in Spanish, but this has not been proved properly.

(112)

Dichoso el árbol // que es apenas sensitivo,

'happy the tree which is barely sensitive

y más la piedra dura, // porque ésa ya no siente;

and more the stone hard because that one anymore not feels

pues no hay dolor más grande // que el dolor de ser vivo,

as not there is pain bigger than the pain of being alive

ni mayor pesadumbre // que la vida consciente.

nor bigger sadness than the life aware'

(Rubén Darío)

Summarising, the main findings point to Spanish verse as basically syllable-timed. This is supported by constraints on poetic structure, that is to say, a poem is constructed using a fixed number of syllables in each line.⁴⁴ Syllable count is thus the major determinant of how a lyric poem is structured. However, it might be the case that, in the phonetic realisation of the syllables in spoken poetry, the phonology of Spanish, including metrical constraints, interacts with a higher level syllable-timing constraint similar to an isochronic stress. In any case, it seems plausible to say that the rhythm of English verse is really different from that of Spanish verse, a conclusion which reinforces the idea that English and Spanish are two languages with different prosodic systems.

3.2. Verse prosody in Spanish

This section gives an overview of the historical evolution and characteristics of the Spanish verse prosodic system. Section 3.2.1. examines the prosodic characteristics

⁴⁴ Navarro Tomás (1944) highlights the fact that each language has a preference for a specific number of syllables in its speech utterances. In the case of Spanish, the preference range is in between five and ten syllables, with a predominance of the utterances formed by seven or eight syllables. Taking this into account, it is not surprising that seven or eight syllable lines or hemistichs have always been predominant in the history of Spanish poetry (Sosa 1999:37).

of Classical Latin verse as well as the changes that it underwent parallel to the evolution of the language, until the various Romance languages and prosodic systems emerged. In sections 3.2.2., 3.2.3. and 3.2.4. I explore the ruling rhythmic principles of Spanish poetry, namely isosyllabism, syllable counting and its related devices and stress placement.

3.2.1. Introduction: from Latin quantitative verse to Spanish syllabic verse

In neolatin languages like Spanish, there exists a common metrical basis derived from the versification system used in Medieval Latin, which was, in turn, inherited from the rules of versification in Classical Latin. Classical Latin used a model of quantitative verse, according to which lines were designed in accordance to patterns of long and short syllabic values or, more accurately, vocalic values. Quantity in Latin often correlated with syllable weight, and the weight of a syllable determined whether it could be stressed or not: in this way, heavy syllables were normally stressed, while light ones were prototypically unstressed. This did not mean that a heavy syllable had to be stressed necessarily, but, rather, that the syllable in question qualified as stressable when language was put into verse; on the other hand, some heavy syllables which were stressed when the corresponding word was taken in isolation, lost their stress when the poetic composition required it. Let us have a look at the following poetic fragment (Horace, in Gasparov 1996:86), glossed and translated in (3) in section 2.2.1.1.:

(113)

Ínteger ví-tae / scelerísque púrus

Non éget Máu-ris / iáculis neque ár-cu

Nec vene-ná-tis / grávida sagít-tis

Fúsce, pharét-ra...

In Classical Latin, the rhythm marked by the stresses determined how a poem would be read; however, the real metre corresponded to the arranged patterns of longs and shorts. We could say that in authentically quantitative languages like Latin, quantity plays a crucial role both in speech and verse prosody, while in languages such as English, where quantity is phonemic and is, therefore, a key element in

speech prosody, quantity does not determine stress and, in fact, gets relegated in favour of the latter when it comes to versification. As a consequence of this, conflicts between prosodic stress and lyric stress are observed as artistic devices in the former, while in the latter they become very violent. In the latter, those conflicts get relegated to a fairly limited number of widely accepted poetic devices.

As regards poetic devices proper, synalepha – more specifically, elision – was an extended practice in Classical Latin verse, especially when a final vowel, or a vowel followed by *-m*, appeared before a word beginning with a vowel – or *-h*. See (114), taken from Virgil’s *Aeneid* (in Raven 1965:27):

- (114) *conticuer(e)* *omnes intentiqu(e)* *ora*
 Shut up-past, 3rd pl. all-nom.intentus-nom. pl. chin-acc. pl.
tenebant
 have-past, 3rd pl.
 ‘They all shut up and paid attention with their hands on their chins’

As would happen later in Spanish, this practice was particularly common in the colloquial verse of comedy, although Virgil also practices it; it was less common in lyric poets. Elision could even happen between lines – we will see that this practice would become common in the fifteenth century in Spain. This was called ‘hypermetric elision’, and can be observed in the following fragment from Virgil’s *Georgicas* (in Raven 1965:27):

- (115) *Si non tanta quies iret frigusque*
 if not so much peace-nom. go-imp.sub. 3sg.cold-acc.
 and
 ‘If there wasn’t so much peace and
caloremqu(e)
 heat-acc.
Inter,
 Between,
 between the cold and the heat,’

Latin poets applied the same process in reverse – a device called ‘aphaeresis’ – with the word *est*, which would lose its initial *-e* after a final vowel or *-m*.

(116) *usa (e)st* (‘she uses’) (in Raven 1965:28)

The opposite process, hiatus, could take place when a final vowel or an *-m* remained present before another vowel. Hiatus was common in Homer, and some Latin poets used it, as Virgil in his *Georgicas* (in Raven 1965:28). In any case, as happened with elision, hiatus was more common in comedy than in lyric verse.

(117) *Glauco/ et Panopeae/ et Inoo Melicertae.*

All this shows that the most common devices used in Romance languages such as Spanish, were already common in Classical Latin.

The transition from a quantitative system to a syllabic one started around the third century, as a consequence of some major changes in the language itself, which involved the loss of the opposition between long and short vowels. For a period of time, the two metrical systems cohabited. On the one hand, in spite of the phonological changes in the language, the system of quantitative metrics survived through the Middle Ages and the Renaissance, thanks to some cultivated authors who continued composing verse based on the old opposition between long – heavy – and short – light – syllables. On the other hand, some poets of the Christian Church, such as Saint Augustine, started to compose hymns that could be understood by non-cultivated people; those hymns were the first instances of syllabic verse written in Medieval Latin (Gasparov 1996:88ff.). At this time, two terms appeared to differentiate the two systems: the old quantitative system was called ‘metrical’, while the syllabic system acquired the name ‘rhythmical’. As the syllabic system became generalised, other devices were introduced to compensate for the loss of syllabic quantity; among those, the most salient device was rhyme (Gasparov 1996:97).

In Romance languages, quantity does not play any prosodic role. Correspondingly, these languages have not produced any instances of quantitative verse. In fact, the essential elements in Romance verse have come to replace the role

of the quantitative combinations of classical metre. Those elements consist, on the one hand, of a formal disposition of the number of syllables, rhyme and stanza and, on the other, of the ordering of the rhythmic markers.

In spite of their common roots, the attitude of Romance languages has not entirely coincided with regard to the development and benefiting from the Latin inheritance and its subsequent transformations. If we focus on specific languages, the French system was the earliest consolidated Romance system – the first works were written around the tenth century –, something which explains that its technique became highly elaborated throughout the years. The Spanish system was consolidated later, in the twelfth century, and was subsequently enriched specifically by means of the cultivation of rhythm – see, for further details, Pensom (1998). According to Gasparov (1996:119-120), the evolution of the French and Spanish systems, which run parallel to the Italian system, can be related to three stages of increasing strictness. In the first stage, which coincides with the beginnings of the Middle Ages, there was a vacillation as to what system, accentual or syllabic, would be adopted after the loss of length contrasts. The second stage was characterised by the appearance of syllabic poetry rooted in medieval Latin verse, and the use of full rhyme. In the third stage, around beginnings of the Renaissance, the syllabic system and the use of full rhyme was consolidated, and blank verse appeared.

By looking at the beginnings and evolution of Romance verse, we can conclude that the character of Romance verse seems to have evolved in intimate connection with the phonological conditions of each language. Thus, for Spanish, given that there is no length contrast in the phonological system any more, syllabic quantity does not play any role in the structure of modern Spanish verse. Rather, verse shape and limits are determined by means of the number and combination of syllables in the line, the distribution of accents, and, less importantly, the placement of pauses.

3.2.2. The principle of isosyllabism in Spanish verse

The two main principles that govern both the composition and scansion of Spanish verse are syllable counting, on the one hand, and rhythmic accent placement, on the other. This means that the virtually perfect correlation between the timing of language – syllable-timing in the case of Spanish – and the type of verse in the

corresponding language – for Spanish, we would predict syllabic verse, that is, verse that relies exclusively on the counting of syllables per line – is not always univocal or, at least, not totally so.⁴⁵ In spite of the existence of clear exceptions to the quoted law of a perfect correlation between language and poetry, there are languages in which this is indeed the case, that is, some syllable-timed languages are rendered into poetry using syllabic verse – for instance, Modern Greek.

Syllabic verse is governed by rules of composition and scansion which take the syllable as their rhythmic basic unit. Lines are measured according to the number of syllables which they contain, and there is a strong tendency for them to be made up of a regular number of syllables. This essential principle is called ‘isosyllabism’.

While some languages, such as French and Italian, adopted the isosyllabic system from the early days of their literary tradition, some others, like Spanish, show a partial reluctance to accept the isosyllabic rule totally. Poetry written in Spanish seems to follow two clearly different paths, depending on the tradition from which it stems. According to the first tendency, which is clearly strong and pervading in most cases, lines are isosyllabic. The second tendency comprises non-metrical lines. This second tendency, which separates itself from isosyllabism, includes the poems defined as ‘non-isosyllabic’, where lines have a non-fixed number of syllables.

⁴⁵ Spanish poetry is by no means rigidly syllabic. In Duffell’s (1999:45) words,

Some modern reference works on poetics and metrics for the general reader state simply, but inaccurately, that Romance (or Spanish) verse is syllabic [...] This sweeping generalisation is a compliment to the influence that French literature and literary theory have exerted on the poets of other Romance languages, but it ignores the Spanish accentually-based metrics that has for centuries contested French influence and syllabism.

(118)

Isosyllabic (7 syllables/line)	Non-isosyllabic (indefinite number of syllables)
<i>Va la brisa reciente</i>	<i>Vayan los mandados por los que nos deven aiudar</i>
'Goes the breeze recent	'Go the orders for which they us must help
<i>por el espacio esbelta,</i>	<i>los unos a Xérica y los otros a Alocad,</i>
through the space slender	some to Xérica and some to Alocad
<i>y en las hojas cantando</i>	<i>desí a Onda e otros a Almenar,</i>
and in the leaves singing	some to Onda and others to Almenar
<i>abre una primavera.</i>	<i>los de Borriana luego vengan acá.</i> ⁴⁶
it opens a spring'	the lens for Borriana then come here'

(Luís Cernuda)

If we follow Baehr (1970:37) and look at the history and evolution of the rules of isosyllabism, we will observe that Spanish versification started being governed by those rules after the Middle Ages, when they were introduced by influence of the French and Italian writers of the fourteenth century. Those writers had acquired this regularity principle from Latin medieval poetry. This meant that, from that century onwards, lines were counted as acceptable or unacceptable depending on whether they had the same number of syllables in a specific composition.

Before isosyllabism was generally accepted, there was a certain degree of tolerance towards subtle formal differences, which did not affect the essential rhythm of lines. In fact, when syllables started being counted literally, a conventional kind of prosody was imposed on verse, in such a way that certain vacillations and discrepancies became common, and the linguistic aspect of the poetic works was minimised. Non-isosyllabism, on the other hand, was indeed very common during the Middle Ages, both in folk versification and in literary poetry, and its gradual elimination from the so-called 'cultivated poetry' only occurred throughout the

⁴⁶ From 'Cantar del Mío Cid', in Duffell (2002:75). The system used to compose this poem remains a mystery for metrists, who cannot decide whether it is isosyllabic or isotonic (Gasparov 1996:135).

fifteenth and sixteenth centuries.⁴⁷ Nevertheless, we can still observe traces of its existence in contemporary traditional poetry, as well as in the kind of cultivated poetry that consciously shows a popular inspiration.⁴⁸

Looking now at specific literary examples, following Baehr (1970:54-57), the first instance of poetry in Spanish which deliberately sought isosyllabism appeared in 1240, in the first lines of the historical poem *Libro de Alexandre*. Those lines allow us to point once again to the fact that, at the same time that cultivated poetry turned towards isosyllabism, there was a popular kind of poetry which remained non-isosyllabic. The poetry of Gonzalo de Berceo – first half of the thirteenth century – represents the first culmination of the isosyllabic Spanish poetry in the Middle Ages. The continuation of the isosyllabic tradition, started by the *Libro de Alexandre* and by Berceo, is linked, at least until the end of the fourteenth century, to the ‘cuaderna vía’, and therefore, with the so-called ‘mester de clerecía’.⁴⁹ But even the poetry which did not belong to the ‘mester de clerecía’ had a similar evolution, tending towards isosyllabism, although in a less systematic, if more natural, way. The poet that achieved the best instances of isosyllabism was, together with Berceo, Pero López de Ayala (fourteenth century), with *Rimado de Palacio*. Towards the end of the fourteenth century, the so-called ‘major art line’, where the number of syllables per line is regular, spread all over Spain.

⁴⁷ In the fifteenth century, ‘major art’ poems were widely composed. They were compositions of variable syllabic size, with lines of two hemistichs (from 4 to 6 syllables each), two stresses per hemistich and arbitrary endings. It is said that ‘major art’ constituted a step back to a less strict syllabism, while strict syllabism kept pushing into the whole of the Spanish versification system. ‘Major art’ disappeared in the sixteenth century, and was replaced by a new type of long verse borrowed from Italian poetry (Gasparov 1996:138-139).

⁴⁸ Most contemporary poetry shows non-isosyllabism – and lack of rhyme – as a stylistic marker.

⁴⁹ The first forms of cultivated Spanish poetry written in the Romance language appeared towards the middle of the thirteenth century. Those works were written by clergymen and followed Medieval Latin models, with some French influence. Their topics were very varied and usually of a serious character – religious, historical and novelesque. This corpus of literary works was to survive throughout the following century. The name ‘mester de clerecía’ – ‘mester’ means ‘profession’ and comes from Latin ‘ministerium’ – refers to the group of clergymen that wrote those works, as well as to the works themselves. The most famous writers of the ‘mester de clerecía’ were Gonzalo de Berceo and Juan Ruiz, Arcipreste de Hita. The most characteristic stanza of the ‘mester de clerecía’ was the ‘cuaderna vía’, written in Alexandrine verses of fourteen syllables – seven plus seven – rhyming AAAA. The ‘mester de clerecía’ was opposed to the ‘mester de juglaría’, a popular kind of literature which appeared in the twelfth century. The works of the ‘mester de juglaría’ were characterised by the secularity of its topics, their oral quality – it was never written down – and by the fact that they were anonymous – their authors and performers were called ‘juglares’, which means ‘minstrels’.

At this moment in time, Spanish lines were classified into ‘simple lines’ (*versos simples*) and ‘complex lines’ (*versos compuestos*). The former were, in turn, divided into two subgroups, namely ‘minor art simple lines’, which contained from two to eight syllables, and ‘major art simple lines’, which comprised lines from nine to eleven syllables. Lines with more than eleven syllables were classified as ‘complex’. This classification has a correlate in the phonetic system of Spanish: according to Quilis (1984:54-55), an average phonic group in Spanish is formed by eight to eleven syllables.

In the second half of the fifteenth century, Juan de Mena and Marqués de Santillana made popular a dodecasyllabic line. Marqués de Santillana, a follower of Dante’s, states a theory of isosyllabism, this time under the Italian influence (Duffell 1999:53). Short lines started losing their original irregularity. Only romances kept a certain irregular taste until the sixteenth century. In that same century, the poets Juan Boscán and Garcilaso de la Vega implanted a series of lines and stanzas of Italian precedence, something which meant the final victory of isosyllabism in the Spanish cultivated poetry (Duffell 1999:37).

As Gasparov (1996:140) explains, from the sixteenth century to the very end of the nineteenth century, there were two overwhelmingly dominant metrical patterns in Spanish versification, namely the hendecasyllabic line – used in cultivated poetry, combining syllabic and accentual elements – and the octosyllabic line – used in folk poetry and song, with a purely syllabic rhythm.

In sum, the development of the syllabic system in Romance languages stemmed from the natural evolution of the languages, as well as from their poetic traditions.

3.2.3. Syllable counting and its related devices

Syllabic measurement in contemporary Spanish has a lot to do with isosyllabism and, given that achieving the same number for each line is sometimes hard, there is a series of metrical devices put into practice to regularise syllable counts and thus make lines isosyllabic. Also, Spanish poetry shows a tendency for stresses to fall at specific points in the line, depending on the number of syllables; according to this, in an octosyllabic line, there will always be a compulsory stress on the seventh syllable – Spanish poetry always counts the last syllable as extrametrical – and then a series

of optional stresses on the first, third and fifth syllables. The same metrical devices that contribute to making lines isosyllabic serve the purpose of forcing stresses to fall on their corresponding syllables, thus rendering the line metrical. Two of those devices, synaeresis and dieresis, happen at the word level, while the other two, synalepha and hiatus, happen at any level beyond the word. Given that synaeresis and synalepha follow a natural tendency of the Spanish language, the closer poetry comes to the natural pronunciation of Spanish, the more synaeresis and dieresis are likely to happen.

There are certain rules which regulate the use of each of the four devices at the poetic level. According to Baehr (1970:42), in cultivated poetry from the Renaissance period, the following criteria are applied to the use of synaeresis:

- a) Synaeresis happens naturally in proparoxytones where the two vowels involved follow the stressed syllable.

(119) *héroe* ('hero')

- b) It is less frequent if one of the vowels /a, e, o/ carries a full tonic stress.

(120) *creémos, aéreo* ('we believe')

Synaeresis is less frequent when the vowel /e/ comes before the tonic vowel.

(121) *león, real* ('lion', 'royal')

- c) When in the group of vowels the second one is a tonic /i/ or /u/, poetic synaeresis is considered forced.

(122) *raíz, saúco* ('root', 'elder')

Whereas synaeresis follows a natural tendency of spoken Spanish, dieresis is in opposition with the common tendency of spoken language. In poetry, it has two

main functions, namely the general function of achieving the desired number of syllables in a line which is defective, or the more specifically poetic function of highlighting the artistic nature of the line where dieresis appears. The norms that regulate the use of dieresis are the following (Baehr 1970:42-43):

- a) It is not normally used in those diphthongs which are considered inseparable, namely /'ie/ and /'ue/, which come from vulgar Latin open /ε/ and /o/, respectively.

(123) *tiérra, puénte* ('soil', 'bridge')

If the diphthong is in a post-tonic syllable, then dieresis is used more often.

(124) *cámbio, injúria* ('change', 'insult')

- b) There are some diphthongs that can be easily dismantled with dieresis, namely /ui/, /'ais/, /'eis/ – the last two in the second person plural of the present tense.

In addition to synaeresis and dieresis, there are other metrical devices that take place within the level of the word (see Baehr 1970:43-44). The first group has to do with the reduction of the syllables in a word by means of its suppression; these are aphaeresis – suppression of the first syllable of a word,

(125) *ciano* for *anciano* ('elderly person')

syncope – suppression of a middle syllable,

(126) *redor* for *alrededor* ('around')

and apocope – suppression of the last syllable of a word.

(127) *hi* for *hijo* ('son')

The second group has to do with the augmentation in the number of syllable in a word; these are prosthesis – addition of a syllable at the beginning of a word,

(128) *arradio* for *radio* ('radio')

epenthesis – addition of a syllable in the middle of a word,

(129) *coronista* for *cronista* ('columnist')

and paragoge – addition of a syllable, typically constituted by the vowel /e/, at the end of a word.

(130) *cantare* for *cantar* ('to sing')

These six devices, used artistically in the composition of poetry, belong to the level of colloquial or even uncultivated use of language in the study of linguistic registers.

The two devices which take place between different words are synalepha and hiatus. These two follow a series of rules which were introduced in Spanish after the first half of the fifteenth century. Synalepha is the most important device used to help isosyllabism, and has proved to be idiosyncratic of the Spanish language; it is also called 'Syllable Merger' (Roca 1986:350). In spite of its extended use in poetry, it is important to note again that synalepha is a common phenomenon in Spanish speech. Syllable Merger is a "resyllabification rule which coalesces a string of adjacent vowels not carrying main stress" (Roca 1986:350). According to Roca (1986:355), Syllable Merger becomes synalepha when it occurs in a syntactic environment. Piera (1980:210) reinforces the idea that, in Spanish, synalepha is a side-effect of a phonological process: "[s]ynalepha is a phonological rule in Spanish, which then becomes a metrical rule (property of the line)".

The rules that regulate poetic synalepha are the following (Piera 1980:10):

- a) Adjacent vowels belonging to different words correspond to a single metrically relevant position.

- b) All the remaining metrical positions correspond one to one to the vowels in the line, up to and including the rightmost stressed one.
- c) 1, 2 can be overruled in specific cases by hiatus, both word-internally and between words.

Synalepha can be considered the normal form in the oldest literary works, which is supported by the fact that its poetic manifestations coincide with the realisation of the Spanish language. As a result, synalepha must have been used since the very beginning of the most popular forms of Spanish poetry (Baehr 1970:59). One can suppose that synalepha, at least between unaccented vowels, would be a common phenomenon in the thirteenth century, and that it was definitely developed from the fourteenth century onwards (Navarro Tomás 1956:82). As for cultivated poetry, synalepha appears in Berceo – thirteenth century. The disappearance of the ‘mester de clerecía’ at the end of the fourteenth century contributed to an even more salient use of synalepha as a natural device in Spanish. The introduction of Italian forms in the fifteenth and sixteenth centuries only supported the extended use of synalepha, because this device is utterly common in Italian, too (Baehr 1970:59-60).

As pointed out in sections 2.3.2. and 3.2.3., synalepha is not the same as elision but rather, the latter is a subtype of the former. In Latin poetry elision proper was not practised. In Spanish poetry, like in Greek poetry, the use of such a device is optional (Balaguer 1954:14). Elision in Spanish only takes place when two alike vowels come into contact, that is, when the same vowel appears in two contiguous syllables in different words.

(131) *vá a América* (‘she goes to America’)

Punctuation marks and caesurae do not help synalepha from happening (Baehr 1970:47). In Old Spanish, synalepha was reflected orthographically (*de + él = dél*), but this is not the general case anymore, although cases where synalepha has a spelling counterpart can be observed in the transcription of songs or popular poetry.

The appearance of synalepha and hiatus is determined by (i) the way in which vowels get into contact, (ii) the placement of stresses and (iii) syntactical reasons. They are in a proportion of 1 to 2 thirds in favour of synalepha – in contemporary

poetry there are 96-99 instances of synalepha for 1-4 cases of hiatus (Baehr 1970:48-49):

- a) Synalepha always occurs between two identical vowels. Stresses are taken into account only in the case that the second of the two vowels carries a main stress; in such a case hiatus is preferred.
- b) Synalepha is compulsory when two non-tonic vowels get into contact. Conjunctions *y* and *o* ('and', 'or') – and their variants *e* and *u* – are not considered to be vowels, and so they require the application of hiatus.

If a tonic vowel gets into contact with a non-tonic one, synalepha is regulated as follows (Baehr 1970:49):

- a) It is compulsory when the tonic vowel precedes the non-tonic vowel.

(132) *echó el cimient*o ('he laid the foundation')

- b) It normally happens when the non-tonic vowel precedes the tonic one.

(133) *Subido he* por tu pena ('I have climbed up your sorrow')

- c) It is generally applied when two tonic vowels come together.

Synalepha is incompatible with the final pause in a line or with caesurae in compound lines. This general principle has an only exception, constituted by the so-called *estrofas de pie quebrado* ('broken-foot stanzas'), which make the final vowel of the last word on a line belong to the next line (Balbín 1962:79). The following lines (in Balbín 1962:79) constitute an example of this type of stanza and the corresponding realisation of synalepha:

(134)

Quien se inclina

'Who himself inclines

a la muy fina

to the very fine

dulze flor de clavellina

sweet flower of carnation'

[...]

(Fernán Pérez de Guzmán)

Balbín (1970:80-81) explains that the configuration of synalepha between different lines presents the following characteristics from the fifteenth century:

- a) It arises between a line and the immediately following line. The second one is usually shorter and its length is in proportion to that of the first one; the second line is a 'broken-foot line' (*verso de pie quebrado*).
- b) It is very usual that, between the lines which enter synalepha, there is symmetry in the location of lexical stresses.
- c) The word affected by synalepha in the second line is – with few exceptions – an 'empty word', that is, a clitic or a prepositional prefix.

As previously mentioned, hiatus is for synalepha what dieresis is for synaeresis, that is, the reverse process. Baehr (1970:49-51) formulates the rules of the application of hiatus as follows:

- a) Vowels /i/ and /u/, conjunctions *y* and *o* – and their equivalents *e* and *u* –, which follow in between vowels, are considered consonants.
- b) When the last stress in a line is preceded by a vowel, hiatus is usually applied.
- c) In a phrase, when two words hold a close morphological relationship, if the second one starts with a stressed syllable and the first one is a grammatical word, there is a tendency towards hiatus. This tendency has numerous exceptions, namely (i) if the vowel that precedes the stressed syllable is an *e*, even at the end of a poetic line, (ii) when two alike vowels happen to occur as explained above.

- d) The application of hiatus and synalepha is not systematic when two stressed vowels come into contact. In these cases, hiatus is usually the norm, although it only happens consistently when the stress of the second vowel is more obvious due to the rhythmic accentuation of the line.

Summarising, in Spanish poetry – and also in song –, when two syllables P and Q form a sequence PQ in which P ends in a vowel and Q begins with a vowel, the sequence PQ tends to be associated with a single metrical position.

As we have seen in this section, isosyllabism in poetry is achieved by applying a series of different devices to poetic composition and scansion. Those devices have undergone a specific process of evolution throughout the history of Spanish, and currently follow certain rules. Of all them, synaeresis and synalepha respond to a natural tendency of the Spanish language, that of making two or more adjacent vowels belong in the same syllable; this is one of the most salient idiosyncrasies of Spanish rhythm and metre.

3.2.4. The displacement of stress in Spanish poetry

The construction of rhythm in Spanish poetry, be it isosyllabic or non-isosyllabic, is based not only on the counting of syllables, but also on the periodic distribution of prosodic stresses. This relies on three fundamental principles (Balbín 1954:122-123):

- a) The centre of rhythmic movement is always the stanza.
- b) The distribution of prosodic stresses follows the principle of alternating stressed and unstressed syllables.
- c) Together with the main stanzaic rhythm, there is also a secondary rhythmic system, held at the level of the line.

The previous statements point to the fact that the beauty of the line and the stanza depends to a great extent on the placement of the rhythmic accents. Their placement often coincides with the position of lexical stress in individual words. At this point it is important to take into account that stress is totally objective and determined by the rules of stress assignment of specific languages. In addition to the stressing of a certain syllable in each word, the lines in each kind of poetic composition carry a number of rhythmic accents on specific syllables. This does not mean that rhythmic accents must coincide with lexical stresses in order for the line to be acceptable. In fact, there are two specific devices that have to do with the deliberate shifting of the

position of stress within a word; these are systole – the stress goes back to the previous syllable,

(135) *ímpio* for *impío* ('pitiless')

and diastole – the stress goes forward to the next syllable in a word.

(136) *oceáno* for *océano* ('ocean') (Baehr 1970:44-45).

On top of the number of lexical stresses found in a line, there are a number of fixed accents which must appear in that line in order for it to belong to a Spanish poetic category. For instance, a fixed rule for Spanish poetic composition is that all lines must carry an accent on the penultimate syllable – *acento estrófico* ('stanzaic accent'). The remaining accents are called 'rhythmic accents' proper. Let us have a look at the following example (in Quilis 1984:35):

(137)

Pastór que con tus sílbos amorósos

'Shepherd who with your whistles loving

me despertáste del profúndo sueño;

me woke up from the deep sleep

Tú, que hiciste cayádo de ése léño,

you that made cane out of that piece of wood

en que tiéndes los brázos poderósos,

in which you put the arms powerful'

[...]

(Lope de Vega)

In the above stanza, all the accents in the first and second lines are rhythmic accents – additionally, they coincide with the lexical stresses in the corresponding words. As for the third and fourth lines, the ones that fall on the first syllable (*tú*) and on the third (*hiciste*) syllables are extrarhythmic – they are lexical stresses which

nevertheless do not fulfil any prosodically functional position. If an extrarhythmic accent occurs next to a rhythmic accent, giving rise to a clash, the former is called 'antirhythmic accent' (Quilis 1984:33-36).

The lack of correspondence of stresses and rhythmic accents is a usual phenomenon in Spanish poetry. It is said (see Balaguer 1954) to have been inherited from the time when song and poetry were the same thing. As Balaguer (1954:223-226) observes, if we take into account the influence of music on poetry, it is not surprising that the rhythmic accent may change in certain lines of the poem. Indeed, the displacement of prosodic accent may be a similar device to that used in popular song, where poetic stress is subsidiary to musical accent. It could be said that the more music and poetry run parallel, the more stresses get displaced.

This theory about a basic correspondence between poetry and song proves the important point that any simple instance of verse has, apart from the final accent – often manifested on the penultimate syllable –, another rhythmic marker situated on one of its first syllables, which would correspond to the first beat in a bar, always the primary accent. Rhythmic markers usually fall on strongly accentuated syllables, but not every prosodic accent gets a rhythmic marker, in the same way that not every rhythmic marker corresponds necessarily to a prosodic or etymologic accent.

To illustrate this, the grids in Table 23 portray the metrical organisation of several musical templates according to different time signatures. As can be observed, there is always a rhythmic marker on the first beat of the bar; moreover, the last accent always corresponds to the main or secondary beat in the bar, depending on whether the time signature is simple (2/4, 3/4), or complex (4/4, 6/8). Poetic structures can be made to correspond to those same patterns, in such a way that some lexical stresses would fall on the main beat of the bar, while some others would be made to map onto a weak position, thus giving rise to what we have called 'stress displacement':

2/4:

X	X
X	
X	

3/4:

X	X	X
X		
X		

4/4:

X	X	X	X
X		X	
X			

6/8:

X	X	X	X	X	X
X			X		
X					

Table 23: Musical metrical grids.

In the (138), the organisation of stresses gets distorted when a poetic fragment is set into music.

(138)

5/8 (main beat = 1) (secondary beat/s = 2)

A esta puerta hemos llegado

(1) (2) (1)

'To this door we have arrived

con intención de cantar;

(2) (1) (2) (1)

with the intention to sing

denos licencia la niña

(2) (1) (2) (1)

grant us permission the girl

*para ver de comenzar.*⁵⁰

(2) (1) (2) (1)

in order to try and start'

⁵⁰ 'A esta puerta hemos llegado', in Manzano Alonso (2003:366).

But leaving aside the possible influence of song on the accentuation of poetry, there are hints that allow us to think that Spanish poets enjoy a certain degree of freedom in order to shift stresses around in accordance with the needs of the rhythm. This may have been inherited from Latin poetry, where rhythmic accent did not necessarily coincide with stress, as shown in section 2.1.1.

Following this historical perspective, the displacement of stresses is by no means exclusive to modern or contemporary poetry. Indeed, there are many instances of accent displacement due to rhythm need in the *Siglo de Oro* ('Golden Age') poets – seventeenth century. This practice is also found in the nineteenth century, as well as in contemporary poets (Balaguer 1954:232-234). Let us have a look at the rhythmical arrangement in a fragment written in the seventeenth century (fragment quoted in Quilis 1984:193):

(139)

Lexical stresses

¡Oh excélsó muro, óh tórres coronádas

Ah sublime wall ah towers crowned

de honór, de majestád, de gallardía!

with honour with majesty with elegance

¡Oh grán río, grán réy de Andalucía,

ah big river great king of Analusia

de arénas nóbles, yá que nó dorádas!

of sands noble if not golden'

Rhythmic accents and synalephas

¡Oh excelso muro, oh torres coronadas

de honor, de majestad, de gallardía!

¡Oh gran río, gran rey de Andalucía,

de arenas nobles, ya que no doradas!

(Luis de Góngora)

The above stanza comprises four hendecasyllabic lines, with a compulsory accent on the penultimate – tenth – metrical syllable; this accent coincides with the lexical stress in the last word of each line. All the accents which fall on an even metrical syllable – second, fourth, sixth, eighth, and tenth – are rhythmic accents. The accents on the first and fifth syllables of lines one and three are antirhythmic accents in that they give rise to an accent clash. The only perfect instance in terms of the rules of the

hendecasyllabic line is line four, which carries rhythmic accents on the second, fourth, sixth, eighth and tenth syllables.

Summarising, the most common types of displacement in Spanish poetry are those which follow metrical needs, that is, those which have to do with the need to re-establish the line's rhythm (Balaguer 1954:228-230).

This allows us to conclude that rhythm in Spanish poetry is achieved not only through isosyllabism – as has long been thought – but also through the periodic organisation of rhythmic accents. These accents may or may not coincide with lexical stresses, depending on the requirements of the poetic template used; in the case that they do not coincide, we have to talk about 'stress displacement', which seems to be the most widely used rhythmic device in the history of Spanish poetry.

3.3. Verse prosody in English

The following sections are devoted to the study of English verse prosody. In section 3.3.1. I analyse the evolution of the English prosodic system, from OE to PdE, emphasising the parallelisms between the evolution of the language and that of its system of versification. Sections 3.3.2. and 3.3.3. examine the regulating principles of verse composition in English, as well as the use of such devices as synalepha.

3.3.1. From Old English alliterative verse to Present-day English accentual-syllabic verse

The principles of English verse composition and scansion have undergone a radical evolution from the first records in OE to current practices. This evolution is intimately linked to that of the language itself, together with the influences from other traditions which have been incorporated into English throughout the centuries.

Before the appearance of OE verse, the Germanic system of versification had suffered a major change from "hypothetical Indo-European syllabic verse" (Gasparov 1996:36), again as a consequence of the changes in the language, which ranged from the shift and intensification of stress to the reduction of unstressed syllables. In spite of the difficulties in reconstructing Old Germanic verse, scholars agree that it evolved "from syllabic into tonic [or accentual]" (Gasparov 1996:36) in the three Germanic areas, namely Scandinavia – and later Iceland, too – Britain and

what is nowadays Germany.⁵¹ Following Gasparov's (1996:37) taxonomy, the main contrast between syllabic and accentual metre is that, while the former relies on the basis that "any syllable is equivalent to any other syllable (regardless of its length, stressing and so on)", the latter is regulated by the fact that "any group of syllables united by one stress is equivalent to any other group of syllables also united by a stress (regardless of the number of syllables in the groups)".

In Britain, the OE period – from mid sixth century to early twelfth century – witnessed the flourishing of a highly complex system of verse composition, called 'alliterative verse'. This system encoded a series of principles which had to be applied to any poem that aimed at being labelled 'metrical'. As a general rule, OE poetry was composed in hemistichs – half lines – also called 'verses'. Those half lines were joined together into long lines by means of structural alliteration. The number of unstressed syllables in between stressed ones was variable and, as a result, so was the length of hemistichs and lines. Alliteration was, thus, the surface basis of OE verse composition. Its working principles were the following: in the second hemistich, the first metrical beat should alliterate with either one or two metrical beats in the first hemistich. Let us observe how alliteration works in the following fragment from *Beowulf* – the segments which alliterate are in bold:

(140) (from Gasparov 1996:39):

*Beowulf wæs **breme** / – **blæd** wide sprang –*
*Scyldes eafera / Sceldelandum in.*⁵²

In spite of the fact that alliteration has been analysed as the most idiosyncratic feature of OE poetry, some metrists (McCully and Anderson 1996:44) regard it as a superficial feature, and not as an essential building element in OE poetry. For those metrists, OE alliteration was a stylistic marker, used to distinguish verse from prose in accentual languages, and so its role could be compared to the role of rhyme in Romance and later English poetry. Gasparov (1996:38) argues that Old Germanic

⁵¹ The information that we have about Old Germanic verse goes back only as far as the ninth century AD.

⁵² Translated by Heaney (1999:3): *Shield had fathered a famous son: / Beow's name was known through the north.*

languages did not develop rhyme instead of alliteration because of the suprasegmental characteristics of those languages. They had a strong stress at the beginning of the word, so that roots were more prominent than the reduced suffixes in speech and, as a result, in poetry. Rhyme works in the opposite way, as it is applied to the coda of the last stressed syllable in the word onwards, somehow shifting the prominence from the first to the last segments. Compare (140) with the following fragment from *King Horn* (c. 1225), the first courtly romance in English (rhyming segments are in bold):

(141) (from Gasparov 1996:182):

*A song ychulle ou **singe***
*Of Allof þe gode **kyng** –*
*Kyng he wes by **weste**,*
*Þe whiles hit **yleste** –⁵³*

The comparison between (140) and (141) points at both alliteration and rhyme as verse-building devices which employ similarity of sound as their basis and mark line ends; on the other hand, we observe that, while alliteration binds hemistichs into lines, rhyme binds lines into stanzas, a fact that is linked to the opposition between accentual metre and syllabic metre.

A second observation on alliterative verse, made by Duffell (1996:211), is that, although OE verse had a fixed number of stresses and a variable number of syllables, it would not be accurate to state that it was purely accentual, because the number and position of the unstressed syllables was regulated to some extent, as we will see in the next paragraphs.

Metrists have for years tried to give a clear-cut account of the types of lines that existed in OE.⁵⁴ In spite of the ongoing controversy, the most widely accepted taxonomy of verse types is the one established by Sievers in 1885. Sievers used a set of stress patterns as a starting point, and claimed that the designs of stressed and

⁵³ Translated by Weston (in Gasparov 1996:182): *A song I now will sing / Of Allof, crowned king; / He reigned in the West / While he with life was blest.*

⁵⁴ In an insightful article, McCully (1996:42-58) revises Sievers's theory and explores the possible verse-demarkation devices in OE poetry, highlighting the importance of the relationship between stresses and syllables at the end of the verse.

unstressed syllables within each hemistich could be organised into five types – A through E. Cable (1974, 1991) expands Sievers’s theory, basing it on the counting of stressed and unstressed syllables to form positions. It is important to note that simple syllable-counting is not a property of early Germanic verse; as mentioned above, Cable counts positions, which are prototypically but not exclusively occupied by one syllable each. According to this, a half-line will have four syllables; the exceptions to this will be a series of unstressed syllables – from two to five – which could substitute for one or the first two syllables in the half-line.⁵⁵

A	/ x / x
B	x / x /
C	x / \ x
D	// \ x
E	/ \ x /

Table 24: OE verse types.

According to Sievers, every hemistich in OE will correspond to one of the stress patterns above.

In the following example, we can see the scansion of several OE half-lines from the poem *Beowulf* – seventh-tenth century AD – both in terms of the counting of syllables – Cable’s theory, indicated with / for stressed syllables, and x for unstressed ones – and the counting of positions – Sievers and Cable’s theories, indicated with numbers. (142) marshals the classification of some lines from *Beowulf* according to Sievers’s taxonomy.

⁵⁵ Crucial to this theory are the concepts of ‘dip’ and ‘lift’ (Beare 1957:105-106). The former was a weak position (\) typically occupied by a single syllable, but sometimes occupied by more than one weak syllable. The latter was a strong position (/) typically occupied by a main stress, but sometimes by a secondary stress – in types A, B, C in Table 24 – or a stressed plus an unstressed syllable.

(142) (from Cable 1996:10):

Type of line	Scansion	Line in <i>Beowulf</i>
A	1 2 3 4 / x x / x <i>weox under wolcnum</i>	(8a)
B	1 2 3 4 x x x / x / <i>forðam he manna mæst</i>	(2645a)
C	1 2 3 4 x x x / \ x <i>Þæt hie in beorsele</i>	(482a)
D	1 2 3 4 / / \ x <i>leof landfruma</i>	(31a)
E	1 2 3 4 / \ x / <i>sægrunde neah</i>	(564b)

A theory of OE metre would not be complete if it only accounted for the stressed and unstressed positions in each hemistich and line. The four possible combinations of stresses within each hemistich rendered something much more interesting than simple metrical patterns, namely intricate melodic patterns (Table 25). In Cable's (1974:94) words, 'given four metrical positions and certain constraints upon those positions, a model of metre will produce exactly eight contours', of which the last three were not permitted in OE poetry.

1:	1 \	2 /	3 \	4
2:	1 /	2 \	3 /	4
3:	1 /	2 \	3 \	4
4:	1 \	2 \	3 \	4
5:	1 \	2 \	3 /	4
(6):	1 /	2 /	3 \	4
(7):	1 /	2 /	3 /	4
(8):	1 \	2 /	3 /	4

Table 25: Melodic contours in OE hemistichs.

The implications of this approach go far beyond metrics, in the sense that it actually seems to downgrade the importance of stress in favour of pitch movement. As Cable (1974:96) puts it, “the metrical basis of OE poetry was the melodic formula, a contour of pitch drawn from a set of five contours of a possible set of eight, to which words are fitted according to stress rules”. In this view, metrical patterns are subsidiary to melodic paradigms, used by poets to decide upon which words could fit a specific paradigm. Cable (1974:96) himself used musical notation to illustrate the melodic paradigms found in *Beowulf*. Despite the plausibility of this theory, in later works Cable (1991:55) restates it, arguing that the melodic paradigms “were the result of a metrical paradigm and not the engine that was driving it”. What is clear is that, in OE, metre interacted with melody in a natural way, something which should not be surprising if we take into account that OE poetry was meant to be recited, chanted or sung. Unfortunately, no OE songs have survived in musical notation. The earliest English songs preserved with music are St. Godric’s, probably from the middle of the twelfth century, which nevertheless have echoes of the classical Anglo-Saxon verse form (see Baltzer *et al.* 1991).

We know that the most important elements in OE metre were stress, alliteration, quantity, resolution and suspension of resolution, measures, breath groups, foot divisions, and inferred pitch contours. Under the rules of OE versification, the wrong placement of a stressed position, or the wrong placement of an unstressed syllable, would make the whole verse unmetrical.

The Germanic tradition of alliterative verse did not stay the same in Britain after the Norman Conquest, in 1066. Between the tenth and the fourteenth centuries, English verse experienced a series of innovations, helped by both Germanic and Latinate influences. Two new types of verse emerged in this period, each of them rooted in a different tradition. On the one hand, the Germanic inheritance continued to be cultivated in the form of accentual verse, which still relied on alliteration, though relaxing the principles of OE alliterative verse. Lines were still composed of two half-lines, each generally of two stresses, and many half-lines were linked by alliteration. However, half-lines carried a greater number of unstressed syllables than in OE. More specifically, the first half-line almost always had two or more unstressed syllables before and after the first stress. In (143), extracted from *Sir*

Gawain and the Green Knight, we can observe that the presence of two unstressed syllables before the first stressed one, which was legitimate in ME, would have made the half-line unmetrical in OE:

(143) (from Cable 1996:9):

x x / x x x /

And he granted and hym gafe

A number of new forms indebted to both Medieval Latin and Anglo-Norman models come into play at the end of the twelfth century. The Latin quantitative foot and the English accentual foot were connected through French, which was purely syllabic. A hexasyllabic line in French, characterised for having six syllables and rhyme at the end, could respond, in English, to five types of metre, depending on whether the beat fell on alternate syllables – either odd- or even-numbered –, or on every third syllable – beginning either with the first, second, or third syllable. The same would happen with a heptasyllabic line, and so on, as we can see in the alternative scansions rendered in (144).

(144)

Trochaic (/ x): *Féar no/ móre the/ lightning/ flásh*

(Shakespeare, *Cymbeline*)

Iambic (/ x): *Féar/ no móre/ the light/ning flásh*

In any case, the classical names for those five types of metre, which in classical Latin responded to length contrasts, were kept in English to designate stress contrasts, hence the denominations ‘trochaic’, ‘iambic’, ‘dactylic’, ‘anapaestic’ and ‘amphibrachic’ (see Beare 1957:33).⁵⁶

The second important innovation is rhyme, which starts to be used as a structural principle, borrowed probably from Medieval Latin, but also due to the

⁵⁶ Early prosodists not only applied the foot taxonomy borrowed from Latin to English verse, but also identified length with stress, a confusion that has survived until the present day. It was only in the eighteenth century that prosodists such as Bysshe (*Art of English Poetry*, 1702) and Johnson (*Dictionary*, 1818), rejected this idea and started to dissociate length and stress (Beare 1957:34).

influence of French, particularly certain Anglo-Norman forms (see Preminger 1974:355; Gasparov 1996:41). At the beginning, internal rhyme breaks the alliterative long line into a short alliterative couplet, until eventually it supplants alliteration altogether. In general terms, the introduction of rhyme marks a shift from accentual verse to accentual-syllabic verse, where both the number of syllables and the number and position of stresses in the line become relatively fixed.

In spite of the Romance influences, alliteration continued to be a widely cultivated structural principle throughout the thirteenth and fourteenth centuries and, indeed, the fourteenth century witnessed a revival of this principle, in a movement which has come to be called 'Alliterative Revival'. There are also some remains of the interaction between the Germanic and the Romance tradition, like, for instance, Layamon's *Brut* (ca. 1200) or the anonymous poem *The Owl and the Nightingale* (ca. 1200), two poems which are completely different with regard to internal organisation (Gasparov 1996:182). The latter is a good example of the attempt, on the part of English medieval poets, to conform to a model – the syllabic one – which was alien to their native accentual system. Let us observe (145), written following the French principle of syllabism:

(145) (from Beare 1957:33):

þo stóð on óld stóc þár bi-side
þár þo úle song hire tíde

It is not until the fourteenth century that the attention of poets such as Chaucer shifts completely towards strictly French and Italian forms of composition, considered more elegant and cultivated. Chaucer imitated the hendecasyllabic line – eleven syllables per line –, which was being cultivated by French and Italian poets like Dante, Petrarca, Boccaccio, but he transformed it into the iambic pentameter, a truly accentual-syllabic kind of metre composed of five feet, with prototypically two syllables (unstressed – stressed) per foot. As a result, any triple time line was eliminated – in other words, only double-timed feet were kept. The following example (146) of iambic pentameter lines is taken from the General Prologue of the *Canterbury Tales*:

(146) (from Gasparov 1996:184)

*Whan that Aprille with his houres soote
The droghte of March hath perced to the roote
And bathed every veyne in swich licour,
Of which vertu engendred is the flour, (...)*

After Chaucer's major shift toward Romance models, the modern period – sixteenth and seventeenth centuries – witnessed the definitive decay of accentual and alliterative forms, which nevertheless persisted up until 1550. The iambic pentameter of Chaucer and his successors, transformed by Petrarchan French and Italian influences, had by then become the most widely used pattern of English verse. Shakespeare used the iambic pentameter line extensively, and indeed carried it a step further, complicating it by the addition of stresses which increased the density of lines. In (147), extra stresses are indicated in bold:

(147) (from Guthrie 1991:90)

*Be thou the tenth **mu**se **ten** **ti**mes more in worth*

It can be said that, around this time, English verse gradually becomes what it currently is, that is, an accentual-syllabic type of verse. The Romance influence grows in such a way that Elizabethan poets even manage to adapt the patterns of Classical Latin to English, replacing quantity with stress as the marker of the pattern. Throughout the seventeenth, eighteenth and nineteenth centuries, poets like Donne, Milton or Browning contribute to the definitive adoption of Romance and Latin models by English poets, and finally arrive at the Present-day accentual-syllabic model of composition and scansion.

3.3.2. Stresses and syllables in contemporary English 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 placing of stresses in the line responds to a specific pattern, and is thus fixed. This principle was extensively used by Renaissance poets such as Shakespeare, as seen in (144), quoted below as (148b):

(148)

Verse pattern	Verse instance
a. Iambic (x /)	<i>To réar/the có-/ lumn ór/ the árch/to bénd</i> (Pope, <i>Epistle to Burlington</i>)
b. Trochaic (/ x)	<i>Féar no/ móre the/ lightning/ flásh</i> (Shakespeare, <i>Cymbeline</i>)

This entails that, whereas in pure accentual verse “the stresses must all be true speech-stresses” (Preminger 1974:771), in accentual-syllabic verse, the mind, following a metrical pattern, will adjust the stressing to fit via the strategies of promotion and demotion of syllables – observe, in the iambic instance in (148a), the metrical promotion of ‘or’, an otherwise unstressed syllable.

The second principle states that the number of syllables per foot must be identical. If we look at (148) above, we observe that all the iambic feet in the first line have two syllables, and the same goes for the trochaic sequence, with the exception of the last foot, which has the second (weak) position empty (*flásh – Ø*).

With these two principles at work, English verse is considered accentual-syllabic and not only accentual. Nevertheless, the type of accentual verse which was inherited from Germanic can also be found in English. It is the type for much popular verse, folk poetry oral poetry and also for the English and Scottish ballads. As Preminger (1974:770) points out, in such compositions,

isochronous timing (although stresses do not fall on fixed positions like in accentual-syllabic verse), the use of metrical rests, and the phenomenon known as ‘dipodism’ are all characteristic of such verse and reveal its direct origin in or close association with song.

The fact that the system of English verse can be both accentual and accentual-syllabic depending on the form of composition, has to do with the traditions merged in English. On the one hand, the native tradition of Old Germanic metres requires the occurrence of four syllables carrying major stress in each line, while the number of syllables in the line is relatively less constrained. On the other hand, the Romance metres imported in the Middle Ages regulate the number of syllables strictly. From the fourteenth century onwards there was a clear proliferation of Romance poetic

forms, since the English imported Romance verse forms to map onto a shifting morphological base which had been built upon Germanic and Latin strata (see Preminger 1974:779).

Table 26 marshals the possible combinations between syllables and stresses in different stages in the evolution of English versification:

		Position of stress	
		Fixed	Free
Number of syllables	Fixed	Accentual-syllabic verse E.g.: Present-day English	Syllabic verse E.g.: Medieval Latin, French
	Free	Accentual verse E.g.: Old English	Free verse E.g.: Present-day English

Table 26: Verse systems.

The accentual-syllabic metre adapted to English first by Chaucer and again by Wyatt and Surrey from French and Italian syllabic models was during the first centuries of its use in England principally a syllable-counting metre with secondary constraints on stress, and only later developed into a really accentual-syllabic verse, with constraints both on syllabic counting and also on stress placement. As Preminger (1974:786) well indicates, the requirement that each line had to have a definite number of syllables entailed the adoption of specific devices in order to adjust the number of syllables in the line. Of those devices, the most salient one was synalepha, first imitated from French and Italian by Chaucer, and then cultivated by latter poets, such as Milton, Donne or Jonson.

3.3.3. Synalepha in English: a comparison with Spanish

In sections 2.2.3. and 3.2.3. I showed that Spanish stress at the level of the utterance is not completely free since it depends on the overall rhythm of the utterance; for each utterance, there must be a number of rhythmical leaning points at specific intervals. This entails the existence of certain compression devices, the most important of which is synalepha. So, in Spanish, synalepha is primarily a phonological rule with a predominantly rhythmical function, and manifests itself in the prosody of everyday speech.

Let us summarise the principles that regulate synalepha in Spanish, as seen in section 3.2.3. The realisation of synalepha is regulated by two non-phonological conditions, namely tempo, which must be fast for synalepha to occur, and register, which must be colloquial. There is also a phonological condition for the occurrence of synalepha in Spanish, namely that the two vowels combined into a single syllable form a continuous rise or a continuous fall in prominence.

The domain of synalepha is the syllable across words. It normally covers two words, but sometimes it extends over three. The maximum of vowels joined together under synalepha is five, but the most common synalepha groups two vowels.

There are three types of synalepha (Canellada 1987:54), namely diphthongisation,

(149) *mi amigo* → *mja-mí-go* ('my friend')

reduction,

(150) *se acabó* → *s(e)a-ka-bó* ('it's over')

and elision.

(151) *se acabó* → *sa-ka-bó*

As for the combinations which can give rise to synalepha in Spanish, it is possible to establish an acceptability gradient, whereby synalepha among vowels of non-stressed syllables is always realised,

(152) *y estaban* → *jes-tá-ban* ('and they were')

(153) [a] prevails over [i] in *mi amigo* → *mja-mí-go*

(154) *si unían* → *sju-ní-an* ('they united')

while synalephas among vowels of stressed syllables are more complex and do not always occur.

(155) *echó el cimientó* → *e-chóel-ci-miént-to* ('he laid the foundation')

(156) *Subido he por tu pena* → *Su-bí-doé-por-tu-pé-na* ('I've climbed your sorrow')

As Piera (1980:210) observes, synalepha in Spanish is not only a prosodic device, but a "phonological rule [...], which then becomes a metrical rule (property of the line)". As a device used for metrical composition and scansion, synalepha works very productively for Romance languages such as Spanish, Provençal, Italian and Portuguese, where "the adjustment of the verse line to the required number of syllables is achieved by synalepha [...] by elision or by hiatus" (Preminger 1974:715).

The status of synalepha in English is not that of a phonological device which then acquires a metrical function. Rather, synalepha in English is a metrical device with non-phonological status. In Napoli's (1975:403) words, "[s]ynalepha is an optional convention in English verse". This fact points to a structural difference between Spanish and English in terms of suprasegmental phonology, but also in terms of segmental phonology, because it explains the relationship between vowels in the two languages. In Spanish, the normal tendency is for consecutive vowels to form a diphthong; on the other hand, diphthongs are not entities separate from the vowel system, but rather the result of the joining together of two vowels. In English, consecutive vowels avoid diphthongisation, and diphthongs behave as independent units with specific characteristics, not as the result of the joining together of two single vowels. We are presented, thus, with a dichotomy between those languages which have synalepha encoded as a phonological as well as a metrical rule, and those which observe synalepha as an exclusively metrical device.

Generative metrics sees synalepha as one of the so-called 'correspondence rules' – rules that relate abstract patterns to surface lines – and makes it

interdependent with the constraint against stress maxima appearing in W positions (Napoli 1975:401).

When Halle and Keyser (1971) proposed a series of correspondence rules for iambic pentameter in English, they defined synalepha as “a metrical convention by which we may assign a sonorant sequence incorporating at most two vowels (immediately adjoining or separated by a sonorant consonant) within a line of verse to a single metrical position (S, W or X)” (Napoli 1975:402).

Related to this is the idea that certain segments constitute a more powerful syllable boundary than others. In Youmans’s (1996:191) words, “syllabic boundaries are weaker when nuclei are separated by single [+sonorant] segments and stronger when separated by [–sonorant] segments”. According to the sonority scale, thus, the weakest boundaries are the glides [j, w] and the segment [h], followed by liquids [l, r], nasals [m, n, ŋ] and voiced fricatives [v, ð, z]. At the other end of the scale are the voiceless fricatives and the stops, which constitute powerful syllable boundaries, and therefore prevent synalepha from occurring.

It is obvious that poetic syllabification can be understood as an abstraction, where the number of syllables would remain constant, or as an idiosyncrasy of each single poem written in a given language. The difference between the first group of languages and the second group is that the former have a series of phonological principles encoded in their system, which regulate the possible realisations of synalepha, and rule out those cases which do not respond to a phonological tendency of the language. In the latter, synalepha is just a device *ad hoc*, with much weaker regulating devices, which never correspond to the phonology of the language.

In fact, the latter do not show a correspondence between the metrical pattern and the actual scansion of a verse instance, or between the scansion of a verse instance and its pronunciation. In Halle and Keyser’s (1971:171) words,

The assignment of syllables to positions is, of course, strictly metrical. It does not imply that the syllables assigned to a single position should be slurred or elided when the verse is recited. The correspondence rules are not instructions for poetry recitations. They are, rather, abstract principles of verse construction whose effect on the sound of the recited verse is indirect.

The history of synalepha in English is one of idiosyncratic adaptation of a foreign device, which in principle responded to the segmental and suprasegmental characteristics of Romance languages. In English, synalepha is indeed a common feature of seventeenth and eighteenth century prosody – Milton and Pope used it conspicuously –, where it is used to maintain the syllabic conformity of lines. Indeed, many sixteenth and seventeenth century poets thought it a beauty of verse to use the same word twice in a single line or in adjacent lines, once contracted, once in full. As Preminger (1974:786) indicates, compression and decompression devices could not always be signalled by spelling in sixteenth to seventeenth century English verse. Preminger (1974:787) also points out that it was not until the time of the Romantics that English poets started to admit to the familiar decasyllabic line extra unstressed syllables which did not allow of contraction, whereas poets of the sixteenth, seventeenth and eighteenth centuries – non-dramatic poets, at any rate – seem on the whole not to have written lines containing extra syllables – the feminine ending excepted –, which could not be resolved phonologically.

As happens in Spanish, English also has rules for the realisation of synalepha. The most general tendency is for synalepha to assign two unstressed or weakly stressed syllables to a single position.

(157) *Wholly in heavenly things my mind is set.*⁵⁷

(158) *'Tis virtue alone, or nothing that knits friends.*

It does not assign two fully stressed syllables to a single position. There are instances where synalepha joins together an unstressed or weakly stressed syllable and a fully stressed syllable.

(159) *But though I have found thee, and thou my thirst has fed.*

(160) *For that is first required, a man be his own.*

⁵⁷ Examples (157), (159), (161) by Donne (*Sonnet 17*), (158), (160) by Jonson (*Epistle to a Squib*), and (165) by E.B. Browning (*Sonnets from the Portuguese, XXI*), all borrowed from Fabb (2002). Examples (162) by Rossetti (*Song*), (163) and (166) by Donne (*Sonnet 17*), (164) by Bécquer (*Rimas*).

Fabb (2002:8) gives a clear account of synalepha using a theory of projection of syllables from the metrical pattern to the verse instance. As a general principle, every syllable in a line projects onto the grid and is symbolised by an asterisk:⁵⁸

(161)

Since she whom I loved hath paid her last debt
 * * * * * * * * *

Fabb (2002:9) also elaborates a series of rules which regulate non-projection, that is, realisation of synalepha in verse lines. They go as follows:

Non-projection rule a: Do not project a syllable at the (right-hand) end of the line which is unstressed or weak in stress and which comes after a strongly stressed syllable.

(162)

And if thou wilt remember
 * * * * * * 0

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.

Optional rules like b apply where necessary in order to bring the number of syllables to the right number. This is the equivalent of Spanish synalepha. The difference between Spanish and English with regard to this specific non-projection rule is that, in Spanish, the application of synalepha is the unmarked option, as it corresponds to a phonological rule of the language, while in English, it is the marked option, brought forward as an exclusively poetic artifice alien to the phonology of the language. Compare examples (163) and (164):

(163) *And her soul early into heaven ravished,*

⁵⁸ Fabb (2002) develops an overall theory of scansion. This first projection principle is complemented by a series of other principles, so that, in the end, a grid with asterisks and brackets as the main visual tools is built.

(164) *Del salón en el ángulo oscuro,*
 'Of the hall in the angle dark'

In (164), we see a synalepha realised metrically in order for the line to conform to a verse pattern, which regulates that the line must have a specific number of syllables; this is the same for (163). However, while the pronunciation of the synalepha in (164) responds to a phonological rule of the Spanish language – that synalepha would be realised in everyday speech – the one in (163) is only a metrical convention, which does not correspond to the pronunciation of those two words – or any other two words with similar segmental structures – in connected speech.

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 (Fabb 2002:10):

(165)

<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	*	*

Fabb himself (2002:45) observes later in his book that

The vowel e, which is often cut off by an apostrophe in the word the, and in syllables before r, as dang'rous, gen'rous, &c. ought to be preserved in the pronunciation, because the syllable it forms is so short as to admit of being sounded with the preceding syllable, so as not to increase the number of syllables to the ear, or at all hurt the harmony.

This confirms, once again, that non-projection rules have to do with the abstract metrical pattern, not with the pronunciation of verse instances.

In the examination of the phenomenon in question presented by Halle and Keyser (1971, in Napoli 1975:401-403), instances like (165) above would also qualify as synalepha in English. In Spanish, synalepha can only happen between vowels without any intervening consonants; no other segments can enter the process of synalepha in Spanish.

Non-projection rule d: Optionally, do not project a syllable which is a grammatical monosyllable preceding another grammatical monosyllable (i.e. preposition, article, pronoun, auxiliary verb, modal verb, conjunction, etc.) (Fabb 2002:27):

(166) *To nature, to hers and my good is dead.*

In Fabb's (2002:11ff.) view, the choice to project or not to project is related to the linguistic representation, and not to the performance of the line. Projection has nothing to do with pronunciation: the syllables which are not projected can be pronounced or not pronounced. So synalepha is seen as a fact about the abstract metrical representation of the poem. In other words, projection is a way of ensuring that the line fits the metrical grid which is built from the linguistic representation. Independently of this, there is also a choice to pronounce or not to pronounce. In terms of scansion, a decision has to be made on which syllables to project and then the grid is construed from the projected syllables. If the grid fails, the process starts again by projecting different syllables.

In conclusion, synalepha is a universal metrical device – it has been attested in languages as diverse as Sanskrit, Persian, Arabic, Italian and English, among others –, encoded as a correspondence rule which has the function of making the verse instance project as many syllables as required by the metrical pattern, and leave extra syllables not projected.

However, in some languages, such as Spanish and Italian, synalepha is primarily a phonological rule which is realised phonetically in everyday speech, which then becomes a metrical rule. One could hypothesise that, in English, synalepha was not even a correspondence rule until the Middle Ages, when influence of Italian writers started to be clearly observed in the methods of composition of English poets, for instance Chaucer. Later on, in the nineteenth century, poets such as Rossetti and Browning further adapted Italian poetic forms to English, thus increasing the Romance influence on English Poetics and Metrics.

3.4. Summary

Chapter 3 is devoted to the analysis of verse. There is evidence that the origins of certain metrical patterns used in verse derive from prosodic patterns of phrases uttered in speech. This makes even more sense if we consider the fact that the manifestation of metrical patterns themselves is intended for spoken language. This does not mean that speech prosody and verse prosody can be equated, given that verse has a conscious aesthetic dimension which common speech lacks. In this sense, verse prosody is self-conscious in formal terms, while speech prosody is not. For this reason, the study of specific verse prosodies can shed some light on the rhythmic principles that underlie the corresponding speech prosodies and, more broadly, the ways in which different languages parameterise and articulate the general rhythmic capacity.

Verse prosody is also known as 'metrics', which is usually taken to refer to the study of the art of versification. As happens with speech prosody, metre can be subclassified into three main types, depending on the element that recurs and establishes the rhythm of a specific composition. The first type is 'syllabic metre', where each line is made up of a specific number of syllables. This metrical type corresponds to the versification of the so-called syllable-timed languages in speech prosody, whose rhythm, as explained in chapter 1, is based on the isochronic repetition of syllables. The second type is 'accentual metre', where stresses fall at regular intervals in the line. The third type is called 'accentual-syllabic metre', which is a combination of accentual and syllabic metres, and pays attention to the more or less regular alternation of stressed and unstressed syllables in a line. These two types of metre can be related to the speech prosody of stress-timed languages, whose rhythm is based on the recurrence of stresses at equal intervals in time.

Section 3.1. examines the relationship between speech prosody and poetic rhythm and metre in depth, paying particular attention to the languages analysed in this dissertation, namely Spanish and English. Sections 3.2. and 3.3. explore the historical evolution and specific characteristics of Spanish and English verse prosody.

As explained in section 2.2., Spanish is a syllable-timed language. According to the correlation between speech prosody and verse prosody advocated in this

dissertation, Spanish verse is syllabic. 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.

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. In spite of this, accents in Spanish arguably occur at specific intervals. Throughout history, poetry written in Spanish has followed two clearly different paths. The first tendency is isosyllabism, while the second one comprises non-metrical lines. This second tendency, separated from isosyllabism, includes all the poems defined as ‘non-isosyllabic’, where lines have a non-fixed number of syllables.

Isosyllabic poetry uses a series of metrical devices in order to regularise syllable counts. Curiously, those same metrical devices also serve the purpose of forcing stresses to fall on their corresponding syllables, thus rendering the line metrical. Two of those devices, synaeresis and dieresis, happen at the word level, while the other two, synalepha and hiatus, happen at any level beyond the word. As analysed in section 2.2.3., synaeresis and synalepha follow the tendency of Spanish to make contiguous vowels belong in the same syllable in a process of syllable unification. Other metrical devices that take place within the level of the word are aphaeresis, syncope and apocope, and prosthesis, epenthesis and paragoge.

Although Spanish stress is determined by rules of assignment, the lines in each kind of poetic composition carry a number of rhythmic accents on specific syllables, which does not mean that rhythmic accents must coincide with lexical stresses in order for the line to be acceptable. In fact, there are two specific devices that have to do with the deliberate shifting of the position of stress within a word; these are systole – the stress goes back to the previous syllable – and diastole – the stress goes forward to the next syllable in a word. Last, on top of the number of lexical stresses found in a line, there are a number of fixed accents which must appear in that line in order for it to belong in a Spanish poetic category.

Summarising, Spanish verse prosody is syllabic in that a significant amount 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 synaeresis 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. This is helped by the devices quoted above.

As observed in section 2.3., English is a stress-timed language, which would make us think of its verse prosody as accentual. Nevertheless, English verse is accentual-syllabic. The characteristics of PdE verse prosody have to do with its hybridity, inherited from a fruitful interaction between the native traditions and the Romance one. Old Germanic verse evolved from syllabic into accentual. The OE period – from mid. sixth century to beg. twelfth century – witnessed the flourishing of a highly complex system of verse composition, called ‘alliterative verse’. At the end of the twelfth century a number of new forms indebted to both Medieval Latin and Anglo-Norman models came into play. The Latin quantitative foot and the English accentual foot were connected through French, which was purely syllabic. The second important innovation was rhyme, which started to be used as a structural principle, borrowed probably from Medieval Latin, but also due to the influence of French, particularly certain Anglo-Norman forms. It was not until the fourteenth century that the attention of poets such as Chaucer shifted completely towards strictly French and Italian forms of composition, considered more elegant and cultivated. It can be said that, around this time, 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 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 salient one. However, synalepha in English is not a phonological device which then acquires a metrical function. It is, rather, a metrical device with non-phonological status. This has to do with the fact that, in English, vowels avoid diphthongisation, and diphthongs behave as independent units with specific characteristics, not as the result of the joining together of two single vowels. In fact, the latter do not show a correspondence between the metrical pattern and the actual scansion of a verse instance, or between the scansion of a verse instance and its pronunciation. Synalepha in English 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.

In conclusion, synalepha is a universal metrical device attested in many languages, which works in English as a correspondence rule which has the function of making the verse instance project as many syllables as required by the metrical pattern, and leave extra syllables not projected.

4. The relationship between speech, verse and music

In chapters 1, 2 and 3 I explored the connections between speech prosody and verse prosody. In this chapter I examine the triple relation between speech, verse and music, in order to discern whether there exists any kind of correspondence between speech prosodic typologies, verse prosody and musical rhythm. Section 4.1. analyses the interaction between verse prosody and musical rhythm, arguing that the characteristics of the former have important consequences on the latter. Section 4.2. deals with the main subject of this dissertation, namely text-setting, framing its study within OT. This section is intended as an introduction to the theoretical principles and research methodology which will work as the fundamental bases for the corpus analysis carried out in this dissertation. Each of the seven subsections contained in section 4.2. is devoted to the summary and critical review of an article dealing with OT as applied to metrical analysis, on the one hand, and to text-setting analysis, on the other. Section 4.3. introduces the main object of study in this dissertation, namely folk song. Section 4.3.1. establishes a definition for folk song and summarises its main characteristics. Section 4.3.2. explains the transcription and classification criteria applied to the corpus of songs analysed in this dissertation. Section 4.3.3. briefly introduces the topic of the importance of lyrics in the study of folk song, and section 4.3.4. complements this by exploring the general characteristics of text-tune relationships in folk song. Section 4.3.5. explains the reasons for the choice of a specific folk song corpus, namely Manzano Alonso's (2003) *Cancionero de Burgos* and Kennedy's (1984) *Folk songs of Britain and Ireland*. In section 4.3.6. I present some preliminary conclusions extracted from the analysis of mismatches in the corpus of 444 Spanish folk songs and 239 English folk songs (see also Appendices IV and V), while section 4.3.7. contains a deep analysis of two specific folk songs extracted from the above-mentioned corpus, as well as the conclusions that emerge from it. Section 4.4. is intended as a complement to sections 4.2. and 4.3., as it seeks to extend the theories explored in the former and then applied in the latter by checking them against art song. The internal structure of this section is parallel to that of section 4.3., in this case always referring to art song.

4.1. The interaction between verse prosody and musical rhythm

As has been argued in this dissertation, verse and music share a quality which natural speech lacks – “whereas music characteristically functions as art, language does not” (Lerdahl and Jackendoff 1983:7) – where ‘language’ stands for ‘ordinary spoken language’. As already stated, speech has a dominant instrumental dimension; in other words, it is normally used as a tool for transmitting meanings ‘outside’ itself, while verse, like music, shows a clear aesthetic dimension, that is, its main goal is calling the auditor’s or observer’s attention towards its form. Structuralism held the view that poetry and music are autonomous in that their meaning does not depend on the exterior world:

So far, this is simple and straightforward enough: practical language finds its justification outside itself, in the transmission of thought or in interpersonal communication; it is a means and not an end [...] Poetic language, on the other hand, finds its justification (and thus its entire value) in itself; it is its own end, and no longer a means (Todorov 1988:12).

Beyond the assertion that both poetry and music are art forms, poets have always accepted and even vindicated the inherent musical quality that poetry has:

Poetry is a composition of words set to music. Most other definitions of it are indefensible, or worse, metaphysical. The amount or quality of the music may, and does, vary; but poetry withers and ‘dries out’ when it leaves music, or at least an imagined music, too far behind it (Pound 1917:90-91).

Rhythmic devices articulate the flow of time in an intentional way, hence the aesthetic dimension that both poetry and music share: “Though the material of poetry is verbal, its import is not the literal assertion made in the words, but the way the assertion is made, and this involves the sound, the tempo... and the unifying, all-embracing artifice of rhythm” (Langer 1953:260-61). As Langer observes, the element that makes poetry and music such a close pair is rhythm. The pre-eminence of rhythm as an abstract property of measuring time not only shows the existence of a real affinity between poetry and music, but it also proves that poetry is the only real analogue for music in the realm of human language: “the principal definition of

musical objects lies in their rhythmic articulation. More than any dimension, rhythm articulates time and emphasises its forward movement” (Barry 1990:65).

The link between verse and music is bidirectional. Composers have shown an overwhelming predilection for using poetry, rather than prose, as the verbal material for their vocal works. If one has a look at the different existing vocal genres, one will find that one of the very few forms where poetry is not used is in the canticles of the Christian tradition of sacred Western music.

As regards the most characteristically vocal genres, the use of poetry is especially overwhelming in songs, which constitute an eminently lyrical genre. In a song the poetic articulations of metre and rhyme of a traditional lyric match and reinforce the musical articulations of bar and phrase. Thus, songwriters value the properties of pure sound of a lyric at least as much as its meaning: the patterns of stress suggest a metre, the rhymes define a melodic phrase and the phonetics inspires melodic shape. This is also true for opera, a dramatic genre in which composers often use libretti written in verse. In fact, one of the most intriguing relations between language and music in a genre like opera is the one which emerges from the setting of a text into a specific musical construct. Depending on the specific operatic subgenre, the setting of texts is utterly different. Just to mention two of the most important subgenres of opera, ‘recitative’ is a type of vocal writing, normally for a single voice, which follows the natural rhythms and accentuation of speech and its pitch contours; ‘aria’, on the other hand, refers to a song, either independent or part of a larger work, which results from the setting of poetry to a musical tune. Arias are the dominant elements in opera, while recitatives work as a vehicle for dialogue, that is, an essential means of expression for passages in a libretto for which lyrical treatment is inappropriate, as well as a connecting link between arias.

The rhythmical hierarchy that mediates between poetic text and music is metre. The most significant similarity between poetic metre and musical metre is that they are based on the alternation of strong and weak events, which can be stresses – as in English poetry – or syllables – as in Spanish poetry. According to this, the process of text-setting has to do with a matching between the metrical pattern of the chosen poem and that of the corresponding musical composition. Nevertheless, there is an essential difference between metre in poetry and in music which has to do with

the arrangements of strong and weak positions. As discussed in chapter three, in English there are a number of possible arrangements of stressed and unstressed syllables that result in a variety of feet. Unlike poetic foot divisions, musical bar lines make no distinction between iambic-trochaic, or anapaestic-dactylic, metres. Music avoids this apparently unreal dilemma by automatically inserting bar-lines, which always fall before the main beats (Attridge 1982:20).

As pointed out in section 1.1., the metrical organisation of a musical composition is signalled in the time signature and measured by the bar lines. When poets set verse to music they must take this into account: "Stressed and unstressed syllables in 'sung' forms line up in mandatory ways with strong and weak beats in the musical metre" (Cureton 1992:46-47). According to this, the musical quality most widely used to reinforce linguistic stress is metrical position. Primary counts, that is, the first and third pulse in 4/4, and the first one in 3/4, are usually made to correspond to stressed syllables; secondary counts, that is, the second and the fourth beats in 4/4, and the second and third beats in 3/4, are made to correspond either to stressed or unstressed syllables. Last, submetrics, that is, positions which do not coincide with a count, correspond to unstressed syllables.

In sum, text and music are associated with metrical patterns, which are hierarchical organisations of strong and weak elements. The condition on the association of text and music is that their metrical patterns be congruent. One must be superimposed on the other without conflicts of metrical positions and stress. Nevertheless, there is always a controlled level of tension between stress patterns and poetic metrical schemes, which is important not only for poetical analysis – as was explained in section 3.1., good poetry always shows a certain degree of tension – but also because of its implications for a musical setting – as we will discuss in 4.2., 4.3. and 4.4., tension can be a device used on purpose in vocal music.

In what follows, I shall focus on the analysis of text-setting in two languages that, as observed in this dissertation, show different rhythmical characteristics, both in ordinary speech and in verse. According to the hypothesis stated at the beginning of this dissertation, the different speech rhythms of Spanish and English condition the verse prosody in these two languages and, in turn, the different approaches to sound and rhythm in English and Spanish poetry may be seen to have a correlation to

the differences in the processes of musical setting of the two cultures. As a first step, I introduce the theoretical framework that will later guide the empirical analysis of songs. This framework is related to OT, more specifically to the OT constraints explored for metre, on the one hand, and for text-setting, on the other. Thus, section 4.2.1. – Janda and Morgan (1988) “El acentó dislocadó – pues cantadó – castellanó” and Morgan and Janda (1989) “Musically conditioned stress shift in Spanish revisited” –, although not really framed within OT, summarises the contents of the only two articles which, to my knowledge, deal with text-setting in Spanish. Section 4.2.2. – Hayes and Kaun (1996) “The role of phonological phrasing in sung and chanted verse” –, section 4.2.3. – Hayes and MacEachern (1996) “Are there lines in folk poetry?” –, section 4.2.4. – Hayes and MacEachern (1998) “Quatrain form in English folk verse” – and section 4.2.8. – Kiparsky (2006) “A modular metrics for folk verse” – are devoted to the analysis of OT metrical and grouping constraints as observed in English verse and song lines, while sections 4.2.5. – Hayes (in press) “Text-setting as constraint conflict” –, 4.2.6. – Dell and Halle (in press) “Comparing musical text-setting in French and English songs” – and 4.2.7. – Halle (forthcoming) “Text, tune and metrical form” – are specifically focused on the study of text-setting constraints mainly, but not exclusively, in English songs. Section 4.2.9. summarises the principle findings and conclusions that will guide the empirical analyses in the following sections.

4.2. The text-setting problem in the existing literature

The study of verse and music, as well as of their interaction in song, has contributed to the evolution of suprasegmental phonology. Metrical Phonology (Lieberman 1979) was the first to devise an interdisciplinary methodological apparatus which used tools borrowed from music in order to analyse speech prosody. On the other hand, generative linguistics became the basis for one of the most groundbreaking theories of tonal music (Bernstein 1976; Lerdahl and Jackendoff 1983), which in turn added some fruitful theoretical and methodological elements to the original, thus becoming a precursor to OT (Prince and Smolensky 1993). Since the appearance of Prince and Smolensky’s (1993) *Optimality Theory: constraint interaction in Generative Grammar*, OT has gradually gained ground as one of the major theoretical

frameworks in segmental phonology. More recently, it has expanded to cover the often marginal field of suprasegmental phonology. In the last ten years several important papers on the relationship between linguistic prosody, verse prosody and musical rhythm have been written by linguists such as Hayes and Kaun (1996), Hayes and MacEachern (1996, 1998), Hayes (in press), Halle and Dell (in press), Halle (forthcoming) and Kiparsky (2006). The cited papers share several important characteristics. First, they all deal, to varying degrees, with issues related to text-setting or, put differently, with the conflicts that arise when trying to align text and tune. Second, with the exception of Halle (forthcoming), which approaches the matter from a purely generative perspective, all of them use OT as their theoretical frame. Third, they are primarily concerned with folk verse and song as the object of their empirical study and, with the exception of Dell and Halle (in press), which presents a comparative analysis between folk songs in English and in French, they all analyse folk verse and song in English exclusively. The fact that most of these papers deal overwhelmingly with English verse and songs makes their theoretical claims biased and not powerful enough. Therefore, it is crucial to analyse materials in other languages in order to be able to check the validity of the general theoretical assumptions made by OT in the fields of metrics and text-setting. Fourth, they share the main working theory underlying text-setting studies, namely that, in vocal music, three tiers of rhythmic structure come into play: (i) linguistic prominence and grouping, (ii) verse metre and grouping and (iii) musical rhythm and metre. The questions which naturally arise from such a theory are, among others, (i) what is the relation between the three tiers of rhythmic structure? (ii) which constraints are applied to text-setting in different languages? (iii) are they ranked equally in different languages?

An OT approach to text-setting clearly aims at examining whether the setting of verse to music responds to some kind of underlying rhythmic constraints common to language prosody, verse prosody and music, as well as checking whether those constraints are ranked differently from language to language, and what implications this may have for the study of prosodic investigation. Vocal music can, thus, contribute to a deeper knowledge and understanding of issues related to the long-discussed dichotomy between stress- and syllable-timed languages and, more

generally, the generally neglected consequences of rhythm in language variation and change (see Schlüter 2005).

Before going on with the discussion, it is necessary to clarify two concepts which tend to recur in every linguistic analysis of text-setting. The first key concept is that of ‘song’, defined by Dell and Halle (in press) as “a composite which combines two objects each with its own structure, a linguistic object – text – and a musical object – tune”. The relationship between those two objects is mediated by text-setting, a concept explored by Halle and Lerdahl (1993), which has to do with how lines of linguistic texts are arranged in time against a predetermined rhythmic pattern in sung verse.

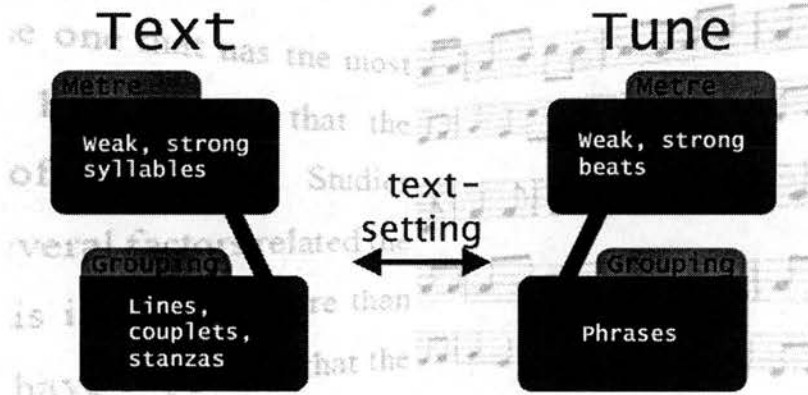


Table 27: Text-setting mediating between text and tune.

As can be observed in Table 27, both text and tune share the characteristic of being arranged hierarchically into two main levels, namely a metrical level – where strong and weak units alternate – and a grouping level – where small units are arranged into bigger constituents. Text-setting operates on those two levels simultaneously, facilitating the interaction between text metre and tune metre, on the one hand, and text grouping and tune grouping, on the other. As pointed out by Hayes and Kaun (1996:10), the existence of a series of constraints (CON) is central to any OT study of text-setting. In OT terms, a series of CON are applied at the metrical level, thus regulating how text metre and tune metre can interact to render a well-formed metrical composite. This interaction will be different in different languages, and will have a direct impact on the ranking of CON at the grouping level, which regulate how text groups and tune groups are arranged to yield a well-formed composite. In

turn, the ranking of CON at the grouping level will have direct consequences on the ranking of metrical CON.

In the following sections, I review the most important aspects in the above-mentioned papers, paying particular attention to the set of metrical and grouping CON which are relevant for the analysis of text-setting in different languages. By reviewing the most recent literature in the topic, I will be in the position to highlight both its values and its flaws in order to then be able to offer a more comprehensive approach to the matter at hand.

4.2.1. Janda and Morgan (1988) “El acentó dislocadó – pues cantadó – castellanó” and Morgan and Janda (1989) “Musically conditioned stress shift in Spanish revisited”

The study of text-setting from a linguistic point of view started to be given serious consideration in the late eighties, after the publication of two papers by Janda and Morgan (1988), and Morgan and Janda (1989). After an exhaustive process of research on the topic, Janda and Morgan (1988:151) list the first observations of what they call “stress-shift” in Spanish music. In 1887, Eduardo de la Barra (1887:19, in Janda and Morgan 1988:151) was the first to point to a stress-beat mismatch in a song, namely that of the preposition *cóntra* (‘against’) in the Chilean national anthem. Almost a century later, the Real Academia Española’s new grammatical sketch (1973:66(7), in Janda and Morgan 1988:151) briefly mentioned the mismatch phenomenon in Spanish song, defined in that text as an “anomaly”. In their 1988 article, Janda and Morgan try to develop a systematic, non-impressionistic analysis of such a phenomenon in order to determine the linguistic consequences of the systematic mismatch between stresses and beats in Spanish music. They elaborate a fully comprehensive list of mismatches found in a variety of songs (1988:153), written from every century since the fifteenth century, original from twenty-two countries, and from a variety of musical styles. According to the authors (1988:160), there are two main consequences of stress-beat mismatches, namely phonetic and phonological/morphosemantic consequences. Within the latter, Janda and Morgan (1988:160) distinguish two major effects of musically-conditioned stress-shift. On the one hand, stress-marked distinctions are levelled; on the other, there is a violation of the conditions on syllable structure and stress proposed by

Harris (1983), according to whom antepenultimate stress is not supposed to happen in words whose penult contains a branching rhyme, which is possible in vocal music.

Janda and Morgan (1988:161-168) put forward four hypotheses for the phenomenon:

a) Null hypothesis: whether or not a language allows musical stress-shift is a parameter of linguistic variation which must be arbitrarily stipulated (1988:161).

b) First hypothesis: the notions of 'syllable-timing' and 'stress-timing' play a crucial role (1988:161-162). This cannot be fully proved, given that there is shift in stress-timed languages, too – for instance, Brazilian Portuguese, Catalan, Russian.

c) Second hypothesis: the crucial consideration in the presence or absence of musical stress-shift in a language is the relative availability of certain word-types which facilitate or at least accommodate the matching of texts to tunes (1988:163). This refers, for instance, to the realisation or not of schwa in Lusitanian Portuguese and French, or the apocope of final vowels in Italian.

d) Third hypothesis: the fact that stress is predictable, and thus recoverable, makes stress-shift a possible device, even in a language like Spanish, where stress minimal pairs are much more common than in English. Also, the frequency of polysyllabic words forces the appearance of stress-shift (1988:167-168).

Morgan and Janda's (1989) paper revises and extends the observations made in 1988 and, more importantly, re-defines the role of speech prosody in musically-conditioned stress-shift. The authors (1989:274) make a slight change in their approach to the topic and assert the importance of the syllable (if not syllable-timing) in the process of stress-shifting in song.

One of the most important differences between linguistic and musical stress is, according to Morgan and Janda (1989:277), the lack of cues for the latter. They allege that pitch, duration and loudness do not work as cues for stress in music, and that the only cue is timing itself: "ordinarily unstressed syllables which coincide with the musical downbeat are not perceived as being stressed as long as the normally stressed syllable has a musical/rhythmic accent which is at least as pronounced".

In spite of the observed plethora of mismatches in Spanish song, Morgan and Janda (1989:282) point to the fact that “the correspondence of musical stress to linguistic stress is demonstrably desirable even in Spanish, given both the other linguistic processes (synaeresis and dieresis, for example) which are employed in order to secure it and the rejection of certain deviant stresses by some speakers”. A further consequence of this has to do with the question “of whether [...] the correspondence of musical phrases to real syntactic units is more important than accentual integrity” (1989:282).

The difference between English and Spanish prosody seems to have an effect on the quantity and quality of mismatches allowed in both languages. As Morgan and Janda (1989:284) observe,

In order to sing an English text successfully, it is important only that intralexical stress not clash with the metrical pattern of the tune at whatever division or subdivision of the whole melody is under consideration [...] Monosyllables can be squeezed into strong or weak positions as necessary. In Spanish [...] we are able to ignore metrical tree structure of texts even within words. Text/tune matching begins with erasure of all lexical stresses and proceeds mechanically by left-to-right, one-to-one association of syllables with notes of the tune.

The conclusion drawn from these two papers has to do with the prosodic conditioning of text-setting. As already mentioned, the main rhythmic unit in Spanish is the syllable, while the main unit in English is stress.

4.2.2. Hayes and Kaun (1996) “The role of phonological phrasing in sung and chanted verse”

Hayes and Kaun’s (1996) paper represents an attempt to link phonological structure to metrics and, more importantly for the purposes of the paper, to text-setting. As the authors (1996:4) point out, “[t]he ideas of metricality and complexity carry over fairly straightforwardly to sung and chanted verse”.

The authors (1996:2) start by stating the fact that rhythmic beats in singing or chanting are isochronous in nature. This statement is, nevertheless, quite inaccurate from a musical point of view, for two main reasons. On the one hand, the isochronic nature of rhythmic beats in singing, just like in instrumental music, depends on the

musical era and genre. Thus, for instance, saying that Gregorian chant shows 'isochronous beats' would certainly be incorrect, and the same would apply to certain forms of contemporary art-music, as well as certain styles of folk song across the world. On the other hand, the authors do not really explain what they mean by 'chanted verse'. By this, we assume that they are referring to a type of flat, inexpressive declamation of a line in order to make its positions fall onto the metrical positions of an abstract grid. In this respect, chanted verse and sung verse are different in nature. The former would be the result of deleting everything but the beat from sung verse – that is, neglecting its melodic contour, expressive devices, accompaniment, and so on, while the latter would be the result of the interaction of all the afore-mentioned musical and textual elements. What the authors are testing in this paper is the degree of metricality of specific instances of predominantly chanted, rather than sung, verse. By 'metricality' they mean the extent to which a specific deliverance of a line corresponds to a pre-designed grid, which, for the purposes of Hayes and Kaun's (1996) paper, is always binary. In other words, the degree of metricality of a specific text-setting is its degree of well-formedness. Although the authors are eliminating any possibility of sharp variation in the settings yielded by the informants – as pointed out, they only accept binary settings –, they acknowledge that setting a text in a specific way is a human choice, which implies that "the variability in text-setting usually represents a spontaneous artistic decision" (1996:5). In this respect, when singers improvise text-settings, they are selecting from the well-formed options in that specific tradition. Hayes and Kaun make the strong assumption that text-setting is a universal ability. The argument used by the authors (1996:5) to prove that every individual has text-setting abilities is that "music publishers often provide the text-setting of only the first verse of a multi-verse song", which implies that the remaining verses must be set by the singer at the moment of performance. This is arguable, since the skill of a singer to set a text to a particular tune depends on many more variables than an innate ability comparable to that explored by linguists for natural language.

As mentioned above, Hayes and Kaun study text-setting as a phenomenon independent of melody, leaving the issue of how the tune influences text-setting for further research. As already pointed out, it seems misleading to think that text-setting

occurs at two distinct levels, namely the metric level and the melodic one. I think that metre and pitch interact in such a way that the analysis of text-setting will not be complete if we isolate one of these two components. In fact, interval and pitch itself are important for the metre, and the melodic contour of songs is decisive sometimes in order to judge the well-formedness of a setting. As Tovey (1979:212) observes,

even with the full muscular strength of modern musical rhythms, the strong beats of the bar constitute only one form of accent, which may be easily counteracted by the length of a note, by its height, by its harmonic colour, and by incidents in the accompaniment.

In order to test their hypothesis that native speakers of a specific folk tradition have intuitions about text-setting, Hayes and Kaun choose a group of informants to whom they give some lines in order to be chanted. As readers, we are confronted with the doubt of whose intuitions the authors are checking. The informants used in this piece of research are a group of twentieth-century – probably middle class – Americans exposed to lines extracted from a collection of English and American folk songs collected at the beginning of the twentieth century by a group of ethnomusicologists of that time, such as Cecil Sharp. In this respect, I would also like to observe that the authors acknowledge the biased character of the original songs, since they were collected and transcribed by ethnomusicologists, such as Sharp and Karpeles, who imposed a grid on an otherwise ‘shapeless’, i.e. not necessarily metrical, chain of pitches. It is a well-known fact that traditional singers do not always – in fact, most of the time they do not – sing metrically – hence my previous observation in relation to Hayes and Kaun’s statement that all music is organised isochronously. If Hayes and Kaun accept that the original transcriptions of the songs which they use respond to Sharp’s own intuitions when he transcribed them, which clearly differed from the original singers’ intuitions, I wonder whether one can be certain about the validity of the authors’ strong hypotheses about the native nature of text-setting abilities. Sharp adapted folk song for a salon, middle-class audience, adding, for instance, piano accompaniments – and the consequent metrical tightening – to an otherwise purely vocal musical genre, which would most probably have been sung in an *ad libitum* fashion. Evidence for this can be found in the various sound

archives collected by people such as Alan Lomax and Peter Kennedy in the second half of the twentieth century.

As pointed above, another important observation in relation to Hayes and Kaun's experiments has to do with the requirement that all informants set the lines in question to a binary rhythm. Real intuitions are thus clearly biased, to the extent that "ternary scanners were gently corrected" (1996:8-9), that is, only those informants who set the lines to a binary metrical grid were taken into consideration for the purpose of the paper. The fact that the authors had to 'correct' some informants contradicts their assertion that "no obvious effects of gender or musical education" (1996:8) were found in the text-setting patterns. Surely, a musically educated informant would have known what a binary rhythm, as opposed to a ternary one, is, whereas some of the non musically-educated informants would have found it problematic to differentiate between a binary and a ternary rhythm and, therefore, to succeed in their binary settings. (167) and (168) illustrate the difference between a binary and a ternary rhythm.

(167) Binary rhythm: the odd beats of each binary unit are stronger than the even beats; the first beat of the highest unit is always the strongest one.

2/4:

X	
X	X

4/4:

X			
X		X	
X	X	X	X

(168) Ternary rhythm: the first beat of every group of three is strong, while the other two are weak.

3/4:

X		
X		
X	X	X

In conclusion, we cannot really say that this paper is checking innate intuitions. The authors seem to be judging whether the chant given by a specific informant is really intuitive or just too affected. Also, they seem to use the informants to prove a theory which they have previously articulated without actually needing to resort to any kind of field work.

If we refer to the notation used, the settings yielded by the informants are notated in such a way that the reader visualises the text-setting structure, the weight of stressed syllables, the stress pattern and the phonological structure (P-structure, a term borrowed from Selkirk 1984), which comprises the hierarchisation of the components into words, clitic groups, phonological phrases and intonational phrases (see Hayes and Kaun's example in 1996:8).

There are a number of key ideas that the authors explore in order to judge the degree of metricality of the settings yielded by their informants. The first central idea in relation to the well- or ill-formedness of text-settings is the distinction between full and truncated text-settings – truncation is related to SALIENCY⁵⁹ –, a concept explored further in a later paper by Hayes and MacEachern (1998). Truncated settings are those in which grids do not have all their strong metrical positions filled with syllables, which makes them salient. There are four main classes of truncated settings, of which the two that are most often used in English song are the following:

- (i) A setting with no syllables between the third and fourth strong positions, so that the syllable in the third position is sung quite long. Hayes and Kaun (1996:9) call this the 'Green O' truncation and exemplify it as follows:

⁵⁹ OT has adopted the typographical convention of capitalising the names of constraints. I am adopting this convention in the sections that deal with text-setting from an OT perspective.

(169)

		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
A-	móng		the		léaves		so		gréen					O

- (ii) A setting with the fourth strong position unfilled, in such a way that the line-final syllable is sung very long, or alternatively, followed by a substantial pause. Hayes and Kaun (1996:10) call these 'three beat' (or '3') lines, with a reminder to the reader that the silent fourth beat remains detectable in the timing:

(170)

		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
I	sówed		the		séeds		of		lóve					

Hayes and Kaun associate the non-filling of metrical positions in the grid with the concept of 'extrametricality', thoroughly explored in verse prosody studies. Nevertheless, it is essential to bear in mind that this concept, which depends on theories of prosody, cannot be equated – as it is in this paper – to musical feminine endings, which have the shape S W, just like extrametrical structures, but where the W beat is completely metrical.

The second nuclear idea in the paper is related to the MAXBEAT CON and, more specifically, with MATCHSTRESS (1996:10), whereby the rises and falls of stress within the line are matched to the rises and falls of the metrical pattern. In all studies of text-setting in English, the correspondence between linguistic stresses and musical strong beats is absolutely essential when it comes to judge the well-formedness and, therefore, the quality of a specific text-setting. Nevertheless, as any reader with a knowledge of folk songs can predict, there are many exceptions to that

one-to-one correspondence between stresses and beats, something of which Hayes and Kaun are aware. I will explore some of these exceptions in the next pages.

As a sort of complement to the idea of truncation in combination with MATCHSTRESS, Hayes and Kaun (1996:11) adapt the so-called Zero Provision rule – used in a musical context by Lerdahl and Jackendoff (1983) – as follows:

when a metrical position is unfilled, or – equivalently – filled by the phonemic continuation of the preceding syllable, treat it as being filled by a stress level that is weaker than that of a stressless overt syllable.

This constitutes a systematic explanation of the fact that any unrealised position must obviously belong in the weakest metrical level in the grid.

This means that no unfilled position can be assigned a stronger beat than a preceding filled position, as we can see in examples (169) and (170) above.

The novelty of this paper is the overt link that Hayes and Kaun (1996:13) draw between metrics and musical prosody, on the one hand, and language P-structure, on the other. In fact, they clearly state that P-structure influences the metrics of the language in question. They summarise this hypothesis by observing that in all metrical traditions there is a requirement that line boundaries coincide with high-level breaks in P-structure, that is, the correspondence between speech, verse and tune grouping – put more precisely, the correspondence between P-phrases, lines and musical phrases – must be as high as possible. As will be observed in Hayes and MacEachern (1996, 1998), Dell and Halle (in press), Halle (forthcoming) and Hayes (in press), the concept of ‘line’ is central to any study of text-setting. While this requirement can vary in strictness, they also observe that, in folk song, the agreement between P-structure and metrics is high. With regard to the interaction between P-structure (grouping) and MATCHSTRESS (metre), Hayes and Kaun (1996:14-15) state that “metrical rules refer to P-structure in just the same ways that rules of phonology do”, and that, a priori, the principle known as “beginnings free, endings strict” applies to text-setting processes, too. As its name explicitly indicates, this principle states that the correspondence of linguistic material to a metrical pattern tends to be free at the beginnings of units, and strict at the ends, something tested for written verse. However, its validity in the field of song needs to be tested.

Although Hayes and Kaun deal exclusively with metrical issues, leaving aside all other variables, they mention durational effects in text-setting, where phonemic duration is said to be determined by weight – a phonological variable –, stress level – particularly the stress distinctions that determine vowel reduction –, segment identity – lower vowels tend to be longer – and location of the syllable within P-structure. In this respect, Hayes and Kaun (1996:16) establish a foundational rule schema for text-setting, called ‘Syllable Duration Rule’.

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

According to this, P-structure relates to syllable duration in the following way: the higher-ranking the right edge in P-structure at which a given syllable appears the more it is lengthened. In this respect, “a high level P-structure right edge (such as that of an I-phrase) should induce substantial lengthening on the preceding syllable, and induce a strong preference for that syllable to receive multiple beats” (1996:17). Hayes and Kaun base this idea on previous studies by Ladd and Campbell (1991) and Wightman *et al.* (1992). Summarising, Hayes and Kaun (1996:24) defend the view that P-structure is a principal determinant of phonetic syllable duration. Owing to the Syllable Duration Rule, syllables in English text-setting are characteristically matched to appropriate numbers of beats. This seems to be too strong a statement, since its total validity would imply that no short vowel would ever be set to a long-value note in music, just as no long vowel would be set to a short-value note in music, something which is clearly not the case. Let us have a look at one of the numerous exceptions found in Kennedy (1984:348):

(171)

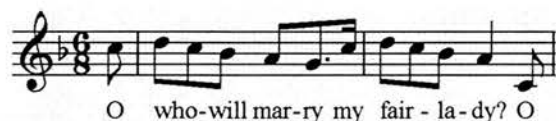
The cuc-koo is a pret-ty bird, he sings as he flies

In (171) we observe that a light syllable like *the*, with a schwa as its nucleus, is made to correspond to a crotchet – the equivalent of two metrical positions –, while a heavy syllable like *-koo* in *cuckoo*, with a long /u:/ as its nucleus, is made to correspond to a quaver – the equivalent of one metrical position. On the other hand,

it is generally true that the ends of musical phrases tend to show longer note-values than any other position – in this example, we can see that the last syllable of the line, *flies*, is assigned to a minim, which is the longest note-value found in the phrase.

Hayes and Kaun analyse a number of settings which represent an exception to MATCHSTRESS. All of those exceptions are related to the concept of ‘mismatch’ as defined in Dell and Halle (in press), which Hayes and Kaun (1996:36) call ‘lexical inversions’. A lexical inversion is defined as “a word internal /oó/ sequence that is mismatched against the metre” (1996:36). As mentioned above, the focus of Hayes and Kaun’s study is “quasi-evenly” bracketed binary structures, such as those of iambic pentameters. According to their observations, “the great majority of lexical inversions in iambic pentameters occur line-initially, and of the remainder, most occur in the third or fourth foot” (1996:36). This seems to confirm the aforementioned rule of “beginnings free, endings strict”. Nevertheless, in folk song there are instances where the lexical inversion happens at the very end of a line – or musical phrase, for this matter. Let us have a look at (172), recorded in Kennedy (1984:653):

(172)



As we can see in (172), the last lexical word of the line in question, *lady*, constitutes an example of a lexical mismatch, since its stressed syllable, *la-*, falls on a weak, single metrical position, while the unstressed syllable *-dy* falls on a strong position, and it is made to correspond to two metrical positions. In this case, neither stress nor duration seems to be reflected as Hayes and Kaun predict in their model.

The explanation to the lexical mismatch phenomenon is linked in Hayes and Kaun (1996:42) to the afore-mentioned concept of ‘truncation’, which typically occurs at the ends of lines. As already pointed out, in a musical setting, extrametrical syllables do not make sense, since there is an isochronous, constant flow of metrical positions. As a result, unfilled positions are freely tolerated at the ends of lines, and the surplus of positions gives the text setter some freedom not available to the spoken

verse poet. Parallel to this, there seems to be a higher degree of mismatch freedom at the ends of lines than at the beginnings. All these facts make us conclude that theories of verse prosody are not the same as theories of musical prosody. A macro-theory which explains the interactions between the two is clearly needed.

As their last observation, Hayes and Kaun (1996:43) insist, once again, that sung verse is performed in isochronous rhythm, with the following consequences:

- a) Substantial constraints are imposed in sung verse requiring a match-up of natural syllable durations with beat count. Non-isochronous spoken verse permits great flexibility of syllable durations.
- b) The metrical pattern for sung verse is in principle continuous.
- c) There is no such thing as an extrametrical syllable in sung verse.

Since I have already manifested my disagreement with some of the above conclusions, I do not consider it necessary to repeat my observations. However, I would like to refer to an interesting paper by Fitzgerald (1998), where a study of the metre of Tohono O'Odham songs from southern Arizona reveals that in certain musical traditions the beginnings of lines are not lax at all but, on the contrary, very strict indeed (1998:31). In any case, a revision of typological claims based on criteria related to the edges of lines seems to be necessary, as well as a separation between the requirements of purely poetic lines and those of lines set to music.

Hayes and Kaun (1996:43) end up summarising their findings in OT terms: in sung verse most candidate lexical inversions are said to be beaten out by rival text-settings that slide the lexical inversion over into a matching position.

Although the originality of this approach is unquestionable, the fact that sung and chanted verse are regarded as equal points to an excessively positivistic view of a clearly cultural phenomenon, such as folk song. There is no cultural allusion whatsoever in this paper, something which can actually be academically beneficial in certain linguistic circles. But there is no doubt that some of the phenomena observed would benefit immensely from a broader, more interdisciplinary perspective. For instance, what Hayes and Kaun call 'lexical inversion' is in fact a stylistic marker of English folk song, and, therefore, an idiomatic way of building allowed mismatches. This is indeed not universal and not specific to English music in general; it is simply a folk song stylistic marker. For a linguist/musician, escaping the cultural side of any

study on folk song, folk verse, or text-setting in general, is misleading. As an illustration to this point, Hayes and Kaun use the term ‘Scotch snap’ for a specific type of mismatch where an S position is displaced – in musical terms, the realisation of a syncope. Nevertheless, the use of that term is musically inaccurate, since a Scotch snap is too quick to sing anything to it, as it is a dance figure.

(173)



A more accurate name would be ‘Lombard rhythm’ – although this term is sometimes used as a synonym for Scotch snap –, a syncope figure used in eighteenth-century opera – and, therefore, a vocal figure.

This paper offers an original, although not scientifically convincing, view of what text-setting really is, as well as which mechanisms are at work when singers of a specific music tradition set lyrics into music.

4.2.3. Hayes and MacEachern (1996) “Are there lines in folk poetry?”

Hayes and MacEachern’s (1996) paper is “part of a research program to study the metres and stanza structures of English folk verse” (1996:1). The authors start with their nuclear hypothesis about the existence of the line as a structural element not only of art-verse, where literate poets use lines as part of their conscious intentions – the evidence for this is that poets arrange lines typographically as units on the page – but also of folk verse. In order to prove the existence and structural importance of the line in folk verse, the authors analyse a corpus of 1028 Southern Appalachian folk songs described in a later paper by Hayes and MacEachern (1998).

In methodological terms, this paper resembles Hayes and Kaun (1996). The method used in their analysis of the songs is based on the metrical grids proposed as a basis for rhythmic analysis by Liberman (1979) and Lerdahl and Jackendoff (1983: chaps. 2-4). For each song in the corpus, Hayes and Kaun establish a grid which can accommodate the rhythm of the song. Once again, silent beats are observed and notated along filled positions, as seen in examples (169) and (170), and the concept of truncation remains essential since “line-final truncation serves to highlight specified levels of metrical constituency” (1996:2).

The hypothesis for the existence of lines put forward by Hayes and MacEachern (1996:3) is the following:

We are claiming that lines exist; that is, that they are authentic structural units of the system. We are also claiming that each line possesses its own distinct stretch of grid, and that a song-specific choice is made for what this stretch is. None of these claims is obviously true, because in the audible form of a song the grid is essentially continuous.

What the authors intend to show is that the virtual chopping of the quoted continuous grid is not arbitrary but follows specific rules.

Hayes and MacEachern's (1996:3-4) hypothesis relies on the "presence of large breaks in the P-structure". In this way, each intuitive line of a song corresponds to a phonologically unified string. According to the authors (1996:4), if one takes almost any song and aligns it with the correct "infinite" rhythmic grid, it will be possible to cut the grid into perfectly equal pieces such that the beginning and the end of each piece will coincide with a large break in the P-structure of the associated verse. Each line of each quatrain could plausibly be thought of as a separate Intonational Phrase. The text of a line is usually bounded by that line's associated grid, which, from a musician's perspective, would be easily explained by saying that each line in a stanza corresponds to the repetition of the same musical phrase. As a second observation, the authors (1996:4) point out that "there seems to be some pressure for the beginning of the line actually to *abut* the line boundary". They observe that grids tend to be filled-up to the very edge at the beginning of lines, in which case the grid starts with a weak position, be it with stressless or stressed syllables, so that the grid starts with a strong position. After analysing the four verses of the traditional "Shady Grove", which are arranged to start on a strong metrical position even if this means to drop certain syllables at the beginning of the line, they conclude that "[l]ine-initial function words are dropped in all cases where it would not lead to unintelligibility, probably because English phrases characteristically start with stressless function words" (1996:6-7). An example of this would be the common dropping of the pronoun *I* at the beginning of lines. The opposite phenomenon, that is, the insertion of an extra weak syllable to fit the grid, can also take place. Such is the case of the archaic verbal prefix *a-*, which is extensively used

in folk song in order to fill certain beats at the end of bars. However, as observed in Hayes and Kaun (1996), there is a small minority of lines that might conceivably start with the strong position, thus failing to fill the left edge of the grid. In any case, these exceptions always allow for a redistribution of the positions to make the strong syllable fall on the first strong position in the grid. In sum, “a song organised according to a particular grid will typically avoid placing the syllables of a linguistic line – defined by phonological phrasing – outside the grid space allocated to that line” (1996:7). What is more, there is a strong tendency for the linguistic line to begin at a point precisely coinciding with the beginning of the line grid. Again, musicians may not be so sure about this statement since it overlooks the high number of anacrusis found not only in folk song but also in art song. I will come back to this in 4.2.6., devoted to Dell and Halle (in press).

As mentioned in Hayes and Kaun (1996), the right edge of the grid seems to be more problematic than the left edge. Hayes and MacEachern (1996:7) wonder why there is not similar pressure for lines to end at the ending of the line grid. The explanation given for the lax nature of the right edge of the grid in this paper (1996:7) goes back to Hayes and Kaun (1996), who argue that, since sung verse tends to match the number of beats assigned to a syllable to the natural phonetic duration of a syllable and line-final syllables tend to be phonetically long, it will be logical to place line-final syllables not too close to line ends. Nevertheless, this statement seems to be an overgeneralisation since there are plenty of examples where longer syllable duration does not necessarily correspond to a higher number of beats, and vice versa. Let us have a look at (171) again, repeated below as (174):

(174)



The cuc-koo is a pret-ty bird, he sings as he flies

Hayes and MacEachern complement their explanation with a short allusion to the theory of “rhythmic cadences”, developed further in Hayes and MacEachern (1998), which allocates the end of the line grid to silent beats, a hint that serves to highlight the constituency of the quatrain. Summarising, the length of line-final

syllables, together with the action of rhythmic cadences, would tend to block the filling of the right edge of the line. For these reasons, Hayes and MacEachern (1996:7) conclude that “we cannot expect the ends of lines to abut the ends of their grids”.

Summarising the first part of the paper, the authors (1996:10) observe, once again, that the distinct behaviour of songs using grids from the same infinite pattern, both in their characteristic line beginnings and possible line endings, indicates a reality for the line as a tacit unit of folk verse composition. Secondly, the method of inspecting the alignments of grids and P-structure seems to be the most precise for locating line boundaries, a statement reinforced by the fact that rhyming positions turn out to be overwhelmingly line-final. Actually, as observed in Hayes and MacEachern (1998), the organisation of silent beats at the ends of lines seems to follow certain fixed patterns. In spite of the strength of their hypothesis, Hayes and MacEachern (1996:10) recognise that, as regards line ends, “[n]either of these diagnostics locates the line boundary with the precision obtained from inspecting line beginnings, but they virtually never diagnose a different parse into lines”.

In the second major section of their paper, Hayes and MacEachern (1996:10) analyse lines in the metrical hierarchy, where they “form only one layer in a quite deep constituent structure that embraces whole stanzas”. They point out that the vast majority of songs analysed in their study could be assigned a structure where all the layers are binary: [quatrain [couplet [line [hemistich [dipod [foot [beat b] f [b b]] d] h [d [f [b b] f [b b]] d [f [b b] f [b b]]]]] l [h h]] c [l [h [d [f [b b] f [b b]] d [f [b b] f [b b]]] h] l [h [d [f [b b] f [b b]] d [f [b b] f [b b]]] h [d [f [b b] f [b b]] d [f [b b] f [b b]]]]]]]

Hayes and MacEachern point out that the evidence used to argue for such a structure coincides with the evidence for the existence of lines: “agreement of metrical constituency with phonological constituency, the patterning of silent beats at the ends of constituents, and rhyming” (1996:11). The hierarchical structure seen above allows us to avoid the “infinite grid” as an element of the theory of metrics, and makes it possible to attribute the similarity of line grids to “shared constituent structure” (1996:11).

Following these observations, Hayes and MacEachern (1996:13) accept the existence of exceptions associated with those clear-cut statements. They observe that, while certain songs show a statistically dominant partitioning, there can be one or two places in which P-structure and line boundaries disagree. The authors argue that the exceptional cases involve what they call the “theft” of positions by a specific linguistic line either from the preceding or the following metrical line. Their view that such cases are exceptional and non-threatening to their theory is supported by their statistical rarity in the corpus that they use. Among the 1028 Southern Appalachian folk songs in their corpus only 107, that is, 10.4%, involved those “thefts” of metrical positions. On the other hand, the cases of songs in which the line boundaries were so fluid as to defy the establishment of a grid norm for line division were really rare. Another argument used to qualify the existence of exceptions is that there is a limitation on how much linguistic material may be stolen: “Leftward theft never takes more than the last strong beat of the preceding line, and rightward theft virtually never reaches even the first strong beat of the following line” (1996:14). Let us observe this phenomenon in (175), taken from Kennedy (1984:503):

(175)

There was an old man and he lived in the east And his

In (175), the syllables *And his*, which belong in the second line of the lyrics, are “stolen” and made to belong in the last bar of the first musical phrase. However, we could argue that, given that the first line begins with an anacrusis, which cannot have been stolen from any other position, in this case the grid is structured from the very beginning of the song so that the first syllable(s) of each line corresponds to the last beat – or, exceptionally, to the last two beats – of each 3/4 bar.

The authors (1996:15-16) go on to explain that the patterning of ‘theft of positions’ can be characterised more precisely. According to them, there are two basic types of ‘theft’. The first one is the theft of a strong position from the preceding line, a pattern that is consistent throughout all the lines of the song in which it occurs. The second type is the theft of a weak or medium-strength position from the

following line; this type is sporadic, even within the song in which it occurs, and its primary function is to accommodate specifically problematic texts, such as linguistic lines with feminine endings or lines with more initial pretonic syllables than the grid can handle – such is the case of (175) above. As we can see in (175), the beginning of the second line of the first stanza does not coincide with the beginning of a bar. The first two syllables of this stanza are made to follow on the last beat of the previous bar. In Hayes and MacEachern's terms, those two syllables, both weak, are 'stolen'.

To the question of why theft of strong positions only occur in the leftward direction, Hayes and MacEachern (1996:16-17) argue that, in English at least, "it is only in the leftward direction that the positions are there to be stolen". This idea is linked to that of truncation, whereby a number of beats are left unfilled at the end of lines, with the consequence that syllables from the following line can be used to fill those empty positions. On the other hand, as already observed, English seldom leaves the first position of grids empty, so theft from this position sounds less natural.

Concluding this section, Hayes and MacEachern (1996:17) repeat their idea that "characteristic correspondences between linguistic phrasing breaks and particular metrical locations indicate that the line is an authentic, well-defined metrical constituent of the sung-verse grid". They acknowledge the existence of counterexamples to their claim, involving theft of positions in both directions. However, given that interline thefts are statistically unusual compared to the mainstream cases and that the number of beats that may be stolen is limited, they conclude that they have strong evidence in support for the line as a neatly bracketed structural unit.

The very last section of the paper is devoted to what they call "ambiguous line levels", which refer to the possibility to parse higher constituents into lines "at more than one metrical level" (1996:17). Hayes and MacEachern claim that the basic principles of metrical form in English folk verse are relatively abstract, and involve "requirements for strict identity of constituents of equal rank, isochronism of rhythmic beats at all levels, and a strong preference for binary and ternary branching" (1996:19). At such a level of abstraction, they accept that there is no such

thing as a ‘line’ as a formal primitive. Nevertheless, there are certain intuitions about what the lines are, such as the following (1996:20):

- a) The line level seems to be the level at which the ‘echoing’ of metrical bracketing with P-structure bracketing is the clearest – an idea borrowed from Attridge (1982:107).
- b) There seems to be a rough maximum for the internal metrical complexity of lines. Restricting our attention to binary metres, we find that a metrical grid with 16 terminal positions typically suffices to provide a rhythmic location for every syllable in the line.
- c) Stanzas tend to be long enough to permit at least four ‘lines’.

Hayes and MacEachern’s (1996:20) corpus-based conclusion is that the lines diagnosed in their paper “serve important additional structural purposes in defining rhyme schemes and patterns of line-final empty positions”. Finally, they point at the addition of two difficulties to their study, namely a discussion of the unusual cases where P-structure boundaries disagree with line boundaries, and the cases where more than one level of the binary hierarchy could serve as the line level. As a major conclusion, they state that neither of those two difficulties threatens their results.

4.2.4. Hayes and MacEachern (1998) “Quatrain form in English folk verse”

Hayes and MacEachern’s (1998) paper constitutes an attempt, loosely framed within OT, to find out exactly what rules regulate the patterns of truncation at the end of lines in English folk verse. The concept of ‘truncation’, as already observed, was also central to Hayes and Kaun (1996) and Hayes and MacEachern (1996). Truncation is defined by Hayes and MacEachern (1998:473) as the “nonfilling of metrical positions at the end of lines”. According to the authors, there are 26 possible truncation patterns for English folk quatrains. The function of truncation, as observed by Hayes and MacEachern (1996) is to render certain layers salient in the natural constituency of the quatrain, namely the line, the couplet or the quatrain as a whole. Given that it is not possible to render those three constituents salient at the same time, the constraints that regulate saliency will be in conflict amongst themselves, as well as with metrical CON, which require, by definition, that all metrical positions be filled with syllables and stresses. The outputs of the multiple ways of solving these conflicts among constraints – the solving process has to do with different

prioritisations of the constraints themselves – is the group of 26 well-formed truncated quatrain types.

The authors start by stating the problem. They observe that native speakers of English are able to assess the well- or ill-formedness of a given quatrain. Truncation poses a problem with regard to native intuitions. In a nursery rhyme quatrain, to quote one of the most widely used verse quatrains in English, each line is felt to have four metrical beats, but often the fourth beat is ‘silent’, that is, it is “observed in the isochronous timing of the recitation but not aligned with a syllable” (1998:473). On the other hand, certain beat combinations (*3434; *3444) are immediately judged ill-formed by native speakers (1998:474). As in Hayes and Kaun (1996), the authors assume an innate ability on the part of native speakers to judge the metricality of lines. Hayes and MacEachern (1998:474) give a brief account of previous work done on the topic of native intuitions with regard to folk verse types. They quote Burling (1966), which concluded that children’s verse types from unrelated, geographically distant languages resemble one another in a very striking way, something which is not so noticeable in art-verse.

In their analysis of the quatrain as the main compositional unit of English folk verse, Hayes and MacEachern (1998:475) emphasise the idea, already observed in Hayes and MacEachern (1996), that the fundamental basis of folk verse is a binary hierarchy. This means that a quatrain is not simply a sequence of four lines, but rather a pair of pairs, with the structure [quatrain[couplet[line line]][couplet[line line]]]. This pair-of-pairs structure is reinforced by rhyme.

To represent the structure of the lines scanned in their study, Hayes and MacEachern use, once again, the grid, which not only represents the alignment of syllables and beats, but also the saliency of certain beats with respect to others.

In order to distinguish the different types of quatrains, the authors observe that the ends of lines work for each of them. For this purpose, Hayes and MacEachern (1998:476) define the use of ‘rhythmic cadences’, a term borrowed from the musical tradition. In musical terms, a ‘cadence’ is an ending within rhythmic or musical form. A rhythmic cadence is thus “the characteristic grid placement of the final syllable or two of the line” (1998:476). The authors list the

most common rhythmic cadences in English folk verse, as seen in Hayes and Kaun (1996) and Hayes and MacEachern (1996):

- a) 3, where there are no syllables after the third strong position of the line. It is the first truncated cadence quoted in Hayes and MacEachern (1996)
- b) 'Green O' (G), which fills the third and fourth strong positions, with no syllables between them. It is the second truncated cadence quoted in Hayes and MacEachern (1996)
- c) 4, where the fourth strong position is filled, but not in the fashion of G, that is, there has to be at least one syllable between the third and fourth strong positions
- d) 3-feminine (3f), where the third is the last strong position filled, but it is followed by an extra weak position.

Cadences, as pointed out in the paper (1998:478), are an essential element in song since they mark the end of constituents. In songs, these cadences are usually repeated in every stanza, that is, the cadence tends not to shift from one type to another. So, for instance, if the first stanza ends off with a '4', the rest of the stanzas will present the same cadence to mark their end points.

The main empirical basis of this paper is the examination of a large corpus of English folk songs compiled in the early twentieth century in rural areas of the English-speaking world: rural England (Karpeles 1974), Newfoundland (Karpeles 1970) and the Southern Appalachians (Karpeles and Ritchie). The materials from the Appalachians, which had already been used in Hayes and MacEachern (1996), were, once again, the ones most-carefully analysed for the purpose of the investigation reported in the paper.

In order to classify the quatrains found in the analysed corpus of songs, Hayes and MacEachern (1998:478-483) use a mixed taxonomy. For some metres, like the 'common metre', they use the taxonomy inherited from the study of Church Hymns, while for others they simply make up names that fit their purpose:

- a) 4343 metre, called 'common metre' according to the traditional taxonomy, referred to as 'couple-marking' quatrain type in this paper
- b) 4443 metre, called by the authors 'quatrain-marking' type

- c) 3333 (GGGG, 3f3f3f) metre, referred to as 'line-marking' construction, known as 'short metre' in Church Hymns studies
- d) 4444 metre, called 'metrically replete' quatrain in this paper, traditionally known as 'long metre'
- e) 3343 (GG4G, 33G3 > 44 ('semiquatrains'), 333f3) metre, called 'short metre' or 'long-last construction'.

Together with these clearly-defined metres, Hayes and MacEachern find two other quatrain categories, namely

- f) Quatrains with 3 different cadences: 3f343, G343, 3f3G3
- g) Quatrain types with free variation: F(or G)3F3.

The classification of the corpus of quatrains relies on their analysis in relation to four groups of CON. The first group refers to CON on saliency domains, and relates to the concepts of cadentiality and saliency. The idea of a rhythmic cadence can be applied to speech as well as to music, meaning a kind of slowing down at the ends of phrasal units. This is what phoneticians call 'final lengthening', an idea explored in Hayes and MacEachern (1996) in the field of folk song. Turning this tendency around, Hayes and MacEachern (1998:483) posit that final lengthening is a cue for phrasehood, that is, a kind of constituency marker. The hypothesis is that "the rhythmic cadences are ranked according to their ability to induce the perception of a group ending (line, couplet, quatrain)" (1998:483). According to this, the cadentiality ranking is represented as follows, from most to least cadential: 3 >> 3f >> G >> 4.

SALIENCY has to do with two main conditions (1998:485), namely (a) the final rhythmic cadence of that constituent is more cadential than all of its non-final cadences, and (b) all of its non-final cadences are uniform. Under this definition, the salient quatrains in the analysed corpus would be 444G, 4443f, 4443, GGG3f, GGG3, 3f3f3f. The salient couplets would be 4G, 43f, 43, G3f, G3, 3f3. As for lines, Hayes and MacEachern (1998:485) see them all as "(vacuously) salient".

The analysis of the corpus according to this type of constraint leads the authors (1998:486) to make the following observations:

- A major goal of the metrical system of English folk verse is to render salient the major structural units: line, couplet, quatrain. This is done by placing the final syllables of lines in appropriate arrangements of cadentiality.

- The most heavily cadential line endings are also the most metrically truncated: to serve the cadential function, they must fail to fill quite a few positions at the end of the grid.

The second major CON is PARALLELISM, also mentioned in Dell and Halle (in press), defined as follows: two couplets with the same structure form a quatrain. What this entails is that “the cadences ending the units of the maximal analysis of a quatrain must be identical” (1998:488), a principle that is never violated in well-formed quatrains. Parallelism is, in turn, induced and reinforced by rhyme in that the cadences of the maximal analysis usually rhyme with each other.

The third main group of CON refer to long-last constructions, and it can be subdivided into the CON PREFER LONG-LAST, PARTIAL LONG-LAST COHESIVENESS and TOTAL LONG-LAST COHESIVENESS. The LONG-LAST CON implies that, in a sequence of groups of unequal length – like 3434 –, the longest member should go last. A quatrain is a long-last construction if (a) its second couplet is salient by the all-or-nothing definition, and (b) both its first and second lines are more salient (by the gradient definition) than the third line (1998:489). An important remark seems in order here. Quatrains satisfying the LONG-LAST criterion will be experienced as line + line + couplet. This is important if we take into account that, in principle, quatrains of the type 3343 would not satisfy it if every line was experienced as such. On the contrary, what Hayes and MacEachern suggest is that the first two lines (3-3) will be experienced as such, while the two last ones (4-3) will be taken to form a single constituent, a couplet. In this way, the last constituent of the quatrain will indeed be longer than the first. At first glance, it is problematic to accept this change in the constituents’ length and shape in order to fit some theoretical statement.

LONG-LAST is directly linked to the degree of cohesiveness of a given quatrain. In Hayes and MacEachern’s (1998:489) words, “a couplet, if salient, is cohesive inversely to the saliency of its first line”. According to this principle, 43, 43f and 4G are fully cohesive couplets, G3 and G3f are less cohesive, and finally 3f3 is the least cohesive one.

The fourth major group of CON are metrical CON, which have to do with the requirement that all positions in the grid be filled with syllables and stresses. This group can be subdivided into three main CON (1998:492):

- a) Fill strong positions (FILLSTRONG): fill the strongest positions in the line
- b) Avoid lapse (*LAPSE): avoid sequences in which no syllable is placed in the interval between any two of the four strongest positions in the line
- c) Match stress (MATCHSTRESS): employ G in feminine endings, 4 elsewhere.

The last main CON dealt with by Hayes and MacEachern is that of STANZA CORRESPONDENCE (1998:492), which states that “in a song, the set of salient domains must be invariant across stanzas”.

As mentioned above, the theoretical tools and methodology employed in this study are those of OT. In order to give an OT account of the analysis, the CON are ranked, and the quatrains in the corpus are analysed in accordance to the different possible rankings. Last, the ideal candidate in each ranking is chosen, and finally, the absolute best candidate is chosen.

The empirical evaluation of the songs is divided into the following steps:

- a) scanning of all the songs
- b) stanzaic pattern of rhythmic cadences for each song
- c) filtering of criteria to know what we consider a quatrain. Total: 627; 1028 songs.

The results are as follows: there are 203 metrically replete verses (4444); 188 couplet marking (4343); 38 couplet marking (4G4G); 35 quatrain marking (4443); 29 quatrain marking (F3F3); 26 are ill-formed (not predicted), 14 of them with refrain, 12 without refrain (1998:25). Refrain can be defined as the verbal material that is invariant across stanzas, that is, it gets repeated through the song.

The conclusions observed by the authors are surprising: there are some attested ill-formed quatrains, and also unattested quatrains which nevertheless appear in certain songs (1998:497). Some refrain examples sound fairly well formed where analogous non-refrain examples seem rather lame. Furthermore, there are structures, such as 4G43, which are ill-formed as stanzas but well-formed as refrains. From a musicological point of view, it is easy to see the reason for this: whereas stanzas are

generally structured into two groups of two lines each, where each line has the same length and rhyme in such a way that a specific stanza could be easily transferred from one song to another one, refrains tend to be more irregular both metrically and in their rhyming schemes, which makes them more idiosyncratic to the song to which they belong. In this sense, it is only natural that the structural requirements of stanzas are much tighter than those of refrains.

The authors try to explain the occurrence of non-predicted quatrain structures and mention a possible memory lapse on the singer's part. As a less individual-bound explanation, they suggest that minor unmetricalities add a certain zest of unexpectedness to a song (1998:498-499). However unscientific this argument might seem, that "adding of a certain zest" is actually one of the main principles that regulate song composition and performance – what Hanson and Kiparsky (1996:295) refer to as "interest parameter": "maximise the aesthetic interest of the verse". If every song responded totally to the principles of metricality as stated in this paper, the corpus analysed by Hayes and MacEachern would be nothing else than a monotonous uninspired repetition of metrical schemata, where the artistic dimension inherent to song would be inexistent. As examples of what they call "minor unmetricalities" the authors quote *43f4G, *3f3f3f4, *434G, plus refrains 4344, 433f3. Second, they try to explain why certain predicted structures, as is the case with GG4G, 3f3f43f, 3f3fF3f, 33F3, 3f3F3, F3fF3f, do not appear in the analysed corpus. They accept that there might be a gap in the chosen corpus, given that the quoted structures are attested in nursery rhymes (1998:499).

The last section of the paper (1998:500) is devoted to highlighting the role of a theoretical frame such as OT in studies of this kind. According to the authors, OT provides a way of taking a set of raw structural preferences and turning it into an explicit grammar as well as a natural account of the variety of quatrain types. The inherent goals of the system are in conflict and each outcome represents a particular resolution of the conflict by assignment of priorities. More importantly, it makes it possible to rule out certain forms without actually formulating a constraint against them.

The ranked set of CON with which Hayes and MacEachern (1998:502) end up is the following:

- (i) PARALLELISM (inviolable)
- (ii) PARTIAL LONG-LAST COHESIVENESS (inviolable)
- (iii) STANZA CORRESPONDENCE (inviolable)
- (iv) COUPLETS ARE SALIENT
- (v) TOTAL LONG-LAST COHESIVENESS
- (vi) FILL STRONG POSITIONS
- (vii) AVOID LAPSE
- (viii) QUATRAINS ARE SALIENT
- (ix) LINES ARE SALIENT
- (x) MATCH STRESS
- (xi) PREFER LONG-LAST

The general categories of quatrains according to the analysis carried out are the following (1998:504-505):

- a) well-formed, well-attested quatrains, such as 4343, derived from statistically likely rankings of CON
- b) well-formed, poorly attested quatrains, derived from fully legitimate but statistically unlikely rankings (3f3f43f)
- c) marginal quatrains, derived from ranking certain CON slightly outside their normal range of strictness (333f3)
- d) ill-formed quatrains, poorly- or unattested, not derivable within the system (3434).

As a major conclusion to this paper, Hayes and MacEachern (1998:505) point out that English folk verse is tightly patterned at the level of the quatrain: the various rhythmic cadences are arranged in non-random, essentially strategic fashion. They emphasise the importance of cadences in order to induce perceived bracketings, which are then employed to structural ends: i) the enhancement of metrical constituents at various levels and ii) the placement of long elements last. Last, they praise the theoretical model of OT, which makes possible the use of primitive analytical ingredients to capture complex descriptive results.

The general feeling of a musically-trained linguist is, as already pointed out, that in order to build up a comprehensive model of folk verse/song analysis, it is

necessary to devise a musical theory that describes metre, pitch, quantity and phrase/cadence, since we are dealing with song.

4.2.5. Hayes (in press) "Text-setting as constraint conflict"

Hayes's (in press) paper constitutes a deeper exploration of text-setting from the point of view of OT. In this case, Hayes focuses on the interaction between the metrical CON in Hayes and MacEachern (1998) and pays special attention to four of them, namely:

- a) matching of stress to strong position
- b) avoidance of long lapses
- c) avoidance of extreme syllable compression
- d) alignment of phonological phrase boundaries with line boundaries.

The pervasive idea repeated in these papers is that people have productive text-setting abilities, which they show by finding the best possible resolution between conflicting metrical principles in order to come up with a well-formed text-setting. The stated resolution between metrical principles involves "the sacrifice of certain principles in order to realise others with a higher priority" (in press).

Hayes introduces his paper by defining the "text-setting problem" (Halle and Lerdahl 1993), as seen in section 4.2., which "arises in the context of sung and chanted verse" (in press). I have already pointed out that equating sung and chanted verse is not accurate since there is much more to sung verse than a succession of beats.

Once again, the chosen representation for text-setting is the metrical grid, following Liberman (1979), Lerdahl and Jackendoff (1983) and others. As Hayes (in press) explains, "in a grid, units arrayed in rows depict series of isochronous beats on a hierarchy of levels, and the columns indicate the strength of individual beats". So, in this model, grids are very explicit in that they represent both constituency and saliency.

The paper emphasises the idea that text-setting, just like natural language, involves well-formedness intuitions (in press). This entails that, for a specific text-setting tradition, "if we are willing to abstract away from a modest amount of free variation, text-setting is predictable" (in press). This statement seems too simplistic, as already noted. The "modest amount of free variation" to which Hayes refers is

precisely what makes certain folk song traditions so different from certain others, and definitely what makes the study of text-setting such an intricate topic.

According to Hayes, the representation of the text-setting process would be as follows:

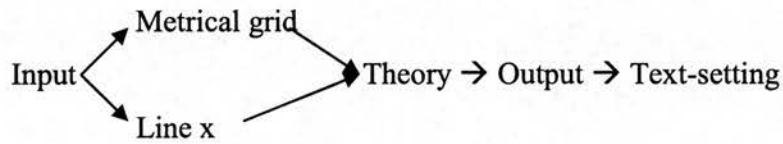


Table 28: Text-setting process.

As a sort of background review, Hayes explores the first attempt made by Halle and Lerdahl (1993) to build up a metrical grammar for text-setting. Their grid looked as follows:

	s	w	m	w	s	w	m	w
S row	x				x			
M row	x		x		x		x	
W row	x	x	x	x	x	x	x	x

Table 29: Text-setting metrical grammar.

Hayes does not agree with the utility of the fourth row in Halle and Lerdahl's grid, which would differentiate between main accents and secondary accents in music, so he eliminates it from his own grid representations.

Halle and Lerdahl base their system on the so-called Syllabic Distribution Algorithm (SDA), which consists of a set of mapping rules and a group of principles of when and how the rules apply. The algorithm works by taking a number of steps in the analysis of a specific instance of text-setting (Hayes in press):

- a) The line in question is scanned from left to right.
- b) When a stressed syllable is found, it is made to occupy the first S position in the grid.
- c) The algorithm iterates, thus repeating steps (a) and (b).

- d) Whenever the assignment of a stressed syllable to S skips over a sequence of stressless syllables, a specific rule is applied: the algorithm counts the number of stressless syllables before the stressed syllable that was mapped, then maps these stressless syllables one-to-one, right-to-left onto the highest grid level having enough marks available to accommodate them.
- e) Any trailing stressless syllables are mapped to the remaining positions, again preferring the highest grid row that can accommodate them, but this time from left to right.

Hayes (in press) sees an important flaw in this algorithm: it does not set all lines correctly. This could be due to a variety of reasons. First, the SDA, based on left-to-right mapping, places the stresses as far to the left as possible, even if this leaves large gaps later on the line. Second, squeezing the stressless syllables – when there are more than three – could be achieved by the (undesirable) creation of an Extra Weak (EW) row. Third, if there are more than four stresses, the SDA causes the additional stresses to be spilled over into the next line. Fourth, if there are consecutive stressed syllables, the algorithm creates enormous lapses. A possible solution to avoid these problems would be promoting and demoting stresses systematically, something which would only serve the purpose of “rescuing the metrics” (in press). This is not desirable, so Hayes suggests an alternative. According to him (in press), in order to determine why the algorithm sets so many lines incorrectly, it is crucial to know its ultimate goal. Such a goal is getting the stressed syllables into the strong positions. Hayes states a need to account for other goals, which are identified with the constraints with which he is dealing in the paper. Such goals will be avoiding lapses, aligning beginnings and ends of Intonational Phrases and ends of lines, and avoiding the extreme compression of syllables. So, the next step in the paper is to ‘translate’ these intuitive constraints into OT language and rank them in such a way that the system can yield well-formed text-settings. The CON are the following (Hayes in press):

- a) A stressed syllable “wants” to be in S position
- b) If not, M preferred over W
- c) An empty syllable is preferred following the scale W – M – S.

- d) Directionality – as seen in Halle and Lerdahl (1993) is not necessary, unless we want to refer to the phenomenon that S units tend to be long as well.

Hayes (in press) adds three more CON to the above list, namely:

- e) *LAPSE: never three empty positions in a row
f) *FILL EXTRA WEAK: no EW position
g) *RUN-ON: never more than the established number of syllables per line.

These CON would be ranked as follows (Hayes in press): *FILL EW > *RUN-ON > *LAPSE > *NULL IN S > *STRESSLESS IN S > *STRESSED IN W > *STRESSLESS IN W > STRONG IS LONG > *STRESS IN M > *NULL IN M.

After stating his theoretic-methodological apparatus, Hayes (in press) presents his results, obtained after devising and testing the ranking as a computer program on 364 lines. 70.6% of the outputs matched the favourite setting of the native speakers, whereas 23.1% matched the SDA. In this respect, Hayes's alternative seems to complete the SDA in order to make the system effective in predicting the well-formedness of text-settings.

Hayes's conclusions have to do with what he calls a "need for more Metrics" (in press), which will yield better results. The specific aspects that should be studied are the following (in press):

- a) Stress involves multiple levels.
b) Importance of both word and phrase boundaries.
c) Text-setting involves duration matching: the number of grid positions assigned to a syllable (all else being equal) tends to match that syllable's natural phonetic duration, as determined by pre-boundary lengthening and syllable weight.

4.2.6. Dell and Halle (in press) "Comparing musical text-setting in French and English songs"

One of the terminological issues addressed in the papers under discussion in 4.2. has to do with the definition of 'song'. Dell and Halle (in press) define it as a composite in which two independent elements, namely a text and a tune, interact in two main ways: on the one hand, in terms of the alignment of stressed syllables and strong positions; and on the other, in terms of the alignment between the grouping of linguistic and musical constituents.

Throughout the analysis of Dell and Halle's (in press) and Halle's (in press) papers we observe that they rely heavily, both theoretically and methodologically, upon Lerdahl and Jackendoff's (1983) groundbreaking *A Generative Theory of Tonal Music*. In fact, when Dell and Halle refer to the two ways in which text and tune interact, they are simply re-naming two of the four main hierarchies developed in Lerdahl and Jackendoff (1983), namely the 'metric hierarchy' and the 'grouping hierarchy' – the other two are 'time-span reduction' and 'prolongation reduction'. As Dell and Halle (in press) explain, alignments or text-settings make use of two formal similarities between language and music, namely the relative prominence of certain units over others and the arrangement of constituents into hierarchical structures. Dell and Halle do not explore prominence in language in a systematic way; in fact, they simply state that this prominence is achieved by means of stress or accent, two terms which they use as synonyms. Of course, anyone familiar with the study of accent and stress in language knows that the former is not a synonym for the latter and that stress is only one of the ways in which prominence can be manifested in language, together with tone and a combination of both pitch and stress – for a thorough analysis of accentual typologies, see section 1.4. In order to avoid terminological mistakes, it is crucial to point out that, when Dell and Halle (in press) talk about the manifestation of prominence in "stressed or accented syllables in language", they are necessarily referring to the action of linguistic stress in languages such as French and English.

Leaving linguistic inaccuracies aside, Dell and Halle (in press) state that, in order for text-setting to be successful, "linguistic groups must match with musical groups on the one hand, and on the other hand certain stressed syllables must fall on strong beats". So, using Lerdahl and Jackendoff's (1983) terms, both the metrical and the grouping hierarchies of a given text and a given tune must coincide in order for the alignment to be considered well-formed. As an obvious consequence of this, any mismatch between linguistic constituent structure and musical constituent structure will render ill-formed alignments. So, as a general rule, "[t]he edges of linguistic groups must coincide with those of musical groups" (in press). Again, Dell and Halle are not precise in their explanations of linguistic terms. In this case, they solve the vagueness of the term 'linguistic groups' by saying that they use it as a synonym of


In the first main section of their paper, Dell and Halle briefly introduce the difference between the traditionally-called ‘syllable-timed languages’ and ‘stress-timed languages’, though they refrain from using these terms and simply explain that “in languages in which certain syllables are more stressed than others, there is a preference for songs to associate stressed syllables with strong metrical positions” (in press), something that they call “stress-to-beat matching” (in press). In these languages, such as English, certain stress-to-beat mismatches are prohibited in all environments. The reader thus wonders which mismatches are allowed and which are not, and on what grounds this gradient occurs. Are these mismatches phonologically, morphologically or syntactically motivated? Is it the case that this gradient only has to do with the prosodic characteristics of the language in question?

Before going on to define a stress-beat mismatch, Dell and Halle (in press) explain what they consider a crucially important concept in their theory, namely the concept of ‘tactus’. They adopt Lerdahl and Jackendoff’s (1983) notion of the tactus as the regular beat into which listeners naturally fall when they clap their hands to that tune. The notion of tactus is, nevertheless, much more problematic than Dell and Halle assume. An idea the neither Lerdahl and Jackendoff nor Dell and Halle seems to embrace is that the process of perceiving the tactus is culturally determined, as proved by ethnomusicological research carried out in different parts of the world (see Gray 1998:617-627). As Gray (1998:620ff.) shows, if a Western audience was exposed to Ugandan traditional songs and asked to clap to the music, they would probably define the placement of the tactus on the main downbeat of each bar, while the Ugandan people would clap on the upbeat, which would mean their perceived tactus falls on the opposite beat to the Western-perceived tactus. Let us have a look at an example of a musical piece which starts on the upbeat – also called ‘anacrusis’ – and then one which starts on the downbeat:

(177) Piece that starts on the upbeat (from Kennedy 1984:503):

There was an old man and he lived in the east And his

(178) Piece that starts on the downbeat (from Kennedy 1984:606):



I've got a coat and a nob-by nob-by coat

In both (177) and (178), the tactus falls at the level of the crotchet, which is signalled by 4 in the time signature.

The definition of tactus is essential for Dell and Halle's (in press) explanation of what a stress mismatch is in English:

Let S and s be two syllables occurring in any order, one accented (S) and the other unaccented (s). A stress mismatch occurs if S and s are adjacent within the same line and the metrical position associated with s belongs to the tactus, but not that associated with S.

So, for Dell and Halle, a mismatch would be roughly the same as what Hayes and Kaun (1996) call "lexical inversion". This paper, just like all the papers co-authored by Hayes, identifies lexical mismatches as a sign of ill-formedness in text-setting. We could say that this identification is valid, but only to a certain extent. In almost all European traditions one can observe a certain degree of mismatch between music and text in song. Actually, mismatches are also common in the works of art-music composers of all times. In his outstanding analysis of the early songs of the fifteenth-century composer Dufay, Boone (1999:82) states that, in the process of imbrication or overlapping between musical and poetic rhythms, there is always a "seeming disjunction between poetic and metrical (musical) stresses [...] commonly prompted by the presence of formulaic rhythmic modules in the musical setting". A certain degree of disjunction is in fact necessary, as it prevents songs from becoming monotonous and inartistically predictable.

As explained before, 'accented' means 'stressed' for Dell and Halle. What is more important is that in this definition Dell and Halle suddenly introduce the line – a verse prosody concept – as an essential structural unit in the determination of stress mismatches. The line is not one of the prosodic constituents established by Hayes (1989), whom Dell and Halle quote. It seems to be the case that this paper does not draw a clear difference between theories of speech prosody, verse prosody and

theories of text-setting. In that regard, the theoretical frame used by the paper is too ad hoc to be applied systematically to the study of text-setting in different languages.

After analysing the setting of specific English texts into a chosen tune, Dell and Halle (in press) come up with a series of mismatches which they label “illicit” for English. They state their observations as follows:

A composite is ill-formed if it contains a stress mismatch which meets one of the following conditions:

- the two syllables are not separated by a word boundary or
- the two syllables are separated by a word boundary and s precedes S.

Again, Dell and Halle’s level of accuracy in the linguistic side of their investigation is not as high as would be desirable. As stated by the authors, (a) applies to simple words whose stressed syllable falls on a weak beat – they give *dinner*, *balloon*, *rebate* as examples – but does not account for compounds, while (b) applies to sequences of a grammatical word plus a lexical one – *the boy* – but does not account for the opposite order – *see me*. They do not give any further account of the cases not covered by the above statement.

At the end of this section, Dell and Halle turn their attention to the melodic element of tunes, and thus leave aside the idea that mismatches can be explained by relying solely on the metrical side of alignments. They explore which elements are important in order to define ‘melodic contour conservation’ as follows: “for two tunes to be perceived as variants of the same tune, they must have the same melodic contour” (in press). This is a requirement of a purely musical nature, which has to do with the harmonic, melodic and rhythmic skeletal structure of a piece of tonal music. The contour of a tune is the sequence of the pitch changes in that tune, abstracting away from the number of notes attached and their timing. So, there is a kind of musical parallelism that has to do with the contour of a tune. On the other hand, there is another kind of parallelism which Dell and Halle call “positional parallelism” (in press): “two alignments are positionally parallel if the distributions of their syllables along the grid are identical”. The implication of this concept for text-setting is that two composites which do not have the same number of syllables cannot be positionally parallel. Nevertheless, this statement needs to be qualified. In a language like Spanish, where syllable compression devices are often applied to both speech

and verse, two lines with different numbers of syllables which can be made to correspond to the same number of beats. Dell and Halle lack a clear explanation of what they mean by 'syllable': are they speaking about syllables in the prosodic hierarchy or in the poetic line?

The second main section in the paper explains what a mismatch is for French, according to the prosodic characteristics of this language. Dell and Halle state that "in traditional French songs, alignments are rigidly constrained at the end of lines", where certain melodies require a form which is phonetically feminine, while others require a form which is phonetically masculine. Again, we can see the importance of the line as a structural and salient component of the grouping hierarchy. Dell and Halle (in press) use 'form' meaning 'the pronunciation of a word'. This has to do with the status of schwa in French, normally corresponding to an e-muet, which can sometimes be pronounced for metrical reasons, while on other occasions it is not pronounced at all.

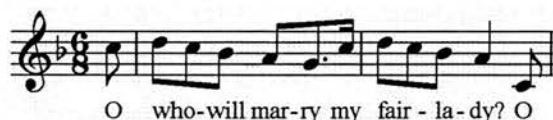
They define the concept of 'local maximum' (in press) as the requirement that "the grid position associated with the last accented syllable in a line must be stronger than the positions associated with adjacent syllables belonging to the same line", parallel to the concept of 'stress maximum' stated by Halle and Keyser (1971). As explained here, local maximum only concerns line ends, although a tendency to match stress and beat is commonly found in traditional French songs. The crucial point about French text-setting is that, while the conflicts between stress and beat at the end of lines are extremely rare and notorious from the point of view of perception, such conflicts are commonplace in other positions, where most of the time they escape the listener's notice. Dell and Halle (in press) also notice that in a French strophic song all stanzas are positionally parallel. According to them (in press), a song with positionally parallel stanzas has the following properties, among others:

- a) All stanzas have exactly the same tune.
- b) If two lines occupy the same position in the stanza, they have the same number of syllables.
- c) The distribution of melismata is the same in all stanzas.

Once again, a musical term is not explained. In this case, the definition of the term ‘melisma’ is missing, something which will certainly make this part of the paper unintelligible for non-musicians.

Melismata are very common in traditional music, as we can observe in (179), from Kennedy (1984:653):

(179)



In (179), we observe that the syllables *who* and *fair* are both set to two different pitches. They constitute examples of melismata.

Dell and Halle (in press) take the quoted three properties as consequences of combining positional parallelism with Melodic Contour Conservation. According to them, the main difference between English and French is that in the former departures from positional parallelism are common – less syllables or more syllables per line require that notes be deleted or inserted (in press).

In their conclusions, Dell and Halle (in press) emphasise the importance of positional parallelism and arrive at a universal definition for this concept: “as far as possible, identical melodies give rise to positionally parallel settings”. It is only at the end of their paper that Dell and Halle (in press) raise the question of why prominence matching is enforced in a much more rigorous manner in English than in French. They state that, rather than being merely a matter of poetic convention, that difference has a phonological basis: stress has more perceptual salience in English than in French, something which any acute observer could conclude. In this respect, they emphasise once again the importance of acknowledging the fact that the only constraints on the texts are those deriving from the linguistic grammars. Texts on their own, without considering their potential for being set to a tune, do not have “any intrinsic organisation in feet, lines, stanzas, or any other of the units of poetic form as it is conventionally understood” (in press). The chunks analogous to the lines and stanzas of literary verse in which song can be broken down are just by-products of text-to-tune alignment. They are not features of the text itself but features

of the composite text-tune. This seems to be in contradiction with the theoretical assumption made by the authors that the line exists as a structural unit previous to being set to music.

This paper, like Hayes's papers, builds up a theory of text-setting around the idea that a perfect match between linguistic stress and musical metre is the epitome of text-setting well-formedness. However, this seems to be precisely the weakness of all the theories with which the analysed papers deal. As Boone (1999:212) suggests, there have always been two major pitfalls in theories about the relationship between music and text: "the first pitfall is overemphasis on tonic accent as the prime agent of rhythmic activity in linguistic rhythm; the second is overemphasis on a simple alignment between tonic accent and musical accents as the hallmark of 'good' prosody". The view of the relationship between music and language as a stress game is distorted because it does not really tell us much about what is going on in language, poetry and music and, more importantly, yields a simplistic account of the infinite ways of interacting between text and music, be it in the enormously rich and varied traditions of folk song, or in the intentionally imbricated realms of art song.

4.2.7. Halle (forthcoming) "Text, tune and metrical form"

Halle's (forthcoming) paper takes as its point of departure the concept of 'song' put forward in Dell and Halle (in press) and starts by defining it as a "composite made up of two independent objects, a text and a tune" (forthcoming). There are two observations to be made about the interaction of these two objects. On the one hand, each of them must be judged as acceptable within the grammars applicable to the respective cognitive domains. On the other, an acceptable interaction between music and text, while necessary, is not sufficient to determine a well-formed text-setting: "both the text and the tune must be judged acceptable as independent objects in order for the composite to be judged acceptable" (forthcoming). However, this statement seems to be too strong if we think about songs in which words and music have been written by the same person – such is the case of Thomas Campion's or Cole Porter's songs – in a simultaneous sort of way. In this case, it is often the case that the lyrics, without the music, could hardly be considered verse or, indeed, good verse. Consider, for instance, the fragment in (180) by Porter, in which the lyrics would not have been considered well-formed poetry if they had been composed with no tune in

mind – maybe they would, just because of the fact that the lines show quite a predictable rhyme scheme:

(180) It was just one of those things,
just one of those crazy flings,
one of those bells that now and then rings,
just one of those things.
It was just one of those nights,
just one of those fabulous flights,
a trip to the moon on gossamer wings,
just one of those things.

Halle's paper deviates from Dell and Halle in that it does not constitute a theoretical comparative study of two distinct text-setting traditions but rather an attempt to devise a generative model of text-setting for English. In this sense, as already pointed out, it relies on Lerdahl and Jackendoff (1983) as well as on certain theoretical and methodological OT tools.

Halle's main hypothesis on the well-formedness of text-settings relies on the concept of 'similarity metric', defined as "a mechanism which formally describes the class of variants which can be construed as acceptable" (forthcoming). The judgements on well-formedness are made independently for music and text with reference to the similarity metric. Halle (forthcoming) states that two types of constraints are at work in 'strophic song' – a term which will be defined in the next pages –, namely interactive constraints and independent constraints. First, there are three types of interactive constraints, namely stress matching constraints, constituency constraints and constraints on melismatic structure. In order to explain stress matching constraints, Halle resorts to Morgan and Janda (1989), one of many studies which demonstrate that mismatches between stress and metre are judged along a gradient range from natural to unacceptable. Thus, Halle (forthcoming) states that in English the melody gets distorted so that the stresses do not get mismatched, while in Spanish there is no alteration of the tune in spite of the encountered mismatches. Halle (forthcoming) repeats the conclusion stated in Dell and Halle that

in certain vocal traditions positional uniformity – the relation of musical events to the metrical grid found in the original tune – can be maintained in the face of what are defined as stress mismatches. In other text traditions, significant deviations from positional uniformity are required in order to achieve a proper matching of syllabic stress and the metrical grid.

Halle points out the importance of observing constraints which deal with units bigger than the syllable. One of these constraints is constituency matching. The author (forthcoming) suggests text substitution as a test to find evidence that an ‘in phase’ relationship between musical and linguistic constituency tends to result in more natural sounding arrangements – prosodic hierarchies. Halle (forthcoming) concludes that constituency mismatches appear to constitute a more or less exact analogue to enjambment within metrical poetry.

The third set of interactive constraints is that of constraints on melismatic structure. He defines melismata as “syllables assigned to more than one note” (forthcoming) and quotes Stoquerus (1988[1570]) when he states that “whereas the number of notes in a passage can [...] exceed the number of syllables, the number of syllables cannot exceed the number of notes” (Halle forthcoming). This statement is not completely accurate since, in a language like Spanish, the number of phonological syllables can exceed the number of notes – a different issue is how those syllables are realised phonetically, that is, whether there are syllable compression devices that operate in order to adjust the number of syllables to the metrical positions available. Halle quotes Stoquerus’s Rule 4, namely “no two adjacent events assigned to the same syllable may be assigned to the same pitch” – this would not be a melisma, but a repetition (Halle forthcoming). Last, Halle (forthcoming) explains the adaptation of Stoquerus’ rules by Gauldin (1995) and Veltman (2001), which results in the rule that “the continuation of a melisma may not appear in a metrically stronger position than its onset”. The use of Stoquerus’s rules, which date back to the sixteenth century, seems to be anachronistic here since it has been proved that Stoquerus’s rules only apply to the repertoire which he discusses. Also, with reference to Stoquerus’s rule 4, it is important to note that in

early seventeenth Italian monody it was indeed possible to realise a specific subclass of melisma on a note, a figure called ‘trillo’.

Second, another set of constraints is that of independent constraints on strophic song, which can be linguistic or musical. Halle (forthcoming) defines strophic song as a musical genre characterised by “the reiteration of a melody onto which distinct texts – variously referred to as stanzas, verses or strophes – are superimposed”. Given the nature of strophic songs, the disparity of syllables from verse to verse forces significant deviations in the musical form of each; also, some pitch deletions are acceptable – they do not compromise the integrity of the tune –, whereas some others are not (forthcoming). We could say, going back to the concept of ‘contour’ in Dell and Halle, that any pitch deletion or insertion is acceptable as long as it does not disrupt the melodic contour of that specific song, that is, as long as it respects the Melodic Contour Conservation principle.

Halle (forthcoming) introduces the idea that there is a gradient where, with regard to the metrical pattern of the first stanza of a specific song, certain metrical positions allow for either vacancy or occupancy by events; other positions appear to require occupancy by events for the setting to be acceptable and finally, others are required to remain vacant. Section 3.0. of the paper explores the descriptive requirements of similarity metric or metrical parallelism by means of testing text substitutions on well-known tunes. Section 3.1. states the formal characteristics of representation of similarity metric. There are four main cases with their representations, namely, mandatory occupancy, where the position must be occupied in the original and in all acceptable variants – represented with ‘1’; optional vacancy, where the position is occupied in the original, but it may be deleted in acceptable circumstances – represented with ‘-’; optional occupancy, where the position in the original is vacant, and it may be occupied in acceptable variants – represented with ‘+’ and mandatory vacancy, where the position must be vacant in the original and in all acceptable variants – represented with ‘Ø’. According to Halle (forthcoming), “the similarity metric associates each metrical position with one of these categories”. Section 3.2. discusses the derivation of similarity metric. Halle (forthcoming) departs from the idea that “knowing a tune means in addition to having access to the original

form, an awareness of the distribution of optional and mandatory positions which define acceptable variants as these are specified in the similarity metric”.

Halle (forthcoming) suggests a series of similarity metric assignment rules (SMAR), which resemble the kinds of rules developed by Lerdahl and Jackendoff (1983) under the name ‘well-formedness rules’. These rules are mandatory, as opposed to ‘preference rules’, which have to do with a choice on the part of the composer/listener. The SMARs are listed as follows (2005:20-29):

- (a) SMAR 1 (metre): given the original form of a strophic song, assign category 1 to all strong metrical positions occupied in the original. A strong position is assimilated to a “tactus” level position, as defined by London (*Grove Music Online*).
- (b) SMAR 2: if X and Y are adjacent events assigned to different pitches, X precedes Y, and X is stronger than Y, assign category 1 to Y.
- (c) SMAR 3 (group edges): within each group, designate as mandatory (i) the metrical position corresponding to the initial event in the original and (ii) all metrical positions to the right of the final event of a group.
- (d) SMAR 4 (default assignment): designate all non-assigned positions as optional categories – or 3 depending on their occupancy status in the original.

Section 3.5. introduces a novelty with reference to the previous articles. In this section Halle (forthcoming) talks about pitch assignment and makes two important observations, namely that optional locations inherit pitch from the nearest assigned location to the left, and that the pitch assigned to anacrusis will be determined on a case-by-case basis.

Section 3.6. deals with the concept of ‘fusion’. Fusion, as expressed within the similarity metric, requires that any mandatory status initially assigned to reiterations of repeated pitch events may be redesignated as optional so that these attacks can be deleted in the variant:

- (e) SMAR 5 (fusion): given two adjacent identically pitched events X1 and X2 within the same subgroup of an original tune where X1 precedes X2, redesignate the location containing X2 as optionally vacant category 2 (forthcoming).

Finally, section 3.7. gives some restrictions on grouping structure. As a result, SMAR 3 must be revised:

- (f) SMAR 3 (grouping-revised, see (c)): within each group, designate as mandatory (i) the metrical position corresponding to initial event in the original, and (ii) all metrical positions to the right of the final event of a group when this event is maximally long (forthcoming).

In his attempt to build up a generative theory of text-setting, Halle concludes by stating, once again, the importance of similarity metric in text-setting and metrical form, particularly when it comes to judging the acceptability of certain cases located in a somewhat 'grey area' of marginally acceptable or unacceptable settings.

There is no doubt that the approach chosen by Halle represents a different perspective from those explored in the previous papers. Nevertheless, it does not represent a novelty in that it clearly inherits its theoretical framework and methodological tools from Lerdahl and Jackendoff's (1983) generativist views applied to music, a theory which, although groundbreaking, has ultimately proved limited in its conception of music as a series of innate mechanisms manifested in metric and melodic terms.

4.2.8. Kiparsky (2006) "A modular metrics for folk verse"

Kiparsky's (2006) paper constitutes a critical comment, as well as a theoretical and methodological follow-up, to Hayes and MacEachern (1998). Hayes (in press) had already pointed out that the prosodic structure of language (stress) and verse (metre) is governed by separate constraint systems which must be jointly satisfied by well-formed verse. As Kiparsky (2006:7) observes, this statement entails a modular approach to metrics. Kiparsky (2006:7) takes the modularity idea a step further by arguing that "the composer and performer of a song constructs a match between three tiers of rhythmic structure: linguistic prominence, poetic metre, and musical rhythm". These three tiers are organised according to the same kind of principles, as hierarchies of alternating prominence which are visually represented by trees or grids. The crucial point is that, as Dell and Halle (in press) observed, the tiers are autonomous. Kiparsky (2006:7) explains the concept of tier autonomy by saying that "a text has an intrinsic prosodic form independently of how it is versified, a stanza has an intrinsic metrical form independently of how it is set to music, and a tune has

an intrinsic musical rhythm independently of the words that may be sung to it". In other words, "the stress pattern (or other linguistic prominence relation) which determines the intrinsic linguistic rhythm of a song's text is assigned by the language's prosodic system", while "the metre of its stanzas and the rhythm of its tune are normally drawn from a traditional repertoire of rhythmic patterns" (2006:7). How the tiers correspond to each other, and in what ways they can be mismatched and mutually accommodated, is regulated by conventions that evolve historically, though within limits grounded in the faculty of language.

Kiparsky (2006:8) disagrees with Hayes's idea that the metrical form of a verse can be equated to the way in which its text is aligned with the musical beats in performance. He gives several arguments against such identification and in support of the traditional division between metre and music. Although Kiparsky (2006:8) agrees that OT is a good frame to model the groundedness of metrical preferences and constraints and their competition within a metrical system, he also argues that variation is better treated by partial constraint ranking than by stochastic OT, which is the model used in Hayes (in press).

The paper is centred on Hayes and MacEachern's (1998) results with regard to the inventory of quatrains in English folk verse. The conflict between well-formed but unattested quatrains and ill-formed but attested quatrains encountered by the former is resolved by Kiparsky as follows. On the one hand, he treats all unattested quatrain types as unmetrical, except where the existing gap in the corpus can plausibly be considered accidental – this is the case only with refrain quatrains, which are so infrequent as a whole that the data is unlikely to be a full sample. On the other hand, he treats quatrain types attested more than once as metrical. Also, he sets nursery rhymes aside, arguing that their metres are "too diverse to be entirely covered in the same constraint system as folk song quatrains" (2006:9).

In order to review Hayes and MacEachern's conclusions and integrate his own theoretical and methodological additions to the study of text-setting, Kiparsky re-analyses the corpus of folk songs examined in Hayes and MacEachern (1998) and tests the model in a collection of hymns by Watts.

The first part of Kiparsky's paper is devoted to summarising Hayes and MacEachern's (1998) findings, as well as adding some new ideas to enhance the

theory. As already observed, Hayes and MacEachern (1998) classify folk lines into four types on the basis of their rhythmic cadence – 3, 3f, 4 or G –, which has to do with the grid placement of the final two syllables. In a latter appendix to their paper they recognise other types. Kiparsky (2006:12) argues that types G, 3f and 4 can be grouped under the same metrical category, characterised by a monosyllabic last foot, which he calls 3'. The reason is that “with respect to metrical form, type G lines are indistinguishable from type 3f lines, but with respect to musical performance they are more like type 4 lines” (2006:12). This is the first argument used by Kiparsky to defend a separation between metrical and musical constraints. As he points out (2006:12), “[t]he *musical* implementation of such a final monosyllabic foot is appropriate to the tune to which it is sung [...] Within the limits imposed by that constraint the singer is free to decide and variation is free to occur”. The second argument has to do with the well-acknowledged fact that “the same words are commonly sung to different musical measures, yet maintain certain invariant constraints on stanza form” (2006:14). The third argument for separating constraints on musical performance from metrical constraints on stanza form is that “the metrical constraints are applicable also to literary verse that was never meant to be sung or chanted” (2006:15). Actually, as Kiparsky (2006:15) observes, literary inventions follow the same laws of stanza construction as folk poetry, namely parallelism and closure.

So, once again, the difficulty is discovering how text and tune are related in a song tradition. Kiparsky's first assumption is that text and tune have “a more or less firm conventional association” (2006:15), yet they are to some degree independent. The second assumption is that text and tune “can originate and develop separately, and lines, couplets, and entire quatrains can float from one song to another” (2006:15). For all these reasons, Kiparsky doubts that the relation of verse form to musical performance is as close as Hayes and MacEachern claim. In conclusion, Kiparsky (2006:16) argues for the validity of a modular approach, which “makes for a simpler metrical inventory, and correctly predicts why some line types alternate with each other, and why others do not occur at all”.

Kiparsky's paper is almost entirely devoted to grouping CON, that is, SALIENCY and PARALLELISM. With regard to SALIENCY, Kiparsky (2006:16)

states that closure can be achieved either by rhyme or by the presence of a refrain. There are two generalisations to be made about the interaction between these two CON, namely that “a couplet must not have decreasing saliency – saliency is measured by the inverse of length (shorter lines are more salient than longer ones)”, and that “a couplet is parallel if its lines are equally salient – a couplet must be either salient or parallel” (2006:17). So the two main grouping CON cannot be obeyed at the same time. In other words, one must be ranked higher than the other in order for stanzas to be well-formed.

Kiparsky (2006:21) conceives of SALIENCY as a gradient property, meaning that a line or another unit is salient in relation to another similar unit. In descriptive terms, saliency is “the inverse of length, measured in beats: a full foot (F) is non-salient, a reduced foot (f, Ø) is salient” (2006:21). At the theoretical level, SALIENCY is unfaithfulness, because it causes mismatches between metrical positions and the linguistic elements that correspond to them. In this way we can observe the interaction between grouping and metrical CON: “in English verse, saliency results from unfilled positions in the metrical grid, definable as a violation of the faithfulness constraint MAXBEAT” (2006:21). The interaction between the two types of CON is summarised as follows (2006:22):

- Full foot (F): [o o] (perfect match, non-salient).
- Degenerate foot (f): [o] (mismatch, salient).
- Null foot (Ø): [] (maximum mismatch, maximally salient).

At this point Kiparsky diverges from Hayes and MacEachern (1996) by observing that SALIENCY and PARALLELISM work at higher levels of metrical structure in recursive fashion. To those main grouping CON, Kiparsky (2006:23) adds a third one, CLOSURE, which states that the salient couplet of a salient quatrain contains no marked feet. This last constraint indicates that the final couplet of a salient quatrain is not only salient, but unmarked, i.e., it is of the form 43.

An interesting point in Kiparsky’s analysis is his in-depth explanation of the peculiarities of refrain quatrains in folk verse. First, “[r]efrain quatrains [...] do not have a rhyme, probably because the closure of rhyme is served by the refrain constituent” (2006:24). Second, they allow all metrical types, plus a few additional ones. Third, they “do not seem to consist of two couplets, but simply of four

coordinated lines ([4443']), or in some cases even three lines capped by a refrain ([444][3'])" (2006:24).

The main criticism made by Kiparsky to Hayes and MacEachern's account is that the latter follow stochastic OT in order to explain the ranking of CON, whereby "each constraint is assigned a place on a scale of real numbers, which governs its likelihood of outranking other constraints and being outranked by them" (Kiparsky 2006:27). In other words, the stochastic model predicts that the constraints are strictly stratified along the scale. Kiparsky (2006:27) identifies three problems with such an approach. On the one hand, it does not relate frequency to unmarkedness, a fact that seems to contradict the intuition that the most frequent metrical structures tend to be those which are the simplest. On the other hand, Kiparsky (2006:27) mentions "the harmonic bounding problem", which has to do with distinguishing two stanza forms that differ only in strictness. Kiparsky (2006:27) exemplifies this by stating that, in OT terms, "ABCB is harmonically bounded by ABAB". As for variation, stochastic OT posits that each CON has a range of fixed width, within which it can freely vary. The last problem is overgeneration, that is, Hayes and MacEachern's (1998) model predicts too many quatrain candidates. According to Kiparsky (2006:37), those three problems could easily be solved with the introduction of an extra CON, namely FAITHFULNESS: "The effect of Faithfulness is to license any candidate not excluded by higher-ranked constraints as metrical. Constraints ranked above this cut-off point restrict metricality, while constraints ranked below it are inactive" (Kiparsky 2006:33). If FAITHFULNESS outranked the metrical CON – grouped under MAXBEAT –, any input would be accepted, with the consequence that the output would be prose.

The solution given by Kiparsky (2006:39) has to do with what he calls "a partial ranking account", meaning that "rankings can be either free or fixed". In any case, FAITHFULNESS "must be dominated by at least one markedness constraint" (2006:39). According to this, there are three possible grammars for this system, namely:

- a. SALIENCY >> FAITHFULNESS
- b. PARALLELISM >> FAITHFULNESS
- c. SALIENCY, PARALLELISM >> FAITHFULNESS.

Summarising, the single additional restriction that SALIENCY is more important than FAITHFULNESS has a double beneficial effect in the system. On the one hand, it excludes the prohibited couplet types and, on the other, it generates the pattern of preferences among the remaining permissible couplet types. Among grouping CON, SALIENCY is preferred over PARALLELISM at the couplet level, something which can be observed both in the corpus of folk songs and in Watts' hymns. According to Kiparsky (2006:42), the reason for this ranking "lies on the musical side", that is, "it is a feature of stanzas intended to be sung". In contrast to this, PARALLELISM would dominate "in literary verse designed for reading rather than singing" (2006:42). At the quatrain level, PARALLELISM outranks SALIENCY, something which Kiparsky is not able to explain. As a sort of conclusion, the author (2006:44) states that "parallelism supersedes saliency as the dominant organising principle" – actually, "above the level of the stanza, parallelism is almost completely dominant".

4.2.9. Summary

In this section I have introduced the main theoretical and methodological principles on which OT relies for the study of text-setting.

The assumption made by OT in this particular field is that text-setting, that is, how lines of linguistic texts are arranged in time against a predetermined rhythmic pattern in sung verse, is a universal ability. Among the numerous text-setting CON explored in Hayes and Kaun (1996), Hayes and MacEachern (1996, 1998), Hayes (in press), Halle and Dell (in press), Halle (forthcoming) and Kiparsky (2006), there are three essential ones – MAXBEAT, SALIENCY and PARALLELISM – which can be subdivided as follows:

a) Metrical CON: MAXBEAT

a.1) MATCH STRESS:

a.1.1) strong syllables fall on strong beats

a.1.2) a syllable that falls on a S beat must be stressed

a.2) FILLSTRONG: fill the strongest positions in the line

a.3) *LAPSE: avoid sequences with no syllables between any two of the strongest positions in the line

b) Grouping CON:

b.1) SALIENCY (truncation): non-filling of metrical positions at the end of lines

b.1.1) Lines are salient

b.1.2) Couplets are salient

b.1.3) Stanzas are salient

b.2) PARALLELISM:

b.2.1) the cadences ending the units of the maximal analysis of a quatrain must be identical

b.2.2) the onsets of syllables coincide with musical beats.

Both the interaction and the conflicts among the above CON determine and are determined by the type of language that we are dealing with. In a language like English, traditionally classified as stress-timed, the metrical CON MAXBEAT and, more specifically, MATCHSTRESS, regulate the process of text-setting in folk song. In other words, if a text-setting in English is to be qualified as well-formed, the agreement between musical beats and linguistic stresses must be as high as possible. On the other hand, a language like Spanish, classified as syllable-timed, will work in a different way with regard to MATCHSTRESS. As Dell and Halle (in press) observe in the case of French, in Spanish the agreement between beats and stresses is not enforced in strophic song; in fact, the cases where MATCHSTRESS is highly ranked are scarce. As a result of these observations, we can state the preliminary conclusion that the importance of stress-placement as a determinant of text-setting well-formedness in folk song is systematically diminished in syllable-timed languages, while it is enhanced in stress-timed languages. This fact has a clear impact on the behaviour and ranking of grouping CON such as PARALLELISM in different languages. As explained above, positional parallelism requires that the distributions of syllables along the grid are identical for each stanza of a strophic song. Consequently, if lines are not isochronous, they cannot be positionally parallel. In other words, only languages whose verse prosody requires isosyllabism – like Spanish – can render positionally parallel text-setting, while languages where the number of syllables per line is secondary to the position of stresses will hardly require positional parallelism to be judged well-formed. The analysis of the most recent papers on the topic of text-setting confirms that in vocal music there are three

tiers of rhythmic structure – linguistic prominence and grouping, verse metre and grouping and musical rhythm and metre –, which are intimately related. While the constraints applied at the level of speech prosody cannot be directly translated into verse prosody constraints, they determine the ranking of constraints at that level, and the same happens with the influence of both speech prosody and verse prosody constraints on text-setting constraints. As explained in this chapter, in English the link between verse prosody and musical metre is stronger than in Spanish. This entails that MAXBEAT is enforced in English, not in Spanish, while positional PARALLELISM is enforced in Spanish and constantly overlooked in English.

These preliminary conclusions need to be tested against a corpus of folk songs. In ensuing sections I will apply the theories reviewed to specific case studies selected from a corpus of 444 Spanish songs and 239 English songs – analysed in section 4.3.6. and appendices IV and V. The main criteria used to select this specific corpus, which I explore in full in section 4.3.5., could be reduced to two. On the one hand, they are folk songs, passed from generation onto generation, which entails that the matching of lyrics and music responds to some intuitive understanding of text-tune alignment. On the other hand, in opposition to the songs used by the authors reviewed above, the ones analysed here have been minimally edited by the collectors.

4.3. Object of study: Folk song

This section is devoted to discussing the genre of folk song, to which the corpus of songs analysed in my dissertation belongs. Section 4.3.1. explores the meaning and evolution of the terms ‘song’ and ‘folk song’, pointing to the most important characteristics of the genre, as opposed to ‘art song’. Since the goals of my dissertation are linguistic and, to a certain extent, musicological, but not anthropological, it is beyond the purposes of this dissertation to cover every aspect of folk culture. Section 4.3.2. explores the transcription and classification criteria used by the collectors of folk song and, more specifically, of the two collections that I am analysing, namely Kennedy (1984) and Manzano Alonso (2003). I shed some light on the controversies related to ideas of ‘faithfulness’ to the originals as sung by the informants. In section 4.3.3. I carry out a brief analysis of the possible reasons why

lyrics have always been given more importance than melodies in the study of folk song. In connection to the previous issue, section 4.3.4. introduces the topic of the relationship between lyrics and melody in folk song, which is the nucleus of my dissertation, with special emphasis on the concepts of lyric and melodic 'identity' and 'variation'. Section 4.3.5. constitutes a brief introduction to the corpus analysed in my dissertation. Finally, section 4.3.6. is devoted to the analysis of the first sample corpus, made up of two folk songs, and presents some preliminary conclusions on the topic of text-setting in different languages.

4.3.1. Definition and characteristics of folk song

As a musical form, 'song' can be defined as an integrated whole of inseparable functional elements. There is an indispensable melodic element to which lyrics, that is, a verbal text endowed with meaning, is joined. This composite of tune and text can be associated with dance, that is, a series of controlled body movements in accordance with the melody and the rhythm of the song. Songs are often accompanied by sound-producing instruments which enhance the melody or help establishing the rhythm. In this sense, song is a complex cultural manifestation.

As a musical genre, song can be subdivided into two broad but sharp categories, namely art song and folk song. As will be explained in the following pages, the main difference between the two has to do with an exogenous factor, that is, whether the songs have been kept written down or whether they have passed orally from one person to another – mainly in rural environments and in connection to the life and habits of country people – in the course of history, until they have become a part of a specific musical and, more broadly, cultural tradition.

When talking about folk song we need to acknowledge that the term 'folk' has experienced a considerable semantic evolution over the years. As Kennedy explains (1984:3ff.), at first 'folk' referred to unlettered country people or peasants. During the Romantic period, the word gradually acquired another nuance and definitely left its derogatory sense behind. At the end of the nineteenth century, the word 'folklore' was coined in order to refer to local beliefs and customs. It was at that time that the music of the folk, both songs and dances, started to attract the attention of composers and musicians looking for new literary and musical art forms, and by people working in educational fields. The words 'folk song' are a literal

translation of a German word, *Volkslied*. Folk song is, therefore, only one of the multiple manifestations of folk culture, evolved as the behaviour pattern of non self-conscious peoples. As Kennedy (1984:8) points out,

These include [...] those country folk who still depend on an unconscious process of acquiring their traditional craft skills and lore, such as the oral transmission of music, poetry and stories. Traditional songs and ballads reflect the social conditions and ways of life of a community.

There are several issues which make folk song a difficult genre to deal with. First of all, the definition of folk song has been controversial amongst ethnomusicologists of all times. Second, stemming from the difficulty to find a unifying definition, there is the issue of determining the characteristics that make folk song what it is. Third, when we come to deal with the scholarly study of folk song, all sorts of methodological problems emerge, ranging from the recording and collecting of the songs to their transcription and the grouping criteria used to classify them. All these difficulties are directly linked to the relatively short time span to which we can resort in order to hypothesise about folk song. As the Spanish ethnomusicologist Manzano Alonso (2001:31) points out, “the hypotheses about the origins of these songs can only go as far behind as five or six centuries, and never become anything else than hypotheses”. One of the first and most influential folk song collectors and scholars in Britain was Sharp (1859-1924), whose volumes *English Folk Song: Some conclusions* – originally published in 1907 – and *English folk songs, collected and arranged with pianoforte accompaniment by Cecil J. Sharp* – first published in 1916 – set a milestone in the academic study and, indeed, in the folk song revival witnessed at the beginning of the twentieth century in England.⁶⁰ Sharp’s perceptions about the nature of folk song are still taken to be basic for any

⁶⁰ The revival of British folk music occurred in two distinct phases during the twentieth century, namely before World War I and in the period between wars. The most notable folksong collector in the decade before World War I was Cecil Sharp. After World War II and the ensuing period of austerity, there grew a national desire to celebrate British culture and tradition. The BBC undertook to coordinate the fieldwork. Expertise for the project, referred to as the *BBC Folk Music and Dialect Recording Scheme*, was drawn from folk music experts such as Peter Kennedy and Seamus Ennis.

study of the topic. In fact, Sharp's (1954:1) definition of 'folk song' has become classical in the literature on the topic:

In every land we do find music of a distinctive and often of a very beautiful quality, of a beauty and character of its own which colour although they may not be entirely shared by the educated of art-music of the same nation. This spontaneous utterance is called folk-song.

This definition is not without problems. Sharp plays around with certain highly subjective terms which are not always convincing *per se*. What does 'having a beautiful quality' mean? Who are the unlettered classes – or who were they at the time? What are the specific limitations of folk song? The second of these questions is addressed by Sharp (1954:4), who explains that "the non-educated, or 'the common people', are the unlettered, those whose faculties have undergone no formal training whatsoever, and who have never been brought into close enough contact with educated persons to be influenced by them". So, according to Sharp, folk song is defined by its beauty, its common origin and its spontaneity, among other things. Apart from this, Sharp (1954:15) emphasises the "communal, never completed, multiple in form" nature of folk music, as well as the "radical importance" of lyrics in this genre (1954:19).

In the above quotations, the oral quality of folk song is hinted at through an allusion to the 'unlettered people', who were presumably musically as well as linguistically unlettered. Orality is, in fact, the most important defining element of folk song in opposition to art song, which is typically written down. In order to complete Sharp's definition, Lloyd (1967:16) suggests that

in its natural state a folk song is poetry and music *perpetuated by mouth-to-mouth transmission not by print*; it is founded on certain inflexible principles but subject to personal variation; its acceptance and survival depends on how well it accords with the tastes, views and experience of the community (emphasis mine).

Sharp's initial ideas, though undoubtedly valuable at the time when they were stated, have encountered disagreement among present-day scholars. To start with, his emphasis on oral transmission as a defining element of folk song has to do, according to Russell (*Grove Music Online*), with a controversial "post-Darwinian theory for its

evolution based on concepts of continuity, selection and variation”, a view which “led to a preoccupation with identifying songs that fulfilled such criteria”, thus leaving aside songs which could otherwise have qualified as folk songs. Russell (*Grove Music Online*) also points out that Sharp’s conception “led to concentration by scholars on the modal characteristics of folk melodies”, a fact which has raised many problems when dealing with tonal melodies. In spite of all these observations, Russell (*Grove Music Online*) acknowledges that traditional music is indeed oral in the sense that it “does not exist on the printed page but in the performances of individual singers or instrumentalists, in the contexts of family gatherings, singsongs in public houses, meetings of social clubs and other social groupings”.

The issue of authorship has also been a constant topic of debate for ethnomusicologists, who have not been able to reach an agreement as to whether folk songs are composed by a community, or by an individual, like art songs. Manzano Alonso (2001:44) categorically states that

we must abandon the idea that popular music is the result of a collective process of creation. *The creator is the individual, both in popular and cultivated music.* In both cases the individual creates within a group from which she gets the structures used, and then the group accepts [or not] the created work (emphasis mine).

Not only does Manzano Alonso state the individual authorship of folk songs but he also emphasises the professionalism of folk song composers as well as the mutual influence exerted by folk song and art song composers, an idea which gets away from the widely accepted view that folk song composers are uncultivated: “European popular song had to have emerged from musical sources created by more or less professional authors, and vice versa. Professional authors used musical elements taken from popular music” (Manzano Alonso 2001:33). According to Manzano Alonso’s statement, it is not authorship that establishes the radical difference between folk song and art song. As he (2001:34) points out, “folk song and art song have co-existed for many centuries. The former is oral and uses archaic elements, while the latter is written and has left those elements behind some centuries ago”.

Summarising, what are the formal differences between a folk song and an art song? I have already mentioned the presence of certain archaic elements in the

former. But the most important difference between the two song subgenres has to do with the oral quality of folk song. Given that folk song is not written down, the composer is faced with the need to create a short form, which can be remembered by people. Thus, composers of popular music only develop their ideas in a short space, given that any object of oral transmission requires brevity – it is important to note that brevity has to do with the musical structures, which are normally repeated several times in the same song, with different lyrics.

Linked to the brevity of forms are the formal or structural characteristics of folk song, which have to do with simplicity, a clear narrowness of melodic range, as well as the afore-mentioned archaic character of their musical configuration. ‘Archaic’ elements are those which cannot be found in art song in the last four or five centuries but have nevertheless survived through the centuries in folk song. In Manzano Alonso’s (2001:33-34) words, “those elements are (i) modal melodic systems, (ii) limited melodic range, (iii) instabilities and chromatic movements, (iv) freedom in the development of the melodies, (v) irregularities in the rhythmic formulae”. The issue of modal melodies is a tricky problem, as the *Volkslied*, which gives its name to the whole genre, never ceased to be sung by all classes in Germany, and thus changed stylistically over the centuries, so that a *Volkslied* of 1820 sounded like a simple song by an art-composer of that period. The English folk song sounded archaic to its collectors around 1900 because it had been forgotten by educated people and thus had scarcely changed since the late Middle Ages. The same happened with Spanish folk song.

What Manzano Alonso shares with traditional ethnomusicologists is the idea that folk song has its origin in rural areas: “Popular music of oral transmission is linked to the lives and costumes of mostly, though not exclusively, rural people” (2001:29). This link to the lives of people has to do with the function of folk song. While cultivated music enjoys a time and space which we call ‘concert’, a performer and a hearer, popular music has always been integrated in people’s lives, at any time, and has never had specific protagonists or receivers – with the exception of certain examples of instrumental music. Nevertheless, with the loss of rural life and the coming of urbanisation, the relationship between music and function has somehow been lost (Manzano Alonso-Alonso 2001:63-65).

A further, not uncontroversial, characteristic of folk song as opposed to art song is the fact that most of it is unaccompanied, that is, performed by a solo singer – sometimes accompanied by simple percussion instruments. Russell (*Grove Music Online*) observes that, although “unaccompanied singing has been the most common form of traditional singing that has been recorded during the past century and a half, this does not presuppose that it had always been the case”.

The last essential characteristic of folk song is that it is strophic, that is, it is organised into musical stanzas which share the same melodic contour and structure – often with a certain amount of variation – to which the lyrics, also organised into stanzas, are set over and over again. This characteristic is very important for my purposes, as it directly affects the process of text-setting in different folk traditions.

4.3.2. Transcription and classification criteria

A good part of the folk song repertoire has been collected from the nineteenth century in what we call ‘folk song collections’ – Spanish *cancioneros*. The notation used to transcribe the folk songs that make up these collections is just another one of the controversial aspects in the study of this genre. Traditional ethnomusicologists are reluctant to render folk song transcriptions in music notation, arguing that this type of notation does not have enough means to account for the innumerable melodic and rhythmic nuances of folk song. It is obvious that the only way of remaining faithful to the original songs is to rely completely on field recordings. As Parfrey (1984:13) observes in a short introduction to Kennedy’s (1984) compilation, when researchers render collected songs into conventional musical notation, they must acknowledge a certain amount of ‘unfaithfulness’ to the sung version. As a consequence, the researcher has to accept that ordinary music notation can only serve as a basic guide to an idiomatic original,

for whereas conventional music notations can show the basic essentials of the tune, a really accurate scientific notation of the source singer’s performance would mean using some kind of continuous graph tracing of the sound which could follow all the subtleties and idiosyncrasies of the original performance by the traditional singer.

In spite of all the alleged unfaithfulness of folk song transcription, it is important to bear in mind that the ultimate goal of transcribing popular music is to make it

possible for the researcher to extract conclusions about its basic elements, namely, its melodic system, its intervals, rhythmic organisation and melodic structure. As Manzano Alonso (2001:232) emphasises,

Although a live performance and/or direct recording will always be the most valuable documents in order to analyse a specific type of music, it is only through the analysis of a written score that songs can be studied, classified, analysed and related to one another.

In sum, traditional music is transcribed to be analysed in terms of melody and texts, and not necessarily to be performed.

In the two song collections used in my dissertation – Kennedy (1984) and Manzano Alonso (2003) –, transcriptions have been made departing from recorded documents, according to the following criteria:

- a. Each song is written as if it was to be sung, so the register and key signature are chosen to be easily intelligible.
- b. The key signatures used keep the melody's character. There is a tendency not to use complex key signatures.
- c. Time signatures are used like in art-music. There are no open bars or suppressed bar lines. Time signatures tend to be appropriate to the basic rhythmic structure and the tempo of the piece.
- d. Melodic variants: when the melody changes, the transcribed melody is the most frequently used for the song.

Most ethnomusicologists follow a classification which has also been used to classify oral poetry. The most common categories would be, according to Finnegan (1977:12-13), the following:

- a. Epic: narrative character, long.
- b. Ballad: sung, narrative poem, shorter than epics, concentrated on one episode.
- c. Lyric: non-narrative.
 - c.1. Love lyrics.
 - c.2. Psalms/Hymns.
 - c.3. Dancing songs.
 - c.4. Political/Topical verse.
 - c.5. War songs/ Initiation songs.

c.6. Spirituals/Laments.

c.7. Work songs.

c.8. Lullabies.

In order to classify the corpus of songs analysed in my dissertation, the collectors used a mixed criterion, taking into account not only the functional aspect of the songs, which is the basis of Finnegan's classification, but also the musical aspects and the ones dealing with the contents of the texts.

4.3.3. The importance of lyrics in the study of folk song

Nearly all the popular literature written before the twentieth century that we know of comprises song lyrics. However, Russell (*Grove Music Online*) observes that “[f]or the folksong collectors it was the beauty of the melodies that was the first priority, rather than the texts or performance styles of the singers from whom the songs had been notated”. Despite this undeniable fact, there is still no discipline in present-day academic organisation that studies songs as unitary wholes which comprise a musical element, a linguistic element and a literary one. The musical side of songs has often been neglected or minimised by literary scholars in order to pay attention to their area of expertise, that is, lyrics. Among linguists, both diachronic and synchronic, the in-depth analysis of the musical setting of a text or a string of sounds has seldom been acknowledged as an essential step for the study of the history of a specific language or of its prosodic characteristics at a specific point in time. The scientific protection of the lyric side of songs in contrast to the neglecting of their music, together with the fact that throughout history the mechanisms for song transmission have proved to be naturally more efficient for texts than for melodies, has resulted in the unsurprising outcome that we have kept many more popular lyrics than melodies. As a consequence, scientific studies have not paid much attention to the study of song as a three-dimensional phenomenon, with a text, a tune and the result of the interaction between those two.

Ethnomusicologists have always emphasised the importance of lyrics in the genre of folk song. Kennedy (1984:10) questions the musical status of the songs, saying that the tunes “are nearly always completely subservient to the story or emotion, rather than forming the basis of any great musical achievement”. We could say that the importance of the study of folk songs from a linguistic perspective has to

do with the fact that their musical side was neatly linked to their linguistic and literary side in such a way that the interaction between certain characteristics of the language and their musical setting respond to a series of unified rhythmic criteria, which presumably cannot be violated gratuitously. The fact that the most salient characteristic of folk songs is their strong rhythmic vitality points at a deep connection between language prosody and musical rhythm and metre.

In spite of the apparent differences between art and folk song, the two genres share a series of characteristics with regard to lyrics, namely

- a) linguistic simplicity for a successful communication
- b) the most commonly used devices are the most elementary and powerful, such as repetition, antithesis, hyperbole, irony
- c) the use of linguistic formulae – starting, greeting – is pervading
- d) the style of versification has to do with the use of simple stanzas and rhyme.

4.3.4. Text-tune relationships in folk song

As repeatedly mentioned in this dissertation, a song is a composite of two main elements, namely text and tune. The process of connection between the two elements can take place in different ways, which is worth taking into account not only when we deal with the differences between art song and folk song but also when we analyse the principles of text-setting in general. In art song and other vocal genres of art-music, what usually happens is that the composer creates a melody for a pre-existing text – be it a well-known text, a text created by the composer herself or a text specifically written in order to be set to music. In any of the previous cases, the text exists prior to the music and the music is created as a melodic declamation of the text, so to speak. The process in popular music is completely different. Very often we do not know whether a text existed long before the melody was created and, more important, the relation between text and tune is not bi-univocal: different texts can be set to the same melody, and vice versa. Thus, it is not accurate to think that each song has a specific text which can only be associated with that tune, as would be the case in an aria, for instance. The same text often appears joined to different melodies and vice versa, a melody is adapted to fit different texts. The possibility of interchange between texts and melodies does not mean, however, that none of the texts has been created for a specific melody. We can presuppose that each text was

originally associated with 'its tune'. Nevertheless, it seems correct to assume that, given the oral character of folk song, as time passed texts and melodies started to be dissociated and re-associated with new tunes and texts, respectively, until the original connection was completely lost. Sharp (1954:25-26) explains this process in terms of variation stages that lead onto the creation of new songs:

A frequent source of variation arises when the singer, having partially forgotten his words, has substituted corrupt and unmetrical lines. The attempt to adapt the tune to these irregularities will often lead to the inventions, unconscious of course, of interesting melodic changes [... Another] source of variation springs from the attempt to adapt an old tune to new words of slightly different metre [...] If, as must frequently happen, the metre of the word is not exactly the same as that of the tune, they will alter and adapt the latter to meet the new conditions.

This possibility for interchanging texts and melodies in popular music is an essential factor to be taken into account when analysing text-setting. This factor gives rise to a very singular relationship between text and tune, often comprising a series of frictions and lack of adjustments that popular music solves in very idiosyncratic manners.

The question arises of why it is relatively easy to set virtually any given folk text to any folk tune. The reason for this is metrical:

The usual stanza of poetry contains four measured lines, not necessarily of equal length, though forming a just balance; normally, of four phrases, the points of division being marked off by means of cadences (Sharp 1954:72).

As Manzano Alonso (2001:66) explains, following Sharp's statement, "the reason for this is that the poetic measure of most texts is the same". As will be observed in the analysis of the case studies (4.3.6.), most Spanish texts have the same syllabic and stanzaic measurement. English traditional song is also based on the stanza and sung to a musical cycle or tune that is repeated. As Russell (*Grove Music Online*) explains, stanzaic songs

usually have a four-line isometric structure with a simple rhyme scheme (usually ABAB, AABB, or ABCB), each line organised into alternate

lines of four and three feet (ballad or common metre). Long metre (four lines of four feet) is also prevalent. Some songs contain refrain lines; others have a more formal chorus section, notably the monumental chorus of the later parlour ballads. Such items demonstrate a more complex and varied structure.

The only exceptions to the interchangeability of texts and melodies are refrains, which are usually less flexible than stanzas because they normally have more peculiar poetic forms and they individualise each song from the rest (Manzano Alonso 2001:67). As an illustration to this, in the corpus of Spanish songs analysed in my dissertation, 70 per cent of the texts are octosyllabic quartets – or hexadecasyllabic lines which can be divided into eight plus eight syllables –, 20 per cent are *seguidillas* – a very widely-used stanza in Spanish poetry and music, composed of four lines alternating seven (or six) and five syllables, rhyming –a–a – and 10 per cent are really varied poetic measurements and stanzaic formulae. There can be no doubt that the overwhelming presence of the octosyllabic quartet (*cuarteta octosilábica*) and the *seguidilla* in the lyrics of the songs conditions the rhythmic patterns and the melodic developments of the songs.

This interchange phenomenon, which is nowadays exclusively associated with folk song, was also accepted as common practice in art song. It was common in the eighteenth century for new words to be written to existing operatic arias; there were whole operas (called *pasticci*, ‘jumbles’, *Grove Music Online*) in which old music was used, sometimes with its original words, and with new words on other occasions.

4.3.5. The choice of a specific folk song corpus

The following two sections contain a brief summary and explanation of the contents of the two folk song collections which compose the corpus analysed in this dissertation.

4.3.5.1. Manzano Alonso's *Cancionero de Burgos*

This volume is the fifth in a collection of six volumes where Manzano Alonso transcribes hundreds of songs collected in the Spanish region of Burgos. Each of these six volumes is devoted to a specific topic related to the function of the songs. The fifth one, entitled *Canciones del ciclo anual y vital* (‘Songs of the annual cycle

of life'), comprises all those songs which refer to and complement traditions which are not linked to liturgical or religious rites. These are all profane songs related to the cycle of field works and/or *fiestas*. In this dissertation, I have analysed 444 songs in terms of the OT metrical constraint MATCHSTRESS.

4.3.5.2. Kennedy's *Folk songs of Britain and Ireland*

This is a unique volume which contains 360 songs collected by Kennedy throughout Britain and Ireland. It is divided into sixteen parts, classified following two different criteria. On the one hand, there are six parts grouped according to the language and the region where the songs were collected. Thus, we find I Songs in Scottish Gaelic, II Songs in Irish Gaelic, III Songs in Welsh, IV Songs in Manx Gaelic, V Songs in Cornish and VI Songs of the Channel Islands. On the other hand, there is a vast collection of songs in British English, which are subdivided according to the function of the corresponding songs. In the group of songs in English we find the following subgroups: VII Songs of Courtship, VIII Songs of False love and true, IX Songs of Seduction, X Songs of Uneasy Wedlock, XI Songs of occupations, XII Songs of country life, XIII Songs of good company, XIV Songs of diversion, XV Songs of newsworthy sensation, XVI Songs of the travelling people. In this dissertation, I have analysed the 239 songs in English. As with the Spanish corpus, I have applied the metrical CON MATCHSTRESS to this corpus, in order to be able to draw a statistical chart which would work as an indicator of tendencies.

4.3.6. Analysis of the folk song corpus: overview and general observations

Given the high number of songs contained in the chosen corpus – a total of 683 –, for the purposes of this dissertation I have opted for a clear-cut analytical method, which neatly shows the ways in which music and lyrics interact in English and Spanish. As stated in the introduction, one of my objectives is to define the nature and properties of linguistic stress in each of these two languages, which implies observing how syllables are aligned with musical beats in song. The best way to observe the interaction between syllables and beats is applying the metrical constraint MATCHSTRESS to each of the 683 songs in my corpus.⁶¹ After applying MATCHSTRESS to both song collections, we can conclude, first of all, that both the Spanish and the English songs show lexical and grammatical mismatches. However,

⁶¹ The detailed analysis of the mismatches in the 683 songs appears at the end of this dissertation, as Appendix IV (Spanish songs) and Appendix V (English songs).

the number of stress mismatches in lexical words in the English corpus – that is, nouns, verbs, adjectives, adverbs – is comparatively low. Moreover, the cases where a lexical stress is mismatched can be grouped into three categories. First, we encounter a group of disyllabic nouns with initial stress, which are nevertheless set so that the musical beat falls on the second syllable, a fact that gives rise to lexical mismatches. The mismatched nouns end with a <y> – or its spelling variant <ie>, both of which correspond to the phoneme /ʌ/. This is a peculiar phoneme, known as “neutralised /ʌ”, which shows two unique characteristics, namely (i) it is neither short nor long and (ii) it is tense. In other words, this phoneme cannot be reduced to schwa (/↔/), which might be a reason why it can be stressed when it is set to music.

The nouns which undergo this type of mismatch are *pen-nie*, where *-nie* falls on the second beat of a 6/8 bar, which is also the last beat in the song, and therefore, of a musical phrase; *la-dy*, where *-dy* falls on the second beat of a 6/8 bar, which is also the last bar in the line; *coun-te-rie*, *-rie* falls on the first beat of a 6/8 bar, which corresponds to the end of a line; *North Coun-te-rie*, where *-rie* falls on the second beat of a 6/8 bar. In all these cases, mismatches coincide with the end of the corresponding musical phrase. This type of mismatch is idiosyncratic, to the extent that it can be said to be a style or genre marker. Something similar happens in the word *mor-ning*, where *-ning* falls on the first beat of a 3/4 bar, which corresponds to a minim and is, once again, the last figure in the song. In sum, this first type of idiosyncratic lexical mismatch which occurs in disyllabic words with initial stress at the end of musical phrases can be called “*Country mismatch*”.

Second, there are two instances of mismatches occurring in compound nouns, namely *shoe-ma-kers*, where *-kers* falls on the last beat of a 4/4, which is a secondary beat, and therefore gives rise to a very gentle mismatch; *game-kee-per*, where *-per* falls on the second beat of a 6/8. In these cases, the fact that the mismatch does not occur on the main beat of the corresponding musical bar minimises its impact and makes it acceptable.

Last, there is only one instance of mismatch which occurs on a verb, namely *be-came*, where *be-* falls on the third beat of a 3/2, which makes it a relatively acceptable mismatch.

Summarising, lexical mismatches are rare in the English corpus. Most of them belong in the class that I have called “*Country mismatch*”, where a tense vowel that cannot be reduced is made to fall on a strong beat. As mentioned above, this is a style marker.

Moving on to grammatical mismatches, these are more numerous than lexical mismatches in the English corpus, and respond to what Giegerich (1978) calls “Hierarchy of stressability of function words”. According to this, certain function words can be stressed without further consequences, while others cannot be stressed under practically any circumstances. According to Giegerich’s scale of stressability – reproduced below –, prepositions and conjunctions are never stressable, followed by auxiliaries, pronouns, modals, wh- words and demonstratives. The number of mismatched words in the corpus appears under the corresponding category:

less stressable	more stressable
others < PREP, CONJ < last AUX < PRO < MODALS < WH- < DEM	
Count: 49 12 7 (5) 23 (2) 2 3 2	

As we can observe, the number of mismatched function words in each category coincides, to a certain extent, Giegerich’s scale. Prepositions are frequently mismatched (49 instances), followed by pronouns (25), conjunctions (12) and auxiliaries (12). On the other hand, demonstratives, wh- words and modals are hardly ever mismatched in the corpus. This could also respond to Giegerich’s (1978) second hypothesis for the stressing of function words, the so-called “Rhythmic Stress Change”, whereby a monosyllabic function word can be assigned stress (Table 30).

			[+DEM]		
			[+WH-]		
			[+MOD]		
[-stress] →	[+stress] /	[-stress] #	[+PRO] #	[-stress]	
			<AUX>a	<-AUX>b	
			[+PREP]		
			[CONJ]		

Table 30: Rhythmic Stress Change

In the English folk song corpus, most mismatched function words are monosyllabic, with the exception of the prepositions *unto*, *into*, *over*, the adverb *until* and the auxiliary verb *could* with the negative particle attached, *couldn't*, each of which appears mismatched only once.

As a summary, below is a statistical summary of all the mismatches observed in the English folk song corpus. If the same mismatch happens more than once in the same song, it counts as one mismatch only.

Number of songs: 239 Songs with mismatches: 62 = 25.94%

Total number of mismatches: 118

1. On lexical words: 8 = 6.78%

a) Nouns: 7 b) Full verbs: 1

2. On grammatical words: 110 = 93.22%

a) Prepositions: 49 b) Conjunctions: 12

c) Pronouns (rel.): 23 (2) d) Determiners (art./dem.): 13 (2/2)

e) Auxiliaries (mod.): 7 (2) f) Particles: 6

The Spanish corpus is much richer and complex in terms to syllable/stress/beat mismatches. There are nearly as many lexical mismatches as grammatical mismatches, although the latter are more frequent and often appear several times in the same song. Prepositions, determiners, conjunctions and pronouns are the most widely mismatched grammatical words, while nouns and verbs are, by far, the most commonly mismatched lexical words. The general tendencies are the same as in the English corpus, but the number of mismatches is much higher in the Spanish corpus.

Unlike in the English corpus, where lexical mismatches can be grouped and accounted for as idiosyncratic, in the Spanish corpus they do not respond to any single stylistic criterion. Indeed, the act of mismatching lexical words constitutes the style. By observing the behaviour of mismatches in this corpus (see Appendix IV), one could easily conclude that Spanish folk song allows for all kinds of lexical mismatches. This brings on an important issue, which has to do with the subgenre of folk song to which most songs in this corpus belong. As can be observed in Appendix IV, the majority of the songs contained in this specific corpus are dance

songs, that is, they were traditionally performed in order for people to dance. This implies that both the tempo and the metre of the song are bound to be rigid – definitely more so than in a ballad, which is performed *ad libitum*, often by a solo singer. As a result, we can presume that words will be mismatched on even more occasions than in another type of song. However, this does not necessarily mean that all languages allow for more mismatches in dance songs; in English, lexical mismatches are generally disallowed regardless of what function the song fulfils.

The mismatches found in the Spanish corpus are summarised below:

Number of songs: 444

Songs with mismatches: 292 = 65.77%

Total number of mismatches: 945

On primary counts: 920

On secondary counts: 25

1. Lexical words: 449 = 47.51%

- | | |
|-------------------|----------------------|
| a) Nouns: 264 | b) Full verbs: 128 |
| c) Adjectives: 36 | d) Adverbs: 11 |
| e) Numerals: 7 | f) Tonic pronouns: 4 |

2. Grammatical words: 494 = 52.28%

- | | |
|------------------------|--------------------------|
| a) Prepositions: 156 | b) Conjunctions: 87 |
| c) Pronouns: 69 | d) Determiners: 154 |
| e) Auxiliary verbs: 12 | f) Relative pronouns: 15 |
| g) Particles: 2 | |

In the next two sections, I develop an in-depth analysis of two case studies, one extracted from the English corpus, and one extracted from the Spanish corpus. These analyses are intended as detailed applications of the metrical and grouping constraints explored in the literature review, which will lead to an informed view of the nature of stress and the syllable in the two languages.

4.3.7. Case studies

One of the difficulties in choosing two songs for the case studies had to do with the criteria underlying such choice. Both song collections comprise dance and non-dance songs, songs with a slow tempo and songs with a fast tempo, and the time signatures vary from song to song. The chosen songs have a number of common characteristics, namely (i) they are strophic songs, (ii) they are not dance tunes and they are meant to be sung without any accompaniment (iii) they are written in binary time signatures (4/4 and 2/4, respectively), (iv) they have roughly the same number of bars (16 and 15). In both cases, the constraints applied for the analysis of text-setting are MATCHSTRESS, *LAPSE, FILLSTRONG, PARALLELISM and SALIENCY.

4.3.7.1. Case study I: “The farmer’s boy”

“The farmer’s boy” is a strophic song collected in Kennedy’s (1984) *Folksongs of Britain and Ireland*, within 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. The song is a narrative divided into four four-line stanzas, each of which is followed by a refrain (in italics in Table 31) 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 31 below, the number of syllables per line goes from six to nine. There are two instances of synalepha (underlined in Table 31), 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.⁶² The case of *little*, pronounced /^llɪtl/, corresponds to Fabb’s (2002:10) Non-projection rule c (quoted in section 3.3.3. and repeated here for convenience):

⁶² The non-phonological status of Non-projection rule b is arguable since 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 to this, observe how in 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.

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.

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 (quoted in section 3.3.3. and repeated here for convenience):

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 (xx/) 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* (4th line), *I* and *you* (5th line), *will* and *me* (6th line), *if* and *me* (9th line), *this* (12th line), *her* (18th line), *what* (19th line), *for* (28th line) – as well as the demotion of certain stressed syllables – *ask* (5th line), *give* (6th line). Promotion and demotion of syllables is a thoroughly standard rhythmic device in English poetry and can be compared to the phenomenon of stress shift in English prosodic phonology. As for rhyme, even lines (trimeters) rhyme, while odd ones (tetrameters) do not rhyme.

Line	Syllables	Rhyme	Feet (accents)
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 to the farmer's door.	6	a	x/x/x/Ø
<i>May I ask you, if any there be</i>	9	–	x/x/x/xx/
<i>That will 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 if that thou won't me 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 this 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 her five children small	6	e	x/x/x/Ø
And what is worse for mother still	8	–	x/x/x/x/
I'm <u>the</u> eldest 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 for 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 31: Scansion of "The farmer's boy" lyrics.

The fact that the number of syllables changes from line to line in English is not unimportant since 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 that one of the two synalephas realised in the scanned version of the lyrics – the one between *the* and *eld-* in the 20th line – is eliminated in the musical setting 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.⁶³ In this song, as in most songs, the only factor that must remain constant is the occurrence of accents at equal intervals, something which is automatically achieved in vocal music since 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 setting of the text is perfectly aligned with the scansion given in Table 31 above, 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 thus say that the agreement between verse prosody and music rhythm and metre is virtually total in this song and, presumably, in English folk song.

⁶³ In order to complement the written analysis of the songs, I have included a CD with the recordings of the four case studies – two folk songs and two art songs – examined in my dissertation.

The sun went down, be - yond yon hills. A - cross yond rear - y moor When wear - y and lame, a
 And if that thou won't me em - ploy One thing I have to ask Will you shel - ter me, till
 6 My fa - ther's dead, my mo - ther's left With her five chil - dren small And what is worse for
 In course of time, he grew a man. The good old far - mer died And left the boy the
 boy there came Up to the far - mer's door May I ask you, if an - y there be That will give me em -
 break of day From this cold win - try blast? At break of day I'll trudge a - way Elsewhere to seek em -
 12 mo - ther still I'm the eld - est of them all Though little I be, I fear not work If thou wilt me em -
 house now has And his daugh - ter for his bride The boy that was, the farm now has He thinks and smiles with
 ploy? To plough and sow, to reap and mow And to be a far - mer's boy.
 ploy To plough and sow, to reap and mow And to be a far - mer's boy.
 ploy To plough and sow, to reap and mow And to be a far - mer's boy.
 joy Of the luck - y day he came that way For to be a far - mer's boy.

Figure 1: Musical score of “The farmer’s boy”⁶⁴

As explored in 4.2., in text-setting there are two groups of constraints at work, namely metrical and grouping constraints. The first metrical constraint, MATCHSTRESS, 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* (4th line), *be* (8th line).
- Second stanza: *if* (1st line, as seen in (181)), *me* (1st line, see (181)), *you* (3rd line), *this* (4th line), *elsewhere* (7th line).
- Third stanza: *her* (2nd line), *what* (3rd line), *the* (4th line).
- Fourth stanza: *the* (7th line), *he* (7th line), *to* (8th line).

There are also mismatches on secondary counts:

- First stanza: *you* (5th line), *be* (5th line), *me* (6th line)
- Second stanza: *thou* (1st line), *me* (3rd line)
- Third stanza: *of* (4th line), *be* (5th line), *me* (6th line)
- Fourth stanza: *for* (4th line), *that* (7th line)

⁶⁴ 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.

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




The sun went down, be - yond yon hills. A -
And if that thou_won't me em- ploy One

The musical notation shows a treble clef with a key signature of one flat and a 4/4 time signature. The melody consists of eight notes: a quarter note, a quarter note, a quarter note, a quarter note, a quarter note, a quarter note, a quarter note, and a quarter note. The lyrics are aligned with the notes, with a duration mismatch between the word 'to' and the note it is set to.

Curiously, mismatches are not always signalled in terms of musical accent, but also in terms of note value. 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 (182), where the function word *to*, realised as /τ↔/ 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 assigned to a quaver. In this sense, we could say that the mismatch of the grammatical word *to* is doubly salient.

(182) Duration mismatch



boy there came Up_ to___ the far-mer's door
break of day From this___ cold win-try blast?

The musical notation shows a treble clef with a key signature of one flat and a 4/4 time signature. The melody consists of eight notes: a quarter note, a quarter note, a quarter note, a quarter note, a quarter note, a quarter note, a quarter note, and a quarter note. The lyrics are aligned with the notes, with a duration mismatch between the word 'to' and the note it is set to.

Duration mismatches such as the one above contradict Hayes and Kaun's (1996:16) Syllable Duration Rule (quoted in section 4.2.2. and repeated here for convenience):

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

The second metrical constraint, FILLSTRONG, 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 since 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 (183), where we can observe that the fourth position of the second line is left empty:

(183) Non-filling of strong positions

	x					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	x
A-	cross	yon	drea-	ry	moor		Ø	

This phenomenon is linked to truncation, which 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 not only that the unfilled strong position is forced to be linked to the previous one(s) – in this song, the end of each musical phrase is marked with a long note value – but also that the very last beat might be left without any corresponding prosodic position in that line – the very last beat of each even bar in this song is filled with a syllable from the next poetic line. This is what Hayes and MacEachern (1996:16-17) call “theft” of positions (184), which, in English, happens only to the left: “it is only in the leftward direction that the positions are there to be stolen” (see section 4.2.3.).

(184) 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	x
{A-	[cross	yon	drea-	ry	moor}		Ø	When]

FILLSTRONG 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.

The third metrical constraint, *LAPSE, states that sequences in which no syllables are 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 since 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.

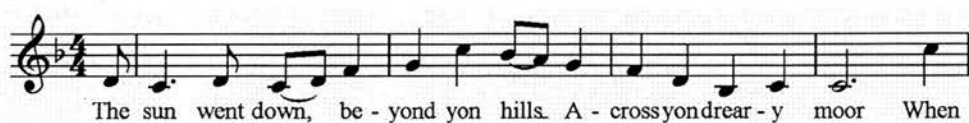
(185) Presence of lapses

	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 and section 4.2.2.). As explained above, 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 first couplet of the first stanza (43):

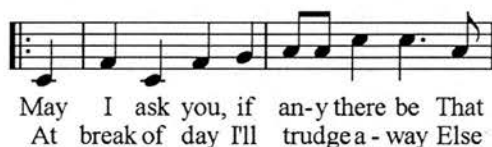
(186) SALIENCY

	x					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	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. At the strophic level, positional parallelism (Halle and Dell in press) 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.

(187) Positional PARALLELISM



	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	x	
A-	cross	yon	drea-	ry	moor		∅	(3)

In general, we observe that 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 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 (188):

(188) 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

In conclusion, 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 takes place 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.

4.3.7.2. Case study II: “Alegría, caballeros”

“Alegría, caballeros” is a strophic song collected in Manzano Alonso’s (2003) *Cancionero popular de Burgos*, more specifically, in the subsection entitled “Cantos de aginaldo en la fiesta de los Reyes”. In this type of festive song, the singers 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 this mechanism in the notated score (Figure 2), where the two stanzas that configure the song are set to exactly the same tune, with no added or eliminated note values.

The musical score is written in 2/4 time with a key signature of one flat (B-flat). The melody consists of eighth and quarter notes. The first line of music is followed by the lyrics: "A-le - grí-a ca-ba - lle - ros, no-ble fies-ta de los Re- yes. Los Re- yes ya son ve - ni - dos, los Re Los pas - to - res, que su - pie-ron queel ni- ¿Qué me quie-res de-cir, ni - ño, con e-". The second line of music starts with a measure rest of 8, followed by the lyrics: "yes ya son ma - ña - na, la pri - mer fies-ta del a - ño que se ce - le-braen Es - pa - ña. A-le ñoes-ta-baen Be - lén, han de - ja - do las o - ve - jas yhan a - pre-ta - doa co-rrer. se de-do pi - na - do? — ¿Me quie - res lle-var a jui- cio?Per-dó - na-me los pe - ca - dos."

Figure 2: Musical score of “Alegria caballeros”⁶⁵

The fact that the melodic and rhythmic contours have to remain untouched has several prosodic consequences – we could also say that the prosody determines the melodic and rhythmic contour, but we will leave the matter of directionality aside for the moment – one of which is the preservation of isosyllabism. In Figure 2, we observe that each note value corresponds to a prosodic syllable – with the exception of the melismata on the seventh syllable of each line –, 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. It is necessary to make clear that ‘beat’ means the same as ‘tactus’, defined as the level at which a listener would clap.⁶⁶ In an instance of 2/4 metre, as in “Alegria, caballeros”, the tactus is located at the crotchet level. On the other hand, a 2/4 bar can comprise various note values, of which the most common ones are two crotchets combined, a series of four quavers,

⁶⁵ Track number 2 in the attached CD comprises two songs. I analyse the first one, which lasts 1:09.

⁶⁶ Although ‘beat’ and ‘tactus’ are nowadays used as synonyms, the latter is better applied to music written before the bar era, while the former is applied to music organised into bars.

eight semiquavers, or any combination of these. In this song, the most common combination is four quavers or three quavers plus two semi-quavers per bar. Some syllables fall on the beat – *-grí-* in *alegría* – but have a smaller value than the beat itself – the beat in 2/4 is the crotchet –, whereas others, while falling on the beat, take a bigger value, either realised on the same musical pitch – *-yes* in *Reyes* – or on two pitches (melisma) – *-lle-* in *caballeros* or *-ña-* in *mañana*.

Line	Syllables	Rhyme	Rhythmic accents
Alegría, caballeros,	8	-	3, 7
noble fiesta de los Reyes.	8	-	1, 3, 7
Los Reyes ya son venidos,	8	-	2, 5, 7
los Reyes ya son mañana,	8	a	2, 5, 7
la <i>primer</i> fiesta del año	8	-	3, 4, 7
que se celebra <u>en</u> España.	8	a	4, 7
Alegría, caballeros,	8	-	3, 7
noble fiesta de los Reyes.	8	-	1, 3, 7
Los pastores, que supieron	8	-	3, 7
<u>que el niño estaba en</u> Belén,	7+1	b	1, 4, 7
han dejado las ovejas	8	-	3, 7
<u>y han apretado a</u> correr.	7+1	b	1, 4, 7
Alegría, caballeros,	8	-	3, 7
noble fiesta de los Reyes.	8	-	1, 3, 7
Qué me quieres decir, niño,	8	-	1, 3, 6, 7
con ese dedo pinado?	8	c	2, 4, 7
Me quieres llevar a juicio?	8	-	2, 5, 7
Perdóname los pecados.	8	c	2, 7

Table 32: Scansion of “Alegría caballeros”.⁶⁷⁶⁸

⁶⁷ In the English case studies, the third column corresponds to the analysis of the foot structure of lines, which implies analysing their rhythmic accents. Spanish poetry is not organised into feet, so for the Spanish case studies, I have only specified the rhythmic accents in each line.

⁶⁸ Translation: ‘Joy, gentlemen/noble party of the Kings./The Kings have come,/the Kings come tomorrow,/the first party in the year/that is celebrated in Spain./Joy, gentlemen.../The shepherd, who knew/that the boy was in Bethlehem,/have left the sheep/ and have hurried up./Joy, gentlemen.../What

As can be observed in Table 32, the eighteen lines that configure the recorded version of the song have eight syllables each. The main compression device used in this song is synalepha, both between two different vowels, in line 6 – *celebra en* –, and between different vowels, in line 10 – *que el*. All the instances of synalepha are kept in the song. This song shows an instance of apocope, a much less common compression device. As explained in section 3.2.3., apocope is the truncation of a word by eliminating its last syllable, as in the case of the metrically-conditioned disyllabic word *primer* in line 5, which would otherwise be a trisyllabic word, *primera*.

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 2/4 metre, naturally falls on the first beat of each bar – primary count – and on the second beat of the bar, corresponding to the second crotchet or the third quaver – secondary count. This causes an effect of total disagreement between verse prosody and music rhythm in the song, with the latter winning over the former. In Table 33 we observe the musically accented syllables (in bold), which often do not coincide with the prosodically accented syllables (in italics).

are you trying to tell me, boy,/with that straight finger?/Do you want to take me to court?/Forgive me for my sins.'

Line	Rhythmic accents
Alegría, caballeros,	3, 7
noble <i>fiesta</i> de los <i>Reyes</i> .	1, 3, 7
Los <i>Reyes</i> ya <i>son</i> <i>venidos</i> ,	2, 5, 7
los <i>Reyes</i> ya <i>son</i> <i>mañana</i> ,	2, 5, 7
la <i>primer</i> <i>fiesta</i> del <i>año</i>	3, 4, 7
que se <u>celebra</u> en <i>España</i> .	4, 7
Los pastores, <u>que</u> <i>supieron</i>	3, 7
<u>que</u> el <i>niño</i> <u>estaba</u> en <i>Belén</i> ,	1, 4, 7
han <i>dejado</i> las <i>ovejas</i>	3, 7
y <u>han</u> <u>apretado</u> a <i>correr</i> .	(1), 4, 7
¿ <i>Qué</i> me <i>quieres</i> <i>decir</i> , <i>niño</i> ,	1, 3, 6, 7
con <i>ese</i> <i>dedo</i> <i>pinado</i> ?	2, 4, 7
Me <i>quieres</i> llevar a <i>juicio</i> ?	2, 5, 7
Perdóname los <i>pecados</i> .	2, 7

Table 33: Musical accent versus prosodic accent.

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.

In order to make the above metrical and grouping observations systematic, it would be possible to translate them into constraints. As seen in the introduction to this chapter, the first group of constraints occurs at the metrical level. In “Alegría, caballeros”, MATCHSTRESS – the constraint that requires that lexical stresses and musical beats are aligned – is often violated. In the grid notation of the first two lines of the song, we can observe that sometimes prominence falls on unstressed syllables – in *caballeros* –, while stressed ones are relegated to a secondary level by the musical arrangement of the sentence – in *noble*.

(189)

		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
A-	le-	<i>grí-</i>	a			<i>ca-</i>	ba-			<i>lle-</i>					<i>ros</i>	
		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
no-	ble-	<i>fies-</i>	ta			<i>de</i>	los			<i>Re-</i>					<i>yes</i>	

Second, FILLSTRONG 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. Example (190) gives an alternative setting for the same lyrics, where the first and fourth strong positions are left empty. As a result, the setting of the line becomes unacceptable.

(190)

		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
*A-	le-		<i>grí-</i>	a			<i>ca-</i>	ba-			<i>lle-</i>				<i>ros</i>	

Third, *LAPSE is violated within stanzas, as there are empty positions between the third and fourth strong positions. Nevertheless, violating this constraint does 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.

(191)

		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
A-	le-	grí-	a			ca-	ba-			lle-	Ø			ros	
		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
no-	ble-	fies-	ta			de	los			Re-	Ø			yes	

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 level to which constraints apply is the grouping level. As mentioned in the introduction to this chapter, SALIENCY is directly related to the concept of rhythmic cadence, which in turn concerns how the last or two last positions in the line are fulfilled. In “Alegria, caballeros”, the refrain and first and third stanzas share their rhythmic cadence structure, with four ‘Green O’ lines. This means, as explained in section 4.2.4., that the four strong positions in each line, including the last one, are filled with syllables, but there are no weak positions filled between the third and fourth strong positions – as seen in (191). In this case, neither couplets nor stanzas are salient in cadential terms. There is another 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.

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 the most important constraint in Spanish.

4.3.7.3. Conclusions

The analyses carried out in this section point to a correspondence between the timing typologies of language and rhythmic typologies of music. As I have shown, both English and Spanish show inconsistencies or mismatches between speech prosody, on the one hand, and verse and music rhythm, on the other. 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 language there is a natural counterpoint or dialogue between speech prosody and musical rhythm, in the second type this counterpoint is considered non-rhythmic 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 I 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 each constraint is allowed to incur.

4.4. Stretching out the theory: Art song

In section 4.3. I analysed two folk songs, one in English and one in Spanish, and explored the similarities and differences between the rules of text-setting in these two languages. As mentioned in the introduction, folk song is the best locus for this kind of analysis, because one can presume that its composition is governed by a definite number of constraints which are native to that specific tradition. In this sense, the potential link between speech prosody, verse prosody and text-setting can reveal which elements of the phonological system of a specific language are structurally essential, and which elements are less central or even completely marginal when it comes to realising a specific utterance or verse instance. There is, nevertheless, a potential problem with assuming this unidirectional relation between speech, verse and music if the analysed corpus is reduced to a single idiom within a genre. Indeed, it could easily be argued that the observed behaviour of metrical and grouping constraints in folk song is completely idiomatic and cannot be generalised as a phonological fact. That is why the theories explored and applied in the previous sections of this chapter need to focus on at least a sample corpus of art song, a subgenre of song which is built on similar principles to those of folk song but differs from the latter in several important aspects. Section 4.4.1. is devoted to explaining the basic characteristics of art song, in such a way that the differences between the former and folk song become apparent. Sections 4.4.2. and 4.4.3. introduce and contextualise the two songs analysed in section 4.4.4.

4.4.1. Definition and characteristics of art song

Art song differs from folk song in that it has a known author and it is always written down in order to be performed in front of an audience, normally with a piano accompaniment.

The relationship between text and music in art song is very similar to that found in folk song. As a major characteristic shared by both subgenres, there is always a tension between the prosody of the verse and the rhythm of the music. The difference with folk song has to do with the awareness of composers, who often try to avoid stress mismatches as much as possible, tending towards a high level of agreement between speech prosody and music. In spite of this conscious effort, that original conflict is never completely solved. As Kramer (1984:160) points out,

The primary fact about song is what might be called a topological distortion of utterance under the rhythmic and harmonic stress of music: a pulling, stretching, and twisting that deforms the current of speech without negating its basic linguistic shape. The art song as a genre is the exploitation of this expressive topology – its shaping both as a primary experience and as a reflection of the contest between musical and poetic meanings.

Summarising, any song, be it a folk song or an art song, is characterised by a tension between the versification laws that govern the formation of the lyrics, and the musical laws that regulate the flowing of the tune. Moreover, those two can conflict with the prosodic rules of the language itself, more specifically with stress assignment rules.

4.4.2. Art song in Britain

Outstanding composers of art song in Britain, namely Ralph Vaughan Williams, Percy Grainger and George Butterworth, among others, were also actively involved in folksong collecting, which became an invaluable resource for their compositions (see Barlow 1997; Stevens 1960). Generally, the folk songs collected by these composers were transcribed and published in the same format as art song, that is, in vocal scores with piano accompaniment, arranged to be performed by trained singers. It is worth noting that their lyrics normally underwent a process of ‘stylisation’ whereby the often erotic lyrics of the original were made subtler or toned down.

Butterworth (b. 1885, d. 1916) used two main elements as the bases of his compositional technique, namely folk music and the poetry of A.E. Housman. Butterworth was one of the key figures in the folk music revival which took place before the First World War, after having met Vaughan Williams and Cecil Sharp. The composer joined the Folk Song Society in 1906, eight years after its foundation (see Barlow 1992:72ff.) and started collecting folk songs all over England. He collected around 300 folk songs, a figure only comparable to the 810 collected by his fellow composer and famous folk song collector Vaughan Williams. In 1911, he became a founder-member of the English Folk Dance Society. Butterworth was a great believer in the power of folk music to endow art-music with a distinctive flavour, to the extent that “[f]olksong influence can readily be discerned in the

melodic lines and accompaniments of [Butterworth's] songs" (Barlow 1997:95). As he once noted, "if ever the opportunity occurs for a truly national production of ballet or opera, the success of the undertaking will rest in the hands of those who have mastered the technique and absorbed the spirit of our English dances and songs" (Banfield, *Grove Music Online*).

Butterworth's most famous vocal pieces were a series of settings of six of the eleven poems that form A.E. Housman's *A Shropshire lad* (1911), arranged for voice and piano. Housman (b. 1859, d. 1936) wrote *A Shropshire lad* in 1896. This collection attracted an enormous following among composers such as Sir Arthur Somervell, Vaughan Williams, John Ireland, Ivor Gurney and C.W. Orr, among others (see Stevens 1960:151-173). According to some researchers (see Barlow 1997:55), the popularity of Housman's poetry among English composers of the last decade of the nineteenth century and the first decades of the twentieth century was not a coincidence. As Quinlan (1959:137-138) observes, "the brevity of the lines, their essential Englishness, their pastoral atmosphere, their rhythm, and their simple spontaneity of feelings were contributing factors". In other words, composers were drawn by the folk song simplicity of Housman's language and metre, together with the fact that his poetry is itself a combination of the styles of the traditional English ballad and classical verse. In Butterworth's case, his settings of Housman's poems have been said to outshine other settings of the same materials for their "simplicity and directness" (Barlow 1997:55). Butterworth's musical setting of *A Shropshire lad* stands out in that it is one of the few collections where the distinction between folk and art song gets blurred, and thus the native language of song composition can be best observed. In Barlow's (1997:56) words, "the influence of folksong undoubtedly helped Butterworth to illustrate the essential spirit of Housman's words, thereby creating a distinctively lyrical form of English song".

4.4.3. Art song in Spain

Following German influences, the genre of art song extended to the whole of Europe, including Mediterranean countries like Spain. Many Spanish composers of the late nineteenth century and early twentieth century, such as Falla, Granados and Mompou, cultivated this genre, which was ideal for the recreation of the various traditional styles of Spanish singing and dancing. Federico Mompou (b. 1893, d.

1987) was essentially a piano miniaturist and songwriter. As happened with Butterworth, the most salient characteristics of Mompou's songs are their surface simplicity and their reliance on distinctive symbols or gestures drawn from the folklore of the different Spanish regions. For instance, he uses modes and figures typical of Andalusian and other regional idioms, but more often his melodic writing is rhythmically and structurally suggestive of Catalan folksong, to the extent that, occasionally, authentic or quasi-authentic Catalan melodies are also used (see Mellers 1987:81). In sum, in Mompou's vocal works, the structural and modal idiosyncrasies of folk music are indeed pervading. As Mellers (1987:81-82) puts it,

He is a composer of refinement, who finds in the folk songs and dances of his race an aspect of his own identity. He 'absorbs' (...) the country that made him, in a way that was possible for him in Catalonia, as it wasn't for Holst and Vaughan Williams in industrial Britain.

The collection *Becquerianas* (1971) sets to music six poems collected in the work *Rimas*, by the Andalusian poet Gustavo Adolfo Bécquer (b. 1836, d. 1870). The titles of these poems are "Hoy la tierra y los cielos me sonríen" (*Today the earth and heavens smile on me*), "Los invisibles átomos del aire" (*The invisible atoms of the air*), "Yo soy ardiente, yo soy morena" (*I am burning, I am brown*), "Yo sé cuál es el objeto" (*I know what the object*), "Volverán las oscuras golondrinas" (*Dark swallows will return*), "Olas gigantes" (*Gigantic waves*). Although the brevity and metrical orthodoxy of Bécquer's poems make them particularly apt for being set to music, their highly romantic character has always been regarded by composers themselves as a practical difficulty for writing a musically worthy setting of those poems. Spanish composers like Albéniz, Falla, Turina, among many others, have all worked with texts by Bécquer, but most of those vocal works have been forgotten for one reason or another. Mompou's collection was written as a commission to celebrate the poet's death centenary, and became the best cycle of settings of Bécquer's poems ever written.

4.4.4. Case studies

In this section, I explore a second sample corpus in the light of the theoretical apparatus applied to the study of folk song in section 4.3. This second corpus comprises two art songs. In section 4.4.4.1. I analyse Butterworth's setting of

Housman's poem "Loveliest of trees". In section 4.4.4.2. I study Mompou's text-setting of Bécquer's poem "Olas gigantes". The goal of these case studies will be either confirming or counter-acting the conclusions drawn in section 4.3.6.3., in such a way that the principles of text-setting established for folk song will be either confirmed as idiosyncratic principles of Spanish and English folk song, or re-defined as general principles of text-setting for any vocal genre in Spanish and English.

4.4.4.1. Case study III: "Loveliest of trees"

"Loveliest of trees" is the first, and probably the most famous one, out of the six songs that form the collection *Six songs from 'A Shropshire lad'*, a cycle completed in 1911. The other five are entitled "When I was one-and-twenty", "Look not in my eyes", "Think no more, lad", "The lads in their hundreds" and "Is my team ploughing?". "Loveliest of trees" became particularly well-known after it was turned into an orchestral rhapsody by Butterworth in 1912.

The poem "Loveliest of trees" is organised into three stanzas of four lines each. Each stanza is delimited by punctuation and syntax rather than by rhyme. As can be observed in Table 34, rhyme works at the couplet level – aa, bb, cc, and so on –, while punctuation and syntactic closure both work at the quatrain level. This disagreement is not significant, as the intelligibility of the structure of the poem is not jeopardised by the rhyme scheme. Within each line, there are four feet, and those feet are predominantly iambic – that is, they follow the pattern unstressed-stressed, which means that there are four stressed syllables per line. This pattern gets challenged in the first foot of lines one and five, where we can observe a substitution of a trochee (stressed-unstressed) – a phenomenon known as 'inversion' (in bold) – and in lines 4, 6 and 10, which do not follow the metrical pattern established for the other lines, as the second foot has a weak position missing, something which renders those lines, in principle at least, unmetrical – in the metrical analysis of the song we will explore the different musical solutions given to these lines. Since all the lines in the poem have four feet, it is logical to think that they will all have the same number of syllables. This is indeed the case, as each line has eight syllables, with the exception of lines 4, 6 and 10, which have only seven syllables each.

Line	S.	R.	Feet (accents)/Type of line
Loveliest of trees, the cherry now	8	a	Inverted iambic tetrameter
Is hung with bloom along the bough,	8	a	Iambic tetrameter
And stands about the woodland ride	8	b	Iambic tetrameter
Wearing white for Eastertide. Ø	7	b	Inversion + missing weak syll.
Now, of my three-score years and ten,	8	c	Inverted iambic tetrameter
Twenty will not come again, Ø	7	c	Inversion + missing weak syll.
And take from <u>seventy</u> springs a score,	8	d	Iambic tetrameter
It only leaves me fifty more.	8	d	Iambic tetrameter
And since to look at things in bloom	8	e	Iambic tetrameter
Fifty springs are little room, Ø	7	e	Inversion + missing weak syll.
About the woodlands I will go	8	f	Iambic tetrameter
To see the cherry hung with snow.	8	f	Iambic tetrameter

Table 34: Scansion of “Loveliest of trees”.

Bearing in mind the high degree of metricality of the poem, we can presume that the composer did not want to upset its pristine structure. Figure 3 gives us a general idea of Butterworth’s setting of the poem into music.

Molto moderato, sempre rubato e con espressione

Love - liest of trees, the cher-ry now Is hung with bloom a -
 8 *poco rit.* *p a tempo* *cresc.*
 long the bough, And stands a -bout the wood-land ride. Wear-ring white for
 15 *f* *p*
 East - er tide. Now, of my three-score
 24 *cresc.*
 years and ten, Twen-ty will not come a-gain, And take from seven-tysprings a score, It on - ly leaves me
 30 *poco rit.* *p a tempo* *cresc.*
 fit - ty more. And since to look at things in bloom Fif - ty springs are lit - tle room,
 37 *f largamente* *dim.*
 A - bout the wood - lands I will go To see the cher - ry hung with snow.

Figure 3: Musical score of “Loveliest of trees” (melody)

Before analysing the metrical arrangements of the musical setting of Housman’s poem, it is important to note that in “Loveliest of trees”, like in any other art song, the metrical analysis goes hand in hand with the stylistic analysis. In this sense, the first observation to be made about this song is that it constitutes a perfect blending between art song and folk song, a fact which can be observed in the metrical and melodic structure of the composition.

The first stanza describes melodic contours typical of art song. The first melodic idea (193) is introduced – first on the piano and then on the voice – with a sequence of *senza misura* bars – bars 1-6 –, that is, the shape of the bars and the points where bar lines fall are relatively meaningless in metrical terms in that they only work as a loose guide for the performers. For this reason, we could say that the notated metre is different from the ‘real’ metre, which is loose, nearly *ad libitum*.

(193)

Musical score for (193) in G major, 3/4 time. The score consists of two systems. The first system has a vocal line and a piano accompaniment. The vocal line starts with a rest, then has notes for 'Love - liest of trees, the'. The piano accompaniment features a delicate, flowing line with a *p* dynamic and *espressivo* marking. The second system continues the vocal line with 'cher-ry now' and the piano accompaniment. Pedal points are indicated with 'Ped.' and asterisks.

In the first musical phrase, the delicacy of the piano accompaniment is remarkable. At several other points throughout the song, the piano is not present at all – bars 4-5 in (193) – and very often its role is reduced to articulating the odd note to keep the overall sense of form. On the other hand, on the few occasions where the piano's line is relatively fulsome and melodic, it never overshadows the vocal part.

The looseness of the first musical phrase (194a) is evident in that the listener's metrical intuition tends towards a 3/2 time signature (194b).

(194a)

Musical score for (194a) in G major, 3/4 time. It shows a single vocal line starting with a rest, then notes for 'Is hung with bloom a - long the bough, -'. The tempo marking is *poco rit.* and there is a fermata over the final note.

(194b)

Musical score for (194b) in G major, 3/2 time. It shows a single vocal line starting with a rest, then notes for 'Is hung with snow a long the bough, -'. The time signature is 3/2.

After the first phrase, a series of crotchets establishes a sense of metre – bars 9-17 in (195) – which is in agreement with the notated time signature and bar divisions.

(195)

8 *poco rit.* *p a tempo* *cresc.*
long the bough, And stands a - bout the wood - land ride Wear - ring white for
15 *f*
East - er tide.

The clear-cut musical rhythm goes on until bar 17, where the accompaniment turns into triplets (196), thus re-establishing the initial metrical freedom. After the triplets on the piano, there is a repeat of the first melodic idea – the main theme of the song – in crotchets and minims. The combination of the triplets with the crotchets and minims gives a grand character to this part.

(196)

cresc. *ff* *meno f*
Ped *

The piano part in (196) also works as a transition into the second stanza, achieved by means of a modulation which is, nevertheless, ambiguous. The melody in the second stanza – bars 22-31 – is in C sharp minor, while the accompaniment remains in the original key of E Major. At this point we can observe an imitation of folk song both at the rhythmic level, where the song becomes totally metrical.

(197)

Musical score for (197) showing a vocal line and piano accompaniment. The key signature is E major (one sharp) and the time signature is common time (C). The score starts at measure 22. The vocal line begins with a piano (*p*) dynamic and the lyrics: "Now, of my three-score years and ten, Twen-ty will not come a-gain,". The piano accompaniment features a steady rise in the melody, marked with a piano (*p*) dynamic and a fortissimo (*pp*) dynamic.

The third stanza shows a steady rise to the melodic and lyric climax – bars 32-37 in (198) –, which takes place on the syllable *wood-* – bar 38. This rise is marked by the crescendo and the ascending intervals in the melody – observe the first or second beat of each bar in (198): B, C sharp, D sharp, E.

(198)

Musical score for (198) showing a vocal line and piano accompaniment. The key signature is E major (one sharp) and the time signature is common time (C). The score starts at measure 30. The vocal line begins with a *poco rit.* marking and a piano (*p*) dynamic, followed by *a tempo* and a *cresc.* marking. The lyrics are: "fil - ty_ more. And since to look at things in bloom Fif - ty springs are lit - le room, A - bout the wood - lands I will go To see the cher - ry hung with snow." The piano accompaniment features a *f largamente* marking and a *dim.* marking.

After the climax, there is a return to the rhapsodic rhythm of the beginning, marked by an art song character – bars 38-42 in (198). Again, bar lines are not meaningful metrically, which means that the accent on the syllable *I* and the unstressed position of *cher-* are not significant in terms of mismatches.

The melodic ending – bars 41-42 in (198) – is inconclusive, as the melody ends on a C sharp, while the piece is in E Major. This is another trait of the folk song style of minor modes, where C sharp indeed might be the tonic. Nevertheless, the piano completes this ambiguous structure by returning to E Major – bars 43-48 in (199) –, thus cancelling the modal character of this part and re-establishing the initial tonic.

(199)

In conclusion, the song could be divided into two main parts. One comprises the first and third stanzas, which are set in a rhapsodic, art song style. In this part, the actual metrical notation is not rigid, that is, the bar lines are not musically or prosodically significant, so the few stresses which fall on weak syllables do not constitute real mismatches. The second part comprises the second stanza, which is composed in a folk song style. It follows the prosody of the words and, at the same time, achieves perfect metricality. These two parts are combined in such a way that the song represents a perfect blend of folk song and art song.

Grouping does not only work regarding the macrostructure of the song, but also has an impact at the microstructural level. The two main stylistic parts can be further subdivided, as the surface structure of the melody shows a series of subtly realised parallelisms which bring it close to the basic principles of folk song composition. If we divide the melodic line into stretches comprising two verse lines each, we can make the following observations. First, there is an instance of parallelism between the first two lines (200a) and the last two lines (200b). In both cases this parallelism leaves out three syllables – the last three syllables of line one (-long the bough) and the first three syllables of line eleven (a-bout the).

(200a)

(200b)

Musical score for (200b) in G major, 4/4 time. The score starts at measure 37. The melody is written on a treble clef staff. The lyrics are: "A-bout the wood - lands I will go To see the cher - ry hung with snow." The tempo markings are *f largamente* and *dim.* There are dynamic markings *f* and *dim.* and a fermata over the final note.

Second, line 3, which is almost bare harmonically, is linked to line 4 to form the second musical phrase, which is loosely parallel to the musical phrase formed by lines 9 and 10. In (201) we can observe that lines 3 and 9 are practically identical melodically. On the other hand, the parallelism between lines 4 and 10 is less clear. Line 10 repeats the note-values and intervals of line 9, and is less conclusive, within its musical phrase, than line 4. Nevertheless, it has the same function as line 4 in that they both establish a sense of closure in their respective phrases.

(201)

Musical score for (201) in G major, 4/4 time. The score starts at measure 8. The melody is written on a treble clef staff. The lyrics are: "long the bough, And stands a -bout the wood-land ride. Wear-ring white for East - er tide. fil - ty more. And since to look at things in bloom Fif - ty springs are lit-tle room." The tempo markings are *poco rit.*, *p a tempo*, and *cresc.* There are dynamic markings *p* and *f*.

Third, lines 5 and 6 are melodically parallel.

(202)

Musical score for (202) in G major, 4/4 time. The score starts at measure 24. The melody is written on a treble clef staff. The lyrics are: "Now, of my three-score years and ten, Twen-ty will not come a-gain." The tempo marking is *p*.

Last, lines 7 and 8 are also loosely parallel. As can be observed in (203), the first and third bars are in terms of the note-values, and so is the second bar. As for the fourth one, given that it closes the musical phrase formed by these four bars, it

introduces new note-values, which nevertheless do not disrupt the feeling of parallelism achieved by the previous three bars.

(203)

come a - gain, And take from seven-ty springs a score, It on - ly leaves me fif - ty more. —

In conclusion, the setting of the poem describes a chiasmic structure, illustrated in Table 35, where the extreme lines mark the beginning and, more importantly, the sense of closure of the song, while the middle lines develop the musical and lyric elements in such a way that the listeners perceive a sense of direction in the composition.

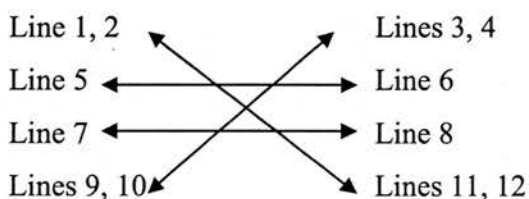


Table 35: Musical structure.

At the metrical level, the song displays a series of interesting mismatches between the prosody of the words and the note values, a phenomenon which is related to the art song character of the first and third parts. Although Butterworth consciously simplified the piano accompaniment in order not to upset the flowing of the poem realised in the melody, at certain points the musical setting seems to intentionally 'disagree' with the verse prosody, mostly in terms of the arrangements of durations. The poem's very first line – bars 4-6 – is a good example of this rearrangement of durations (see (200a) for the musical notation).



All the syllables that are set to long note-values have a long vowel or a diphthong as their nucleus, so they are not examples of what I referred to as ‘Durational mismatches’ in the analysis of English folk song (section 4.3.). This setting seems to get as close to a mismatch as it is possible in English, but without actually mismatching anything.

Summarising, there is a series of beautiful effects achieved by mismatching certain lexical words in terms of duration and pitch. The overall effect is that of a mismatching in the normal enunciation of the phrases, which are contradicted in ways that the listener nevertheless understands. These are not proper mismatches.

The relative durational freedom applied to the setting of Housman’s poem solves the apparent unmetricality of lines 4, 6 and 10, which, as we observed in Table 34, start with inverted foot, while the second foot has a missing weak position. Butterworth’s solution for this metrical problem branches as follows:

- In line 6, *Twenty will not come again*, the durationally stretched syllable is *twen-*, and the same happens in line 10, *Fifty springs are little room*, where the elongated note corresponds to the syllable *fif-*.
- For line 4, *Wearing white for Eastertide*, the music makes it metrical by inserting a long note in *white*, thus compensating for the missing weak syllable.

The two musical realisations of the defective lines are summarised in Table 36:

(x)	/	x	/	(x)	/	x	/
-----	Twen-	ty	will	not	come	a-	gain
-----	Fif-	ty	springs	are	lit-	tle	room
Ø	Wear-	ing	white	-----	-----	for	East-

Table 36: Musical settings of defective lines.

There are a few instances of stress-beat mismatches in this song. A few of them happen on primary counts:

- a. First stanza: *about* (3rd line; bar 10), a mismatch that follows the verse prosody, not the phonological prosody.
- b. Third stanza: *I* (3rd line; bar 39), which has to do with the crossing between real and notated metre.

The vast majority of those mismatches occur on secondary counts, that is, on the third beat of the bar:

- c. First stanza: *loveliest* (1st line; bar 4)
- d. Second stanza: *of* (line 2nd line; bar 22), *will* (3rd line; bar 25)
- e. Third stanza: *since* (1st line; bar 32), *about* (line 3rd line; bar 37)

As can be observed, most of the mismatches found in the text-setting of Housman's poem correspond to grammatical words, and most of them fall on secondary counts. They are minor mismatches, which are acceptable in English text-setting and, therefore, cannot be judged as ill-formed.

In sum, "Loveliest of trees" presents the grouping and metrical characteristics of a prototypical song in English, that is, a rigid alignment between beats and syllables at the metrical level, whereby stress mismatches are practically ruled out, and a relative flexibility at the grouping level, where the different lines do not necessarily follow a pre-established structure. In this case, the song blends the looseness of art song bar division and the exactness of folk song phrasing. Overall, the analysis of this song points to the confirmation of the hypothesis that the principles applied to folk song are not idiosyncratic, but respond to a general set of constraints common to verse and song composition.

4.4.4.2. Case study IV: "Olas gigantes"

"Olas gigantes" is one of the six poems by Bécquer set to music by Mompou under the generic title *Becquerianas*, in 1971. Like most of Bécquer's poetic works, "Olas gigantes" is a highly regular composition in terms of metre and rhyme. The poem is subdivided into four stanzas of four lines each, of which the fourth is much shorter than the other three (heptasyllabic versus hendecasyllabic lines). Synalepha (underlined in Table 37) is realised whenever possible – for the rules on the realisation of synalepha, see 3.2. – so that the number of syllables per line remains

stable. Using Hayes and MacEachern's (1998) taxonomy of lines, we can observe how the fourth line in each quatrain is a type 3 line, while the other three are type 4, which means that the quatrain is salient. The rhyme scheme remains untouched throughout the whole poem, where even lines rhyme, while odd lines do not.

Line	Syllables	Rhy.	Stressed positions
Olas gigantes <u>que os</u> rompéis bramando	11	–	1, 4, 8, 11
en las playas desiertas y remotas,	11	a	3, 6, 10
envuelto <u>entre</u> la sábana <u>de</u> espumas,	11	–	2, 6, 10
¡llevadme con vosotras!	7	a	2, 6
Ráfagas <u>de</u> huracán <u>que</u> arrebatáis	11	–	1, 6, 11
del alto bosque las marchitas hojas,	11	a	2, 4, 8, 10
arrastrado <u>en</u> el ciego torbellino	11	–	3, 6, 10
¡llevadme con vosotras!	7	a	2, 6
Nubes de tempestad que rompe <u>el</u> rayo	11	–	1, 6, 8, 10
y <u>en</u> fuego <u>ornáis</u> las desprendidas orlas,	11	a	2, 8, 10
arrebatado <u>entre</u> la niebla <u>oscura</u> ,	11	–	4, 8, 10
¡llevadme con vosotras!	7	a	2, 6
Llevadme por piedad a donde el vértigo	11	–	2, 6, 10
con la razón <u>me</u> arranque la memoria.	11	a	4, 6, 10
¡Por piedad! ¡Tengo miedo de quedarme	11	–	3, 4, 6, 10
con mi dolor a solas!	7	a	4, 6

Table 37: Scansion of "Olas gigantes".⁶⁹

Stress-placement structures are also consistent throughout the poem. There is a series of structures which gets repeated. Of these, the most salient one is that in the fourth line of each quatrain, which realises only two stresses (on the 2nd or 4th and 6th

⁶⁹ Translation: 'Gigantic waves that break crying/in the desert, remote beaches,/tuck into the foamy sheet./take me away with you!/Hurricane blasts that steal/from the high wood the faded leaves,/carried away in the blind whirlwind/take me away with you!/Rainclouds that the thunder breaks/and in a fire you adorn the unpinned trims,/carried off among the dark fog,/take me away with you!/Take me away, I beg you, where vertigo/with reason tears off my memory,/I beg you! I am afraid of being left/alone with my pain!

positions), while the rest show three or even four stresses. Table 38 shows all the combinations of stressed positions in the poem (numbers signal the positions which are stressed).

	2				6				10	(x3)
	2				6				10	(x2)
		3			6				10	(x2)
	2		4				8		10	(x2)
1			4				8		10	
1					6		8		10	
		3	4		6				10	
1					6				10	
			4		6				10	
			4				8			
			4		6					

Table 38: Rhythmic accents.

Table 38 illustrates a certain degree of parallelism in stressed positions. There is a tendency towards the occurrence of a stress in the middle of the line, in such a way that the overall structure of most lines is five syllables plus stressed syllable (6th) plus five syllables. Apart from this, there is a fixed stress on the last-but-one (10th) syllable of each line, something which is a rule in Spanish poetry, as explained in section 3.2.

If we analyse the position of stresses line by line, we will conclude that all the stresses on the first position happen in the first line of a stanza, while there is a tendency towards second or third position stress in second lines. The same tendency (with one exception, stress on the fourth position) can be observed in third lines, while fourth lines follow the pattern 2, 6 and, in the very last line, 4, 6.

As a song, Mompou's "Olas gigantes" is different from "Loveliest of trees" in that it has an operatic, rather than pastoral, character, meaning that its declamatory style imitates the rhetoric of operas (see Figure 4).

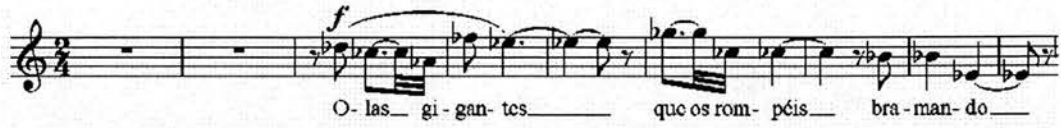
O-las gi-gan-tes que os rom-péis bra-man-do En las
 10 pla-yas de-sier-tas y re-mo-tas, En-vuel-to en-tre las sà-ba-nas dees-pu-mas, Lle-vad-me
 20 con vo-so tras. Rá-fa-gaschu-ra cán-quaca-rro-ba-táis Del ul-to bos-que las mar-chi-tas ho-jas.
 30 A-rras-tra-do cie-go tor-be-lli-no, Lle-vad-me con vo-so - tras.
 40 Nu-bes de tem-pes tad-que rom-peel-ra - yo, Yen fue-goor-náis las des-pren-di-das or -
 50 las, A-rra-ba-ta-do en-tre la nie-blaos-cu-ra, Lle-vad-me con vo-so-tras.
 60 Lle-vad-me por pie-dad Por pie-dad a don-deel
 68 vér-ti-go Con la ra-zón mea-ran-que la me-mo-ria, Por pie-dad ten-go
 77 mie-do de que-dar-me Con mi do-lor a so-las.

Figure 4: Musical score of “Olas gigantes” (melody)

The setting of the poem does not fully conform to that of a lyrical song, but explores the limits between this genre and opera. In spite of this operatic character, which normally corresponds to a high degree of metricality, some of the proportions are not metrical in themselves – observe, in (206), the long note half way through the

first line, on the syllable *-tes*, then on *-péis*, but not at the very end of the line, on the syllable *-do*.

(206)



In most cases, the metricality of the words is used to render the song's rhetorical style visible – in (206), we can observe the setting of the synalepha *que os* to the first beat of bar 6, corresponding to a high pitch (G flat), which makes this syllable doubly salient.

There are many numerous examples of this interaction between the character of the song and the metrical setting of the words. Bars 14-18 (207) are a build-up towards the climax, following from a chromatic rise in the melody and a dense piano part in the accompaniment. From bars 16 to 21 (207), the rhetorical style of the song is evident, with high notes and lengthening of notes altogether. The syllable *-vad-* in *llevadme* is the dramatic climax of the phrase.

(207)



The line *llevadme con vosotras* has a role analogous to that of the refrain in a folk song. Three of the four stanzas in the poem finish off with this line, which demarcates the quatrains – as previously mentioned, this last line is heptasyllabic, whereas the previous three are invariably hendecasyllabic. The setting of this climactic line is not that of a typical refrain, but follows the rhetorical character of the song. In each of the three instances where it appears – bar 18, bar 34, bar 55 (see 208) –, the line in question is set to a different musical phrase. We could say that the composer is imitating how an actor would not deliver the same the second time s/he speaks a specific line. In any case, whatever the composer does, he always matches the three syllables in *vosotras*.

(208)

Bars 18-22:

amplitud

Lle-vad - me con vo - so tras. —

Bars 34-37:

ff

Lle-vad me con vo - so - tras.

Bars 55-57:

p più lento

Lle- vad - me con vo - so- tras.

In terms of grouping constraints, although the internal structure of “Olas gigantes” seems to be more complex than that of “Loveliest of trees”, the former follows the latter’s tendency towards macrostructural parallelism. In this respect, we can observe that the first and last stanzas – bars 3-22 and 63-84 (209) – are musically parallel, and the same happens with the second and third stanzas – bars 23-37 and 40-57. Bars 63-84 are what in musical terms is called a ‘reprise’ of the theme exposed in bars 3-22. This clear instance of parallelism stretches along approximately 20 bars.

(209)

Bars 1-22:

O- las gi- gan- tes que os rom- péis bra- man- do En las

10 pla- yas de sier- tus y re- mo- tas, En- vuel- to en- tre las sí- ba- nas dees - pu- mas, Lle- vad - me

20 con vo - so tras. —

Bars 63-84:

60 LLe- vad - me por pie dad. Por pie-dad a don-deel

63 vér - ti go. Con la ra - zón me - rran-que la me - mo - ria, Por pie- dad ten-go

77 mie-do de que - dar - me. Con mi do- lor a so - las. *rit.* *a tempo*

The setting of “Olas gigantes” describes, as happened with “Loveliest of trees”, a chiasmatic structure:

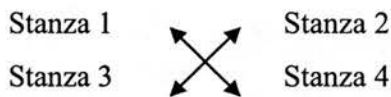


Table 39: Music structure.

The enforcement of parallelism partially explains the mismatches found in these bars, as the composer is using identical metrical structures and laying the syllables under the corresponding notes, almost without rearranging the consequent mismatches. This can be observed in bar 63, where the half-line *llevadme por piedad* is made parallel to bar 3, where the half-line *Olas gigantes* is set (210). The fact that the former has one more syllable than the latter forces the composer to introduce an extra note-value in bar 63. The mismatches found in bar 3 do not appear in line 63, but a new one occurs in the latter.

(210)

Ó-	las	gi-	gán-	tes
Lle-	vád-	me	por	pie- dád

O- las_ gi- gan- tes_____

Lle- vad - me por pie- dad_____

The preposition *por* is mismatched several times, as the line *por piedad* is repeated – something which does not happen in the poem. This indicates that the mismatch is highly consciously done.

The poem shows twelve instances of synalepha (see Table 37). The musical setting of the poem keeps synalepha in eleven cases, making two syllables correspond to a single note value. Such is the case of the synalepha between the relative pronoun *que* and the reflexive pronoun *os* in the first line.

(211)



In bar 52 (212), the composer is forced to divide either of the two prosodic synalephas so that the line ends in two crotchets. The setting of the two contiguous vowels in *arreatado entre* to two different note values breaks the synalepha and turns it into an instance of hiatus – for a definition of this term, go to section 2.2.3.

(212)



Mompou's setting of Bécquer's poem shows a series of stress-beat mismatches, both on primary counts – first beat of each bar – and on secondary counts – in this case, given that the time signature is 2/4, the second beat of each bar is a secondary count. Let us have a look at the mismatches on primary counts first:

- First stanza: *que os* (1st line; bar 6)
- Second stanza: *de hu-* (1st line; bar 24), *las* (2nd line; bar 28), *arrastrado* (3rd line; bar 30)
- Third stanza: *tempestad* (1st line; bar 41), *desprendidas* (2nd line; bar 48), *entre* (3rd line; bars 52-53)
- Fourth stanza: *piedad* (1st line; bar 64), *a* (1st line; bar 67), *de el* (1st line; bar 67), *por* (3rd line; bar 75), *mi* (4th line; bar 80)

There are numerous mismatches on secondary counts:

- First stanza: *olas* (1st line; bar 3), *en* (2nd line; bar 9), *y* (2nd line; bar 12), *entre* (3rd line; bar 15) – *entré* with the stress on the last syllable would be the 1st person singular of the past tense of *entrar* (*to come in*); this is what Janda and Morgan (1988:160, see 4.2.1.) refer to as the levelling of stress-marked distinctions, where the context disambiguates the meaning – *sábanas* (3rd line; bar 16), *con* (4th line; bar 20), *vosotras* (4th line; bar 21)
- Second stanza: *arreatáis* (1st line; bar 25), *hojas* (2nd line; bar 29), *torbellino* (3rd line; bar 32), *con* (4th line; bar 35)
- Third stanza: *y en* (2nd line; bar 46), *arreatado* (3rd line; bar 51), *con* (4th line; bar 56)
- Fourth stanza: *donde* (1st line; bar 67), *vértigo* (1st line; bar 68), *la* (2nd line; bar 69), *la* (2nd line; bar 72), *de* (3rd line; bar 77)

The metrical mismatches in this song have to do with the position of the stressed syllables, which often fall on weak beats, while unstressed ones fall on strong beats. However, mismatches are emphasised by pitch, that is, on many occasions the mismatched syllables can be set to high pitches, which makes the mismatches even more salient. An example of this phenomenon can be found in bar 24 (213), where the metrical mismatch in *de huracán* is set to a high note, while the heavy accompaniment on the first beat of the bar makes it explicitly a mismatch. The same happens in bar 25 – *arreatáis* –, which is parallel to the former.

(213)

Rá-fa-gas dehu-ra-cán

On other occasions, mismatches are realised exclusively by pitch and duration, in such a way that these two contradict what the metrical arrangement

establishes. Such is the case of the last syllable in *arrastrádo* – bar 30 (214) –, which is set to the weakest beat in the bar, and to a melisma at the same time. In terms of metre, the syllable *-do* is not mismatched, but the presence of the melisma on this syllable makes it salient in terms of pitch and duration.⁷⁰

(214)

30 *più espr.*

A - rras - tra-do cic-go tor-be - lli- no,

This observation contradicts Morgan and Janda’s (1989:277) statement that pitch, duration and loudness do not work as cues for stress in music: “ordinarily unstressed syllables which coincide with the musical downbeat are not perceived as being stressed as long as the normally stressed syllable has a musical/rhythmic accent which is at least as pronounced”. As observed in the analysis of mismatches in “*Olas gigantes*”, a syllable that falls on a musical downbeat is indeed perceived as stressed in the unmarked case, while any other extra musical accents are perceived as marked. On the other hand, pitch and duration act as secondary stress cues in text-setting, and can reinforce the effect of the downbeat or contradict it.

In this instance of Spanish art song, there are as many lexical as grammatical mismatches, something which does not happen in the English song analysed in 4.4.1. Given that the setting has a time signature of 2/4, many of these mismatches, particularly those with long words involved, are virtually unavoidable. This is indeed one of the reasons that Janda and Morgan (1988:167-168, in 4.2.1.) give in order to explain the recurrence of lexical mismatches in Spanish song. In their words, “the frequency of polysyllabic words forces the appearance of stress-shift”.

As observed by these authors – for a full review of their papers on the topic, see 4.2.1. – the fact that stress is predictable, and thus recoverable, makes stress-shift

⁷⁰ The vocal score places *-do-* on the second quaver of the bar, as shown in (214). However, in the recorded version, the performer realises a melisma on *-tra-*, the stressed syllable, which occupies three quavers, while *-do-* is made to correspond to the last quaver in the bar and not realised as a melisma. Whether this is a circumstantial decision of the performer or a widely-accepted ‘correction’ made to the original score is beyond my knowledge.

a possible device, even in a language like Spanish, where stress minimal pairs are much more common than in English.

The process of text-setting in Spanish differs from that in English in that, in the latter, the prosody of the words can be slightly altered in order to fit a verse pattern, but the musical setting must preserve that pattern without altering the prosody of the words any further. In Spanish, the prosody of the words can be altered to conform to a specific verse design, but that design is not enforced when the text is set to music. Actually, that intermediate stage – from speech to verse – can be practically neglected, and the prosody of the words can be directly rearranged to fit a specific tune. In Janda and Morgan's (1989:284) words, "[in Spanish, t]ext/tune matching begins with erasure of all lexical stresses and proceeds mechanically by left-to-right, one-to-one association of syllables with notes of the tune".

In conclusion, the analysis of "Olas gigantes" confirms the hypothesis that, in Spanish, music rhythm and metre do not agree with verse prosody or speech prosody, but work independently of both. This is facilitated by the fact that the preservation of word stress is not crucial for the acceptability of a musical setting. That is the reason why metrical constraints are not enforced in Spanish text-setting, while grouping constraints – chiefly, PARALLELISM – are always observed. This is not surprising if we take into account that the primary building principle in Spanish verse is isosyllabism – go to section 3.2.2. Spanish text-setting, both in folk song and art song, proves that the syllable is the most important unit of rhythm, in such a way that the parallel arrangement of syllables constitutes the essence of rhythmic perception in this language. Stress signals the division between the parallel rhythmic units, but where that stress falls is not important for determining the acceptability of a setting.

4.4.4.3. Conclusions

In the introduction I argued that folk song represented the best locus for the analysis of the relation between speech, verse and music at the prosodic level. The reason was that we could presume that folk song composition and performance is governed by a definite number of constraints native to a specific tradition. In section 4.3. I concluded that this was indeed the case, and that metrical and grouping constraints worked differently in Spanish and English due to the nature and prosodic functions of stress in both languages. However, this conclusion might have been slightly

ambitious, as the musical analysis was restricted to a specific subgenre of song. The observed behaviour of metrical and grouping constraints in folk song might be described as idiomatic, in which case it could not have been generalised as a phonological fact. This is the reason why the initial hypothesis needed to be tested on another subgenre of song, in this case, art song. The essential difference between folk song and art song is that the latter has a known author or authors and is composed in order to be performed in front of an audience. We could say that art song composers do not necessarily apply 'native' linguistic constraints when they set a text to music. They often try to avoid stress mismatches as much as possible, tending towards a high level of agreement between text and music.

In the first case study analysed in 4.4.1., we observed that a prototypical art song in English responds to a rigid alignment between beats and syllables at the metrical level, and a relative flexibility at the grouping level, where the different lines do not necessarily follow a pre-established structure. The analysis of "Loveliest of trees" confirmed the hypothesis that the principles applied to folk song are not idiosyncratic but respond to a general set of constraints common to verse and song composition.

In the second case study, the analysis of a Spanish prototypical art song confirmed the hypothesis proved for Spanish folk song, namely that the syllable is the most important unit of rhythm, in such a way that the position of lexical stress is not essential for the perception of that rhythm, while the parallel arrangement of the syllables in different lines and stanzas is.

The characteristic that the two songs have in common is the aesthetic dimension of every single device applied to the interplay between text and melody, something which is not so clear in folk song, given the 'intuitive' character of text-setting in the latter. Although the constraints applied to art song are the same to those at work in folk song, art song consciously plays with the available constraints, stretching out the possibilities available to the language and thus clearly marking the borderline between what is acceptable and what is not for each specific language. Summarising, art song brings to the surface text-setting 'mechanisms' that are unconsciously applied to folk song, which makes the initial hypothesis of a link between speech, verse and song even more powerful.

5. Summary, conclusions and further implications

The last decade has witnessed a steadily growing interest in the study of music from a linguistic perspective. A significant group of linguists – mainly phonologists, phoneticians and syntacticians – have started to turn their attention towards the musical field, which had been practically neglected during the first half of the twentieth century only to be marginally re-incorporated into linguistic investigation in the fifties, and even more so in the seventies and eighties. It is a well-known fact that the relationship between language and music was always regarded as an relevant topic of investigation in the philosophical, literary and even mathematical fields – from Ancient Greece to the Enlightenment –, but once the modern university system established a radical separation between disciplines and the consequent overspecialisation on the part of researchers, such a relationship became unimportant and was thus relegated to a lesser position among linguists and music theorists.

The interdisciplinary nature of the research enterprise presented here is neither capricious nor accidental. My dissertation intended to follow the tradition of the research done in the fields of phonology and text-setting during the last decade of the twentieth century, taking such work a step further by means of adding both linguistic and musical academic depth to it. As stated in the introduction, the main aim of this dissertation was to explain the potential connections between the rhythm of speech, verse and vocal music, as well as to define the limits of the generalisations to be made about the matching of speech rhythm with poetic metre and with musical rhythm. In other words, the study of vocal music, both at the lyric and melodic level, was intended as a linguistic project in the broad sense, which could shed light on a topic that has proven to be central to phonological investigation.

My project encompassed the application of specific linguistic theories, such as OT, to non-linguistic or partially linguistic objects – songs – as well as the analysis of linguistic prosody from a musical point of view. The choice of OT as the theoretical framework for the study of text-setting was not random, as OT has become a major theoretical framework in the study of the relationship between linguistic prosody, verse prosody and musical rhythm. However, most studies on the topic are preoccupied exclusively with English verse and songs, a fact which makes their theoretical claims biased. In order to solve this problem – at least partially –, the

present work incorporated the analysis of materials in a second language – Spanish – to check the validity of the theoretical assumptions made by OT.

I am aware that the scope of my dissertation is very broad, as it covers not only text-setting in English and Spanish, but also speech and verse prosody in both languages. This fact forced me to devise a simple structure for it, where I would present each of those rhythmic manifestations separately, in order to be able to reach a comprehensive view of the topic.

I started off introducing the nuclear concept – rhythm – which pervades the temporal manifestation of speech, verse and music. In order to illustrate the temporal nature shown by these three rhythmic manifestations, Liberman's (1979:313) observation that "[a]ll temporally ordered behaviour is metrically organised" was quoted as a strong statement which would stimulate the discussion from the very beginning. I also explored the deep structural parallelism between language and music, related to the fact that they share a hierarchically organised metrical structure. The articulation of such a structure depends on the nature of the language and the music genre in question, a fact that forces the researcher to narrow down the set of languages and musical genres to be explored. For this reason, I decided to deal with the two languages I know best – Spanish and English – and the vocal music genre where the relationship between melody and text is strongest – song.

Chapter 1 introduced prosody, the discipline that studies the mechanisms with which languages articulate the flow of sound in a rhythmical way, where weaker segments are grouped around a salient one – what in poetry is called 'metre' proper. Depending on the mechanism/s used to assign prominence to certain segments over the others, languages are classified into stress, pitch-accent or tone languages. As stated in this chapter, English and Spanish, like most languages in Europe, are stress languages, since they make use of stress in order to mark syllable prominence.

Chapter 2 narrowed the inventory of languages to be analysed, presenting the two main subgroups into which stress languages have been divided, namely the subgroup of stress-timed languages and the subgroup of syllable-timed languages. As explained in this chapter, in syllable-timed languages syllables are perceived as being rhythmically equal, while in stress-timed languages, stresses fall at perceptually equal intervals. As a consequence, the syllable is considered the major unit of rhythm

in the former group, while in the latter this role is played by the foot. Once these differences were established, the scope of study was narrowed again to two languages, namely English – which has been traditionally classified as the stress-timed language *par excellence* – and Spanish – which has been classified as one of the prototypically syllable-timed languages. In order to understand the current status of stress and the syllable in those two languages, the second half of chapter 2 presented their evolution, starting from Latin and Old English, respectively, and then gave an overview of some of the current theories of stress assignment in both languages.

PdE stress assignment was explored following Halle and Vergnaud (1987), Giegerich (1992) and Gasiorowski (1998). The most important points about PdE stress for the purposes of this dissertation are, first, that English is a quantity-sensitive language – that is, stress is assigned according to syllable weight, which is, in turn, closely related to vowel length and vowel tenseness – and second, English stress is assigned in accordance with two series of rules, namely phonological rules and morphological rules. The interaction between stress, length and weight is, therefore, essential in the organisation of English speech rhythm.

Regarding stress assignment in Present-day Spanish, two major theories, Harris (1983) and D’Introno *et al.* (1995), were explored and contrasted. The former elaborates a specific analytical apparatus on the observation that the characteristics of the Spanish stress system follow from the morphological structures of words, where etymology plays an essential role, and from the contrast marked vs. unmarked stress placement. The most important observation of the system is that nonverbal stress falls on one of the last three syllables of the word. Beyond the word-level, in spite of its predominantly syllable-timed rhythm, Spanish requires that stresses fall at specific points in the utterance, which work as rhythmic ‘leaning-points’. This entails that, at points where unstressed syllables are very numerous, there will be a need to increase the speed of discourse. As a consequence of this, certain processes of ‘compression’ must come into play. The two compression devices applied in Spanish, synaeresis and synalepha, take place at the syllable level, and affect the length of vowels, but not their quality or their tenseness.

Chapter 3 introduced the analysis of systems of versification for Spanish and English. The point of departure for such an analysis was Abercrombie's (1967:98) observation that "[t]he rhythm of everyday speech is the foundation of verse, in most languages". According to this, the speech prosody of a language correlates to its verse prosody.

English verse is defined as accentual-syllabic, stating that it 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. As happened in speech, in English verse feet are perceptually isochronous. The devices which are used in English in order to adjust the duration of feet do not respond to phonological rules. Synalepha, for instance, is highlighted as a standard metrical device which has to do with the abstract metrical representation of the poem, not with the actual realisation of syllables per line. In summary, English verse prosody is found to respond 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.

Spanish verse is defined as syllabic, which can be proved by the fact that much Spanish verse has a perfectly regular syllable count. Apart from being isosyllabic, Spanish verse also takes metrical accents into account. The most important observation in this regard is that, in order to regularise syllable counts, Spanish isosyllabic poetry uses a series of metrical devices, which coincide with the compression devices used in speech – in fact, we observed that the compression devices used in speech are phonological rules that then get transferred into verse.

Chapter 4 constituted the nucleus of this dissertation, as it examined the triple relation between speech, verse and music rhythm as realised in folk and art song. The chapter started off by presenting a definition of song – “a composite which combines two objects each with its own structure, a linguistic object – text – and a musical object – tune” (Dell and Halle in press). The relationship between those two objects is mediated by text-setting, which has to do with how lines of linguistic texts are arranged in time against a predetermined rhythmic pattern in sung verse.

Given that both text and tune take place in the time line and thus have a rhythm, they share the characteristic of being arranged hierarchically into a metrical level – where strong and weak units alternate – and a grouping level – where small

units are arranged into bigger constituents. In OT terms, a series of CON are applied at the metrical level, thus regulating how text metre and tune metre can interact to render a well-formed metrical composite. This interaction will be different in different languages, and will have a direct impact on the ranking of CON at the grouping level, which regulate how text groups and tune groups are arranged to yield a well-formed composite. In turn, the ranking of CON at the grouping level will have direct consequences on the ranking of metrical CON. Among the numerous text-setting CON explored in Hayes and Kaun (1996), Hayes and MacEachern (1996, 1998), Hayes (in press), Halle and Dell (in press), Halle (forthcoming) and Kiparsky (2006), there are three essential ones – MAXBEAT, SALIENCY and PARALLELISM, which can be subdivided as follows:

a) Metrical CON:

a.1) MATCHSTRESS:

a.1.1) strong syllables fall on strong beats

a.1.2) a syllable that falls on an S beat must be stressed

a.2) FILLSTRONG: fill the strongest positions in the line

a.3) *LAPSE: avoid sequences with no syllables between any two of the strongest positions in the line

b) Grouping CON:

b.1) SALIENCY (truncation): non-filling of metrical positions at the end of lines:

b.1.1) lines are salient

b.1.2) couplets are salient

b.1.3) stanzas are salient

b.2) PARALLELISM:

b.2.1) the cadences ending the units of the maximal analysis of a quatrain must be identical

b.2.2) the onsets of syllables coincide with musical beats.

Both the interaction and the conflicts among the above constraints determine and are determined by the type of language that we are dealing with.

The conclusions reached in chapter 4 were based on the metrical analysis of a corpus of 444 Spanish and 239 English folk songs, and on the metrical and grouping

analysis of four case studies, of which two correspond to folk song in English and Spanish, and two correspond to art song in English and Spanish.

The main conclusion derived from these analyses is that the importance of stress-placement as a determining factor in text-setting well-formedness is systematically diminished in syllable-timed languages, while it is enhanced in stress-timed languages. In English, the link between verse prosody and musical metre is stronger than in Spanish. In OT terms, the metrical CON MATCHSTRESS has an enormous impact on the process of text-setting in English. According to this, if the text-setting of an English song is to be qualified as well-formed, the agreement between musical beats and linguistic stresses must be as high as possible. On the other hand, a syllable-timed language like Spanish will ignore MATCHSTRESS in order to concentrate on the one-to-one correspondence between syllables – be it stressed or unstressed – and musical beats – be it strong or weak. In other words, Spanish observes the grouping CON PARALLELISM, which follows from the fact that Spanish verse is predominantly isosyllabic. In English, where the number of syllables per line is secondary to the position of stresses, the observation of PARALLELISM is not required for lines to be judged well-formed.

The analyses carried out in this dissertation point to a correspondence between the timing typologies of language and rhythmic typologies of music. Both English and Spanish show 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 between speech prosody and musical rhythm, in the second type this counterpoint is exceptionally rare. The small mismatches in the English corpus could be regarded as nearly insignificant instances of compositional clumsiness that signal or imitate the folk song style. On the other hand, the continuous radical stress mismatches observed in Spanish folk and art song would make nonsense of an English text because English word stress is vital to the rhythmic arrangement of the language at all levels. 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 I 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. This entails that the rhythmic nature of English and Spanish is essentially different. Whether or not we believe in the old dichotomy between stress- and syllable-timed languages, we must admit that there is a structural difference in the way that these two languages articulate rhythm.

The initial aim of this work of research was providing new insights into the study of linguistic prosody and text-setting. I firmly believe that this aim has been achieved. Moreover, new research paths have been opened for the future.

At the theoretical level, there is still much work needed in order to build a consistent theory of text-setting. Let us mention the potential application of the methods of analysis developed for these data to new folk song and art song corpora, which could be chosen from traditions as diverse as the Celtic, the Japanese, or the sub-Saharan tradition. This would imply dealing with what linguists have called 'mora-timed' languages, such as Japanese, and reaching beyond stress languages in order to explore tone languages.

As a previous step, given that this study has focused on a qualitative analysis of the data, it is still necessary to apply quantitative methods of analysis in order to be able to re-analyse the data collected from a statistical point of view, something which will add to the strength of my conclusions.

Last, but not least, certain analytical flaws in the current state of the text-setting theory must be addressed. These flaws have to do with the lack of analysis of two elements which are essential in the analysis of text-setting in any musical tradition, namely the role of pitch and the placement and realisation of melismata. Up until now, the focus of text-tune alignment studies has been the relationship between syllables and beats from a metrical point of view, in such a way that the melodic element has been systematically neglected in favour of the rhythmic/metric element. On the one hand, it is essential to determine the role of pitch in the perception of prominence, as well as the interaction between stress and pitch in vocal music taken from different traditions – analysing the setting of a tone language such as Mandarin will certainly bring really significant insights into a theory which has exclusively

dealt with stress languages. If we accept that the three major cues for linguistic stress are pitch, loudness and duration, we cannot accept an integral theory of text-setting which focuses exclusively on the interaction between the strength (loudness) of beats and syllables, leaving pitch and duration aside. On the other hand, in practically all text-setting traditions, melisma is a widely used device. The placement of melismata interacts in interesting ways with the assignment of pitch and duration to specific notes/syllables. I believe that, to a certain extent, melisma can be made parallel to the linguistic and poetic phenomenon of synalepha, so that explaining the constraints that regulate the occurrence of melismata could help us understand the nature of synalepha in language.

In short, some of the potential specific steps which could be taken in order to develop new research relate to

- a) the definition of the role of pitch in the realisation and perception of prominence in the corpus analysed in this dissertation
- b) a statistical analysis of rhythmic and pitch mismatches in English and Spanish
- c) the selection of a new corpus: Ugandan folk music, Japanese traditional music, Scottish traditional music
- d) the analysis of the new corpus using OT implemented constraints, focusing on beat-syllable mismatches, pitch mismatches, melismatic mismatches, closure
- e) an in-depth comparison with the results in the old corpus
- f) a statistical classification of the songs in the new corpus
- g) the discussion of the new results
- h) the exploration of linguistic and musical implications.

In sum, this dissertation opens up new possibilities of research within the field of phonology, thus contributing to the development of traditional disciplines such as prosody or metrics. There are also potential applications for it in disciplines stemming from the findings of music theory and theoretical linguistics, but also applied linguistics, sociolinguistics, ethnomusicology and even medicine. One of the most promising currents of practical research to which my research could contribute involves the use of songs in the teaching and acquisition of a second language, as well as the positive impact of rhythm and melody on pupils with learning disabilities.

References

- Abercrombie, David (ed.) (1973[1964]) *In honour of Daniel Jones*. London: Longman.
- Abercrombie, David (1967) *Elements of general phonetics*. Edinburgh: Edinburgh University Press.
- Abercrombie, Lascelles (1923) *Principles of English prosody, Part 1*. London: M. Secker.
- Adams, Stephen (1997) *Poetic design: an introduction to meters, verse form, and tables of speech*. Ontario: Broadview.
- Allen, George D. (1975) "Speech rhythm: its relation to performance universals and articulatory timing". *Journal of Phonetics* 3:75-86.
- Allen, William S. (1973[1964]) "Quantity and quantitative verse". Eds. David Abercrombie and Dennis Butler Fry. *In honour of Daniel Jones*. London: Longman, 3-15.
- Allen, William S. (1973) *Accent and rhythm*. Cambridge: Cambridge University Press.
- Aske, Jon (1990) "Disembodied rules versus patterns (testing the psychological reality of Spanish stress rules)". *Berkeley Linguistics Society* 16:30-45.
- Attridge, Derek (1982) *The rhythms of English poetry*. London: Longman.
- Baehr, Rudolf (1970) *Manual de versificación española*. Madrid: Gredos. Spanish translation by K. Wagner and F. López Estrada.
- Balaguer, Joaquín (1954) *Apuntes para una historia prosódica de la métrica castellana*. Madrid: Consejo Superior de Investigaciones Científicas.
- Balbín, Rafael de (1962) *Sistema de rítmica castellana*. Madrid: Gredos.
- Baltzer, Rebecca, Thomas Cable and James Wimsatt (eds.) (1991) *The union of words and music in medieval poetry*. Austin, Texas: University of Texas Press.
- Banfield, Stephen, "Butterworth, George". Ed. L. Macy. *Grove Music Online* (Accessed 15 May 2006), <http://www.grovemusic.com>
- Bárkányi, Zsuzsanna (2002) "A fresh look at quantity sensitivity in Spanish". *Linguistics* 40 (2):375-394.
- Barlow, Michael (1997) *Whom the Gods love: the life and music of George Butterworth*. London: Toccata Press.

- Barnwell, Thomas P. (1971) "Relation of syllable boundaries to vowel durations". *Journal of the Acoustical Society of America* 50:117.
- Barry, Barbara R. (1990) *Musical time: the sense of order*. New York: Pendragon Press.
- Beare, William (1957) *Latin verse and European song*. London: Methuen & Co.
- Beckman, Mary E. (1986) *Stress and non-stress accent*. Dordrecht: Foris.
- Bécquer, Gustavo A. (1963[1871]) *Rimas*. Madrid: Espasa-Calpe.
- Bell, Alexander Melville (1876) *Visible speech*. London: Simpkin, Marshall & Co.
- Bernstein, Leonard (1976) *The unanswered question*. Cambridge, Mass: Harvard University Press.
- Bertinetto, Pier Marco and Michele Loporcaro (eds.) (1988) *Certamen phonologicum: papers from the 1987 Cortona phonology meeting*. Torino: Rosenberg & Seller.
- Bertinetto, Pier Marco (1989) "Reflections on the dichotomy 'stress- vs. syllable-timing'". *Revue de Phonétique Appliquée* 91/93:99-130.
- Blacking, John (1973) *How musical is man?* Seattle: University of Washington Press.
- Bloomfield, Leonard (1933) *Language*. New York: H. Holt & Company.
- Bolinger, Dwight (1958) *Interrogative structures of American English: the direct question*. Alabama: University of Alabama Press.
- Bolinger, Dwight (1965) *Forms of English: accent, morpheme, order*. Tokyo: Hokuou.
- Bolinger, Dwight (1981) *Two kinds of vowels, two kinds of rhythm*. Bloomington, Ind.: Indiana University Linguistics Club.
- Bond, Zinny S. and Joann Fokes (1985) "Non-native patterns of English syllable timing". *Journal of Phonetics* 13:407-420.
- Boone, Graeme M. (1999) *Patterns in play*. Lincoln, Neb.: University of Nebraska Press.
- Burling, Robbins (1966) "The metrics of children's verse: a cross-linguistic study". *American Anthropologist* 68:1418-1441.
- Burzio, Luigi (1994) *Principles of English stress*. Cambridge: Cambridge University Press.
- Butterworth, George (1911) "Loveliest of trees". *Six songs from 'A Shropshire lad'*. London: Augener Ltd.

- Byshee, Edward (1968[1702]) *The art of English poetry*. Menston: Scholar Press.
- Cable, Thomas (1974) *The meter and melody of Beowulf*, Urbana: University of Illinois Press.
- Cable, Thomas (1991) "The meter and musical implications of Old English poetry". Eds. Rebecca Baltzer, Thomas Cable and James Wimsatt. *The union of words and music in medieval poetry*. Austin, Texas: University of Texas Press, 49-71.
- Cable, Thomas (1996) "Clashing stress in the metres of Old, Middle, and Renaissance English". Eds. Christopher McCully and John Julian Anderson. *English historical metrics*. Cambridge: Cambridge University Press, 7-29.
- Cabré, Teresa and Pilar Prieto (2003) "Vowel contact resolutions across word boundaries in Catalan". Eds. María Josep Solé, Daniel Recasens and Joaquim Romero. *Proceedings of the 15th international conference of phonetic sciences*, Volume II, Barcelona: Universidad Autònoma de Barcelona, 1687-1690.
- Canellada, Josefa and John Kuhlmann Madsen (1987) *Pronunciación del español*. Madrid: Castalia.
- Catford, John C. (1977) *Fundamental problems in phonetics*. Edinburgh: Edinburgh University Press.
- Cernuda, Luis (1996) *Antología poética*. Ed. Philip Silver. Madrid: Alianza Editorial.
- Chomsky, Noam and Morris Halle (1968) *The sound pattern of English*. New York: Harper and Row.
- Chomsky, Noam (1971) *Selected readings*. Eds. J. P. B. Allen and Paul Van Buren. London: Oxford University Press.
- Classe, André (1939) *The rhythm of English prose*. Oxford: Oxford University Press.
- Coleman, C. L. (1974) *A study of acoustical and perceptual attributes of isochrony in spoken English*. Ph.D. Dissertation. Washington D.C.: University of Washington.
- Cooper, Grosvenor and Leonard B. Meyer (1963) *The rhythmic structure of music*. Chicago: University of Chicago Press.
- Couper-Kuhlen, Elizabeth (1986) *An introduction to English prosody*. London: Edward Arnold.
- Couper-Kuhlen, Elizabeth (1993) *English speech rhythm: form and function in everyday verbal interaction*. Amsterdam: Benjamins.

- Crystal, David (1969) *Prosodic systems and intonation in English*. Cambridge: Cambridge University Press.
- Cummins, Fred and Robert Port (1998) "Rhythmic constraints on stress timing in English". *Journal of Phonetics* 26:145-171.
- Cureton, Richard D. (1992) *Rhythmic phrasing in English verse*. London: Longman.
- Cureton, Richard D. (1993) "The auditory imagination and the music of poetry". Ed. Peter Verdonk. *Twentieth century poetry: from text to context*. London: Routledge.
- Dauer, Rebecca (1983) "Stress-timing and syllable-timing reanalyzed". *Journal of Phonetics* 11:51-62.
- Delattre, Pierre (1951) *Principes de phonétique française à l'usage des étudiants anglo-américains*. Middlebury, Vt.: École Française D'Été, Middlebury College.
- Delattre, Pierre (1965) *Comparing the phonetic features of English, French, German and Spanish: an interim report*. Heidelberg: Groot.
- Delattre, Pierre (1966) *Studies in French and comparative phonetics: selected papers in French and English*. The Hague: Mouton.
- Dell, Francois, and John Halle (in press) "Comparing musical text-setting in French and in English songs". Ed. Jean-Louis Aroui and Andy Arleo. *Towards a typology of poetic forms*. Amsterdam: Elsevier.
- D'Introno, Francesco, Enrique del Teso and Rosemary Weston (1995) *Fonética y fonología actual del español*. Madrid: Cátedra.
- Dogil, Grzegorz (1979) *Autosegmental account of phonological emphasis*. Edmonton: Linguistic Research.
- Dowling, W. Jay and Dane L. Harwood (1986) *Music cognition*. Orlando: Academic Press.
- Downing, A. Thomas (1995) *Music and the origins of language*. Cambridge: Cambridge University Press.
- Duffell, Martin (1996) "Chaucer, Gower, and the history of the hendecasyllable". Eds. Christopher McCully and John Julian Anderson. *English historical metrics*. Cambridge: Cambridge University Press, 210-218.
- 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.

- Duffell, Martin (2000) "The Santillana factor: the development of double audition in Castilian". Ed. Alan Deyermond. *Papers of the Medieval Hispanic Research Seminar* (PMHRS) 28:113-28. London: Queen Mary and Westfield College.
- Duffell, Martin (2002) "Mio Cid studies: 'some problems of diplomatic' fifty years on". *Papers of the Medieval Hispanic Research Seminar* (PMHRS) 42. London: Queen Mary and Westfield College.
- Eddington, David (2000) "Spanish stress assignment within the analogical modelling of language". *Language* 76 (1):92-109.
- Eliot, Thomas Stearns (1942) *Music of poetry*. Glasgow: Jackson.
- Fabb, Nigel (1997) *Linguistics and literature*. Oxford: Blackwell.
- 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.
- Finnegan, Ruth (1977) *Oral poetry: its nature, significance and social context*. Bloomington and Indianapolis: Indiana University Press.
- Fitzgerald, Colleen M. (1998) "The metre of Tohono O'Odham songs". *International Journal of American Linguistics* 64(1):1-36.
- Fraisse, Paul (1964) *The psychology of time*. London: Eyre and Spottiswoode.
- García-Bellido, Paloma (1993) "La cantidad fonológica en el español actual". Ed. Ralph Penny. *Actas del primer congreso anglo-hispano*, Volume I. Madrid: Castalia, 151-168.
- Gärding, Eva (1977) *The Scandinavian word accents*. Lund: CWK Gleerup.
- Gasiorowski, Piotr (1998) *Optimal stress for English*. Poznan: Motivex.
- Gasparov, Mikhail Leonovich (1996) *A history of European versification*. Oxford: Clarendon Press. English translation by G.S. Smith and Marina Tarlinskaja.
- Gauldin, Robert (1995) *A practical approach to sixteenth/eighteenth-century counterpoint*. Prospect Heights, Ill.: Waveland Press.
- Giegerich, Heinz J. (1978) "On the rhythmic stressing of function words: a modest proposal". *Work in progress*. Edinburgh: Department of Linguistics, University of Edinburgh, 42-51.

- Giegerich, Heinz J. (1985) *Metrical phonology and phonological structure: German and English*. Cambridge: Cambridge University Press.
- Giegerich, Heinz J. (1992) *English phonology: an introduction*. Cambridge: Cambridge University Press.
- Gili Gaya, Samuel (1950) *Elementos de fonética general*. Madrid: Gredos.
- Glass, Leon and Michael C. Mackey (1988) *From clocks to chaos: the rhythms of life*. Princeton: Princeton University Press.
- Grammont, Maurice (1933) *Traité de phonétique: avec 179 tables dans le texte*. Paris: Delagrave.
- Gray, Catherine T (1998) "Temporal heterogeneities in Kiganda music". Eds. Costin Miereanu and Xavier Hascher. *Les universaux en linguistique: actes du quatrième congrès sûr la signification musicale*. Paris: Sorbonne, 617-627.
- Guthriw, Steven (1991) "Meter and performance in Machaut and Chaucer". Eds. Rebecca Baltzer, Thomas Cable and James Wimsatt. *The union of words and music in medieval poetry*. Austin, Texas: University of Texas Press, 72-100.
- Halle, John, and Fred Lerdahl (1993) "A generative text-setting model". *Current musicology* 55:3-23.
- Halle, John (forthcoming) "Text, tune and metrical form".
- Halle, Morris and Roman Jakobson (1964) *Fundamentals of language*. The Hague: Mouton.
- Halle, Morris and Samuel J. Keyser (1971) *English stress: its form, its growth, and its role in verse*. New York: Harper & Row.
- Halle, Morris and Jean-Roger Vergnaud (1987) *An essay on stress*. Cambridge, Mass.: The Massachusetts Institute of Technology Press.
- Halliday, Michael (1967) *Intonation and grammar in British English*. The Hague: Mouton.
- Halliday, Michael (1973) *Explorations in the functions of language*. London: Edward Arnold.
- Halliday, Michael (1985) *Spoken and written language*. Victoria: Deakin University.
- Halliday, Michael (1994[1985]) *An introduction to functional grammar*. London: Arnold.
- Hamer, Enid (1966[1930]) *The metres of English poetry*. London: Methuen.

- Hanson, Kristin and Paul Kiparsky (1996) "A parametric theory of poetic meter". *Language* 2/2:287-335.
- Harris, James W. (1983) *Syllable structure and stress in Spanish*. Cambridge, Mass.: The Massachusetts Institute of Technology Press.
- Hayes, Bruce (1995) *Metrical stress theory*. Chicago: University of Chicago Press.
- Hayes, Bruce (in press) "Text-setting as constraint conflict". Ed. Jean-Louis Aroui and Andy Arleo. *Towards a typology of poetic forms*. Amsterdam: Elsevier.
- 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.
- Heaney, Seamus (1999) *Beowulf*. London: Faber and Faber.
- Heffner, Rose-Merrill (1949) *General phonetics*. Madison, Wis.: University of Wisconsin Press.
- Hogg, Richard and Christopher B. McCully (1987) *Metrical phonology: a coursebook*. Cambridge: Cambridge University Press.
- Housman, A.E. (1896) *A Shropshire Lad*. London: K. Paul, Trench, Treubner.
- Hualde, José Ignacio and Mónica Prieto (2002) "On the diphthong/hiatus contrast in Spanish: some experimental results". *Linguistics* 40 (2):217-234.
- Huggins, A.W.F. (1972a) "Just noticeable differences for segment duration in natural speech". *Journal of the Acoustical Society of America* 51:1270-1274.
- Huggins, A.W.F. (1972b) "On isochrony and syntax". Eds. Gunnar Fart and Marcel André Tatham. *Auditory analysis and perception of speech*. London: Academic Press, 455-464.
- Hyman, Larry M. (1975) *Phonology, theory and analysis*. New York: Holt, Rinehart, and Wilson.
- Hyman, Larry M. (ed.) (1977) *Studies in stress and accent*. Los Angeles, Ca.: University of Southern California.
- Hyman, Larry (1985) *A theory of phonological weight*. Dordrecht: Foris Publications.

- Jackendoff, Ray and Fred Lerdahl (1983) *A generative theory of tonal music*. Cambridge, Mass.: The Massachusetts Institute of Technology Press.
- Jackendoff, Ray (1989) "A comparison of rhythmic structures in music and language". Eds. Paul Kiparsky and Gilbert Youmans. *Phonetics and phonology, Volume 1, rhythm and meter*. San Diego, Ca.: Academic Press, 15-44.
- Jakobson, Roman (1960) "Closing statement: linguistics and poetics". Ed. Thomas A. Sebeok. *Style in language*. Cambridge, Mass.: The Massachusetts Institute of Technology Press, 350-377.
- Jakobson, Roman (1966) *Poetics*. The Hague; Paris: Mouton Warszawa, Polish Scientific Publishers.
- Janda, Richard D. and Terrell A. Morgan (1988) "El acentó dislocadó – pues cantadó – castellanó: on explaining stress-shift in song-texts from Spanish (and certain other Romance languages)". Eds. David Birdsong and Jean-Pierre Montreuil. *Advances in Romance linguistics*. Dordrecht: Foris, 151-170.
- Johnson, Samuel (1818) *A dictionary of the English language*. London: Longman.
- Jones, Daniel (1950[1909]) *The pronunciation of English*. Cambridge: Cambridge University Press.
- Jones, Daniel (1956[1917]) *English pronouncing dictionary*. New York: E.P. Dutton.
- Kant, Immanuel (1884) *Critique of pure reason*. London: George Bell. English translation by J.M.D. Meiklejohn.
- Kennedy, Peter (1984[1975]) *Folk songs of Britain and Ireland*. London: Oak Publications.
- Kiparsky, Paul (2006) "A modular metrics of folk verse". Eds. Elan B. Dresher and Nila Friedberg. *Formal approaches to poetry*. Berlin, New York: Mouton de Gruyter, 7-52.
- Kramer, Lawrence (1984) *Music and poetry: the nineteenth century and after*. Berkeley: University of California Press.
- Ladd D. Robert and W. N. Campbell (1991) "Theories of prosodic structure: evidence from syllable duration". *Actes du XIIIème congrès international des sciences phonétiques*, Volume 2. Aix-en-Provence: Service des Publications, Université de Provence, 290-293

- 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.
- Ladefoged, Peter (1975) *A course in phonetics*. New York: Harcourt Brace Jovanovich.
- Langer, Suzanne (1953) *Feeling and form: a theory of art developed from philosophy in a new key*. London: Routledge.
- Lass, Roger (1994) *Old English*. Cambridge: Cambridge University Press.
- Laver, John (1994) *Principles of phonetics*. Cambridge: Cambridge University Press.
- Lehiste, Ilse (1970) *Suprasegmentals*. Cambridge, Mass.: The Massachusetts Institute of Technology Press.
- Lehiste, Ilse (1977) "Isochrony reconsidered". *Journal of Phonetics* 52:53-263.
- Lerdahl, Fred, and Ray Jackendoff (1983) *A generative theory of tonal music*. Cambridge, Mass.: Massachusetts Institute of Technology Press.
- Liberman Mark and Alan Prince (1977) "On stress and linguistic rhythm". *Linguistic Inquiry* 8:249-336.
- Liberman, Mark (1979) *The intonational system of English*. New York: Garland Publishing.
- 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.
- Lloyd, A.L. (1967) *Folk song in England*. London: Laurence and Wishart.
- Lloyd, Paul M. (1993) *Del latín al español (I. Fonología y morfología históricas de la lengua española)*. Madrid: Gredos. Spanish translation by Adelino Álvarez Rodríguez.
- London, Justin, "Rhythm". Ed. L. Macy. *Grove Music Online* (Accessed 1 March 2006), <http://www.grovemusic.com>
- Luce, Gay Gaer (1971) *Biological rhythms in human and animal physiology*. New York: Dover Publications.
- MacPherson, Ian (1975) *Spanish phonology: descriptive and historical*. New York: Barnes and Noble.

- Major, Roy C. (1981) "Stress-timing in Brazilian Portuguese". *Journal of Phonetics* 9:343-351.
- Malof, Joseph (1970) *A manual of English meters*. Bloomington, Ind.: Indiana University Press.
- Manzano Alonso, Miguel (ed.) (2001) *Cancionero popular de Burgos: I Rondas y canciones*. Burgos: Diputación Provincial de Burgos.
- Manzano Alonso, Miguel (ed.) (2003) *Cancionero popular de Burgos: V Canciones del ciclo anual y vital*. Burgos: Diputación Provincial de Burgos.
- McCully, Christopher B. (1996) "Domain-end phenomena and metrical templates in OE verse". Eds. Christopher McCully and John Julian Anderson. *English historical metrics*. Cambridge: Cambridge University Press, 42-58.
- Mellers, Wilfrid (1987) *Le jardin retrouvé*. York: The Fairfax Press.
- Miller, M. (1984) "On the perception of rhythm". *Journal of Phonetics* 12:75-83.
- Minkova, Donka (1997) "Constraint ranking in Middle English stress-shifting". *English language and linguistics* 1:135-175.
- Mompou, Federico (1980) "Olas gigantes". *Canciones becquerianas*. Barcelona: Tenora edicions musicals.
- Morgan, Terrell A., and Richard D. Janda (1989) "Musically-conditioned stress shift in Spanish revisited: empirical verification and nonlinear analysis". Eds. Carl Kirschner and Jannet Ann Decesaris. *Studies in Romance linguistics, selected proceedings from the XVII linguistic symposium on Romance languages*. New Jersey: Rutgers University, 273-288.
- Napoli, Donna Jo (1975) "A note on synalepha and stress maxima". *Poetics* 4:401-410.
- Navarro Tomás, T. (1956) *Métrica española*. New York: Syracuse University Press.
- Navarro Tomás, T. (1965[1918]) *Manual de pronunciación española*. Madrid: Gredos.
- Nespor, Marina and Irene Vogel (1986) *Prosodic phonology*. Dordrecht: Foris Publishers.
- Núñez-Cerdeño, Rafael A. and Alfonso Morales-Front (1999) *Fonología generativa contemporánea de la lengua española*. Washington D.C.: Georgetown University Press.

- Olsen, C. L. (1972) "Rhythmical patterns and syllabic features of the Spanish sense-group". *Proceedings of the ninth international congress of phonetic sciences*. The Hague: Mouton, 990-995.
- Penny, Ralph (1991) *A history of the Spanish language*. Cambridge: Cambridge University Press.
- Pensom, Roger (1998) *Accent and metre in French*. Bern: Peter Lang.
- 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.
- Pike, Kenneth L. (1948) *Tone languages: A technique for determining the number and type of pitch contrasts in a language, with studies in tonemic substitution and fusion*. University of Michigan Publications in Linguistics 4. Ann Arbor: University of Michigan Press.
- Plato (1998) *Republic*. Oxford: Oxford University Press. English translation by Robin Waterfield.
- Pointon, Graham E. (1980) "Is Spanish really syllable-timed?". *Journal of Phonetics* 8:293-304.
- Pound, Ezra (1917) "Vers libre and Arnold Dolmetsch". *Egoist* July:90-91.
- Preminger, Alex (ed.) (1974) *Princeton encyclopedia of poetry and poetics*. London: Macmillan.
- Prince, Alan and Paul Smolensky (1993) *Optimality Theory: constraint interaction in Generative Grammar*. New Brunswick, NJ: Rutgers University.
- Quilis, Antonio (1984) *Métrica española*. Madrid: Ariel.
- Quilis, Antonio (1993) *Tratado de fonología y fonética españolas*. Madrid: Gredos.
- Quinlan, John (1959) "A.E. Housman and British composers". *The musical times*. Volume c, 1393:137-138.
- Raven, David Sebastian (1965) *Latin metre, an introduction*. London: Faber and Faber.
- Roach, Peter (1982) "On the distinction between 'stress-timed' and 'syllable-timed' languages". Ed. David Crystal. *Linguistic controversies*. London: Arnold, 73-79.

- Roca, Iggy (1986) "Secondary stress and metrical rhythm". *Phonology Yearbook* 3:341-370.
- Roca, Iggy (1988) "Theoretical implications of Spanish word stress". *Linguistic Inquiry* 19:393-423.
- Roca, Iggy (1997) "On the role of accent in stress systems: Spanish evidence". Eds. Fernando Martínez-Gil and Alfonso Morales-Front. *Issues in the phonology and morphology of the major Iberian languages*. Washington DC: Georgetown University Press, 619-664.
- Roca, Iggy and Wyn Johnson (1999) *A course in phonology*. London: Blackwell.
- Ross, John R. (1972) "A reanalysis of English word stress". Ed. M. Brame. *Contributions to generative phonology*. Austin, Texas: University of Texas Press.
- Russell, Ian, "England (i) II: Traditional music". Ed. L. Macy. *Grove Music Online* (Accessed 16 May 2006), <http://www.grovemusic.com>
- Schlegel, August Wilhem von (1883) *Lectures on dramatic art and literature*. London: G. Bell and Sons. Translated by John Black.
- Schlüter, Julia (2005) *Rhythmic grammar*. Berlin, New York: Mouton de Gruyter.
- Schwitters, Kurt (1993[1926]) *Poems, performance pieces, proses, plays, poetics*. Philadelphia: Temple University Press. English translation and edition by Jerome Rothenberg and Pierre Joris.
- Scruton, Roger (1997) *The aesthetics of music*. Oxford: Clarendon Press.
- Selkirk, Elisabeth O. (1984) *Phonology and syntax: the relation between sound and structure*. Cambridge, Mass.: The Massachusetts Institute of Technology Press.
- Sharp, Cecil (1954[1907]) *English folk song*. London: Methuen and Co.
- Sonnenschein, Edward Adolf (1925) *What is rhythm?* Oxford: Blackwell.
- Sosa, Juan Manuel (1999) *La entonación del español. Su estructura fónica, variabilidad y dialectología*. Madrid: Cátedra.
- Stevens, Denis (1960) *A history of song*. New York: The Norton Library.
- Stoquerus, Gaspar (1988[1570]) *De musica verbali libri duo: two books on verbal music*. Edition, translation and introduction Albert C. Rotola. *Greek and Latin Music Theory* 5. Lincoln: University of Nebraska Press.
- Sweet, Henry (1878) *Handbook of phonetics*. Oxford: Clarendon Press.

- Todorov, Tzvetan (1988) *Literature and its theorists: a personal view of twentieth-century criticism*. London: Routledge and Kegan Paul. English translation by Catherine Porter.
- Tovey, Donald Francis (1979[1949]) "Words and music: some *obiter dicta*". *The main stream of music and other essays*. New York: AMS Press, 202-219.
- Trager, George L. and Henry Lee Smith (eds.) (1951) *Outline of English structure*. Norman, Okla.: Battenburg Press.
- Trubetzkoy, Nikolai S. (1969) *Principles of phonology*. Berkeley: University of California Press.
- van Leeuwen, Theo (1999) *Speech, music, sound*. London: Macmillan.
- Veltman, Joshua (2001) *Textual accent and mensural rhythm in sixteenth-century setting of the 'Pater Noster'*. Master's thesis: The Ohio State University.
- Weil, G. (1960) "Arud". *The encyclopedia of Islam, new edition*, Volume I. Leiden, London: H.A.R. Gibb *et al.* (ed.), 675ff.
- Weinreich, Uriel and André Martinet (eds.) (1954) *Linguistics today*. New York: Linguistic Circle.
- Wenk, Brian J. and François Wioland (1982) "Is French really syllable-timed?". *Journal of Phonetics* 10:177-193.
- Wightman, Colin, Stephanie Shattuck-Hufnagel, Mari Ostendorf and Patti J. Price (1992) "Segmental durations in the vicinity of prosodic phrase boundaries". *Journal of the acoustical society of America* 91:1707-1717.
- Woodrow, Herbert (1951) "Time perception". Ed. Stanley Smith Stevens. *Handbook of experimental psychology*. New York: Wiley, 1224-1236.
- Yavas, Mehmet and Brian Goldstein (1998) "Phonological assessment and treatment of bilingual speakers". *American journal of speech-language pathology* 7:49-60.
- Youmans, Gilbert (1996) "Reconsidering Chaucer's prosody". Eds. Christopher McCully and John Julian Anderson. *English historical metrics*. Cambridge: Cambridge University Press, 185-209.

Appendices

Chapters 1-4:

Appendix I: Glossary of terms and symbols

a) Terms:

Accent: suprasegmental mechanism displayed by languages in order to signal prominence. Accent can be subdivided into three main classes, one of which is stress, the other two being tone and pitch.

Allophone: a phonetic variant of a phoneme in a particular language.

Ambisyllabicity: phenomenon by means of which a consonantal segment working as the onset of an unstressed syllable can also function as the coda of the preceding syllable if this is stressed.

Amphibrachic: a metrical foot consisting of a long/stressed syllable with one short/unstressed syllable either side of it.

Anapaest: foot formed by two unstressed positions and a stressed one, in this order.

Aphaeresis: suppression of the first syllable of a word.

Apocope: suppression of the last syllable of a word.

Bar line: the vertical line placed on the staff to divide the music into bars.

Beat: regular pulsation; a basic unit of length in musical time.

Clausal intonation: rise and fall of the voice pitch.

Coda: component of the syllable rhyme that includes the segments that follow the syllabic nucleus.

Compression: the process and result of reducing the number of syllables in a word, phrase, clause or verse line by joining the nuclei of two or more of those syllables (i.e. the vowels) together.

Coronal: consonant articulated by raising the blade of the tongue.

Couplet: a stanza consisting of two successive lines of verse; usually rhymed.

Dactyl: a metrical foot consisting of one accented syllable followed by two unaccented or of one long syllable followed by two short.

Decompression: the process and result of undoing compression.

Degenerate foot: a foot that lacks one or more positions.

Diacritic: a graphic mark added to a letter to indicate a special phonetic value or distinguish words that are otherwise graphically identical.

Dieresis: the separation or resolution of one syllable into two; the opposite of synaeresis.

Directionality: the serial order in which linguistic phenomena take place.

Diastole: shifting of the position of stress within a word, whereby the stress goes forward to the next syllable in a word.

Dipod: the basic foot of dipodic verse, consisting (when complete) of an unaccented syllable, a lightly accented syllable, an unaccented syllable, and a heavy accented syllable, in that succession.

Diphthong: union of two vowel sounds pronounced in one syllable.

Elision: tendency to elide the unstressed vowel whenever there are two equal vowels next to each other.

Epenthesis: addition of a syllable in the middle of a word.

Euphony: agreeable sound, especially in the phonetic quality of words.

Extra metrical: a syllable which does not count as a metrical position when scanning a verse line.

Foot: the basic unit used in the scansion or measurement of verse.

Fixed stress (languages): languages which place lexical stress always on the same syllables of words.

Fricative: a consonant produced by the forcing of breath through a constricted passage. Also called *spirant*.

Gerund: the term generally is used to refer to the present participle in English and the verbal present participle of Spanish. These are the verb forms that end in '-ing' in English and *-ando* or *-(i)endo* in Spanish. In both languages, the gerund is used to form the progressive or continuous tenses. The English gerund is frequently used as a noun.

Glide: a vowel-like sound that serves as a consonant; semivowel.

Hiatus: two vowels which appear in succession and are pronounced in separate syllables.

Hemistich: a half line of verse, especially when separated rhythmically from the rest of the line by a caesura.

Iambic: a metrical foot consisting of one unaccented syllable followed by one accented syllable.

Ictus: the accent that falls on a stressed syllable in a line of scanned verse.

Isochrony: regular occurrence in time of an event or events.

Leader-timed (languages): languages that create right-headed stress groups where the tonic joins with the preceding unstressed syllables to create the stress group.

Lexeme: a minimal unit (as a word or stem) in the lexicon of a language.

Lexicon: the lexemes of a language considered as a group; a language user's knowledge of words.

Margin: the coda of a syllable, which comprises the consonant sounds that follow the syllabic nucleus, which is usually a vowel.

Middle English (ME): name given to the diverse forms of the English language spoken between the Norman invasion of 1066 and the mid-to-late 15th century.

Medial: in the middle of a word.

Melisma: a passage of several notes sung to one syllable of text.

Metre: any kind of patterning, either consciously used and delimited by rules, as in the case of verse, or unconsciously used and loosely regulated, as in the case of normal speech; the basic pattern contained in a particular piece of verse; the template plus a specified amount of tension.

Mode: any scalewise arrangement of pitches; more generally, the term refers to the patterns upon which medieval music was structured, which preceded the development of major and minor scales and tonality.

Mora: the minimal unit of metrical time in quantitative verse.

Morpheme: the smallest meaningful linguistic unit, which cannot be divided into smaller meaningful parts.

Nominative: the category of nouns serving as the grammatical subject of a verb.

Nucleus: the point of the syllable when the airstream is least obstructed – the sonorant peak, normally a vowel.

Old English (OE): an early form of the English language that was spoken in parts of what is now England and southern Scotland between the mid-fifth century and the mid-twelfth century.

Oxytone: a word that has a heavy stress accent on its last syllable.

Paragoge: addition of a syllable, typically constituted by the vowel /e/, at the end of a word.

Paroxytone: a word with stress on the penultimate syllable.

Pitch accent (languages): those languages that assign prominence to an accented syllable or mora of a word using variations in pitch (frequency).

Present-day English (PdE): the English language as is spoken nowadays.

Phonemic: serving to distinguish phonemes or distinctive features.

Quatrain: a stanza or poem of four lines.

Rhythm: that property of a sequence of events in time which produces on the mind of the observer the impression of proportion between the duration of the several events or groups of events of which the sequence is composed.

Scansion: the process of measuring verse, that is, of marking accented and unaccented syllables, dividing the lines into feet, identifying the metrical pattern, and noting significant variations from that pattern.

Schwa: a mid-central neutral vowel (/ə/), typically occurring in unstressed syllables.

Sliding: the formation of a glide.

Sonorant: a usually voiced speech sound characterized by relatively free air flow through the vocal tract and capable of being syllabic, as a vowel, liquid, or nasal.

Sonority: the degree of resistance of an articulation to free airflow through the vocal tract, where higher resistance equals lower sonority.

Stochastic OT: view of OT where constraints have a range of ranking positions and one can read off the actual frequency of different variants.

Stress: a phonetic peak or climax occurring typically in every word.

Stress-timed: a type of rhythm where stressed syllables tend to occur (at least perceptually) at equal intervals in time.

Stop: a consonant produced with a complete closure in the vocal tract which interrupts the airflow. Stops may be subdivided into oral stops, essentially plosives and affricates, where there is no nasal airflow, and nasal stops, commonly just called nasals, where the soft palate is lowered and the airstream escapes via the nasal cavity.

Suppression: all the phenomena that result in the elimination of a syllable within a word.

Syllable: a phonological structure composed of speech sounds. The syllable is the domain of association for such phenomena as accent, stress and lexical tone. Syllables

are generally considered to be composed of a number of constituents: onset, rhyme, nucleus and coda. Words are made up of syllables.

Syllable-timed: a type of rhythm where syllables tend to have equal duration. French is an example of a language said to be preponderantly syllable-timed.

Syllable weight: the weight of a syllable depends on the make-up of its rhyme constituent. If the rhyme consists of a long vowel (or diphthong) followed by zero or more consonants (a so-called strong cluster), the syllable is heavy. If the rhyme consists of a short vowel followed by a maximum of one consonant (a weak cluster), then the syllable is light.

Synalepha: process that takes place when two or more vowels which belong to different words that come into contact in a rhythm group are combined to form a single syllable.

Synaeresis: process that takes place when two or more vowels which belong to the same word but different syllables are combined to form a single syllable.

Syncope: suppression of a middle syllable.

Syncopate: to shift the accent in music by putting it on a normally unaccented beat.

Systole: shifting of the position of stress within a word, whereby the stress goes back to the previous syllable.

Tactus: term for a beat, i.e. a unit of time measured by a hand movement, used in the 15th and 16th centuries. In theory, each tactus (representing a downward and an upward motion of the hand) stood for a semibreve/whole-note in normal tempo, and was approximately equal to the pulse of a man breathing normally (between 60 and 70 times per minute).

Tempo: the speed at which music is or ought to be played, often indicated on written compositions by a descriptive or metronomic direction to the performer.

Tenseness: vowel quality.

Timbre: the quality of a musical note or sound that distinguishes different types of sound production or musical instruments.

Tone language: a language, as Mandarin or Somali, in which pitch or the pitch contour distinguishes the meanings of words that are otherwise the same phonologically.

Tonic syllable: stressed syllable.

Trochee: a metrical foot consisting of one accented syllable followed by one unaccented syllable.

Truncation: the shortening of a metrical verse line by omitting a syllable or syllables (usually unstressed) from the full complement expected in the regular metrical pattern.

Typology: the study or systematic classification of languages that have characteristics or traits in common.

Trailer-timed (languages): languages that create left-headed stress groups where the tonic joins with the following unstressed syllables to create the stress group

Trillo: a musical trill.

Unmetrical: a line or poem that does not conform to the established foot or syllable patterns.

Vocalic phoneme: vowel.

Voiced: uttered with vibration of the vocal cords.

Vowel length: the perceived duration of a vowel sound.

b) Symbols:

* : incorrect, unacceptable or non-existing form.

- : syllable division.

‘ : concept or term; English translation.

/ : strong position.

\ : weak position.

> : ‘leading to’; ‘more open than’.

< : ‘coming from’; ‘less open than’.

() : translation into English.

[] : phonetic transcription.

+ : ‘plus’.

>> : ‘ranked higher than’.

// : phonemic transcription.

x : metrical position

0 : non-projected syllable.

Bold : mismatched syllables.

Italics : non-English terms.

Underlined : synalephas.

Chapter 2: Speech prosody

Appendix II: Instances of synalepha and synaeresis in a corpus of 444 Spanish folk songs (Manzano Alonso 2003)

1. Synalepha:

1.1. Between unstressed vowels:

1.1.1. Mismatched (falling on strong beats): *man-da-do un, ca-sa es, can-to a, la has, u-na es-tre-lla, ha-bi-ta en, que es-ta-mos, co-mo es, ca-sa el, pa-ra que el, ce-le-bra en, que el, y a, que ha-bi-ta, ca-sa he, me ex-tien-das, duer-ma us-ted, y al, tri-go y, no-che en-tra-rá, li-cen-cia he, se en-fri-a, la han, su-be y, to-do el, he en-tra-do, ten-go a, la en-ho-ra-bue-na, gra-cia en, pa-dri-no y, li-bre y, que a-sí, que a-yer, pi-san-do es-pi-nas, te es-tá, mar-cho a, que es, hier-ba en, si al-gu-no, mi a-bue-la, bo-rran-do a*

1.1.2. Not mismatched: *pa-ra ha, ca-ba-lle-ro hon-ra-do, que el, a-le-grí-a y, ha-bi-ta en, ce-le-bra en, Se-vi-lla a, dí-ga-le us-ted, cor-te-sí-a ha, ban-de-ra en, pi-dién-do-le a, pa' a-pa-gar, no-che en-tra-rá, a-e-no-ja a-e-no-ja, la han, pa-ra a, si es, la en-ho-ra-bue-na, pa-ra em-, y a, por-que al, fres-qui-ta y, do-ce a-pos-to-lo-rum, y el, y un*

1.2. Between an unstressed and a stressed vowel:

1.2.1. Mismatched: *y u-na, de a-ño, o-tro a-ño, puer-ta he-mos, si es*

1.2.2. Not mismatched: *y u-na, que es-ta, pa-sa-rá us-ted, San-ta Á-gue-da, ce-ba' a-je-na, a es-ta, de e-lla, a es-ta*

2. Synaeresis:

Des-a-ho-gue-mos

Spanish example:

		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
queel	ni-	ñoes-	ta-		baen		Be-		lén								Ø

LONG-LAST: 179, 180, 182, 183

English example: Not attested in case studies (see Hayes and MacEachern 1998)

Spanish example:

Llebadme por piedad a donde el vértigo	2, 6, 10	3
con la razón <u>me</u> arranque la memoria.	4, 6, 10	3
¡Por piedad! ¡Tengo miedo de quedarme	3, 4, 6, 10	4
con mi dolor a solas!	(2), 4, 6	2 (3)

MATCHSTRESS/MAXBEAT: 165, 166, 167, 181, 203, 204, 205, 206, 207, 219, 223, 267, 268

English example:

		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
Up		to				the	far-		mer's								door



boy there came Up_ to_ the far-mer's door
break of day From this_ cold win-try blast?

Spanish example:

Ó-	las	gi-	gán-	tes
Lle-	vád-	me	por	pie- dád

O- las_ gi - gan- tes_____

I.le- vad - me por pie- dad_____

PARALLELISM: 180, 182, 192, 194, 198, 202, 203, 204, 205, 206, 207, 226, 227, 228, 234, 235, 246, 247, 248, 253, 256, 257, 261, 267, 268

English example:

	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

Spanish example:

		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
A-	le-	grí-		a		ca-		ba-		lle-				ros
		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
no-	ble-	fies-		ta		de		los		Re-				yes

*RUN-ON: 187

English example: Not attested in case studies (see Hayes in press)

Spanish example: Not attested in case studies

SALIENCY: 164, 176, 177, 179, 180, 184, 202, 203, 204, 205, 226, 227, 234, 267

English example:

	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	x

con ese dedo pinado?	4
Me quieres llevar a juicio?	4
Perdóname los pecados.	4

(UN)FAITHFULNESS: 203, 204, 207, 213

English example:

	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	x	
A-	cross	yon	drea-	ry	moor			(3)

Spanish example:

x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
los	pas-	to-		res	que		su-	pie-											ron
x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
queel	ni-	ñoes-		ta-	baen		Be-	lén											

Appendix IV: Mismatches between stresses and beats in Manzano Alonso
(2003)

Notation:

Song number in the collection (subgenre: dance, 'd', or not dance, 'nd')

Time signature

Type of mismatch (lexical or grammatical): word category; mismatched word
(number of times that the word is mismatched in the same song) syllabified with
mismatched syllable/s in bold and instances of synalepha underlined

1096 (nd)

3/8

lexical: noun *ca-pu-chi-na*

1097 (nd)

3/8

lexical: noun *gen-te* (twice)

1098 (nd)

2/4

lexical: full verb *fue-ron* (twice)

noun *hi-jo*

noun *mon-te*

full verb *man-da-do un*

noun *puer-tas*

grammatical: preposition (prep.) *de*

determiner (det.) *el*

1100 (nd)

3/8

lexical: noun *ca-ri-ta*

grammatical: conjunction (conj.) *que*

prep. *de*

1101 (nd)

3/4

lexical: pronoun (pron.) *de-nos* (second semi-quaver of first beat and second beat of bar)

grammatical: conj. *y*

prep. *de* (twice)

1102 (nd)

3/8

lexical: noun *ca-ri-dad*

1104a (nd)

3/8

grammatical: prep. *con*

prep. *de* (twice)

1105 (nd)

2/4

lexical: adjective (adj.) *a-ma-bles* (three demi-semiquavers of second beat)

full verb *can-ta-mos* (same as above)

grammatical: prep. *con*

det. *lo* (twice)

1106 (nd)

6/8, 9/8

lexical: noun, full verb *ca-sa es* (second beat of 6/8)

relative pronoun (rel. pron.) *que* (second beat of 9/8) (twice)

grammatical: prep. *a*

1108 (nd)

2/4

lexical: full verb *su-pli-car* (second beat)

1110 (nd)

6/8

lexical: noun *ca-sa* (second beat) (twice)

full verb *ro-ba*

1111 (nd)

2/4

lexical: full verb *so-li-ci-tan-do*

grammatical: prep. *pa-ra ha*

1112 (nd)

2/4

lexical: full verb *can-to a*

grammatical: prep. *a*

1113 (nd)

2/4

grammatical: prep. *por* (twice)

conj. *y* (three times)

1114 (nd)

2/4

lexical: auxiliary (aux.) verb, full verb *da-do* (first and second beats)

1115 (nd)

2/4

lexical: noun *tri-pa*

noun *ce-le-min*

grammatical: det. **la has**

prep. *de*

1116a (nd)

3/8, 9/8

lexical: full verb **ha-bi-ta en**

full verb **que es-ta-mos**

grammatical: det. demonstrative (demonstrative) **es-te**

conj. **que es-**

prep. *en*

1116b (nd)

3/4, 4/4

lexical: noun **re-yes**

noun, adj. **ca-ba-lle-ro hon-**

grammatical: pron. **me**

det. *los*

1116c (nd)

3/8

grammatical: conj. **y**

1116d (nd)

3/8

grammatical: conj. **y**

1116e (nd)

3/8

lexical: noun **ca-sa**

adj. **con-ten-tos** (twice)

1117 (nd)

4+3+4/8

lexical: noun *no-che*

noun *a-mas*

full verb *pi-den*

noun *-na es-tre-lla*

adverb (adv.) *cuan-do*

grammatical: det. demonstr. *es-ta*

det. *u-na es-*

conj. *co-mo es*

1118a (nd)

2/4

lexical: noun *re-yes*

noun *es-tre-lla* (twice)

grammatical: conj. + det. *y u-na es-*

1118b (nd)

3+2/4

lexical: noun *ca-sa el*

grammatical: prep. *pa-ra*

det. *el*

1119a (nd)

2/4

lexical: full verb *e-cha-do* (twice)

grammatical: conj. + det. *para que-el* (twice)

det. *los*

1119b (nd)

2/4

lexical: adj. *bue-nos*

noun *ma-ña-na*

grammatical: prep. *de a-ño*

det. *los*

1119c (nd)

2/4

lexical: noun *re-yes*

grammatical: conj. *y*

det. *los* (twice)

conj. + det. *que el*

1119d (nd)

3/8

lexical: noun *se-ño-res*

full verb *sa-len*

grammatical: rel. pro. *que* (twice)

det. *los*

conj. *y u-na*

4+3+3/8

lexical: noun *re-yes*

full verb *lu-ce*

2/4

lexical: full verb *e-cha-do*

1120a (nd)

2/4

grammatical: prep.+ det. *del*

1120d (nd)

2/8, 3/8

grammatical: prep. *de*

1120f (nd)

2/4

lexical: noun *re-yes*

1120g (nd)

2/4

grammatical: prep. *de* (twice)

prep. det. *del*

1120h (nd)

2/4

lexical: noun *li-cen-cia*

noun *pas-to-res*

noun *na-ci-mien-to*

full verb *de-jan-do*

1120i (nd)

3/8

lexical: noun *per-jui-cio* (twice)

noun *Be-lén*

noun *ja-rri-ta*

grammatical: prep. + det. *del* (twice)

prep. *a* (twice)

1120j (nd)

2/8, 3/8

lexical: noun *O-rien-te*

full verb *ba-jan*

noun *u-na es-tre-lla*

grammatical: prep. *por*

det. *u-na*

1120k (nd)

3+5/8

lexical: noun *re-yes* (twice)

full verb *ce-le-bra en*

grammatical: det. *la*

1121 (nd)

4/8, 5/8

lexical: noun *a-le-gri-a y*

grammatical: pron. *nos*

1122 (nd)

2/4, 3/4

lexical: noun *O-rien-te*

noun *Je-ru-sa-lén*

noun *u-na es-tre-lla*

grammatical: conj. + prep. *y a*

det. *u-na*

1123 (nd)

2/4

lexical: noun *re-yes*

noun *per-jui-cio*

grammatical: prep. + det. *del*

1124a (nd)

6/8

lexical: noun *re-yes* (twice)

1124b (nd)

6/8

lexical: full verb *que ha-bi-ta en*

grammatical: prep. *en*

1125a (nd)

3/8

lexical: noun *li-cen-cia*

full verb *que ha-bi-ta en*

full verb *can-tar*

grammatical: rel. pron. *que ha-*

prep. *-ta en*

prep. *a*

conj. *que*

det. *los*

1125b (nd)

3/8

grammatical: prep. + det. *del* (twice)

conj. *y*

prep. *a*

prep. *en*

det. *los*

1125c (nd)

6/8

lexical: noun *re-yes* (three times)

full verb *ce-le-bra en*

grammatical: det. *los*

1126 (nd)

2/4

lexical: noun *ca-sa he*

noun *a-le-gri-a y*

grammatical: conj. *co-mo*

1127 (nd)

4+5/8

grammatical: det. *el*

1128 (nd)

2/4

lexical: full verb *ha-bi-ta en*

full verb *na-ci-do* (second beat)

grammatical: det. *e-sos* (second beat)

prep. *a*

conj. *que es-ta*

prep. *en*

1129 (nd)

2/4

lexical: full verb *ba-ja*

noun *na-ci-mien-to*

grammatical: prep. *a*

1130 (nd)

2/4

lexical: full verb *me ex-tien-das*

grammatical: pron. *me ex-*

1131 (nd)

3/8

lexical: noun *cho-ri-zos*

noun *man-za-nas*

grammatical: conj. *si*

1132a (nd)

3/4

lexical: full verb *pa-sa-rá us-*

grammatical: pron. *lo*

1132b (nd)

3/4

lexical: noun *Á-gue-da*

grammatical: rel. pron. *que*

1133 (nd)

3/4

lexical: noun *vís-pe-ra*

adj., noun *San-ta Á-gue-da* (3rd beat)

full verb *sa-li-do*

grammatical: prep. *de*

1134 (nd)

6/8

lexical: noun *re-gi-dor*

full verb *duer-me*

noun *ni-ña*

full verb, pronoun *duer-ma us-ted*

full verb *des-can-se*

full verb *vie-ne*

noun *cua-dri-lla*

grammatical: prep. *de*

1135 (nd)

4/4

lexical: full verb *te-ne-mos*

1136a (d)

6/8

lexical: adj. *San-to* (second beat)

noun *Fran-cia*

noun *ce-ba' a* (second beat)

noun *ni-ños* (second beat)

grammatical: prep. *de* (three times)

rel. pron. *que* (second beat) (three times)

pron. *nos* (twice)

det. *la* (three times), det. *los*, det. *el*

conj. + prep. + det. *y al*

conj. *y* (second beat)

prep. *para* (second beat), prep. *pa'* (second beat)

1136c

6/8

lexical: noun *Se-vi-lla a*

grammatical: pron. *nos* (second beat) (twice)

det. *el* (second beat)

prep. *de* (second beat)

prep. *a*

1137

6/8

lexical: noun *po-co*

full verb *com-para-mos*

grammatical: pron. *nos* (twice)

det. *el*, det. *lo*

1138

2/4

lexical: noun *se-ño-res*

noun *ca-pi-tán*

grammatical: conj. *si*

1139

7/4

grammatical: det. *nues-tro*

1140

3/8

lexical: noun *se-ño-res*

grammatical: prep. *pa-ra*

1141

5/4

lexical: noun *ge-ne-ral*

noun *e-jér-ci-to*

1142a

2/4

lexical: full verb *can-tar*

noun *dí-a*

grammatical: pron. *no-so-tros*

det. *mu-cha*

1142c

2+3+3+2/4

lexical: noun *li-cen-cia* (second and third beats)

full verb *me-ter-nos* (first and second beats)

grammatical: prep. + det. *del*

prep. *a* (second beat)

pron. *nos* (second beat)

1142d

3/4, 4/4

grammatical: pron. *le*

1143

2/4, 3/4

lexical: noun *ga-lli-to*

full verb *da-do* (twice)

grammatical: prep. + det. *del*

1144a

3/4, 2/4

lexical: full verb *pa-gar*

full verb *sa-li-do a*

full verb *de-be-mos*

grammatical: pron. *di-ga-le us-*

1144b

6/4

grammatical: conj. *aun-que*

pron. *nos*

conj. *pe-ro*

1146

2/4

grammatical: det. *la* (twice)

1147

6/8

lexical: noun *gi-ta-nas*

full verb *so-mos*

adj. *fe-as*

full verb *na-ci-do*

noun *chi-me-ne-a*

grammatical: conj. ***que he-mos***

det. *u-na*

1148a

2/4, 3/4

lexical: full verb *vie-nen*

grammatical: conj. *y*

det. *la*

pron. *nos*

1148b

4/4

lexical: full verb *pe-di-mos*

grammatical: conj. *y*, conj. *mas*

1149

6/8

grammatical: det. *el*

1150

2/4

grammatical: pron. *se*, pron. *le*

1151a

6/8

grammatical: det. *las*

1152b

2/4

lexical: noun *Ca-ta-lu-ña*

noun *cor-te-sí-a ha*

grammatical: prep. *pa-ra* (five times)

det. *to-do*

conj. *si* (six times)

1153

2/4

grammatical: conj. *si* (five times)

1154b

3/4

lexical: noun *ban-de-ra en*

1154c

3/4

lexical: noun *tri-go y*

grammatical: prep. *de*

conj. *y*

1154f

3/4

lexical: noun *ban-de-ra en*

1155a

2/8

lexical: full verb *re-su-ci-tó*

grammatical: prep. *pa-ra*

det. *u-na*

1155b

2/8 (2/4)

lexical: full verb *res-su-ci-tó* (second beat)

1158

2/4

lexical: full verb *re-su-ci-ta*

1159

2/4, 3/4

lexical: noun *Jue-ves*

noun *ca-sa*

grammatical: prep. *de* (three times)

conj. *y*

1160

3/8

lexical: noun *cho-ri-zo y*

noun *es-cue-la*

grammatical: prep. *de*, prep. *a*, prep. *pa-ra*

pron. *nos*, pron. *le*

conj. *y*

det. *o-tro a-ño*

1161

2/4

lexical: noun *Se-ñor*

full verb *ve-ni-mos*

full verb *can-tan-do*

full verb *se-rí-a*

noun *cho-ri-zo y*

full verb *qué-de-se*

noun *lon-ga-ni-za*

grammatical: prep. *de*

prep. *a*

pron. *nos*

pron. *la*
conj. *y*
conj. *que*
det. *o-tro a-ño*

1164

3/8

lexical: full verb *que-re-mos*

1167

2/4

grammatical: prep. *pa-ra*

1168

2/4

lexical: full verb *re-su-ci-tó*

grammatical: prep. *con, de*

1169

2/4

lexical: noun *án-ge-les*

1170

3/4

lexical: full verb *pi-dién-do-le a*

noun *to-do*

noun *ca-be-za*

full verb *tra-e-mos*

adj. *car-ga-do*

grammatical: det. *u-na*

det. *las* (twice)

prep. *pa-ra* (twice), prep. *pa'*

prep. + det. *al*
pron. *le*, pron. *nos*
rel. pro. *que*
conj. *y*, conj. *que*

1171

2/4

lexical: full verb *pa' a-pa-gar*

1172

6/8

grammatical: det. *a es-ta*

det. *u-nos*

pron. *den-nos*

1173a

3/8

lexical: noun *li-cen-cia* (twice)

1173d

3/8

lexical: noun *án-gel*

1173e

3/8

lexical: noun *mar-zo*

1173g

lexical: noun *mar-zo*

1173h

3/8

lexical: noun *mar-zo*

1173j

3/8

lexical: noun *mar-zo*

full verb *flo-re-cen*

1173k

3/8

lexical: noun *mar-zas*

1173l

2/8, 3/8

grammatical: pron. *nos*

1173m

2/8, 3/8

lexical: noun *Ma-ri-a*

1174a

3/8, 2/4

lexical: noun *se-ño-res*

noun *me-dia*

1174b

3/8

lexical: noun *me-dia*

grammatical: det. *tu*

1174c

2/4

lexical: noun *mu-**jer***

1175a

3/8

lexical: noun *li-**cen-cia*** (twice)

noun *me-**dia***

1175b

3/8, 2/4

lexical: noun *li-**cen-cia*** (twice)

full verb *-che en-**tra-rá*** (second beat of 2/4)

1175c

3/8

lexical: noun *se-**ño-res*** (twice)

noun *li-**cen-cia***

grammatical: prep.+ det. *del* (twice)

1175d

3/8

lexical: noun *se-**ño-res*** (twice)

1175e

3/8

lexical: noun *li-**cen-cia*** (twice)

1175f

3/8

lexical: noun *li-**cen-cia***

1175g

3/8

lexical: noun *li-cen-cia* (twice)

1175h

3/8

lexical: noun *li-cen-cia* (twice)

1175i

3/8

lexical: noun *se-ño-res* (twice)

1175j

3/8

lexical: noun *li-cen-cia*

1175k

3/8

lexical: noun *li-cen-cia* (twice)

1175l

3/8

lexical: noun *per-mi-so* (twice)

1175m

3/4

lexical: noun *li-cen-cia* (twice)

1175n

3/4

lexical: noun *li-cen-cia* (twice)

full verb *-che en-tra-rá*

noun *me-dia* (twice)

1175ñ

3/4

lexical: noun *li-cen-cia* (twice)

noun *me-dia* (twice)

1175o

3/8

lexical: noun *li-cen-cia*

noun *ve-ci-no*

noun *me-dia*

noun *e-so*

1175p

3/8

lexical: noun *li-cen-cia*

noun *me-dia* (twice)

grammatical: prep. + det. *del*

conj. *y*

1175q

3/8

lexical: noun *me-dia* (twice)

1175r

3/8

lexical: noun *li-cen-cia* (twice)

1175s

3/8, 2/4

lexical: noun *me-dia* (twice)

grammatical: det. *vues-tras* (twice)

1175t

3/8

lexical: noun *li-cen-cia* (twice)

1175u

3/8

lexical: noun *li-cen-cia* (twice)

grammatical: pron. *nos*

1175v

3/8

lexical: noun *li-cen-cia*

noun *me-dia*

1175x

lexical (noun): *li-cen-cia* *he*

1175z

2/4, 3/8, 2/4

lexical: full verb *se en-fri-a*

grammatical: det. *tu*

pron. *se en-fri-a*

pron. *vá-mo-nos*

1179

3/8, 2/4, 3/8

lexical: noun *se-ño-ras*

noun *co-rri-llo* (twice)

full verb *sa-li-do* (twice)

noun *vi-lla-no* (twice)

full verb *-ja a-e-no-ja* (twice)

full verb *pa-sa*

noun *Za-ra-go-za* (twice)

noun *ví-to-res* (twice)

noun *ma-yo* (twice)

noun *Fer-nan-do* (twice)

grammatical: prep. *pa-ra*

prep. *ha-cia*

1180

3/8

lexical: full verb *di-ces* (twice)

noun *va-cas*

grammatical: prep. *con*, prep. *en*

det. *los*

conj. *por-que* (twice)

1181

6/8

lexical: noun *ma-yo* (twice)

noun *ca-lo-res*

noun *ce-ba-das*

noun *cier-nen*

noun *e-na-mo-ra-dos*

1182

6/8

grammatical: det. *los* (twice)

1185

5/8

lexical: adj. *en-ra-ma-di-ta*

noun *ma-ña-na*

grammatical: det. *u-na*

pron. *la han*

1186

2/4

lexical: noun *mo-zas*

full verb *va-mos*

1187

3/4, 2/4, 2/4, 2/4

lexical: noun *ti-na-ja* (three times)

full verb *su-be y* (three times)

grammatical: det. *la*

prep. *con*, prep. *a* (three times)

prep. + det. *del*

conj. *si*

adv. *ya*

conj. *su-be y* (three times)

1188

4/4, 6/8

grammatical: det. *las*, det. *la* (twice)

prep. *a*

1189

6/8, 4/4

lexical: adj. *San*

adverb *cuán-do*

grammatical: det. *la* (twice)

prep. *a*, prep. *con*

1190

6/8, 2/4, 3/4

lexical: noun *pu-che-ro* (three times)

noun *ri-ve-ro* (three times)

noun *dí-a*

noun *Jua-ni-to*

noun *a-mor* (twice)

grammatical: det. *la*

prep. *con*, prep. *sin*, prep. *a* (twice)

rel. pron. *que* (twice)

conj. *si*

1192

Recitado

lexical: noun *per-mi-so*

noun *se-ño-res*

noun *to-do el*

grammatical: det. *el*

1193

3/4

lexical: adj. *po-bre*

grammatical: prep. *en*

rel. pron. *que*

1194

2/4

lexical: full verb *co-mer*

grammatical: det. *el*

1195, 1196

2/4

lexical: adj. *po-bre-ci-tas*

adv. *có-mo*

full verb *que-dan*

grammatical: det. *las*

conj. *que*, conj. *y*

pron. *les*

1200

6/8

lexical: noun *co-ra-zón*

full verb *he en-tra-do*

1201

6/8

grammatical: prep. *a*

det. *las*

1202a

6/8, 2/4-3/4, 3/8

lexical: full verb *po-ne-mos*

noun *li-cen-cia* (four times)

full verb *ve-ni-mos* (three times)

grammatical: pron. *la*

1202b

2/4-3/4, 3+2+3/8, 3+2+3/8

lexical: full verb *po-ne-mos*

full verb *des-pe-di-rás*

adj. *sol-te-ras* (twice)

full verb *se a-pa-gan*

noun *to-dos*

grammatical: conj. *y* (three times)

prep. *con* (twice)

1202(II)c

2/4, 3/4

lexical: noun *li-cen-cia*

full verb *pi-do* (second beat of 2/4)

grammatical: prep. *pa-ra*, prep. *a*

det. *es-te*

1202(II)d

lexical: noun *li-cen-cia* (twice)

1202(II)e

lexical: full verb *ten-go a*

grammatical: det. *to-do*

det. *mis* (twice)

det. *mi* (second beat of 2/4)

prep. *ten-go a*

1202(II)f

lexical: *quie-res*

1202(II)g

lexical: noun *la en-ho-ra-bue-na*

grammatical: det. *la en-*

conj. *y*

1203

3/4, 5/8

grammatical: det. *tu*

1204

2/4

lexical: noun *tré-bol*

full verb *flo-rez-ca*

noun *gra-cia en*

grammatical: prep. *-cia en-tre*

1205

5/8

lexical: noun *in-ten-ción*

noun *li-cen-cia*

1206

3/4/7/8

lexical: noun *la en-tra-da*

adj. *Bien-ve-ni-dos*

grammatical: det. *la en-*

1207a

5/4

lexical: noun *puer-ta he-mos*

grammatical: prep. *a*, prep. *de*

1207b

6/8

lexical: noun *no-via*

noun *pa-dri-no y*

noun *ma-dri-na*

grammatical: rel. pron. *que*

conj. *pa-dri-no y*

1207c

3/4

lexical: full verb *lle-ga-do*

noun *cua-dri-lla*

full verb *sen-te-mos*

numeral *vein-ti-cin-co* (twice)

noun *si-llas* (twice)

1207d

3/4

lexical: noun *cu-ra*

noun *pa-dres*

noun *l'a-com-pa-ña-mien-to y* (twice)

grammatical: conj. *que* (twice)

1208a

5/4

lexical: noun *hier-bas*

1208b

grammatical: rel. pron. *que*

conj. *co-mo*, conj. *co-mo*

1208c

10/8

lexical: noun *Ma-rí-a*

grammatical: prep. *pa-ra*

prep. *sin*

1208e

2/4

lexical: noun *li-cen-cia a*

grammatical: aux. verb *he-mos*

prep. *con*, prep. *de*, prep. *pa-ra*

1208g

3/4, 2/4

lexical: noun *bal-co-nes*

noun *la-en-ho-ra-bue-na*

grammatical: conj. *que*

det. *la-en-ho-ra-bue-na*

1208h

10/8

lexical: noun *po-sa-da*

adj. *gui-a-dos*

adj. *pu-li-da*

1208l

2/4

lexical: noun *Ma-ri-a*

full verb *di-ce*

noun *puer-ta*

adj. *con-ce-bi-da*

grammatical: prep. + det. *al*

prep. *sin*, prep. *de*

pron. *nos*

1209

3/4

lexical: numeral *vein-ti-cin-co*

grammatical: prep. *a*

conj. *que*

1210

3/4

lexical: noun *ma-ri-do*

grammatical: prep.+ det. *al*

1211

4/4

lexical: full verb *des-pí-de-te*

tonic pron. *de e-lla* (twice)

grammatical: det. *la* (twice)

1212

3/4

lexical: noun *pa-re-des*

adv. *don-de*

1213

3/4

grammatical: det. *la*

conj. *por-que*

conj. *que*

prep. *por*

1214a

3/4

lexical: noun *A-ve*

noun *pe-ca-do*

grammatical: prep. *pa-ra*

1214b

3/4

lexical : noun *ga-lán*

tonic pron. *que e-lla*
grammatical: pron. *la*

1215

3/4, 3/8, 5/8

lexical: noun *cu-cha-ras*

grammatical: conj. *y*

1216a

3/4

lexical: noun *to-mi-llo*

noun *la-do*

1216b

6/4

lexical: noun *tie-rra*

noun *to-mi-llo*

adj. *flo-ri-do*

noun *no-via*

noun *la-do*

1217a

2/4, 3/4

lexical: full verb *ten-gan*

noun *com-pa-ñe-ras*

grammatical: prep. *a* (twice)

det. *mis*

1217b

5/8

lexical: adj. *bue-nas*

noun *com-pa-ñe-ras*

1217c

2/4

lexical: adv. *bien*

1217d

3/4

lexical: noun *li-cen-cia*

noun *no-via*

full verb *des-a-ho-gue-mos*

full verb *-ve a-co-bar-dar*

grammatical: prep. *pa-ra a-*

det. *nues-tras*

particle *no*

1217e

4+3+2+3/8

grammatical: prep. *de*

conj. *que*, conj. *y*

1217f

4+5/8

lexical: adj. *bue-nas* (three times)

1217g

4+4+3/8/4+5+3/8

grammatical: conj. *que*

det. *la*

1217h

6/8

lexical: adv. *bien*

1217i

4+2+3/8

lexical: noun *ra-mi-tos*

grammatical: pron. *se* (twice)

prep. *a*

det. *las*

1217j

3/4, 7/8

lexical: full verb *de-ja*

1217k

3/4, 7/8

grammatical: pron. *se* (twice)

1217l

2/4

lexical: noun *com-pa-ñe-ra*

noun *pa-ñue-lo*

full verb *des-pi-das*

grammatical: pron. *te* (twice)

1217m

2/4, 3/4

lexical: noun *in-ten-ción*

full verb *po-de-mos*

full verb *em-pe-zar*

full verb *si es*

grammatical: aux. verb *he-mos*

prep. *de* (twice)

conj. *si es*

1217n

2/4

lexical: noun *in-ten-ción*

full verb *pe-di-mos*

grammatical: det. *a es-ta*

aux. verb *he-mos*

prep. *de*, prep. *pa-ra*

pron. *les*

1217ñ

4/4

lexical: noun *A-ve* (twice)

grammatical: prep. *pa-ra*

1217o

2/4

lexical: noun *puer-ta*

adj. *li-bre y*

grammatical: det. *el* (twice), det. *la*

rel. pron. *que* (twice)

1217p

6/8 (3/4)

lexical: noun *in-ten-ción*

grammatical: det. dem. *a es-ta*

prep. *pa-ra*

1217q

6/8

lexical: noun *ra-mi-tos*

1217r

4/8, 3/8

lexical: noun *li-cen-cia* (twice)

full verb *em-pe-ce-mos*

adj. *fe-li-ces*

adv. *que a-sí*

grammatical: prep. *con* (twice)

conj. *que a-sí*

1217s

X

lexical: noun *se-ñor*

grammatical: det. *la*

pron. *nos*

1217t

2+3/8

lexical: adverb *más que a-yer*

grammatical: det. *la* (twice)

rel. pron. *don-de*

conj. *más que a-yer*

1217u

3/4

grammatical: prep. *con*, prep. *de* (twice)

conj. *que*

1217v

3/4

lexical: full verb *des-pí-de-te*, *des-pí-de-te*

adj. *pri-me-ra*

adj. *úl-ti-ma*

1218

2/4

lexical: noun *Ma-ri-a* (twice)

full verb *lle-gar*

1220

6/4

lexical: noun *pa-dri-no*

noun *con-vi-da-dos*

1221

3/4, 2/4

lexical: noun *ma-ña-na*

noun *a-mor* (twice)

1222a

6/8

lexical: noun *a-rro-yo*

full verb, noun *pi-san-do es-pi-nas*

1222b

6/8

grammatical: det. *tu*

det. *to-das*

2/4

lexical: full verb *á-bran-se*

grammatical: prep. *de*

1222c

6/8

lexical: full verb *de-ja* (twice)

noun *gen-te* (twice)

full verb *pa-sar*

noun *la-do*

full verb *te es-tá* (twice)

grammatical: pron. *te es-tá* (twice)

pron. *lo*

1222d

2/4

lexical: full verb *ten-gan*

1222e

3/4, 7/8

lexical: full verb *po-ner*

grammatical: prep. *de*

1224a

3/4, 5/8

lexical: full verb *in-co-mo-dar*

grammatical: prep. *con*

prep.+ det. *del*

pron. *nos*

1224b

3/4, 5/8

lexical: full verb *su-bes*

grammatical: prep. + det. *al*

prep. *por*

1225

6/8

grammatical: det. *la*

conj. *que*

1226a

2/4, 4/4

lexical: noun *per-mi-so*

grammatical: aux. verb *he-mos*

prep. *con*, prep. *de*, prep. *pa-ra*

1226c

X

grammatical: conj. *y* (three times)

conj. *cuan-do*

1226d

4+3+3/8

lexical: full verb *ve-ni-mos*

grammatical: conj. *y*

prep. *a*

det. *la en-ho-ra-bue-na es-ta*

1229

3/4

lexical: noun *la en-ho-ra-bue-na* (twice)

grammatical: det. *la en-ho-ra-bue-na* (twice)

prep. *a es-tos dos*

det. *to-dos* (twice)

prep. + det. *al* (twice)

1230

6/8, 3/4

lexical: noun *se-ño-res*

grammatical: prep. *pa-ra em-*

particle *no*

1231

6/8

lexical: adv. *a-quí*

grammatical: pron. *les*

prep. + det. *al*

prep *a*

prep. + conj. *y a*

1232

6/8

grammatical: det. *la* (twice)

1233

3/4

lexical: noun *po-sa-da*

full verb *ve-ni-mos* (twice)

adj. *pu-li-do* (twice)

grammatical: det. *la*

prep. *a* (twice)

1234

6/8

grammatical: prep. *pa-ra*

pron. *que*

1235

6/8

lexical: full verb *su-bir* (twice)

grammatical: det. *es-ta* (twice)

1238

2/4

lexical: adj. *don-de-te*

grammatical: pron. *có-me-le*

det. *la*

det. *Su*

1239

2/4

grammatical: det. *la*, det. *el*

1240

3/4

grammatical: det. *su*

1241a

2/4

lexical: adj. *San*

grammatical: pron. *le*

1241b

2/4

lexical: adj. *San*

grammatical: pron. *le*

1242

2/4

lexical: full verb *mar-cho a*

grammatical: prep. *mar-cho a*

det. *tus*

prep. *En*

1243

X

lexical: adj. *tem-pra-na*

grammatical: pron. *duér-me-te*

1244

2/4

grammatical: prep. + det. *al* (three times)

prep. *por*

1245

3/8

grammatical: det. *es-te*

1246

3/4

grammatical: conj. *que*

1247

2/8, 3/8

lexical: full verb *duer-me*

grammatical: conj. + prep. *co-mo en*

conj. *por-que al*

1248

2/4

grammatical: det. *el*

1249

2/4

grammatical: prep. *sin*

1252

2/8

lexical: full verb *lla-man-do* (twice)

grammatical: conj. *que* (twice)

conj. *a*

det. *los* (twice)

pron. *duér-me-te* (twice)

1253

2/4

lexical: adj. *dor-mi-do*

grammatical: pron. *duér-me-te*

1254

2/4

grammatical: pron. *duér-me-te*

1256

3/4

grammatical: pron. *duér-me-te*

det. *la* (twice)

1257a

2/8

grammatical: prep. *de*, prep. *con* (twice)

1257b

2/8

lexical: adj. *fres-qui-ta y*

grammatical: prep. *de*, prep. *con*

conj. *que*

pron. *dé-ja-la*

1259

3/4

grammatical: det. *mi*
det. *las*

1261a

3/8

lexical: noun *pa-ja en*

grammatical: conj. + det. *si el*
prep. *de*
prep. *pa-ja en*

1261b

3/8

lexical: full verb *vie-nen* (three times)

grammatical: det. *los* (twice)

1262a

2/4

grammatical: pron. *me*
prep. *a*

1262b

2/4

grammatical: pron. *me*
prep. *a*

1263

2/8

lexical: full verb *que es*

grammatical: det. *la*
pron. *me*
rel. pron. *que es*

1267

3/4

lexical: noun *a-ve-lla-no*

noun *bue-yes*

noun *la-bra-do-res*

noun *hier-ba en*

grammatical: det. *mi*

prep. *con*, prep. *de*, prep. *hier-ba en*

det. *los*

1268

2/4

grammatical: prep. *con*, prep. *de*

1269

2/4

lexical: adj. *di-chas* (twice)

noun *Je-ru-sa-lén* (twice)

grammatical : det. *los* (twice), det. *las*

1270

2/4

lexical: numeral, noun *do-ce a-pos-to-lo-rum*

full verb *man-da*

grammatical: det. *la*

1271

2/4, 3/4

grammatical: det. *la*

1272

2/4

grammatical: det. *las* (three times)

1273a

2/4, 3/4

lexical: numeral *la u-na*

numeral *on-ce mil*

grammatical: conj. *y* (three times)

conj. *que* (four times)

det. *la u-na*

1273b

2/4, 3/4

lexical: numeral *on-ce mil*

grammatical: conj. *y* (four times)

det. *las*

1274a

2/4, 3/4

lexical: noun *vir-gen* (twice)

1274b

2/4, 3/4

grammatical: det. *las* (three times)

1274c

3/8, 2/8

grammatical: rel. pron. *que* (twice)

1275

2/4

grammatical: prep. *de* (twice)

1276a

3/8

lexical: full verb *es-ta-ban*

1277a

2/4

lexical: noun *rue-da*

1277b

lexical: noun *rue-da*

grammatical: det. *la*

1278

2/4

lexical: full verb *te-ner* (twice)

grammatical: conj. *y* (twice)

1279

2/4

lexical: full verb *ca-sar*

1281

2/4

grammatical: pron. *me* (three times)

1282a

2/4

grammatical: prep. *en*

1282b

grammatical: prep. *en*

1283

2/4

lexical: noun *miér-co-les*

noun *jue-ves*

full verb *ha-cer*

noun *sá-ba-do*

noun *do-min-go*

full verb *pue-de*

noun *ma-ña-ni-ta*

grammatical: det. *el* (five times)

prep. *pa-ra* (twice)

1284

2/4

grammatical: conj. *si* (twice)

1285

3/4

lexical: full verb *vi-nien-do*

adj. *so-li-ta*

noun *pa-ti-tas*

full verb *se-rá u-na*

full verb *ve-rás* (four times)

tonic pron. *si al-gu-no*

adv. *don-de*

grammatical: pron. *se*, pron. *me*

conj. + det. *y el*, conj. *y un*

conj. *si al-gu-no*

conj. + det. *o al-gún*

prep. *con*, prep. *a*

det. *la*, det. *los*

1286

2/4

grammatical: det. *el*

1287

3/8

grammatical: conj. *por-que*

pron. *me*

1292

2/4

lexical: noun *pa-tro-nes*

adj. *lo-cos*

noun *lu-gar*

grammatical: det. *la*

prep. *de* (twice)

prep. *con*

det. *el*

conj. + det. *y el*

1293

2/4

grammatical: prep. *en* (twice)

rel. pron. *que*

1294

3/4

lexical: noun *pro-vin-cia*

noun *par-ti-do*

1295

2/4

lexical: noun *fron-te-ra*
grammatical: det. *la*

1296

2/4

grammatical: pron. *se*
pron. *las*

1297

2/4

lexical: noun *huer-to* (twice)
noun *na-bos*

1298

2/4

lexical: noun *ca-so*
noun *mi a-bue-la*
grammatical: pron. *les*
prep. *a*
det. *mi a-bue-la*

1299

2/4, 3/4

lexical: noun *des-gra-cias*
full verb *e-cha-rás*

1300

2/4

grammatical: conj. *si*
pron. *lo*

1302

2/4, 3/4

lexical: noun *da-mas* (twice)

1304

2/4

grammatical: prep. *con*

1305a

2/4

lexical: full verb *bo-rran-do a* (twice)

grammatical: det. *la*

prep. *con*

1305b

2/4

lexical: full verb *es-tar*

grammatical: prep. *en*

1308

2/4

lexical: noun *vi-no* (three times)

noun *Na-va-rra* (three times)

noun *to-dos*

full verb *a-gra-da*

full verb *a-lar-ga*

1309

2/4

lexical: noun *can-ti-nas*

noun *Chi-na*

adj. *me-jor*

noun *to-da*

grammatical: prep. + det. *del*

1310

2/4

grammatical: det. *la*

conj. *que* (twice)

prep. *a*

1311

2/4

lexical: full verb *ti-ra-mos*

grammatical: det. *el*

1313

6/4

lexical: adj. *pre-pa-ra-da*

noun *sar-tén*

noun *ja-rra*

1314

lexical: full verb *va-ya*

grammatical: conj. *mien-tras* (twice)

Appendix V: Mismatches between stresses and beats in Kennedy (1984)

Notation:

Song number in the collection

Time signature

Word class, word instance (in italics) with mismatched syllable (in bold), beat in the bar where the mismatched syllable falls (times the mismatch occurs in the same song).

Explanation for lexical mismatches.

124

3/4

preposition *un-to* on the last beat of a bar and the first beat of the next bar.

126

4/4

preposition *in* on a first beat.

129

6/8

conjunction *and* on a second beat.

130

3/4

determiner *my* on a first beat.

133

3/2

lexical mismatch (full verb): *be-came*, third beat on crotchet and crotchet.

Reasons: transcription forces it (a single metrical arrangement is repeated throughout the whole piece); singer's own memory (maybe it was not that word; not really important, as it falls on the weakest part of the bar).

134

4/4

conjunction *and* on a third beat.

137

3/4

lexical mismatch (noun): *mor-ning*, third beat of a bar and first two of another (last).

Characteristic of English folk song; it would not occur in art song.

139

3/4

preposition *a* on a first beat.

141

3/4, 4/4

preposition *to* on a first beat.

142

6/8

preposition *at* on a first beat.

preposition *of* on a first beat.

154

4/4

preposition *in* on a first beat.

156

4/4

determiner (article) *the*, determiner *a* on a third beat.

161

3/4

preposition *of* on a first beat.
auxiliary verb *was* on a first beat.

162

3/4, 4/4

preposition *o-ver* on a first beat (twice).
determiner *my* on a first beat.

163

6/8

determiner *my* on a first beat.

168

3/4

preposition *with* on a first beat.
determiner *his* on a first beat.

170

6/8

preposition *of* on a first beat.

172

6/8

auxiliary verb *was* on a first beat.
pronoun *I* on a first beat.

175

3/4

preposition *on* on a first beat (twice).

176

6/8

preposition *on* on a first beat.

177

4/4

preposition *but* on a third beat (twice).

preposition *in* on a third beat.

178

3/4

preposition *in-to* on first and second beats.

181

3/4

preposition *to* on a first beat.

184

4/4

preposition *on* on a first beat.

186

4/4

preposition *of* on a first beat.

determiner *my* on a first beat.

pronoun *she* on a first beat.

188

6/8

lexical mismatch (noun): *Coun-te-rie* at the end of a line.

preposition *of* on a first beat.

preposition *in* on a first beat.

205

6/8

preposition *on* on a first beat.

209

6/8

preposition *in* on a first beat.

pronoun *I* on a first beat.

211

6/8, 9/8

conjunction *if* on a first beat; conjunction *when* on a first beat.

preposition *up* on a first beat.

pronoun *I* on a first beat.

212

3/4

preposition *by* on a first beat.

214

2/4

auxiliary verb *is* on a first beat.

215

4/4, 6/4

pronoun *I*, pronoun *her* on first and third beats.

217

4/4

conjunction *when* on a first beat.

218

6/8

preposition *to* on a first beat (three times).

219

4/4

conjunction *if* on a first beat.

particle *not* on a first beat.

225

4/4

lexical mismatch (noun): *shoe-ma-kers* on a third beat (demisemiquaver) and a fourth beat (dotted semiquaver).

It is fine because it is a compound, so the main stress falls on *shoe*.

226

4/4

preposition *on* on a first beat.

pronoun *I* on a first beat (twice).

228

6/8

preposition *at* on a first beat.

pronoun *I* on a first beat (end of line and pause).

modal *can* on a first beat.

229

6/8

conjunction *if* on a first beat.

230

4/4

pronoun *I* on a first beat.

238

4/4

conjunction *when* on a first beat.
conjunction *if* on a third beat.
auxiliary verb + particle *don't* on a first beat.

242

4/4

preposition *for* on a first beat.
pronoun *I* on a first beat.

248

4/4

preposition *to* on a first beat.
pronoun *I* on a first beat.

253

6/8

preposition *on* on a first beat.

255

4/4

preposition *in* on a first beat.

258

6/8

determiner *my* on a first beat.
determiner (demonstrative) *this* on a first beat.
pronoun *I* on a first beat twice.
pronoun *we* on first beat.
lexical mismatch (noun): *game-keep-er* on first crotchet, first semiquaver, and second crotchet. It is fine because it is a compound, so the main stress falls on *game-*.

261

4/4

preposition *on* on a third beat.

266

4/4

relative pronoun *which* on a first beat.

267

4/4

preposition *in* on a first beat.

preposition *un-til* on a first beat.

pronoun + auxiliary verb *we'll* on a first beat.

268

3/4

preposition *to* on a first beat.

273

3/2, 2/2

preposition *to*, preposition *of* on a first beat.

280

6/8

preposition *to* on a first beat.

lexical mismatch (noun): *one pennie* on first crotchet, first semiquaver, and second crotchet of the last bar.

This (what I call 'country mismatch') is idiomatic of folk song; it usually happens at the end of a line or stanza. Look at the re-spelling of *pennie* < *penny*; other cases of this phenomenon are *countrie* or *companie*. The unstressed, lax (neutralised) vowel /i/ becomes tense, and therefore, it can carry some degree of stress.

lexical mismatch (noun): *North Coun-te-rie*, on a first and a second beat.

289

6/8

determiner *his* on a first beat.

determiner (demonstrative) *this* on a first beat.

pronoun *I* on a first beat (twice).

298

6/8

lexical mismatch (noun): *la-dy* on third semiquaver of first beat and crotchet of second beat.

This is not a dance song. As happened with *pennie*, the lax, neutralised vowel /i/ gets tense.

305

4/4

preposition *for* on a first beat.

auxiliary modal verb + particle *couldn't* on a first beat.

314

3/4, 4/4

pronoun *it* on a first beat.

pronoun *he* on a first beat.

preposition *of* on a first beat.

315

4/4

preposition *on* on a first beat.

317

6/8

preposition *of* on a first beat.

320

6/8

auxiliary verb + particle *won't* on a first beat.

321

4/4

preposition *to* on a first beat.

preposition *be-tween* on a first beat.

relative pronoun *which* on a first beat.

auxiliary verb + particle *won't* on a first beat.

354

2/4

preposition *for* on a first beat (twice).

conjunction *if* on a first beat.

conjunction *as* on a first beat.

determiner *my* on a first beat.

pronoun *you* on a first beat.

359

6/8

preposition *in* on a first beat, melismatic.

preposition *in* on a third beat.

pronoun *I* on a first beat (twice).