Espinosa, G. On the uncertain nature of speech rhythm and its relevance in L2 acquisition. *Argentinian Journal of Applied Linguistics.* Vol 4, No 1, May 2016, pp. 5-22. ISSN 2314-3576 (online).

On the uncertain nature of speech rhythm and its relevance in L2 acquisition

Abstract

This article is aimed at reviewing the different accounts for speech rhythm in the languages of the world. It addresses three main aspects in the literature on rhythm based on an impressionistic description, on a phonological explanation, and on the implementation of metrics to measure speech quantitatively. Some relevant studies on L1 speech rhythm are discussed in order to see the different attempts to categorise languages and to describe linguistic variations. In addition, the comparison of Germanic and Romance rhythm leads to the analysis of non-native speech with implications for the teaching and learning of L2 prosody, specifically in the acquisition of English rhythm by Spanish learners.

Keywords: speech rhythm, L1, L2, English, Spanish

Resumen

El objetivo de este artículo es dar cuenta de las distintas descripciones del ritmo del habla en las lenguas del mundo. Se abordan tres aspectos principales en la literatura sobre ritmo basados en una descripción impresionista, una explicación fonológica, y la implementación de métricas para medir el habla cuantitativamente. Se discuten algunos estudios relevantes sobre el ritmo en L1 con el fin de conocer los distintos intentos de hacer una categorización de las lenguas y de describir la variación lingüística. Asimismo, la comparación del ritmo germánico y románico lleva al análisis del habla no-nativa con implicancias al momento de enseñar y aprender la prosodia de una L2, específicamente en la adquisición del ritmo inglés por hablantes nativos de español.

Palabras clave: ritmo del habla, L1, L2, inglés, español

Generally speaking, rhythm is part of basic human activities like walking, dancing and breathing because of the repeated events that occur periodically in time. This general notion of rhythm has been applied to speech rhythm (or just *rhythm* henceforth) because there are some elements in oral speech –syllables, stresses, etc.– that tend to be produced or perceived as similar. These elements encompass aspects that go beyond vowels and consonants and are related to the way individual sounds are organized in the flow of speech. The overall term used to describe this multi-layered organization is *prosody*. Due to the prevailing lack of consensus over interpretations and definitions of rhythm, and the inconsistent results obtained in speech production and perception, speech rhythm has been one of the most controversial domains in prosodic studies.

The assumption of periodicity associated with the organization of oral speech has led to a categorical rhythmic distinction among natural languages: languages can either be stress-timed or syllabled-timed (Pike, 1945; Abercrombie, 1967). After several unsuccessful attempts to find substantial empirical evidence for discrete rhythm classes, two lines of investigation developed. Firstly, speech rhythm started to be described as a product of phonetic and phonological properties. With this description, the languages of the world are thus classified on a rhythm continuum but not under categorical distinctions (e.g. Dauer, 1983; Bertinetto, 1989; Schiering, 2007). This is supported by a large number of recent studies that have demonstrated the absence of distinct rhythm classes (e.g. Kohler, 2009; White, Mattys & Wigit, 2012; Payne, Post, Astruc, Prieto & del Mar Varnell, 2012). Secondly, the thorough search for empirical evidence on the nature of speech rhythm has motivated researchers to make use of rhythm metrics by computing vocalic and consonantal intervals in order to measure rhythm in first languages (e.g. Ramus, Nespor & Mehler, 1999; Grabe & Low, 2002; Fuchs, 2014), and second and foreign languages (e.g. White & Mattys, 2007; Ordin & Polyanskaya, 2014, 2015; Gabriel & Kireva, 2014; Kireva & Gabriel, 2015).

While most of the literature on rhythm is concerned with the efforts to see if speech rhythm can be categorically classified and if its nature can be attested on an empirical basis, there is a need to reflect on the relevance of this aspect of the language in the process of teaching and learning a second or foreign language[1] (L2) in terms of its communicative function, in the context of, for instance, the acquisition[2] of the Germanic rhythm (e.g. English) when the first language (L1) has a Romance rhythm type (e.g. Spanish). The way second language learners (L2ers) adapt their prosody to the target language has introduced a rich field of research that aims at enlightening the nature of languages and the importance of prosodically trained non-native speakers (e.g. Ordin & Polyanskaya, 2014; Gabriel, Stahnke & Thulke, 2015).

In this article, first the hypothesis of isochrony in rhythm and the phonological approach will be briefly discussed. After presenting rhythm metrics, together with their drawbacks, advantages and their current implementation in research, studies on second languages and language contact will be examined in order to address some implications for the L2 field.

The Isochrony Hypothesis

Lloyd James (1940) employed the metaphor of ‘Morse code’ to refer to languages like English where the stress occurred at regular intervals of time, while he used the metaphor of ‘machine-gun’ to describe languages like Spanish where the syllable was the element in charge of the temporal organization. Pike (1945) and Abercrombie (1967) developed these ideas and suggested a rhythm dichotomy among languages because of the different elements that organize the rhythmic patterns of speech and the isochrony associated with them, i.e. the equal division of intervals of time in a spoken language.

The underlying idea of rhythm in Pike (1945) and Abercrombie (1967) is that it is a structural property of languages. Germanic languages like English and Dutch are classified as stress-timed because it is the stress that seems to be isochronous. In other words, stresses occur at regular periods of time in oral speech and they remain constant, irrespective of the number of syllables. On the other hand, Romance languages like Spanish and French are defined as syllable-timed because the element that occurs regularly in similar intervals of time is the syllable. The duration of the syllables stays essentially constant in the flow of speech and they are responsible for the prosodic organization. Authors like Ladefoged (1975) propose a third type of rhythm in which the rhythmic unit is the *mora*, the unit that builds up the weight of the syllable. For example, a light syllable has one mora and a heavy one has two morae. Japanese is the language that has been generally described as mora-timed (see Warner & Arai, 2001, for an overview on Japanese mora-timing).

With the isochrony hypothesis, the languages of the world are spoken with only one rhythm type, and once the isochronous element is identified, it is possible to categorically classify languages (Abercrombie, 1967). Even though the idea of the syllable, stress or mora as the elements that organize speech rhythm has been constantly challenged in contemporary studies, the traditional terms *stress-*, *syllable-* and *mora-timed* are still employed in the recent literature on rhythm in the attempt to identify the prosodic cues that make languages sound different.

The Phonological Approach to Rhythm

Since Lloyd James (1940), research has examined the different methods and techniques that could confirm the impressionistic differences in rhythm types. However, linguists have abandoned the idea of isochrony and the taxonomy that classifies languages in clear-cut rhythm types, mainly due to the failure of a large number of experiments

trying to corroborate these rhythm classes empirically (e.g. Dauer, 1983; Roach, 1982). Dauer (1983) argues that the rhythmic differences between languages like English and Spanish are not exclusively related to equal intervals of time in speech, but to phonetic, lexical, syntactic and phonological properties of each language. Dauer (1983) and Bertinetto (1989) challenge the idea of rhythm as a structural linguistic property and claim that rhythm is the product or the consequence of the phonological properties of a given language. Some of the most important properties are vowel reduction and syllable structure. Stress-timed languages like English have reduced vowels in unstressed positions –exemplified with the recurrent use of schwa–, a great variety of syllable types and a complex clustering of segments within a syllable. On the other hand, syllable-timed languages like Spanish have the same vowels in unstressed positions and a fairly simple syllable structure.

The new conception of speech rhythm, developed after the 80s, regards rhythm as a linguistic continuum where languages are no longer classified as strictly syllable- or stress-timed, but their timing in speech can be based more on the syllable or on the stress. Consequently, the original dichotomy for Germanic and Romance languages was abandoned and re-interpreted. Now, on the one hand, there is an ideal language which possesses all the phonetic and phonological characteristics of a purely syllable-timed language; on the other, there is another hypothetical language with all the traits of an ideal stress-timed one. In theory, the languages of the world could be placed on this continuum according to the phonological characteristics they show (Bertinetto, 1989; Schiering, 2007) and speech rhythm is based on flexible language-specific patterns (Kohler, 2009).

Rhythm Metrics

A brief historical overview. After the definition of rhythm as the product of phonetic and phonological properties, a battery of attempts have been made in order to measure these properties by means of rhythm metrics. The underlying idea is to identify the acoustic correlates of different rhythm types given that "the syllable-timing/stress-timing dichotomy may well be deeply anchored in the human perceptual system" (Ramus et al., 1999, p. 267). With the aim of comparing durational intervals based on vocalic and consonantal measurements, rhythm metrics have become massively popular as a way of empirically attesting prosodic differences among traditionally defined rhythm classes. They are supported by studies showing that infants, including newborns, can discriminate between languages that are supposed to belong to different rhythm types but not between languages of the same type (e.g. Nazzi & Ramus, 2003; Payne et al., 2012). This means that languages may have some inherent prosodic features that even newborns can perceive.

Ramus et al. (1999) are the first to propose the Deltas, which are metrics that account for the timing of rhythm types. Specifically, in their influential paper they study syllable structure and vowel reduction with three ways to measure them: the metric %V measures the proportion of vocalic intervals, ΔV measures the standard deviation of the duration of vocalic intervals, and ΔC measures the standard deviation of the duration of consonantal intervals. Basically, the purpose of analysing the vocalic and consonantal intervals –considered as the duration of each segment or a cluster of segments– is to corroborate the idea that stress-timed languages like English, German and Dutch show a greater variability in these intervals than in syllable-timed languages like Spanish, French and Italian; this is due to the predominant syllable complexity and vowel reduction in Germanic languages. These authors also give an account for mora-timed languages like Japanese and for languages with mixed rhythm like Polish. They are able

to demonstrate some correlations between metrics and traditional rhythm types and argue that metrics are a useful tool to quantify distinctions in rhythm.

Low, Grabe and Nolan (2000) propose a parallel approach called the Pairwise Variability Index (PVI), which measures the duration of vocalic and consonantal intervals, considering each segment in a temporal succession and calculating the mean of the differences. Grabe and Low (2002) used the PVI metrics in order to analyse 18 languages. They are able to classify only some languages into the traditional rhythm types and they show that in some other languages rhythm types overlap, which is why they do not speak of a categorical distinction of rhythm classes. With the PVIs it is possible to measure syllable complexities in terms of their consonantal clusters and also to quantify vowel reduction. Some modifications of PVI are also meant to neutralise the effect of speech rate; this is why a normalised PVI metric was proposed, where the durations in pairs are divided by the mean duration. Consequently, the raw PVI metric (rPVI) does not control for speech rate, while the normalised version (nPVI) reduces its effects.

Some more criticism has centred on the sensitivity of the deltas to speech rate (Barry, Andreeva, Russo, Dimitrova & Kostadinova, 2003; Dellwo & Wagner, 2003). A negative consequence of the influence of rate is that supposedly stress-timed languages like English could be classified as syllable-timed due to an increase in the speech rate or tempo. In order to adjust metrics and control for this influence, Dellwo (2006) proposed the Varcos: new versions of metrics that neutralise the effect of rate in order to obtain a better discrimination of languages. VarcoC is the standard deviation of consonantal intervals divided by the mean consonantal duration of an utterance, whereas VarcoV stands for the standard deviation of vocalic intervals divided by the mean vocalic duration within an utterance (see Benton, 2010, for an overview on the differences in metrics).

Bertinetto and Bertini (2008) modify the PVIs and offer a new model to distinguish the rhythmic tendencies of languages. This new tool for measuring the rhythmic behaviour is termed Control-Compensation Index (CCI) and, basically, it divides the duration of intervals by the number of segments in them. Controlling languages like Spanish seem to show minor fluctuations in intra and inter-syllabic durations, whereas compensating languages like English exhibit higher fluctuations. Additionally, more metrics have been proposed in an attempt to adjust the values of metrics and to better describe speech rhythm (e.g. Arvaniti, 2012, for a brief description of more metrics).

Some drawbacks of metrics and their current usefulness. Despite the relative success of rhythm metrics in accounting for differences in diverse rhythm classes, there is still a lack of empirical evidence in order to claim that metrics are truly useful, mainly due to methodological shortcomings (e.g. Arvaniti, 2012; Loukina, Kochanski, Rosner, Keane & Shih, 2011, 2013). Loukina et al. (2011) argue that each rhythm metric captures different aspects of the language and that there is substantial variation in the values of each metric for each language. They claim that the way in which languages can be classified depends on the type of metrics being used, casting doubt on the validity of measurements to determine linguistic properties. Arvaniti (2012) gives evidence against the validity of metrics to classify languages in well-defined rhythm classes. Her results show that metrics are largely related to inter-speaker variation, elicitation methods and the syllable composition of materials, which prevents researches from accounting for rhythm classes on an empirical basis. The results obtained by Loukina et al. (2013) suggest that rhythm metrics cannot reflect differences in syllable structure and vowel reduction. These authors claim that the values from metrics are highly influenced by the type of text and the speaker; this is why they suggest that it is necessary to design multi-level models in order to include these variants.

There have been more studies that undermine metrics as a reflex of, mainly, syllable structure and vowel reduction as the heart of rhythm classes. For instance, Barry, Bistra and Koreman (2009) address the relationship between rhythmic impressions and the physical reality that metrics can capture. They claim that rhythm metrics cannot entirely account for linguistic rhythm because there are other parameters, like the fundamental frequency (F0 or melodic change), which contribute to the perception of rhythm. This is in accordance with Prieto, del Mar Vanrell, Astruc, Payne, and Post (2012), where their results indicate that the perception of rhythm is comparatively independent of syllable structure and vowel reduction. In addition, Arvaniti and Rodriguez (2013) argue that language discrimination can be based on speaking rate or F0 differences and not completely on timing.

Even though some studies have explored the weak points of the measure of consonant and vocalic intervals, rhythm metrics have led to a growing body of literature regarding the distinction between different language varieties, which overweighs the drawbacks pointed out by several researchers. In short, lower values of V% and higher values of ΔV, ΔC, rPVI, nPVI, VarcoV and VarcoC are associated with a stress-timed tendency and the opposite values are related to a more syllable-timed rhythm. The results obtained with the use of metrics have helped linguists to quantify different phonetic and phonological aspects such as vowel reduction, syllable structure, syllable complexity and phonotactic constraints (Ramus et al., 1999; Carter, 2005; Dellwo & Wagner, 2003; Dellwo, 2006; White & Mattys, 2007; Mairano & Romano, 2007; Mairano, 2011, among others).

Some research has demonstrated evidence in favour of metrics based on vocalic durations (for example: %V and VarcoV) because these values seem to discriminate between languages in a more accurate fashion than consonantal durations (e.g. White & Mattys, 2007; Russo & Barry, 2008; Knight, 2011; Kireva & Gabriel, 2015). Several

other studies attempt to adjust the way metrics measure speech. For instance, Tilsen and Arvaniti (2013) go beyond temporal durations and show a novel type of metrics to quantify speech rhythm based on syllabic and supra-syllabic information. They refer to these new metrics as *envelope* metrics, which are able to capture information about periodicities of syllables, feet and phrases, and can reflect inter-linguistic differences in rhythmicity. Dellwo and Fourcin (2013) propose an analysis based on voiced and unvoiced intervals, rather than vocalic and consonantal intervals, and suggest that voice is a major contributor to the rhythmic discrimination of languages. Fuchs (2014) measures the influence of F0 in the perception of the duration of intervals and suggests a modified version of PVI: nPVI-V(dur*f0). He claims that it is necessary to include the levels of F0 and to develop a multi-dimensional model to measure rhythm.

Rhythm metrics are continuously being used to describe different languages and to account for language acquisition. Payne et al. (2012) analyse the speech of children acquiring English, Catalan and Spanish as their mother tongue, and they claim that, although there are no clear-cut types of rhythm, it is possible to identify language-specific rhythmic indices. White et al. (2012) found that speakers do not distinguish languages on distinct rhythm classes but on temporal cues coming from speech rate, durations in consonantal and vocalic intervals, and lengthening of utterance finals. This is supported by Brown and Matene (2014) who give evidence of rhythm as a byproduct of phonological properties rather than an inherent property of language. Prieto et al. (2012) argue that the acoustic correlates of timing are gradual, whereas the perception of them could be categorical. All in all, rhythm metrics can be useful to measure the temporal differences in languages, and they are still conceived of as good global measures that can give a general characteristic of linguistic rhythm (Mairano, 2014). Moreover, there are many attempts to control for inter- and intra-speaker variations (e.g. Dellwo, Leemann & Kolly, 2015), and to examine social factors (e.g. Ayed, Hamdani-

Droua, Alotaibi & Selouani, 2013) and dialectal differences of the same language (e.g. Leemann, Dellwo, Kolly & Schmid, 2014; Clopper & Smiljanic, 2015). Finally, it is worth mentioning that the technological development in computational software has greatly facilitated the use of rhythm metrics: Praat (Boersma & Weenink, 2015) and Correlatore (Mairano & Romano, 2010) are the most extensively used freely downloadable programs.

Measuring L2s and Language Contact

After the proposal of Ramus et al. (1999) and the proliferation of metrics, linguists also began to search for qualitative evidence in the speech rhythm of second and foreign languages. Some results diminish the usefulness of metrics by suggesting that their values are not reliable and cannot adequately account for L2 rhythm (e.g. Ferjan, Ross & Arvaniti, 2008) or by suggesting that there might not be such a thing as rhythm but mainly some abstract notions of time that influence the duration of segments and that L2ers have to acquire (Gut, 2012). However, a great deal of research has indicated that the use of metrics can yield several encouraging results. Some studies have focused on language contact and have offered sociological accounts related to social prestige and migration. For example, Carter (2005) demonstrates that the PVI rhythmic measures can quantify the degree in which Spanish speaking immigrants in the USA have possibly accommodated their prosody to a more English-like speech. Torgersen and Szakay (2012) employ PVIs to explore changes in rhythmic patterns in London English due to a cosmopolitan mix.

White and Mattys (2007) make use of metrics to measure some languages, among them English and Spanish as an L1 and L2. They claim that VarcoV and %V are the most adequate metrics which can discriminate L1 productions into stress- or syllable-timed languages. They also show that these metrics can yield halfway results

that calculate the possible prosodic adaptation by L2 speakers. Along these lines, Tortel and Hirst (2010) claim that metrics are useful to distinguish native from non-native prosody and also different levels of non-native speech. This is supported by studies like Kolly and Dellwo (2014), in which temporal intervals are analysed as important cues for the perception of foreign accent. Ordin and Polyanskaya (2014) present a longitudinal study in which they show the development of rhythm by comparing different ages in English as an L1 and different proficiency levels in L2 English. They make use of different combinations of nPVIs, ΔV and ΔC, %V, Varcos, and the mean duration of V and C intervals, in order to show the influence of age and proficiency level in rhythm changes. Ordin & Polyanskaya (2015) carried out another experiment with rhythm metrics aimed at analysing the development in timing patterns of German learners of English. They suggest that even though these languages are rhythmically similar, lower proficiency learners of English sometimes show a syllable-timed rhythm, whereas advanced learners' rhythm becomes more stress-timed.

Gabriel and Kireva (2014) and Kireva and Gabriel (2015) make use of some metrics to analyse the rhythm of Castilian Spanish, L2 Spanish (by Italian native speakers), Porteño Spanish and Italian. They provide empirical evidence of the usefulness of these metrics and suggest that %V and Vocalic-nPVI can adequately distinguish the languages studied. They claim that L2 Spanish, Porteño and Italian cluster together, which is evidence in favour of the hypothesis that Italian immigrants transferred some rhythmic properties of their L1 into the Spanish spoken in Buenos Aires (Benet, Gabriel, Kireva & Pešková, 2012); this new type of Spanish was later acquired as an L1 and became the current Porteño.

There are studies that do not implement rhythm metrics fully but they aim to empirically attest rhythm in L2s in an effort to reveal the nature of speech rhythm by analysing the speech of infants and their rhythmic development (Campfield & Murphy,

2014; Vihman, 2014; Schmidt & Post, 2015a,b, for simultaneous Spanish-English bilinguals; Molnar, Gervain & Carreiras, 2014), the discrimination of native and non-native speech (Li & Post, 2014, for L2 English, L1 Mandarin and L1 German; Gu & Hirose, 2014, for Mandarin; Selouani, Alotaibi, Cichocki, Gharsellaoui & Kadi, 2015, for Arabic), and the different roles and dimensions of speech rhythm (Roncaglia-Denissen, Schmidt-Kassow, Heine & Kotz, 2015, for syntactic ambiguity; Fuchs, 2015, for dialectal discrimination; Fuchs & Wunder, 2015, for the learnability of different dimensions of rhythm). In general, more research is required to see how the different languages of the world organize their "beats" in oral speech and to identify the acoustic correlates of rhythm. In particular, further research is necessary to better understand the aspects of prosody that learners have to acquire and to see the ways in which speech rhythm in an L2 can be quantitatively described in natural and academic contexts, considering different ages and linguistic backgrounds. This is a developing area of research that can yield encouraging results of speech rhythm, especially when languages differ in traditionally defined rhythm classes; for instance, learners of English, Russian, Arabic, Dutch, German or Thai (supposedly stress-timed) whose mother tongue is Spanish, Greek, French or Italian (supposedly syllable-timed).

Implications for L2 Teaching and Learning

When learners acquire an L2, they have to become acquainted with a new grammar, vocabulary and pronunciation. In most cases, pronunciation is the aspect of the language that is assumed to be acquired without much explicit instruction, or, if there is conscious work by the learner, it is generally related to individual sounds. However, there are prosodic aspects that are necessary for effective oral communication. On the one hand linguists continue to analyse the way learners accommodate their prosody when acquiring a second language. On the other, L2 teachers can easily perceive that once the

grammar and the individual sounds of an L2 are learnt or acquired, some adjustments in suprasegmentals have to be made so as to achieve more efficient communication. There is ongoing research that aims at studying the prosodic transfer in L2 learning in an academic setting (e.g. Trouvain & Gut, 2007; White & Mattys, 2007; Gabriel et al., 2015) and the way this transfer can be measured and addressed in the L2 classroom (e.g. Ordin & Polyanskaya, 2014).

The gap between research and practice should be bridged by the collaborative work of scientists and language instructors around the world. While academic research findings are necessary to understand the processes of acquiring L2 prosody, each L2 lesson can serve as a substantial body of evidence coming from learners' strengths and weaknesses. Some promising areas of research are related to the relationship between prosody and functions of the language, the usefulness of a top-down approach in the classroom –when the focus is first on prosody and then on individual sounds–, the role of formal instruction, material design, prosodic intelligibility, and identity issues, among others. As there are few studies based on longitudinal research, it would be of enormous importance if language professionals in any linguistic context, either in instructional or natural settings, measure speech rhythm at different stages in the acquisition process and keep a record of the techniques and courses of action aimed at helping learners with the L2 prosody. This would produce significant results for both academic research and classroom practice.

English as an L2 in Argentina is an interesting case with international implications. In this country, Spanish native speakers learn English as a foreign language mainly through formal instruction with non-native teachers whose first language is also Spanish. Learners have to depart from a syllable-timed language to a more stress-timed one, i.e. they have to adapt their Spanish speech rhythm to English prosody. How is this prosodic adaptation superficially evident? Learners very often

simplify consonantal clusters by eliding consonants or by adding an epenthetic vowel so as to get simpler syllables (e.g. Estebas Vilaplana, 2009); they overstress by having too many prominent syllables in their utterances and they understress by not making a clear difference between stressed and unstressed syllables (e.g. Ordin & Polyanskaya, 2014).

In order to help learners to adapt their prosody to the target language, some of the characteristics of the prosody of the L2 should be instructed in the classroom (e.g. Gilbert, 2008; Campfield & Murphy, 2013, for an analysis of prosodically rich L2 input; Ordin & Polyanskaya, 2014) so that learners can improve their speech production in the target language and consolidate their listening comprehension skills. As Gabriel et al. state, "phonological awareness should be promoted in scholar education, with regards to both the learners and the teachers" (2015, p. 214). Perhaps for Spanish learners of English, and for English learners with a rhythmically similar L1 like French or Italian, the two most important aspects to bear in mind are the emphasis on the complexity of syllable structure and vowel reduction. Learners' attention should be directed to the importance of pronouncing complex consonantal clusters and they should understand the way in which function words like auxiliaries, pronouns, prepositions, articles and affixes are generally non prominent and build up the rhythmic patterns of English (Gilbert, 2008). Spanish speakers of English generally expect to say and hear these words in the same way as content words because Spanish speakers make lexically-unstressed syllables prominent for rhythmic purposes (Hualde, 2012), i.e. in Spanish, the difference between strong and weak syllables is not as noticeable as in English (Hualde, 2014). One way of addressing this difference in the classroom is by having students listen to an input that is as English-like as possible. L2 teachers should not be afraid of compressing syllables and of making a clear contrast between strong and weak forms given that 'most English learners who suffer from inadequate training in listening comprehension complain that "native speakers talk too fast"' (Gilbert, 2008,

p. 6). Pronouncing syllables alike in order to help our students understand teachers better may help them at very early stages of the L2 learning, but more English-like prosody in the classroom –i.e. stress-timed speech– will most probably help them to become better listeners and speakers not only in the classroom but also in encounters with native speakers or non-native speakers with other L1s.

In the L2 lessons there should be activities designed to develop the phonetic and phonological characteristics that contribute to the general rhythm of a language. It is important that learners, especially the ones with an advanced command of the target language, perceive the typical processes of connected speech –elision, liaison, assimilation, compression, etc.– which help speakers to maintain rhythm (Ordin & Polyanskaya, 2014). L2 teachers should help students to be "perceptually sensitive to prosodic (including rhythm) characteristics of the target language" so that they "can develop an internal sense of rhythm in the non-native language, and thus be able to communicate effectively" (Erickson, 2013, p. 156). Furthermore, new technologies have evolved with the aim of improving prosodic features like rhythm in L2 acquisition (e.g. Chun, Hardison & Pennington, 2008, for an overview of prosody in context in L2 research and practice; Hincks, 2015, for a recent description on how to use technology to teach and evaluate pronunciation). If rhythm has a communicative guiding function (Kohler, 2009), this aspect of the target language should not be undervalued.

Conclusions

Rhythm as a linguistic phenomenon has been mainly analysed in terms of its impressionistic classification, its phonological characteristics and its acoustic correlates measured by means of metrics. The research on speech rhythm has centred on language classification, the acquisition of L1 and L2, dialectal variations, and perception. In the recent literature, the idea of isochrony has been abandoned, given that it is not a natural

characteristic of spoken languages. However, the traditional terms *stress-*, *syllable* and *mora-timed* are currently used to distinguish languages in a more flexible fashion. Concerning the search for empirical evidence, rhythm metrics are being employed to quantitatively account for the temporal organization of languages. Some authors do not rely entirely on metrics, but some others believe that with some adjustment in the methodology adopted, it is possible to describe rhythm variations in L1 and L2.

English and Spanish prosodies are unquestionably different in relation to the way these languages organize different levels of pronunciation. In the L2 classrooms, notions of rhythm cannot be left unattended since the organization of speech plays a crucial role in communication. Helping learners to perceive and adapt their prosody to the target language will most probably help them to become more efficient speakers and listeners, especially with practice on the phonetic and phonological characteristics of connected speech. Scientists continue to see whether rhythm is an inherent linguistic property or an epiphenomenon emerging from the phonetic and phonological structure of a language, and they keep on adjusting metrics to empirically measure rhythm and see if languages can be classified into distinct prosodic patterns. While researchers try to reveal the nature of speech rhythm, L2 teachers should devote some time to working on this aspect of prosody given that it plays a significant role in the intelligibility of L2ers.

References

Abercrombie, D. (1967). *Elements of General Phonetics*, Edinburgh University Press.

Arvaniti, A. (2012). The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics*, *40*, 351–373.

Arvaniti, A., & Rodriguez, T. (2013). The role of rhythm class, speaking rate and F0 in language discrimination. *Laboratory Phonology ,4*, 7–38.

Ayed, A.B., Hamdani-Droua, G., Alotaibi, Y.A., & Selouani, S.A. (2013). On the relevance of speech rhythm metrics to characterize social factors. In *GCC Conference and Exhibition (GCC), 2013 7th IEEE*, 606–610. DOI: 10.1109/IEEEGCC.2013.6705848.

Barry, W. J., Andreeva, B., Russo, M., Dimitrova, S., & Kostadinova, T. (2003). *Do rhythm measures tell us anything about language type?* In *Proceedings of the 15th ICPhS* (pp. 2693–2696). Barcelona.

Barry, W., Bistra, A., & Koreman, J. (2009). Do rhythm measures reflect perceived rhythm? *Phonetica* 66, (1–2), 78–94.

Benet, A., Gabriel, C., Kireva, E., & Pešková, A. (2012). Prosodic transfer from Italian to Spanish: Rhythmic properties of L2 speech and Argentinean Porteño . In Q. Ma, H. Ding & D. Hirst (Eds.), *Proceedings of the 6th International Conference on Speech Prosody* (pp. 438 –441). Shanghai, China: Tongji University Press .

Benton, M. (2010). A Preliminary Analysis of the Relationship of Speech Rate to Speech-Timing Metrics as applied to Large Corpora of Non-Laboratory Speech in English and Chinese Broadcast News. In M. Hasegawa-Johnson (Eds.), *Proceedings of Speech Prosody 2010*, Chicago.

Bertinetto, P. M. (1989). Reflections on the Dichotomy ‘Stress’ vs. ‘Syllable-timing’. *Revue de Phonétique Appliquée*, *91–93*, 99–130.

Bertinetto, P.M., & Bertini, C. (2008). On modeling the rhythm of natural languages. In P.A. Barbosa, S. Madureira & C. Reis (Eds.), *Proceedings of Speech Prosody* (pp. 427–430). Campinas, Brazil.

Boersma, P., & Weenink, D. (2015). Praat: doing phonetics by computer [Computer program]. Version 5.4.08, retrieved 24 March 2015 from http://www.praat.org/

Brown, J., & Matene, E. (2014). Is speech rhythm an intrinsic property of language?, *Interspeech 2014* (pp. 1693–1697). Singapore.

Campfield, D., & Murphy, V. (2013). The influence of prosodic input in the second language classroom: does it stimulate child acquisition of word order and function words?, *The Language Learning Journal.* DOI:10.1080/09571736.2013.807864.

Campfield, D., & Murphy, V. (2014). Elicited imitation in search of the influence of linguistic rhythm on child L2 acquisition. *System, 42*, 207–219.

Carter, P. M. (2005). Quantifying rhythmic differences between Spanish, English, and Hispanic English. In R. S. Gess & E. J. Rubin (Eds.), *Theoretical and Experimental Approaches to Romance Linguistics* (pp. 63-75). John Benjamins.

Chun, D. M., Hardison, D. M., Pennington, M. C. (2008). Technologies for prosody in context. Past and future of L2 research and practice. In J. G. Hansen Edwards & M. L. Zampini (Eds.), *Phonology and Second Language Acquisition* (pp. 323–346). Studies in Bilingualism. John Benjamins.

Clopper, C. G., & Smiljanic, R. (2015). Regional variation in temporal organization in American English. *Journal of Phonetics, 49,* 1–15.

Dauer, R. M. (1983). Stress-timing and Syllable-timing Reanalysed. *Journal of Phonetics*, *11*, 51–62.

Dellwo, V. (2006). Rhythm and speech rate: a variation coefficient for deltaC. In P. Karnowski & I. Szigeti (Eds.), *Language and language processing* (pp. 231–241). Frankfurt am Main: Peter Lang.

Dellwo, V., & Fourcin, A. (2013). Rhythmic characteristics of voice between and within languages. *Travaux neuchâtelois de linguistique, 2013, 59*, 87–107.

Dellwo, V., & Wagner, P. (2003). Relations between language rhythm and speech rate. In *Proceedings of the 15th International Congress of Phonetics Sciences* (pp. 471-474). Barcelona.

Dellwo, V., Leemann, A., & Kolly, M. J. (2015). Rhythmic variability between speakers: Articulatory, prosodic, and linguistic factors. *Journal of the Acoustical Society of America, 137*(3), 1513–1528.

Erickson, D. (2013). Speech rhythm in English and applications to second language teaching. *Acoustical Science and Technology, 34*(3), 153–158.

Estebas Vilaplana, E. (2009). *Teach yourself English pronunciation. An interactive course for Spanish speakers.* UNED, Netbiblo.

Ferjan, N., Ross, T., & Arvaniti, A. (2008). Second language rhythm and rhythm metrics. *Journal of the Acoustical Society of America, 123*(5), 3427–3427.

Fuchs, R. (2014). Towards a perceptual model of speech rhythm: Integrating the influence of f0 on perceived duration. In H. Li, H. Meng, B. Ma, E. S.Chng & L. Xie (Eds.), *Interspeech 2014* (pp. 1949–1953). Singapore.

Fuchs, R. (2015). You're not from around here, are you? - A dialect discrimination experiment with speakers of British and Indian English. In E. Delais-Roussarie, M. Avanzi & S. Herment (Eds.), *Prosody and Language in Contact: L2 Acquisition, Attrition and Languages in Multilingual Situations* (pp. 123–148). Berlin: Springer.

Fuchs, R., & Wunder, E. M. (2015). A sonority-based account of speech rhythm in Chinese learners of English. In U. Gut, R. Fuchs & E. M. Wunder (Eds.), *Universal or Diverse Paths to English Phonology* (pp. 165–184). Berlin: Mouton de Gruyter.

Gabriel, C., & Kireva, E. (2014). Prosodic transfer in learner and contact varieties: Speech rhythm and intonation of Buenos Aires Spanish and L2 Castilian Spanish produced by Italian native speakers. *Studies in Second Language Acquisition*, *36*(2), 257–281.

Gabriel, C., Stahnke, J., & Thulke, J. (2015). Assessing foreign language speech rhythm in multilingual learners. An interdisciplinary approach. In H. Peukert (Ed.), *Transfer Effects in Multilingual Language Development* (pp. 191–220). John Benjamins Publishing Company.

Gilbert, J. B. (2008). *Teaching Pronunciation: Using the Prosody Pyramid*. Cambridge University Press, New York.

Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. In C. Gussenhoven & N. Warner (Eds.), *Laboratory Phonology 7* (pp. 515–546). Berlin: Mouton de Gruyter.

Gu, W., & Hirose, K. (2014). Rhythmic patterns in native and nonnative Mandarin speech. In N. Campbell, D. Gibbon & D. Hirst (Eds.), *Social and Linguistic Speech Prosody: Proceedings of the 7th International Conference on Speech Prosody 2014* (pp. 592–596). Dublin, Ireland.

Gut, U. (2012). Rhythm in L2 speech. In D. Gibbon, D. Hirst & N. Campbell (Eds.), *Rhythm, Melody and Harmony in Speech: Studies in Honour of Wiktor Jassem, Special edition of Speech and Language Technology* (pp. 83–94). Poznan.

Hincks, R. (2015). Technology and learning pronunciation. In M. Reed & J. Levis (Eds.), *The Handbook of English Pronunciation* (pp. 505–519). Wiley-Blackwell.

Hualde, J. I. (2012). Stress and Rhythm. In J. I. Hualde, A. Olarrea & E. O'Rourke (Eds.), *The Handbook of Hispanic Linguistics* (pp. 153–171). Blackwell Publishing.

Hualde, J. I. (2014). *Los sonidos del español*. Cambridge: Cambridge Univ. Press.

Kireva, E., & Gabriel, C. (2015). Rhythm properties of a contact variety: Comparing read and semi-spontaneous speech in Argentinean Porteño Spanish. In E. Delais-Roussarie, M. Avanzi & S. Herment (Eds.), *Prosody and Language in Contact: L2 Acquisition, Attrition and Languages in Multilingual Situations* (pp.149–168). Berlin: Springer.

Knight, R. A. (2011). Assessing the temporal reliability of rhythm metrics. *Journal of the International Phonetics Association, 41*(3), 271–281.

Kohler, K. J. (2009). Rhythm in speech and language - A new research paradigm. *Phonetica*, *66*, 20–45.

Kolly, M. J., & Dellwo, V. (2014). Cues to linguistic origin: the contribution of speech temporal information in foreign accent recognition. *Journal of Phonetics, 42*, 12–23. DOI:10.1016/j.wocn.2013.11.004.

Ladefoged, P. (1975). *A course in phonetics*. New York: Harcourt Brace Jovanovich.

Leemann, A., Dellwo, V., Kolly, M. J., & Schmid, S. (2014). Disentangling sources of rhythmic variability between dialects. In N. Campbell, D. Gibbon & D. Hirst (Eds.), *Social and Linguistic Speech Prosody: Proceedings of the 7th International Conference on Speech Prosody 2014*. Dublin, Ireland.

Li, A., & Post, B. (2014). L2 Acquisition of Prosodic Properties of Speech Rhythm: Evidence from L1 Mandarin and German Learners of English. *Studies in Second Language Acquisition, 36*(2), 223–255. DOI: 10.1017/S0272263113000752.

Lloyd James, A. (1940). *Speech signals in telephony*. London: Pitman & Sons.

Loukina, A., Kochanski, G., Rosner, B., Keane, E., & Shih, C. (2011). Rhythm measures and dimensions of durational variation in speech. *The Journal of the Acoustical Society of America*, *129*(5), 3258–3270.

Loukina, A., Kochanski, G., Rosner, B., Keane, E., & Shih, C. (2013). What determines duration-based rhythm measures: text or speaker? *Laboratory Phonology*, *4*(2), 339–382.

Low, E. L., Grabe, E., & Nolan, F. (2000). Quantitative characterisations of speech rhythm: 'Syllable-timing' in Singapore English. *Language and Speech, 43*, 377–401.

Mairano, P. (2011). Rhythm typology: acoustic and perceptive studies. PhD thesis, University of Turin.

Mairano, P. (2014). Applying durational metrics to recorded speech vs. TTS – evidence from 6 languages. *Proceedings of the X Convegno dell'Associazione Italiana Scienze della Voce*. University of Turin.

Mairano, P., & Romano, A. (2007). Inter-Subject Agreement in Rhythm Evaluation for Four Languages (English, French, German, Italian). In J. Trouvain & W. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences* (pp. 1149–1152). Saarbrücken, Germany.

Mairano, P., & Romano, A. (2010). Un confronto tra diverse metriche ritmiche usando Correlatore. In S. Schmid, M. Schwarzenbach & D. Studer (Eds.), *La dimensione temporale del parlato*, Proc. of the V National AISV Congress (pp. 79–100). Zurich, Switzerland.

Molnar, M., Gervain, J., & Carreiras, M. (2014). Within-rhythm Class Native Language Discrimination Abilities of Basque-Spanish Monolingual and Bilingual Infants at 3.5 Months of Age. *Infancy*, *19*(3), 326–337.

Nazzi, T., & Ramus, F. (2003). Perception and acquisition of linguistic rhythm by infants. *Speech Communication*, *41*, 233–243.

Ordin, M., & Polyanskaya, L. (2014). Development of timing patterns in first and second languages. *System*, *42*, 244–257. DOI:10.1016/j.system.2013.12.004.

Ordin, M., & Polyanskaya, L. (2015). Perception of speech rhythm in second language: the case of rhythmically similar L1 and L2. *Frontiers in Psychology*, 6(316), 1–15. DOI: 10.3389/fpsyg.2015.00316.

Payne, E., Post, B., Astruc, L., Prieto, P., & del Mar Varnell, M. (2012). Measuring child speech. *Language and Speech*, *55*(2), 203–229.

Pike, K. L. (1945). *The intonation of American English*. Ann Arbor: University of Michigan Press.

Prieto, P., del Mar Vanrell, M., Astruc, L., Payne, E., & Post, B. (2012). Phonotactic and phrasal properties of speech rhythm. Evidence from Catalan, English, and Spanish. *Speech Communication*, *54*(6), 681–702. DOI:10.1016/j.specom.2011.12.001.

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, *73*, 265–292.

Roach, P. (1982). On the distinction between “stress-timed” and “syllable-timed” languages. In D. Crystal (Ed.), *Linguistic controversies* (pp. 73–79). London: Edward Arnold .

Roncaglia-Denissen, M.P., Schmidt-Kassow, M., Heine, A., & Kotz, S.A. (2015). On the impact of L2 speech rhythm on syntactic ambiguity resolution. *Second Language Research*, *31*(2), 157–178. DOI: 10.1177/0267658314554497.

Russo, M., & Barry, W. J. (2008). Isochrony reconsidered. Objectifying relations between rhythm measures and speech tempo. In *Proceedings of Speech Prosody 2008* (pp. 419–422). Campinas, Brazil.

Schiering, R. (2007). The phonological basis of linguistic rhythm. Cross-linguistic data and diachronic interpretation. *Sprachtypologie und Universalienforschung*, *60*, 337–359.

Schmidt, A-E., & Post, B. (2015a). The Development of Prosodic Features and their contribution to Rhythm Production in Simultaneous Bilinguals. *Special issue, Language and Speech, 58*(1), 24–47.

Schmidt, A-E., & Post, B. (2015b). Language Interaction in the Development of Speech Rhythm in Simultaneous Bilinguals. In E. Delais-Roussarie, M. Avanzi & S. Herment (Eds.), *Prosody and Language in Contact: L2 Acquisition, Attrition and Languages in Multilingual Situations* (pp. 271–291). Berlin: Springer.

Selouani, S-A., Alotaibi, Y., Cichocki , W., Gharsellaoui , S., & Kadi, K. (2015). Native and non-native class discrimination using speech rhythm- and auditory-based cues. *Computer Speech and Language, 31*(1), 28–48. DOI:10.1016/j.csl.2014.11.003.

Tilsen, S., & Arvaniti, A. (2013). Speech rhythm analysis with decomposition of the amplitude envelope: Characterizing rhythmic patterns within and across languages. *Journal of the Acoustical Society of America, 134*(1), 628–639.

Torgersen, E., & Szakay, A. (2012). An investigation of speech rhythm in London English. *Lingua, 122* (7), 822–840. DOI:10.1016/j.lingua.2012.01.004.

Tortel, A., & Hirst, D. (2010). Rhythm metrics and the production of English L1/L2. In M. Hasegawa-Johnson (Eds.), *Proceedings of Speech Prosody 2010*, Chicago.

Trouvain, J., & Gut, U. (Eds.). (2007). *Non-native prosody: Phonetic descriptions and teaching practice*. Berlin: Mouton de Gruyter.

Vihman, M. M. (2014). *Phonological development: the first two years*. Wiley-Blackwell.

Warner, N., & Arai, T. (2001). Japanese mora-timing: a review. *Phonetica, 58*, 1–25.

White, L., & Mattys, S. (2007). Calibrating rhythm: first language and second language studies. *Journal of Phonetics, 35*(4), 501–522. DOI:10.1016/j.wocn.2007.02.003.

White, L., Mattys, S., & Wigit, L. (2012). Language categorization by adults is based on sensitivity to durational cues, not rhythmic class. *Journal of Memory and Language*, *66*(4), 665–679. DOI:10.1016/j.jml.2011.12.010.

[1] The terms ‘second’ and ‘foreign’ languages will be used indistinctively unless there is a need to explicitly state a technical distinction in meaning.
[2] The term ‘acquisition’ will be also used to mean ‘learning’.