

**SOCIAL SCIENCE**  
**CONFERENCES 2020 SPRING**  
**CONFERENCE PROCEEDINGS**

Özgür Öztürk DAKAM YAYINLARI

June 2020, İstanbul.

[www.dakam.org](http://www.dakam.org)

Firuzaga Mah. Boğazkesen Cad., No:76/8, 34425, Beyoğlu, İstanbul

SOCIAL SCIENCE CONFERENCES 2020 SPRING

Conference Proceedings

Conference Coordination: DAKAM (Eastern Mediterranean Academic Research Center)

Cover Design: D/GD (DAKAM Graphic Design)

Design: Özgür Aydoğan

Print: Metin Copy Plus, Mollafenari Mah., Türkocağı Cad. 3/1, Mahmutpaşa/  
İstanbul, Turkey

ISBN: 978-625-7034-02-9

# ON THE CONCEPT OF RHYTHM IN PHONOLOGY

**CAMILO ENRIQUE DÍAZ ROMERO**

Ph.D., Instituto Caro y Cuervo (Colombia)

camilo.diaz@caroycuervo.gov.co

## Abstract

The present text exposes the historical existence of two definitions of rhythm: one based on isochronies and duration, “stress-timed” vs. “syllable-timed” (Monboddo, 1787; 1789; Pike, 1945; Han, 1962; Abercrombie, 1967; Dauer, 1983; Rasmus, Nespov and Mehler, 1999), and other based on word and phrase accentuations (Álvarez, 1785; Thomson, 1904; Donegan and Stampe, 1983; 2004; Donegan, 1993), “falling pattern” vs. “rising pattern”. While the first case is focused on physical measures (vowel duration, Voice Onset Time-VOT), combined with some phonological processes (nasalizations, palatalizations, velarizations, lenitions, elisions), the second adds some morphological (number of morphemes per word) and syntactic properties (object-verb order, determinant-noun order, adjective-noun order) to this concept. Definitions of rhythm create differences on conception of language structures, as illustrated with data of Spanish spoken in Boyaca (Díaz, 2017) and Embera (Páez and Ibarra Zetter, 2018; Sáenz, 2018). Although the two cases are considered “syllable-timed” systems in general terms, embera is less “falling pattern” than the Spanish spoken in Boyaca.

**Keywords:** Isochrony, accentuations, phonological processes, Spanish spoken in Boyaca, Embera

## 1. The concept of rhythm based on duration

The concept of rhythm has been defined in various ways throughout history. For instance, according to Monboddo (1787), it was considered as “number of syllables per verse” (p. iv.), even if he thought that this property is not essential for human languages:

“As to melody and rhythm, they are not essential to language; neither do I think the invention of them near to difficult as of articulation; for I am persuaded, that language began in the fourthern countries, where all the inhabitants were naturally musical, as much as certain birds among us. It was therefore natural, and indeed in some fort necessary, that the men,

who first articulated, should join with it both musical tones and rhythms” (Monboddo, 1787, p.181).

After that, Monboddo (1789) replaced the concept of rhythm as a parameter to measure the number of syllables per verse and, instead of this, he redefined rhythm as the duration of syllable components, distinguishing between short and long syllables:

“Some long syllables, longer than others, and some short syllables, shorter than others. -The Halicarnassian’s account of long and short syllables, that is, of the rhythm of language, more distinct than Cicero’s account.- Of syllables words are made, and of words sentences, with all the variety of periods and members of periods” (Monboddo, 1789, p.xxix).

However, Steele (1775) considered that speech rhythm not only is related to duration, but also with presence/absence of syllable codas (cadence), establishing the distinction between heavy, light and lightest rhythms, as well as longest, long, short and shortest patterns:

“Speech is to consist of melody by slides (acute and grave) with force (loud or soft), and measure or rhythmus of motion and rest distinguished by quantity (longest, long, short, shortest) and emphasis or cadence (heavy, light, lightest)” (Steele, 1775, p.24).

More than a century after, Morris (1925) considered that rhythm have to accomplish with two features: regularity and perception. It is not necessarily related to physical production of sound patterns, but to those patterns are heard and valued as entities measurable in regular intervals:

“Rhythm may be conceived as patterns of sound recurring at exactly equal intervals of time, as symmetry may imply identical space intervals. Or the units may be only perceptibly equal or only perceptibly similar, or they may have values perceptibly equal, though given by different factors operative in the medium of time” (Morris, 1925, p.15).

Pike (1945) sustained that rhythms are syllable sequences whose limits are the pauses, like verses in poetry. However, as Morris expressed in the previous paragraph, perception of duration and stress are the main criteria. Such dispositions vary language-to-language. There are two types of systems: stress-timed, which temporal organization is determined by regular intervals of metrical feet, and syllable-timed, based on the perception of syllables as entities of regular durations, regardless of lexical stress:

“A sentence or part of a sentence spoken with a single rush of syllables uninterrupted by a pause is a RHYTHM UNIT [...] A single rhythm unit from such a sequence of units may be considered the regular or normal type. Because its length is largely dependent upon the presence of one strong stress, rather than upon the specific number of its syllables, it may conveniently be labelled a STRESS-TIMED rhythm unit (a phonemic type in contrast to syllable-timed units to be mentioned below, with both of them on a different level of contrast from the simple versus complex rhythm

types) [...] English also has a rhythmic type which depends to a considerable extent upon the number of its syllables, rather than the presence of a strong stress, for some of its characteristics of timing; in English, however, the type is used only rarely. In these particular rhythm units each unstressed syllable is likely to be sharp cut, with a measured beat on each one; this recurrent syllable prominence, even though the stressed syllables may be extra strong and extra long, gives a “pattering” effect. The type may be called a SYLLABLE-TIMED rhythm unit” (Pike, 1945, pp.34-35).

Two decades after, Han (1962) and Abercrombie (1967) add a third class of rhythm systems: mora-timed. Syllables which had long vowels (CVV) or vocal-coda sequence (CVC) are perceived as the same duration as two syllables with short vowels (CV), and stress are defined in regular intervals of moraic feet instead of syllabic feet.

Donegan (1978) considered isochrony as rhythmical unit, but made clear that it depends on intrinsic duration of particular segments and their arrangement on syllables through beats, similar to musical figures such as eighth or sixtieth notes:

“It is commonly recognized that speech, like other natural human activities (walking, chewing, etc.), is basically rhythmic. The number and the intrinsic durations of the individual segmental articulations which make up speech may to some degree disturb this rhythm, and grammatical factors may also intertere, but it appears that the intended timing of speech, at least, is extremely regular. When the basic rhythmic unit is an accent group, or ‘measure’, so that there is equal timing between accents, the language is said to be iso-accentual, or stress-timed. When the basic rhythmic unit is the syllable, so that each syllable constitutes an equal prosodic interval, the language is said to be iso-syllabic, or syllable-timed. And when there is a distinction made between short (one-beat) and long (two-or-more beat) syllables, so that each short vowel (or syllable-offset consonant) is mapped onto one ‘beat’ or time-interval, and each long vowel is mapped onto two, the language is said to be iso-moric, or mora-timed” (Donegan, 1978, p.53).

Dauer (1983) and Auer (1990) considered that the irruption of terms of Cognitive Grammar such as prototypes and graduality is necessary for describing rhythm systems. Most of world’s languages could classify as exclusively syllable-timed or stress-timed, some of them exhibit all prototypical properties, but others are fluctuating among two types:

“As with English, it must be clear that Italian is a language which comes close to the iso-accentual prototype phonology, but doesn’t actually represent it in ‘pure form’. (For instance, Italian is close to the iso-syllabic type, but contrary to what we would expect, accent placement plays some role in its grammar, and accent placement is much more flexible than in other iso-syllabic languages such as French or Turkish). In the life of a language, both principles, the one for iso-accentual isochrony and the one for isosyllabic isochrony, are present; but usually, one is given preference at the expense of the other. That there may be a mixture of elements from the two

prototypes in many of the actual occurring languages is, not in the least, due to diachronic changes of rhythm type, such as in English from iso-syllabic to iso-accentual. The great advantage of working with rhythmic teleologies is that this makes it possible to develop a holistic picture of a phonology” (Auer, 1990, p.19).

With the origin of corpora studies in the decade of 1980, the rhythm is rediscovered as a physical parameter, measurable in statistical terms. An example of these are nPVI (normalized Pairwise Variability Index) and rPVI (raw Pairwise Variability Index), components which analyze the difference of duration among the media of sound intervals combined with speech tempo (rPVI) or not (nPVI). According to Ramus *et al.* (1999), Gibbon and Gut (2001), Low and Grabe (2002), and Nespov, Shukla and Mehler (2011), if there are low levels of rPVI and nPVI, it infers that this language is considered as syllable-timed, but if there are high levels of rPVI and nPVI, the isochrony recognized is stress-timed.

## 2. Rhythm based on word and phrase accentuations

Thomson (1904) and Omond (1905) asserted that rhythm is based on lexical stress, which associates syllables on contiguity relations of intensity contrasts among them, recovering the idea of metrical feet from Greco-Roman tradition mentioned in Álvarez (1785). According to him, rhythm refers to the patterns of 2-3 syllable clusters around to differences of word accentuations. Nevertheless, in contrast to other 18th century theorists mentioned before, Álvarez, and Thomson after, considered that rhythm is identified in ordinary speech, within words and utterances:

“Take the sentence, “A long meandering road led to the river”, and read it in a natural flowing way, tapping with the finger on reaching the accented syllables *long, and, road, led, riv.* It will be found that the taps and, therefore, the accented syllables occur at equal intervals of time. Here we have what is called the principle of equal periodicity. The recurrence of strong accents at equal distances of time runs, with trifling exceptions, through all verse, and is perpetually asserting itself, in a more or less modified form, in the language of prose and ordinary speech. In other words, practically all verse and the bulk of prose is rhythmical [...] But rhythm, or the regular recurrence of accent, may take on different characters according to the varying distribution of speech sounds between the accents. And it is here that we come into contact with the distinctive basis of Teutonic rhythm. The distribution of time over the syllables between the accents might be quite irregular, and referable to no simple principle. But attentive observation shows that this is not the case. On the contrary, the syllabic distribution of time within the intervals admits of rational and simple measurement” (Thomson, 1904, pp. 11-12).

Sapir (1921) sustains that rhythm depends on the operation of diverse prosodic features among languages (stress contrasts, syllable weight differences, melodic variations, etc.), being the verse the main unit:

“Latin and Greek verse depends on the principle of contrasting weights;

English verse, on the principle of contrasting stresses; French verse, on the principles of number and echo; Chinese verse, on the principles of number, echo, and contrasting pitches. Each of these rhythmic systems proceeds from the unconscious dynamic habit of the language, falling from the lips of the folk. Study carefully the phonetic system of a language, above all its dynamic features, and you can tell what kind of a verse it has developed—or, if history has played pranks with its psychology, what kind of verse it should have developed and some day will” (Sapir, 1921, pp. 246).

According to Chomsky and Halle (1968), rhythm is again established on stress and metrical feet, operating on lexical items, as Thomson (1904) asserted previously: “Rule (39) [V → [1 stress] / \_\_\_\_ C<sub>0</sub>VC<sub>0</sub>VC<sub>1</sub>o] produces alternations of stressed and unstressed vowels. It is thus one of the factors contributing to the frequently observed predominance of iambic rhythms in English” (Chomsky and Halle, 1968, p.78).

The relationship of accentuations as the main rhythmic component with phonological processes is supported by Stampe (1969), founder of Modern Natural Phonology. He rescued the position of Jakobson (1968 [1941]), who sustained that children, during first language acquisition stages, tend to produce sound constituents of stressed syllables from lexical items listened of adults who live with them. Moreover, the first syllables produced by children created their own “internal accentuation”, with high contrasts between vowels and consonants:

“The most extreme processes are usually observable only in infancy: unstressed syllables are deleted, clusters and coarticulated ones are simplified, obstruents become lax stops linguals become coronal vowels merge to [a]. The fullest effect of the innate system is seen in the utterances of what might be called the “post-babbling” period, which, although they are still nonsemantic, characteristically consist of well-articulated sequences of identical and stressed syllables composed of lax stop (or nasal) plus low vowel [dadada], [ɲaɲaɲa], [mamama], or the like” (Stampe, 1969, pp. 445-446).

The influence of accentuations as generators of phonological contexts where operate the activation, restriction or removal of processes continue in Donegan (Miller, 1972a; 1972b). For instance, the process of vowel neutralization or centralization, which begins affecting all syllables in children, is suppressed by adults in different world languages, specially on stressed syllables:

“In languages with certain rare vowel systems, a limited form of neutralization may continue to affect stressed vowels even in the adult language. The systems which result seem to lack distinctions of timbre, though some admit distinctions of height; and they will be mentioned again in the section dealing with timbre. Neutralization, it seems, is (almost) always the first process to be limited or suppressed by children, and it is almost universally limited to unstressed or non-tense vowels in adult language” (Miller, 1972a, p.142).

Bjarkman (1976) affirmed that accentuation is a rhythmic factor which has

influence in segmental processes, such as aspiration in voiceless plosives produced by English speakers in casual speech:

“Processes, on the other hand, are often only optionally retained, relative to the governing style and the speed of articulation. There is nothing as all automatic about certain “sloppy speech” rules of English: viz. aspiration of voiceless stops initially in the stressed syllable (e.g. [prəpér] becomes [prəp<sup>h</sup>ér] «prepare»)” (Bjorkman, 1976, p.82).

Even though Donegan and Stampe (1978, pp.32-33) reassumed the isochrony, they clarify that this unit should review within the domain of intention and auditory perception, together with phoneme and musical junctures (Nathan, 1985; Smith, 1980). However, this aspect of rhythm seems to be depended on *tempo* and accentuations, being, in this way, the first time where they established the thesis of stress assignment as the process whose development could affect other prosodic processes:

“The vowel shortenings, deletions of final consonants, monophthongizations and reversals of diphthongs, metatheses, and change of vowel-nasal sequences into nasalized vowels which occurred in the histories of Romance, Slavic, and other languages might find a unified explanation beginning with a simple change in prosodic mapping. In these instances, the shift would have involved a change from a mapping in which syllables with long vowels or with consonantal closures, formerly mapped onto a double beat, were now mapped, like other syllables, onto single beats in the rhythm of speech. The segmental phonological processes would then eliminate long vowels, postvocalic consonants, and falling diphthongs – all of which would be difficult to pronounce in the time allotted a short syllable” (Donegan and Stampe, 1978, p.34).

According to Donegan and Stampe (1979), rhythm is a pattern which structures segmental material on prosodic constituents, where operate fortitive and lenitive processes. However, even if rhythm is related to accentuation and duration, it is connected with intonation, but not tones:

“*Prosodic processes* map words, phrases, and sentences onto prosodic structures, rudimentary patterns of rhythm and intonation. Insofar as syllabicity, stress, length, tone, and phrasing are not given in the linguistic matter, they are determined by the prosodic mapping, which may most easily be described as an operation in real-time speech processing of which setting sentences to verse or music are special cases” (Donegan and Stampe, 1979, p.142).

It is in Donegan and Stampe (1983) when rhythm, defined on terms of word and phrase accentuations, also seems to influence on the development of lexical tones, morphological types (analytic, synthetic, agglutination, fusión, etc.), word order, isochronies, number and classes of vowel phonemes, metrical feet, among other properties. This is the origin of the two types of patterns: falling systems, present in Munda languages, and rising systems, exhibited in Mon-Khmer languages:

“Accentuation is the reason for clisis and affixation in the first place, in that it



subordinates the unaccented element to the accented, and thus typically the merely grammatical element to the lexical. Thus falling accent encourages enclisis and suffixation; rising accent, proclisis and prefixation” (Donegan and Stampe, 1983, p.343).

Hurch (1988) added a property to the influence of rhythm on the phonological structures, asserting that accentuations not only organise syllabic sequences, but also help to determine how many syllables could belong to words or phrases. For this, he made a comparison between German, stress-timed language which, even with the elisión of segments, preserve the quantity of syllables for obtaining binary metrical feet in words as *haben* ‘to have’, meanwhile basque, as syllable-timed system, allows elisions and resyllabifications without maintaining a specific quantity of syllables or metrical feet:

“One of the typological differences between syllable- and accent-timed languages lies in the clarity of the number of syllables in a given word/utterance. Whereas there can be doubts in accent timed languages, there are never doubts is the syllable-timed type. A situation like the reduction and deletion of the suffix vowel in German *haben* «to have», the subsequent assimilation of the nasal results in a possible phonetic representation of something like [habm] with a syllabic nasal. Phonetically this can, for the lack of the oral release of the stop, equally or even better be analyzed as a preglottalized nasal. But the question of the syllabicity of the final consonant and therefore of the number of syllables remains open. In faster speech-styles this problem is solved by the complete deletion of the oral or glotal element respectively. But in normal colloquial pronunciation there exists uncertainty about the number of syllables. Analogous examples can be cited from virtually all accent-timed languages. This problem is inexistent in Basque. Note that this does not mean that the number of syllables remains constant during the whole course of the phonological derivation. There are processes of vowel deletion (like in *ezazu* but *har (e)zazu*), of diphthongization (like in *ama+iru*→[aj]) and others, with subsequent resyllabification processes, but, first, these processes typically take place in V-syllables and therefore do never increase the complexity of confining syllable-onsets or codas by assigning to them a remaining consonant of a deleted syllable, and second there syllabification process operates unequivocally with assigning the quality of syllabicity” (Hurch, 1988, p.818).

Dressler and Siptár (1989) continued in the theoretical proposals of Donegan and Stampe (1983), and Hurch (1988). For this, they exposed data from phonology of Hungarian, with very similar consequences to basque, trochaic feet, syllable-timed isochrony, without significative differences on vowel duration between stressed vs. unstressed contexts, vowel harmony and morphological agglutination:

“1. Fixed place of accent—on the word-initial syllable—and preference for a trochaic rhythm, realised in Hungarian by alternation between unstressed and secondarily stressed syllables. 2. The phonological and phonetic shapes of accented and unaccented syllables are similar, due to similar application of both prelexical and postlexical processes. 3. Accented vowels

are not significantly more lengthened (or otherwise foregrounded) nor are unaccented syllable peaks significantly more shortened (or otherwise backgrounded) than their contrasting classes. 4. And even backgrounding vowel bleaching to schwa is extremely restricted. This is typical for vowel-harmony languages, because vowel bleaching obscures vowel harmony. Thus it tends to occur only in those vowel harmony languages where there are anyway many other exceptions to vowel harmony. In this way, the typological criteria of syllable timing and agglutination favour each other” (Dressler and Siptar, 1989, p.48).

Donegan (1993) decided that, for a characterization of rhythm, it is appropriate to use terms from musicology such as beats and measures. A word could be shaped of a beat of 2 morae (e.g. in Spanish, the term *mamá* ‘mother’ could be said with a monosyllabic item with long vowel, [ma:]) or two beats, each one of them with 1 mora (e.g. in Spanish, *mamá* ‘mother’ could be pronounced as [mama]). Like in music patterns in music, clusters of beats creates measures, where allow the creation of metrical feet, prosodic domains which are the operation bases of segmental processes, with certain degree of flexibility in use, but, like improvisations executed by jazz or heavy metal musicians, it consists on preserving a mental constant background which enables the utterance comprehension:

“The precise coordination of gestures required in speaking and other intentional pattern depends on a neural metronome that emits a flexible but regular pattern in real time, onto which we map intended words and phrases for articulation. The shortest unit in this inner rhythmic pattern that is relevant in speech, verse and song is the time needed to pronounce a short syllable, which I will refer to by the Greek prosodic term **mora**. But the shortest unit of time required to pronounce an independent word is two moras long, and I will refer to it by the musical term **beat**. As I shall use the term, a beat consists of two moras, the first strong and the second weak [...]. A pair of beats, again ordered strong and weak, combine into a **measure** as in English *báby-sitter*, *báby-sit*, *hóuse-sitter*, *hóuse-sit*. Note that there is a subtle lengthening of *sit* here when it is the only material in a beat” (Donegan, 1993, p.7).

The conception of pulses as constituents of rhythmic organization has been its deepest development in the Poznan School of Phonology, with the leadership of Katarzyna Dziubalska-Kořaczyk. According to her, the syllable is not able to explain the preference of vowels as structural nucleus (1996, p.59). Instead of this, the term pulse could solve this problem, related to consonants by means of *binding laws*, which establish sonority nexus on particular sequences:

“I suggest that the notions of “beat”, “word”, and “foot”, as well as “morpheme”, are sufficient to account for the functions of the “syllable” without requiring the unit of a “syllable”. A basic rhythmic speech skeleton consists of regularly recurring beats. Beats are primary rhythmic entities realized preferably by vocalic figures against a consonantal ground. They do not possess any inherent articulatory characteristics since they are functions rather than units, i.e., they are intentional (in the sense of Baudouin de

Courtenay, 1972 [1895]) and perceptual rather than actually articulatorily in nature [...] Inter-relationships between beats and pre-beat and post-beat consonants and consonant sequences are specified by a set of binding laws which look both at a “micro- level”—constituted by a single beat and consonants surrounding it, and at a “macro-level”—constituted by a sequence of beats with consonants interspersed between them, i.e., a level governed by rhythm. Binding laws operate on a phonological level according to the criterion of sonority. The latter is understood as a default intrinsic property of a phonological segment, i.e., belonging to the level of intention” (Dziubalska-Kořaczyk, 1996, p.59).

Hurch (1996) warned some important about accentuations: the assignation of segmental material of different words on accentual structures is not just prosodic or lexical, it is also sensitive to be used for grammatical functions. For instance, Tongan language uses stress for contrasts between definite and indefinite entities. Also, in basque spoken in Zegama (p. 79), stress represent differences between singular and plural nouns. Nevertheless, it exists accentuation preferences registered in typology and ontogenetic studies, with more cases of trochaic and dactylic feet, which belong to falling rhythm patterns, than iambic and anapaest feet (from rising patterns):

“A falling pattern cannot be divided into smaller units on the very same prosodic level, but a rising one can, i.e., a sequence of an accented syllable plus an unaccented syllable constitutes a single foot, but a sequence containing the reverse accentual pattern can consist of an anacrusis plus foot, and the remaining monosyllabic foot even allows a binary interpretation as one possible foot structure [...] In languages with a rising main accent foot the secondary accent foot is frequently falling; the reverse is rare [...] in a four syllable sequence like *New York City*, the sequence of a iambic plus a trochaic foot is remodelled as a sequence of two trochees, namely as *Néw York City* [...] In language acquisition falling foot structures seem to be preferred over rising feet [...] In English “thus we get /'nænə/ or /'bænə/, for *banana*, but never /bə'næn/, /'teto/ or /'peto/ for *potato*, but never /pə'tet/, /læktək/ or /'wæktuək /, but never /ə'læk/” (Hurch, 1996, pp.89-90).

Auer (1990, p.17) made an objection about the role of rhythm patterns in establishing different syntactic constituents: it seemed not to be appropriate the equivalence of phrase accentuation with the location of operators, modifiers or syntactic comments, which could include verbs. Moreover, he rejected the idea of a language with falling rhythm could change to rising, but the opposite was possible. However, Donegan and Stampe (2004) overcame this situation incorporated the relationship between heads and dependents elements, being the last one more similar to discursive focus. They presented Niger-Congo languages as proof of a language could change from falling pattern to rising and from nucleus-initial to nucleus-final, demonstrating the existence of cycling rhythm changes.

Furthermore, compared to Dziubalska-Kořaczyk (1996), in spite of Donegan and Stampe uses beats as equivalents to morae, they didn't dismiss syllables as prosodic domains, because these structures are affected by accentuation and isochronies in

Munda languages (e.g. Sora), with short vowels belong to stressed and unstressed constituents, as well as cases of vowel and consonant assimilations, and Mon-Khmer languages, with coda insertions, vowel lengthenings and dissimilations which facilitate creations of diphthongs:

“Proto-Austroasiatic had isochronous words of one or two syllables. The disyllable had a rising rhythm, like \**bə'lu* ‘thigh’ [...] In Mon-Khmer, a distinctively short vowel in the final syllable was kept short by inserting a glottal stop *bə'luʔ*, but otherwise could be merged with the corresponding long vowels *bə'lu:*. The short initial syllable invites vowel reduction or deletion, and the long final syllable invites diphthongization, as in Khmer *'pliw* ↓ ‘thigh’ [...] In Munda, the disyllable was given a falling rhythm, \**bəlu* ↘, fitting the final syllable into the beat by shortening its vowel (proto-Munda seems not to have vowel length distinctions), and giving the initial syllable a full though short vowel -often, by harmony, ‘*bulu*” (Donegan and Stampe, 2004, pp.20-21).

Jauregi (2008) endorsed the thesis of syllable as the main rhythmic unit in the interplay of structuration principles such as accentuations, sonority and sound qualities (pp. 346-347). The syllable gains relevance when it helps to understand the modulation of sound sequences and how those are perceived by their users.

Finally, Donegan and Nathan (2015) accounted for the rhythmic changes observed between Mon-Khmer y Munda languages are also found in Romance languages, specially from Latin to French, though with different complexity degrees: meanwhile the change of rhythm patterns among Austroasian languages had affected phonology levels, morphological types and syntactic orders, in the case of change from Latin to French, rhythm just had affected the stress patterns, isochronies and syllable types, resulting in classifications of languages with near-falling or near-rising patterns than languages with categorically falling or rising patterns:

“Changes in rhythm may include changes between falling and rising accent and changes in patterns of isochrony—shifts among so-called mora, syllable, and stress timing. For example, the Chamic languages changed in type (Thurgood, 1999); this branch of Austronesian has become atypically final-accented, and its phonology, morphology, and syntax have changed accordingly. In Old French, vowels diphthongized in stressed open syllables and reduced in unstressed syllables, suggesting a shift from the mora-timing of Latin toward stress-timed rhythm. A subsequent series of changes monophthongized the diphthongs and weakened or deleted syllable-final consonants (e.g. Pope, 1934, pp. 103, 190), leading to an open-syllable pattern that became the syllable-timing of modern French. Changes in rhythm can result in wholesale changes in the application of phonological processes because processes are sensitive to prosodic factors such as duration, accent, and syllabication” (Donegan and Nathan, 2005, p.448).

In summary, following theorists of Natural Phonology, rhythm is a set of perception and production patterns of different accentuations which operate on words and phrases. These could proceeded from prosodic processes or morphological rules, generating domains such as phonological words and

phrases, metrical feet, syllables and beats, starting points for the development of segmental (lenitive and fortitive) processes. There are two rhythmic patterns, which are the poles of a continuum: rising systems, whose prototype is identified in Mon-Khmer languages, with lexical morphemes and syntactic objects in final stressed positions, lexical tones, syllables with long vowels and stressed-timed isochronies; and falling systems, whose prototypes are Munda languages, with trochaic feet, preverbal syntactic objects and lexical morphemes in initial stressed positions, restrictions on tonogenesis, tendencies on syllables with short vowels and syllable-timed isochronies.

### 3. The coexistence of two concepts of rhythm: the cases of Embera and Spanish

There are two languages equivalents in terms of isochrony: Spanish and Embera. The first language is one of the most spoken in the world, being official in countries such as Colombia, Peru, Bolivia, Mexico, Argentina, etc. (Instituto Cervantes, 2019). Embera is language spoken only in Colombia and Panama by 70000 people (Barreña and Pérez-Cauarel, 2017). It belongs to Choco family with Wounaan language.

Respect to Spanish, based on acoustic measures of nPVI (normalized Pairwise Variability Index) and rPVI (raw Pairwise Variability Index), Grabe and Low (2002, p.528) asserts that this language exhibits low levels of these parameters, so that indicates the presence of syllable-timed isochrony. These results are found in the dialect spoken in Mexico (Carter, 2005), as well as in Spain (Prieto *et al.*, 2012), that which be inferred that differences among dialects are not meaningful.

Regarding Embera, Páez and Ibarra-Zetter (2018, p.200) carried out measurements of rPVI and nPVI with sound files of *Global Recordings* (GRN, 2017). Here they obtained that this language exhibits low values of both indices, signs of syllable-timed isochrony, the same as Spanish language.

Spanish and Embera have trochaic feet in words and tendencies to melodic cadence in statements. However, on syllabification, Embera doesn't have syllables with coda, only with CV or CVV sequences (Hoyos, 2000, p.77; Sáenz, 2018), meanwhile, for instance, in Spanish spoken in Boyaca, Colombia (Díaz, 2017), although there are cases of diphthongizations (*e.g.* boca → bueca 'mouth'), there are also syllables with coda (*e.g.* trans.por.ta.dor 'transporter').

Donegan (1993) sustained that the emergence of vowel phonemes, as consequence of suppressing some fortitive processes, is a feature of languages with rising pattern. There are nasalized vowels as phonemes (Hoyos, 2000, p.76), the main result of inhibiting the process of vowel denasalization (Donegan and Stampe, 2009). However, it doesn't occur in Spanish, where the inventory of five oral vowels is constant, regardless of dialect (Hualde, 2014). Lenitive process of vowel nasalization, case of assimilation, regular process in prototypical falling patterns, is active in Spanish and Embera. The same situation with consonant weakening of plossives, which is active in both languages.

In morphology, although there is low presence of synthesis (*i.e.* morphemes representing 1-2 properties) in Spanish and Embera nouns, the situation is different among verbs: meanwhile Spanish has suffixes with more than 2 properties, like -as, with the meaning of 2nd Person Singular, Present tense, indicative mood (*piensas* 'you think', *cantas* 'you sing', etc.), as reported by RAE and ASALE (2010), Embera

doesn't have this class of morphemes (Hoyos, 2000, pp.77-78).

Lastly, regarding syntactic constituents order, Embera presents object-verb (OV) order (Loewen, 1958, pp. 107-108), while Spanish exhibits verb-object (VO) as the most frequent type, even if the OV order appears in some cases of discourse focalization (Ocampo, 1995).

Both languages are falling rhythm pattern, but Spanish is nearer to the prototype than Embera. Meanwhile Spanish differs to Munda prototype in 2 properties, Embera has 4 differences with this austroasiatic language. Table 1 summarizes this.

Language	Spanish	Embera
Property		
Word Accentuation	Trochaic foot (mostly)	Trochaic foot (mostly)
Phrase Accentuation	Melodic cadence on statements	Melodic cadence on statements
Syllables with coda	Yes	<b>No</b>
Context-free fortitive vowel denasalization	Yes	<b>No</b>
Context-dependent lenitive vowel nasalization	Yes	Yes
Context-dependent lenitive consonant weakening	Active with all voiced plosives	Active with all plosives
Lexical tones	No	No
Level of morphological synthesis in nouns	<b>Low</b>	<b>Low</b>
Level of morphological synthesis in verbs	High	<b>Low</b>
Syntactic constituent order between object and verb	VO (mostly)/OV (some focalizations)	OV

**Table 1.** Comparison between Spanish and Embera in terms of Natural Phonology definition of rhythm. The differences respect to falling pattern prototype (e.g. Munda) are in **bold**.

#### 4. Research perspectives

On the one hand, there are a long tradition of rhythm studies in terms of isochrony, in spite of it exists new tendencies to the study in auditory perception (Goswami and Leong, 2013), whence it's necessary to develop more analyses. On the other hand, there are different studies on Natural Phonology since the last 50 years. For instance, some M.A. and Ph.D theses on descriptive Spanish Phonology (Díaz, 2017; Silvestre, 2020), studies of Spanish dialects in intercations (Troncoso, 2015) or languages in contact (Larrazá, 2014), explanation of some sound changes

(Salaberri and Salaberri, 2016), replacement of holistic structure of a language (Lakarra, 2006), even the standardization process of pronunciation (Oñederra, 2016) are developments using this theory.

Nevertheless, there are lots of theoretical and methodological issues for solving towards rhythm in Natural Phonology. Owing to processes are synchronic (Donegan and Nathan, 2015: 431), it could carry out a typological or dialectal studies for accomplishing classifications in terms of rhythmic-holistic systems based on the operation of prosodic, lenitive and fortitive processes, morphological and syntactic features. It could be studies that, using systematic parameters, allow the endorsement, clarification or refutation of some hypotheses of areal divisions, such as Amazonian vs. Andean languages in South America (Dixon and Aikhenvald, 1999, pp. 8-9).

## 5. References

Abercrombie, David, 1967. *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.

Álvarez, E., 1785. *Prosodia*. Cervera: La Real y Pontificia Universidad de Cervera.

Auer, P., 1990. A note on prosody in natural phonology. In: J. Méndez Dosuma, ed. *Naturalists at Krems: papers from the Workshop on Natural Phonology and Natural Morphology*. Salamanca: Ed. Univ. de Salamanca. Pp. 11-22.

Baudouin de Courtenay, J. (1972 [1895]). An Attempt at a Theory of Phonetic Alternations. In: E. Stankiewicz, ed. *A Baudouin de Courtenay Anthology. The Beginnings of Structural Linguistics*. Bloomington: Indiana University Press. Pp. 144-212

Barreña, A. and Pérez-Cauarel, M., 2017. La revitalización de la lengua embera en Colombia: de la oralidad a la escritura. *Onomázein*. Pp. 58-76.

Bjarkman, P., 1976. *Natural Phonology and Loanword Phonology (with Selected Examples from Miami Cuban Spanish)*. Ph.D. Thesis. Miami: University of Florida.

Carter, P. M., 2005. Quantifying rhythmic differences between Spanish, English, and Hispanic English. In: R. Gess and E. J. Rubin, eds. *Theoretical and Experimental Approaches to Romance Linguistics*. Amsterdam: John Benjamins. Pp. 63-75.

Chomsky, N. and Halle, M., 1968. *The Sound Pattern of English*. New York: Harper and Row, Publishers.

Dauer, Richard, 1983. Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics*, 11. pp. 51-62.

Díaz, Camilo, 2017. *El español hablado en Boyacá, Colombia: aspectos fonéticos y (morfo)fonológicos*. Ph.D. Thesis. Vitoria-Gasteiz: Universidad del País Vasco/Euskal Herriko Unibertsitatea.

Dixon, R. and Aikhenvald, A., 1999. Introduction. In: R. Dixon and A. Aikhenvald, eds. *The Amazonian Languages*. Cambridge: Cambridge University Press. Pp. 1-22.

Donegan, P., 1978. *On the natural phonology of vowels*. Ph.D. Thesis. Columbus: Ohio State University. Department of Linguistics.

Donegan, P., 1993. Rhythm and Vocalic Drift in Munda and Mon-Khmer. *Linguistics of the Tibeto-Burman Area*, 16(1). pp. 1-43.

Donegan, P. and Nathan, D., 2015. Natural Phonology and sound change. In: P. Honeybone and J. Salmons, eds. *The Oxford Handbook of Historical Phonology*. Oxford: Oxford University Press. Pp. 431-449.

Donegan, P. and Stampe, D., 1978. The syllable in phonological and prosodic structure. In: A. Bell and J. Bybee Hooper, eds. *Syllables and Segments*, Amsterdam: North-Holland Publishing Company. Pp. 25-34.

Donegan, P. and Stampe, D., 1979. The Study of Natural Phonology. In: D. Dinnsen, ed. *Current approaches to phonological theory*. Indiana: Indiana University Press. Pp. 126-173.

Donegan, P. and Stampe, D., 1983. Rhythm and the holistic organization of language structure. In: J. Richardson, M. Marks and A. Chukerman, eds. *Papers from the Parasession on the interplay of Phonology, Morphology and Syntax*. Chicago: Chicago Linguistic Society. Pp. 337-353.

Donegan, P. and Stampe, D., 2004. Rhythm and the synthetic drift of Munda. *The yearbook of South Asian languages and linguistics*, 7. Pp. 3-36.

Donegan, P. and Stampe, D., 2009. Hypotheses of natural phonology. *Poznań Studies in Contemporary Linguistics*, 45(1). Pp. 1-31.

Dressler, W. and Siptár, P., 1989. Towards a Natural Phonology of Hungarian. *Acta Linguistica Hungarica*, 39(1-4). Pp. 29-51.

Dziubalska-Kończak, K., 1996. Natural Phonology without the syllable. In: B. Hurch and R. Rhodes, eds. *Natural Phonology: The State of the Art*. Berlin: Mouton de Gruyter. Pp. 53-72.

Grabe, E., and Low, E. L., 2002. Durational variability in speech and the rhythm class hypothesis. *Papers in laboratory phonology*, 7. Pp. 515-546.

Gibbon, D. and Gut, U., 2001. Measuring speech rhythm. In *Seventh European Conference on Speech Communication and Technology*.

Goswami, U. and Leong, V., 2013. Speech rhythm and temporal structure: converging perspectives. *Laboratory Phonology*, 4(1). Pp. 67-92.

Global Recordings Network-GRN, 2017. Embera-Chami language. Online at: <http://globalrecordings.net/en/language/cmi>

Han, M., 1962. The feature of duration in Japanese. *Onsei no kenkyuu*, 10. Pp. 65-80.

Hoyos, M. E., 2000. Informe sobre la lengua embera del río Napipí. In: M. Rodríguez de Montes and M. González de Pérez, eds. *Lenguas indígenas de Colombia: una visión descriptiva*. Santa Fe de Bogotá: Instituto Caro y Cuervo. Pp. 73-84.

Hualde, J., 2014. *Los sonidos del español (with Sonia Colina)*. Cambridge: Cambridge University Press.

Hurch, B., 1988. Is Basque a syllable-timed language? *ASJU*, 22(3). Pp. 813-825.

Hurch, B., 1996. Accentuations. In: B. Hurch and R. Rhodes, eds. *Natural Phonology: The State of the Art*. Berlín: Mouton de Gruyter. Pp. 73-96.

Instituto Cervantes, 2019. *El español: una lengua viva. Informe 2019*. Madrid: Instituto Cervantes.

Jakobson, R., 1968 [1941]. *Child Language, Aphasia and Phonological Universals*. The Hague: Mouton Publishers.

Kiparsky, P. (1982). From Cyclic Phonology to Lexical Phonology. En H. van der Hulst y N. Smith (Eds.), *The structure of phonological representations. Vol. I* (pp. 131-175). Dordrecht: Foris.

Jauregi, O., 2008. Sobre la naturaleza de la sílaba y la estructuración de sus elementos en euskera. *ASJU*, 42(2). Pp. 337-352.



Lakarra, J., 2006. Protovasco, munda y otros: reconstrucción interna y tipología holística diacrónica. *Oihenart*, 21. Pp. 229-322.

Larrazza, S., 2014. *Listening to dialectal variation in a second-language: not just unavoidable but necessary*. Ph.D. Thesis. Vitoria-Gasteiz: Universidad del País Vasco/Euskal Herriko Unibertsitatea.

Loewen, J. A. 1958. *An introduction to Epera speech: Sambu dialect*. Ph.D. Thesis. Seattle: University of Washington.

Miller, P., 1972a. Some Context-Free Processes affecting vowels. *Ohio State University Working Papers in Linguistics*, 11. Pp. 136–167.

Miller, P., 1972b. Vowel Neutralization and Vowel Reduction. *Papers from the Eighth Regional Meeting of the Chicago Linguistic Society*. Pp. 482-489.

Monboddo, J., 1787. *Of the Origin and Progress of Language. Vol IV*. Edinburgh and London: J. Bell and T. Cadell.

Monboddo, J., 1789. *Of the Origin and Progress of Language. Vol V*. Edinburgh and London: J. Bell and T. Cadell.

Morris, A., 1925. *The Orchestration of the Metrical Line. An Analytical Study of Rhythmical Form*. Boston: Bruce Humphries, Inc.

Nathan, G., 1985. Towards a Natural Phonology of juncture. In: W. Dressler and L. Tonelli, eds. *Natural Phonology from Eisenstadt*. Padua: CLESP. Pp. 101-110.

Nespor, M., Shukla, M. and Mehler, J., 2011. Stress-timed vs. syllable-timed languages. *The Blackwell Companion to Phonology*. Pp. 1147-1159.

Ocampo, F., 1995. The word order of constructions with a verb, a subject, and a direct object in spoken Spanish. In: J. Amastae, G. Goodall, M. Montalbetti and M. Phinney, eds. *Contemporary Research in Romance Linguistics*. Amsterdam: John Benjamins. Pp. 291-307.

Omond, T., 1905. *Metrical Rhythm: Being an examination of a recent attempt to determine the basis of English rhythm in verse and prose*. Tunbridge Wells: R. Pelton.

Oñederra, M., 2016. Standardisation of Basque: from grammar (1968) to pronunciation (1998). *Sociolinguistica*, 30(1). Pp. 125-144.

Páez, D. and Ibarra-Zetter, K., 2018. Cronometría del ritmo en cuatro lenguas colombianas: emberá-chamí, kamsá, kogui, y wayúu. *Forma y Función*, 31. Pp. 183-205.

Pike, K., 1945. *The Intonation of American English*. Ann Arbor: The University of Michigan Press.

Pope, M. K. (1934). *From Latin to Modern French with especial consideration of Anglo-Norman: Phonology and morphology*. Manchester: Manchester University Press.

Prieto, P., del Mar Vanrell, M., Astruc, L., Payne, E., and Post, B., 2012. Phonotactic and phrasal properties of speech rhythm. Evidence from Catalan, English, and Spanish. *Speech Communication*, 54(6). Pp. 681-702.

Ramus, F., Nespor, M. and Mehler, J., 1999. Correlates of linguistic rhythm in the speech signal. *Cognition*, 73. pp. 265-292.

Real Academia Española-RAE and Asociación de Academias de la Lengua Española-ASALE, 2010. *Manual de la nueva gramática de la lengua española*. Madrid: Espasa.

Salaberri, P. and Salaberri, I., 2016. An introduction to Basque aspiration: the

contribution of onomastics. *Fontes linguae vasconum: Studia et documenta*, 48 (122). Pp. 365-392.

Sáenz, Maiced, 2018. *Grabaciones de encuestas y resultados de pruebas auditivas sobre oclusivas de adolescentes bilingües-español y monolingües en español* (files in formats .wav and .txt). Bogotá, D.C.: Instituto Caro y Cuervo.

Sapir, E., 1921. *Language. An introduction to the study of speech*. New York: Harcourt, Brace and Company.

Silvestre, R., 2020. *Un estudio fonético-fonológico sobre las vocalizaciones en el español hablado en Pedro García, República Dominicana*. M.A. Thesis. Vitoria-Gasteiz: Universidad del País Vasco/Euskal Herriko Unibertsitatea.

Smith, R. E., 1980. *Natural Phonology of Japanese*. Ph.D. Thesis. Manoa: University of Hawaii, Manoa.

Stampe, D., 1968. "Yes, Virginia, ...". *Document presented at 4th regional meeting of the Chicago Linguistics Society*. Chicago: Chicago Linguistics Society.

Stampe, D., 1969. The Acquisition of Phonetic Representation. In: R. Binnick, A. Davison, G. Green and J. Morgan, eds., *Papers from the Fifth Regional Meeting of the Chicago Linguistic Society*. Chicago: University of Chicago Department of Linguistics. Pp. 443-454.

Steele, J., 1775. *An essay towards establishing the melody and measure of speech to be expressed and perpetuated by peculiar symbols*. London: W. Bowyer and J. Nichols.

Thomson, W., 1904. *The basis of English rhythm*. Glasgow: W. and R. Holmes.

Thurgood, G., 1999. *From ancient Cham to modern dialects: two thousand years of language contact and change*. Manoa: University of Hawaii press.

Troncoso, M. A., 2015. *Prosodic and Segmental Accommodation: Andalusian Spanish in Asturias*. M.A. Thesis. Vitoria-Gasteiz: Universidad del País Vasco/Euskal Herriko Unibertsitatea.