Capítulo 20. Inferring Stress-Timed Rhythm from Written Songs: A Phonetic–Phonological Study of Elton John

Rocío Gragera Retuerto

https://orcid.org/0009-0004-6057-6113

Universidad de Las Palmas de Gran Canaria (Spain)

https://doi.org/10.14679/4737

Abstract: This study examines how English stress-timed rhythm can be inferred from written song lines, drawing on a corpus of twenty-five officially published Elton John lyrics processed in UAM CorpusTool. The approach is purely phonetic–phonological and text-based, with no reference to audio or melody. The question is whether simple orthographic cues suggest a natural stress pattern in neutral reading. Three cues are tracked. First, weak forms are treated as normally unstressed, and contractions in spelling are taken as further evidence of reduction. Second, three clear linking contexts are isolated: "and," "of" and "to" followed by a vowel-initial word, sequences that typically invite coarticulation and compression of unstressed syllables between stronger beats. Third, main stress in neutral declaratives is modelled with a practical proxy: the final content word of the line, or the nearest preceding content item if the line ends on a function word. In UAM CorpusTool, each verse line receives four tags: clause type, list of function words, visible contractions and main-stress proxy. The analysis is qualitative and modest in scope. Recurrent weak-form zones, regular and/of/to links and consistent final-content-word prominence together outline a plausible phonetic–phonological profile compatible with stress-timed English, while the corpus tool keeps procedures transparent and replicable.

Keywords: stress, rythm, weak forms, corpus linguistics, phonetics

1. INTRODUCTION

English and Spanish differ in the way stress and rhythm organise speech. Spanish is traditionally described as a syllable-timed language, in which syllables tend to have similar duration and vowels are normally produced with full quality. English, in contrast, is stress-timed: stresses tend to recur at regular intervals, unstressed syllables are compressed between them, and vowels in these positions are frequently reduced to weak qualities such as schwa (Estebas-Vilaplana, 2020; VSCEP, 2025). This rhythmic contrast has important consequences for intelligibility. Spanish learners of English often give every syllable similar weight, avoid vowel reduction, and misplace stress in long or complex words, which together produce a strong foreign-accent effect (Estebas-Vilaplana, 2020; VSCEP, 2025).

Stress in English operates at several levels. At the lexical level, words may carry one, two or even three stresses, with a primary stress that is more prominent than any secondary stresses. Stress placement may also distinguish grammatical categories, as in "record" /ˈrekɔːd/ (noun) versus "record" /rɪˈkɔːd/ (verb), and mark the difference between phrases and compounds, as in "green house" /ˌgriːn ˈhaʊs/ versus "greenhouse" /ˈgriːnhaʊs/ (Estebas-Vilaplana, 2020). At the phrasal level, stress interacts with information

structure. In neutral reading, prominence usually falls on the final content word of the clause, while function words remain unstressed and often appear in weak forms (Halliday & Matthiessen, 2014; VSCEP, 2025). In connected speech, rhythm favours alternation between strong and weak syllables, promotes stress shift in some double-stressed words, and encourages the compression of unstressed material through vowel reduction and contraction (Estebas-Vilaplana, 2020).

Songs offer a clear setting in which to explore these mechanisms. Verse lines normally align one or more stressed syllables with musical beats, while intervening unstressed syllables are grouped into rhythmic feet (Estebas-Vilaplana, 2018). For learners whose first language is syllable-timed, song lyrics and nursery rhymes have long been used to practise the shortening of unstressed syllables and the extensive use of weak forms (VSCEP, 2025). Popular music further adds a strong cultural frame that facilitates engagement and identification (Rojek, 2011). Yet most linguistic work on song rhythm draws on audio recordings and metrical analysis rather than on the written lyrics themselves, and little attention has been paid to the extent to which a stress-timed profile can already be inferred from orthography alone.

The present study addresses this gap through a small corpus of twenty-four officially published songs by Elton John, spanning several decades of commercial success. The corpus is processed in UAM CorpusTool as plain text, with one verse per line. The analysis focuses on phonetic–phonological cues to stress and rhythm that are recoverable from spelling. Three cues are chosen because they are both simple to annotate and central to English stress timing: the distribution of weak forms and visible contractions, the presence of three recurrent linking environments (“and”, “of” and “to” before vowel-initial words), and a practical proxy for main stress in neutral declarative lines, defined as the final content word of the line or the nearest preceding content word when the line ends in a function item.

2. RESEARCH QUESTIONS

The study asks how far a stress-timed rhythm can be recovered from songs when only basic phonetic–phonological cues visible in orthography are considered. English is treated as a system of patterned choices in stress, reduction and linking, which may be more or less compatible with stress timing in neutral reading, especially when contrasted with Spanish (Estebas-Vilaplana, 2020).

The central question is whether the written lyrics alone provide sufficient cues to support a stress-timed interpretation. More specifically, the study asks (i) to what extent weak forms and their orthographic correlates populate inter-stress spaces; (ii) how regularly the linking environments “and,” “of” and “to” occur before vowel-initial words; and (iii) whether neutral declarative lines consistently place their final content word in a position that would naturally attract nuclear stress (Halliday & Matthiessen, 2014).

3. THEORETICAL FRAMEWORK

This section outlines the phonetic–phonological concepts that guide the analysis of stress and rhythm in Elton John’s lyrics. The framework combines a systemic view of phonology with descriptive accounts of English and Spanish stress, weak forms and rhythmic organisation (Estebas-Vilaplana, 2020).

From a systemic functional perspective, phonology is a meaning-making resource: speakers select patterns of prominence, reduction and tonal movement to enact interpersonal relations and to package information, much as they select grammatical options in lexis and syntax (Halliday, 1961; Halliday & Matthiessen, 2014). Rhythm is thus treated as a system of choices rather than an optional performance feature. In stress-timed languages such as English, this system favours alternations of strong and weak syllables, organised into feet whose stressed beats tend to recur at roughly equal temporal intervals (Banks, 2019; VSCEP, 2025). In syllable-timed languages such as Spanish, by contrast, rhythm is anchored to the syllable: each syllable occupies a similar amount of time and vowel quality remains relatively stable, irrespective of stress (Estebas-Vilaplana, 2020).

The contrast begins at word level. Both English and Spanish words contain stressed and unstressed syllables, but English displays a wider range of stress patterns. Spanish stress typically falls on the final, penultimate or antepenultimate syllable, with only marginal cases of pre-antepenultimate. English, in addition to these three patterns, freely allows pre-antepenultimate stress in items such as “decorator” /ˈdekəreɪtə/, and often distributes more than one stress within long words (for example “photographic” /ˌfəʊtəˈgræfɪk/, “organization” /ˌɔːgənaɪˈzeɪʃn̩/, “irresponsibility” /ˌɪrɪˌspɒnsəˈbɪləti/), with a primary stress and one or more secondary stresses (Estebas-Vilaplana, 2020).

Word stress contrasts can also signal grammatical category. In pairs like “to present” /prɪˈzent/ versus “a present” /ˈprezənt/, “to rebel” /rɪˈbel/ versus “a rebel” /ˈrebəl/, or “to record” /rɪˈkɔːd/ versus “a record” /ˈrekɔːd/, final stress marks the verb and initial stress marks the noun or adjective. Similarly, stress distinguishes phrases from compounds: a phrase like “green house” /ˌgriːn ˈhaʊs/ carries a secondary stress on the first element and a primary stress on the second, whereas the compound “greenhouse” /ˈgriːnhaʊs/ concentrates primary stress on the first element (Estebas-Vilaplana, 2020). In connected speech, these regularities interact with rhythmic pressures. To avoid adjacent strong beats, English allows stress shift in double- or multi-stressed words when they precede another stressed item, as in “Japanese” /ˌdʒæpəˈniːz/ versus “Japanese food” /ˈdʒæpəniːz ˈfuːd/, where the main stress of “Japanese” moves from final to initial position (Estebas-Vilaplana, 2020).These stress systems feed into contrasting rhythmic types. English is commonly described as stress-timed: the time between successive stressed syllables tends to be similar, so stretches with many unstressed syllables must be compressed. This compression is achieved by shortening unstressed syllables and weakening their vowels, which often centralise towards [ə] or reduce to other weak vowels [i], [u], [ɪ] and [ʊ] in non-prominent positions (Estebas-Vilaplana, 2020). Spanish is syllable-timed: stressed and unstressed syllables have comparable duration, vowel quality remains full, and there is little or no centralisation in weak positions. The result is an even “machine-gun” rhythm in Spanish and a more “bouncy” alternation of strong and weak syllables in English (Banks, 2019; Estebas-Vilaplana, 2020).

Weak vowels underpin the English weak-form system. Many monosyllabic function words (articles, prepositions, auxiliaries, pronouns, conjunctions and modals) have both a strong form, used in isolation or under emphasis, and a weak form, used in ordinary connected speech (Estebas-Vilaplana, 2020). For example, “for” contrasts [fɔː] in emphatic position with [fə] in neutral contexts; “can” contrasts [kæn] in short answers and contrastive focus with [kən] in unstressed positions; and “have” in perfect constructions frequently surfaces as [əv] with no [h] in sequences such as “could have,” “should have,” “must have” and “would have” ([ˈkʊdəv], [ˈʃʊdəv], [ˈmʌstəv], [ˈwʊdəv]). Existential “there is/are/was/were” likewise reduces to weak forms such as [ðəz], [ðərə],

[ðə wəz] and [ðə wə], while the auxiliary "are" commonly appears as [ə] after pronouns ("we're" [wi ə], "they're" [ðe ə]) (Estebas-Vilaplana, 2020).

Not all function words have weak forms; items such as "off," "on" and "those" typically retain full vowels. Moreover, the choice between strong and weak forms is grammatically and pragmatically conditioned: auxiliaries and modals take strong forms in short answers ("Yes, they are"), under negation ("mustn't" [ˈmʌsnt]), or when bearing contrastive focus, whereas prepositions and conjunctions tend to be strong at clause boundaries ("What is it for?") and weak elsewhere (Estebas-Vilaplana, 2020). These distributional patterns are central to English rhythm and provide a rich set of textual cues in written lyrics, where contractions (for instance "I've," "she's," "there's," "can't") often encode the expected weak-form realisation (VSCEP, 2025).

Prominence in connected speech is closely tied to information structure. In neutral declaratives, both English and Spanish normally place nuclear stress on the last lexical item, as in "Mary's singing a lullaby" or "*María canta una canción*." When only part of the clause carries new information, the languages diverge. Spanish tends to preserve sentence-final prominence by moving the focused element to the end ("*La canta María*"), whereas English keeps the syntactic order and shifts nuclear stress forward ("*Mary's singing*") (Estebas-Vilaplana, 2020). English therefore allows considerable flexibility in nuclear placement without syntactic reordering, a flexibility exploited in contrastive exchanges such as "I TOLD Mary" versus "I told MARY," where accent movement alone marks focus.

The default rule of "last lexical item" in English has important qualifications. When clauses end in function words, nuclear stress stays on the preceding content word ("It's a letter for her," "Where do you come from?"), in line with Halliday's account of tonic placement (Halliday & Matthiessen, 2014). Event sentences form another special group: in announcements such as "The phone's ringing," "The dog's missing" or "This is Glenda calling," neutral intonation places the main accent on the subject rather than the verb, because the process itself is predictable or backgrounded (Estebas-Vilaplana, 2020). Phrasal and prepositional verbs also display systematic prominence patterns: phrasal verbs highlight the particle ("turn it *down," "pick you *up"), whereas prepositional verbs stress the verb ("*deal with it," "*write to her"), a contrast that again interacts with the placement of rhythmical beats.

These properties are highly relevant for song text. Verse lines in popular music are typically aligned with metrical beats, but they are also written in ordinary orthography and draw on the same lexical and grammatical resources as other registers. If weak forms, contractions and function-word chains populate non-beat positions, and if line-final content words regularly coincide with likely nuclear positions, written lyrics can strongly suggest a stress-timed, English-like rhythm even in the absence of audio. Songs also lend themselves naturally to rhythmic rehearsal: clapping to stressed syllables, compressing unstressed material, and mapping spoken rhythm back onto melodic timing (Estebas-Vilaplana, 2018).

For Spanish learners of English, the differences sketched above pose persistent difficulties. Transfer from a syllable-timed system often leads to equal syllable timing, full vowel quality in all positions, and over-stressing of content words in sequences such as "little *old *lady," where English commonly de-stresses the middle item to maintain a smoother beat (Estebas-Vilaplana, 2018). Failure to use weak forms, to exploit linking contexts, or to shift nuclear stress forward when required results in speech that is intelligible but heavily accented.

4. METHOD

4.1. Approach

The study adopts a qualitative, text-based approach grounded in phonetics and phonology. Rhythm is treated as a pattern of stressed and unstressed syllables supported by weak forms, reduction and alternation between strong and weak positions (Banks, 2019). The analysis focuses on three orthographic cues: function words and contractions, selected linking contexts, and a proxy for neutral nuclear stress.

4.2. Tools

UAM CorpusTool is used to store the lyrics as plain-text files, segment them into verse lines, and apply four tags to each line: clause type, function words, visible contractions and main-stress proxy. Preparation and metadata management are carried out using standard text and spreadsheet software.

4.3. Corpus

The corpus consists of twenty-four officially released Elton John songs, selected to represent three phases of the artist's career (1970s, 1980s and 1990s). Selection is guided by cultural and commercial prominence, drawing on chart data, streaming figures and critical rankings (Official Charts Company, n.d.; Spotify, n.d.; Wade, 2025). Lyrics are treated as written scripts rather than performance transcripts.

4.4. Procedure

The procedure unfolds in four stages: preparation, annotation, retrieval and interpretation. In the preparation stage, each lyric is copied into a text editor and cleaned. This involves removing headings such as "Verse 1" or "Chorus," eliminating extra spaces, and standardising apostrophes and quotation marks. Verse lines are arranged so that each occupies a single line in the text file. Spoken interjections that clearly belong to the character voice (for instance "Oh, my," "yeah, yeah") are kept. Once depurated, each file is saved in plain-text format with a consistent naming convention (song title and year) and then imported into a UAM CorpusTool project.

The annotation stage then applies the four-line tags. First, line type is assigned. Lines are labelled "neutral declarative," "question," "imperative" or "other," based on their grammatical form and punctuation. "Other" covers exclamatives, verbless slogans and fragmentary repeats. Only neutral declaratives are later used for the main-stress proxy tag, since the default nuclear placement rule targeted here applies most straightforwardly to that clause type (Halliday & Matthiessen, 2014).

Second, function words are marked and contractions are tagged. The annotation records both standard English contractions ("'m," "'re," "'s," "'ve," "'ll," "n't") and frequent lyric spellings such as "gonna," "wanna," "'cause," "ain't" and similar forms. These items are taken as orthographic signals of reduced vowels or elided consonants in performance, and they often bunch around function words, creating dense weak zones between stressed syllables.

Third, the main-stress proxy is assigned for lines labelled as neutral declaratives. The procedure scans each line from right to left until the last content word (noun, lexical verb, adjective or adverb) is found. That word is recorded as the nuclear-stress candidate. If the line ends in a function word, the proxy automatically "backs up" to the nearest preceding content word; if no clear content word is present (for example, in very short exclamations), the proxy is left blank. This simple rule operationalises the widely described pattern whereby nuclear stress in neutral English clauses falls on the final lexical item (Halliday & Matthiessen, 2014).

The retrieval stage uses UAM CorpusTool's query functions to group lines by tag combinations. Examples include: all lines containing one or more contractions, and all lines where "and + vowel-initial word" occurs. Lists of lines are exported and, when useful, arranged in spreadsheet form by song and decade. The emphasis remains on patterns that are visible directly in the text rather than on deriving frequency counts or statistical measures.

Finally, in the interpretation stage, the retrieved lines are read in context. The analysis looks for recurring configurations that match descriptions of English stress-timed rhythm. Particular attention is given to stretches where chains of function words and contractions precede or follow a salient content word, since these patterns provide natural "bridges" of weak syllables between stronger beats. Lines with multiple "and/of/to + vowel" junctures are examined as potential clusters of liaison-friendly sites, especially when they sit between prominent lexical items. Neutral declaratives are checked to see whether the main-stress proxy tends to fall on words carrying discourse payoff.

For the exploratory early/late comparison, the same kinds of patterns are inspected across the three decade-based subgroups of songs. The aim is not to document precise diachronic change but to see whether the reliance on weak forms, linking contexts and line-final content words persists throughout the corpus, and whether later lyrics show any tendency toward denser packaging of contractions or function-word chains often references to time, place, address or affect—as suggested by accounts of nuclear focus in English (Halliday & Matthiessen, 2014).

5. RESULTS

5.1. Line profile and clause types

The corpus comprises 1,038 verse lines across the three periods. Neutral declarative lines are clearly dominant, with questions and imperatives forming a small but meaningful minority. Table 1 summarises the structural profile by decade.

Table 1. Line, clause and function-word profile by decade.

Decade	Lines	Neutral declaratives	Questions	Imperatives	Tokens	Function words (% of tokens)	Lines with ≥2 consecutive function words	Lines with chains (%)
1970s	709	636	20	53	3,382	48.3%	184	25.9%
1980s	211	188	6	17	1,060	49.9%	69	32.7%
1990s	118	99	2	17	585	46.3%	25	21.2%

Source: Author's own elaboration.

Across all decades, around half of all tokens are function words, a typical profile for English connected text. Neutral declaratives make up roughly nine out of ten lines in each period, providing a strong basis for modelling "neutral reading" stress. Imperatives, often associated with performance directions or stance ("Hold me closer, tiny dancer"; "Don't

go breaking my heart"), account for around 7–8% of lines in the 1970s and remain present in later decades.

5.2. Weak forms, function-word chains and contractions

The high proportion of function words illustrates the raw material available for weak-form realisation. Consecutive function-word chains (for example "to the", "in the", "do n't") appear in roughly a quarter of lines in the 1970s, rising to about one third in the 1980s and remaining frequent in the 1990s. Average chain length is just over two items in all periods. Typical chains include "in the", "to the", "for the" and negative clusters such as "do n't" and "could have", which lend themselves to compression between stressed content words (Estevas-Vilaplana, 2020).

Contractions are also well represented. Overall, there are 64 contracted forms in the 1970s subcorpus, 59 in the 1980s and 30 in the 1990s. Because the later periods contain fewer lines, contraction density per 100 lines is considerably higher in the 1980s and 1990s than in the 1970s. The most frequent clitics are "'s", "'m", "'re" and "n't", with "'ll" and "'ve" also present.

These contractions often cluster within individual lines. Examples such as "I'd buy a big house where we both could live" ("Your Song"), "If I was a sculptor, but then again, no" ("Your Song"), or "You're gonna hear electric music, solid walls of sound" ("Bennie and the Jets") combine several reduction-friendly environments within a single foot or two. In later songs, lines like "I'm still standing better than I ever did" ("I'm Still Standing") and "You'll never know how much I really love you" ("Nikita") show similarly dense packaging of clitics. These patterns support an interpretation of the lyrics as designing multiple weak-form slots between stronger beats, consistent with a stress-timed rhythm.

5.3. Linking sequences "and/of/to + vowel"

Linking contexts built around "and", "of" and "to" followed by vowel-initial words are frequent and contribute further to potential syllable compression and coarticulation. Across the corpus, 79 such sequences are attested: 45 in the 1970s, 22 in the 1980s and 12 in the 1990s, which again represents a higher density per line in the later decades.

The sequence "and I" is especially recurrent and appears in all three periods, often in emotionally salient lines such as "And I think it's gonna be a long, long time" or "And I guess that's why they call it the blues"). These junctures naturally favour linking and reduction in neutral speech ("and I" → [ənˈaɪ]), placing a weak, easily compressible bridge between two prominent content items.

Other sequences, such as "and a", "of a" and "of any", appear mainly in the 1970s material and typically follow a stressed noun or verb ("part of a brand new start"). In reading, these lines align with canonical English patterns where clusters like "of a" are realised with schwa vowels and minimal consonant release, helping to keep the temporal distance between stresses relatively even despite variable syllable counts.

5.4. Main-stress proxies and semantic tendencies

The last content word of each line was taken as a practical proxy for nuclear stress in neutral delivery, with the proxy "backing up" when lines ended in function words. A content-word proxy could be identified for 1,037 of the 1,038 lines. In approximately 16% of lines, the orthographic line ended on a function word ("to", "for", "you"), so the proxy moved one or more steps leftwards; this back-up occurred 122 times in the 1970s, 21 times in the 1980s and 18 times in the 1990s–2020s. These proxies were then grouped into broad semantic fields. Table 2 displays the distribution by decade.

Table 2. Semantic fields of main-stress proxies in neutral declaratives.

Category	1970s (n=709)	1980s (n=210)	1990s–2020s (n=118)
Emotion	41 (5.8%)	23 (11.0%)	5 (4.2%)
Address/relationship	3 (0.4%)	6 (2.9%)	0 (0.0%)
Time	47 (6.6%)	16 (7.6%)	26 (22.0%)
Place	21 (3.0%)	6 (2.9%)	2 (1.7%)
Movement	10 (1.4%)	1 (0.5%)	0 (0.0%)
Other	587 (82.8%)	158 (75.2%)	85 (72.0%)

Source: Author's own elaboration.

Although most final content words fall into a broad "other" category (covering a wide range of lexical items), several tendencies support a stress-timed reading. First, time-related items become increasingly common as line-final proxies, rising from about 7% in the 1970s and 1980s to over 20% in the 1990s. Lines such as "I'm still standing better than I ever did", "Can you feel the love tonight", and "Something about the way you look tonight" all end on time nouns or temporal adverbs ("tonight", "time"). These endings provide natural nuclear-stress targets, while earlier material in the line tends to consist of higher proportions of weak-form function words and reduced syllables.

Second, emotion words regularly appear in line-final position, particularly in the 1980s. Proxies such as "blues", "love", "hurt" and "sacrifice" mark the discourse payoff of the line, with preceding stretches often packed with modals, auxiliaries and pronouns ("I guess that's why they call it the blues"; "It's a human sign when things go wrong"). This pattern aligns with descriptions of English nuclear placement on the most informative content item near the end of the tone unit, while keeping function words in weak, compressed positions.

Third, place-related items and movement verbs, though fewer in number, play a consistent supporting role. In the 1970s, final content words like "road", "town" or "home" ("Goodbye Yellow Brick Road") often close lines whose early portions house chains of function words ("back to the howling old owl in the woods"). These lines read naturally with stresses landing on scenic nouns while prepositions, articles and pronouns shrink rhythmically between beats.

6. DISCUSSION

From a theoretical perspective, the interpretation of rhythm adopted here follows standard descriptions of English as a stress-timed language, in which prominence recurs on stressed syllables while unstressed material is compressed between beats (Wells, 2006; Estevas-Vilaplana, 2020). Nuclear stress placement is interpreted in line with Halliday's model of English intonation, where neutral declaratives typically place tonic prominence

on the final lexical item of the tone unit (Halliday & Matthiessen, 2014). Weak forms, vowel reduction and contraction are treated as core phonetic correlates of this rhythmic organisation (Wells, 2006; Estebas-Vilaplana, 2020).

The results point to a coherent answer to the central question of the study: written Elton John songs do provide enough orthographic evidence to support a stress-timed reading in neutral delivery. Across decades, the songs repeatedly organise content and function words, contractions, and linking contexts in ways that mirror the expected alternation of strong and weak beats in English speech.

The line profile offers a first indication of this. Neutral declarative lines clearly dominate, with questions and imperatives forming a small but stable minority. This distribution supports a neutral prosodic template in which stresses fall on content words, while function words (around half of all tokens in each subcorpus) remain available for weak-form realisation (Estevas-Vilaplana, 2020; VSCEP, 2025).The presence of imperatives, especially in 1970s songs and again in later hits, adds further positions where function words cluster around a prominent verb, encouraging weak forms in forms such as “hold me closer” or “don’t let the sun go down on me”.

The density and organisation of function words strongly reinforce this pattern. Consecutive function-word chains appear in roughly a quarter to a third of lines, depending on the decade, and typically include high-frequency items such as articles, prepositions, auxiliaries and pronouns. Sequences such as “in the”, “to the”, “for the” or “do not” are classic sites for reduction in stress-timed English, both because they are semantically light and because they occur between more informative content words. In the lyrics, these chains usually act as bridges between stressed lexical items, supporting isochrony at the level of stress groups by allowing longer stretches to be compressed into temporal slots comparable to shorter ones (Wells, 2006).

Contractions add another layer of evidence. Across the corpus, “’s”, “’m”, “’re”, “’ll”, “’ve” and “n’t” are frequent, and their density increases in the 1980s and 1990s when measured per 100 lines. These forms mark, in the spelling itself, precisely those weak, unstressed syllables that carry auxiliaries and pronouns. Lines like “The spotlight's hittin' somethin' that's been known to change the weather” illustrate several orthographic signals of reduction in very short spans of text. In a neutral reading, each contraction would naturally be realised with a weak vowel or compressed consonant cluster, helping to keep the time between stressed beats stable even when the number of syllables varies considerably.

The behaviour of “and/of/to + vowel” sequences is consistent with this picture. The corpus contains numerous cases of “and I”, “and a”, “of a”, “of any” and similar junctures, with “and I” particularly prominent across all three periods. These environments are well-known gateways for linking and coarticulation in standard descriptions of English connected speech, where function words in pre-vocalic position are regularly reduced and resyllabified (Wells, 2006; Halliday & Matthiessen, 2014). In lines such as “And I think it’s gonna be a long, long time” or “And I guess that’s why they call it the blues”, the sequence “and I” sits between stressed items (“think”, “time”; “guess”, “blues”) and invites a reduced [ənˈaɪ]-type realisation. Their repeated appearance suggests that the lyrics routinely build in orthographic environments where weak syllables are compressed between stresses, supporting a stress-timed rhythm at the level of connected speech (Estevas-Vilaplana, 2020).

The analysis of main-stress proxies further confirms the alignment between textual structure and expected prosody. In almost all lines, a final content word can be identified,

either at line end or by backing up from a final function word. This last content item is a reasonable proxy for nuclear stress in neutral declaratives, in line with standard descriptions of English intonation, where tonic prominence typically falls on the final lexical item of the tone unit unless overridden by information structure (Halliday & Matthiessen, 2014; Wells, 2006). Its semantic profile is not random. Across the corpus, line-final proxies regularly fall into meaningful fields: time ("time", "night", "today", "tonight"), place ("road", "town", "home"), emotion and stance ("love", "blues", "sacrifice", "hurt"). These are precisely the kinds of lexemes that would be expected to carry discourse payoff in popular lyrics.

The diachronic pattern is especially clear in the increased prominence of time-related nouns in later decades, where items such as "tonight" and "time" appear as line-final proxies with much higher relative frequency than in the 1970s. This does not weaken the stress-timed profile; rather, it shows that, as the repertoire evolves, nuclear stress increasingly aligns with abstract temporal frames structuring the song's emotional scene. Earlier in the line, function words and contractions accumulate, producing precisely the weak-syllable cushion that stress-timed English uses to maintain isochrony between prominent beats (Wells, 2006).

Emotion-related proxies, especially visible in the 1980s, work in a similar way. Lines closing on "blues", "love" or "sacrifice" align nuclear stress with key affective labels, while earlier segments of the line are packed with pronouns, modals and auxiliaries that will bear weak forms in neutral speech. This division of labour between content and function words matches standard descriptions of English nuclear placement and rhythm, as outlined in Halliday's model of intonation and in phonetic accounts of stress timing in English (Halliday & Matthiessen, 2014; Wells, 2006): the content word that carries new or salient information receives the main accent; the surrounding grammatical material is reduced and compressed to accommodate the beat.

The limited contrasts across decades also merit comment. The basic rhythmic design remains stable: content words carry the semantic weight and likely stresses; function words and contractions fill the interstices and invite reduction; "and/of/to + vowel" sequences provide systematic linking points. The main diachronic differences are a tighter packaging of contractions and a relative increase in time-related line-final items in later songs. Both trends arguably make the stress-timed pattern more visible rather than less, since repeated "tonight/time" nuclei provide clearer queues where stresses and weak syllables will fall in neutral delivery (Halliday & Matthiessen, 2014).

These findings support the practical claim that written popular lyrics can serve as effective material for teaching stress and rhythm to Spanish speakers. The songs contain numerous lines that read very naturally with stress-timed timing, especially when function words are de-emphasised and contractions and linking contexts are fully exploited. Activities such as clapping the beats on final content words, underlining function-word chains between stressed lexemes, or practising weak forms in clusters like "and I", "of a", "to another" can help learners experience how unstressed syllables "shrink" between stresses (Estevas-Vilaplana, 2020).

Finally, several limitations must be acknowledged. The corpus is modest and focused on a single artist, and the analysis relies on a practical proxy for nuclear stress rather than measured acoustic data. The assumption of neutral reading ignores performance-style deviations and individual interpretive choices. Even so, the convergence of patterns across songs and decades suggests that Elton John's lyrics are textually organised to favour stress-timed delivery, accumulating weak-form material between content-word

beats, canonical linking sites, and semantically salient nuclear positions. In this sense, the corpus provides a clear example of how mainstream English lyrics can encode stress-timed rhythm through orthography alone (Estevas-Vilaplana, 2020; Halliday & Matthiessen, 2014; Wells, 2006).

7. CONCLUSIONS

The study set out to answer whether English stress-timed rhythm can be credibly inferred from written Elton John songs, using only orthographic information and a small set of phonetic–phonological cues. It also asked how those cues are distributed across three decades of songs, and whether the resulting patterns have practical value for learners of English, especially Spanish speakers.

The first research question concerned the adequacy of text-based cues for reconstructing a stress-timed pattern in neutral reading. The results indicate that the corpus provides consistent and interpretable evidence. Across songs and decades, lines are predominantly neutral declaratives whose organisation of content and function words closely matches descriptions of English stress-timed rhythm. High frequencies of function words, systematic weak-form chains between stressed lexemes, numerous contractions and recurrent "and/of/to + vowel" sequences converge to create predictable zones of reduction and compression between stronger beats (Estevas-Vilaplana, 2020). Line-final content words generally carry the main semantic load and form a robust proxy for nuclear stress (Halliday & Matthiessen, 2014). Even without audio, the lyrics therefore encode a rhythmic design that aligns well with stress-timed English.

The second question concerned how weak forms, contractions, linking junctures and line-final content words behave across the corpus and whether any diachronic tendencies can be identified. Several stable regularities emerge. Function words appear in dense clusters that act as bridges between stressed content words, supplying the weak syllables that a stress-timed system requires. Contractions are widespread and become relatively more concentrated in later songs, marking orthographically those positions where auxiliary verbs and pronouns are likely to be unstressed and reduced. The three targeted linking contexts, particularly "and I", occur frequently and often sit between clear content-word beats, providing classic environments for coarticulation and syllable compression (Estevas-Vilaplana, 2020).

Line-final main-stress proxies show a patterned semantic profile. In all decades, proxies typically fall on nouns and adjectives of time, place, movement, address or emotion. Earlier songs often close lines on concrete locations and narrative images, while later songs show a drift toward temporal and affective lexemes like "time", "tonight", "love" and "blues". These differences do not alter the basic rhythmic architecture: in all periods, the syllable that would naturally carry nuclear stress tends to be filled by a content word with obvious discourse payoff, while preceding stretches carry the bulk of function words and potential weak forms.

The third question addressed the pedagogical relevance of these findings. From a teaching perspective, the corpus offers a compact, authentic resource for working on stress and rhythm with Spanish learners. Many lines lend themselves naturally to beat-clapping on final content words, while the surrounding function-word material illustrates how unstressed syllables can be compressed between stresses. The prevalence of contractions and "and/of/to + vowel" sequences provides ready-made material for practising weak forms, linking and coarticulation. Because the songs are culturally familiar and designed

for wide audiences, they model a form of rhythm that is both recognisably English and accessible.

At the same time, the patterns highlighted by the analysis help to explain typical Spanish-accented difficulties. A syllable-timed delivery that maintains full vowels and equal syllable length would neutralise much of the rhythmic design encoded in the lyrics, placing undue prominence on function words and weakening the contrast between strong and weak beats (Estevas-Vilaplana, 2020). The corpus makes visible how much the natural flow of these songs depends on unstressed, reduced material between stresses.

Overall, the study shows that even in the absence of audio, popular English lyrics can encode a clear stress-timed profile in their written form. In this corpus, that profile is stable across decades and offers concrete, text-based starting points for understanding and teaching English stress and rhythm, particularly in contexts where Spanish syllable-timed patterns predominate.

ACKNOWLEDGEMENTS

We express our sincere gratitude to the University of Las Palmas de Gran Canaria and to the Regional Ministry of Universities, Science and Innovation and Culture of the Government of the Canary Islands, which, together with co-funding from the European Social Fund Plus made the development of this work possible.

REFERENCES

Banks, D. (2019). *A systemic functional grammar of English*. Cambridge University Press.

Elton John. (n.d.). *Official lyrics and publications*. https://www.eltonjohn.com

Estebas-Vilaplana, E. (2018). Teaching stress and rhythm in English pronunciation: A phonetic and phonological approach for Spanish learners. In E. Alcaraz Varó & J. Mateo Martínez (Eds.), *Applied phonetics and phonology* (pp. 143–165). Comares.

Estebas-Vilaplana, E. (2020). *The phonetics and phonology of English pronunciation for Spanish learners*. Routledge. https://doi.org/10.4324/9780429024090

Halliday, M. A. K. (1961). Categories of the theory of grammar. *Word, 17*(3), 241–292. https://doi.org/10.1080/00437956.1961.11659756

Halliday, M. A. K., & Matthiessen, C. (2014). *Halliday's Introduction to Functional Grammar* (4th ed.). Routledge. https://doi.org/10.4324/9780203431269

O'Donnell, M. (2008). *The UAM CorpusTool: Software for corpus annotation and exploration*. Universidad Autónoma de Madrid. https://www.wagsoft.com/CorpusTool/

Official Charts Company. (n.d.). *Elton John: Chart history and discography*. https://www.officialcharts.com

Rojek, C. (2011). *Pop music, pop culture*. Polity Press.

Spotify. (n.d.). *Elton John: Artist profile and streaming data*. https://www.spotify.com

VSCEP. (2025). *Handbook for the Summer Course in English Phonetics* [Online course materials]. University College London.

Wade, S. (2025). *The 50 essential Elton John songs*. Rolling Stone. https://www.rollingstone.com

Wells, J. C. (2006). *English intonation: An introduction*. Cambridge Press.