FACULTAT DE PSICOLOGIA, CIÈNCIES DE L'EDUCACIÓ I DE L'ESPORT BLANQUERNA

ENGINYERIA I ARQUITECTURA LA SALLE

(UNIVERSITAT RAMON LLULL)

(UNIVERSITAT RAMON LLULL)

Màster en Formació del Professorat d'Educació Secundària, Batxillerat, Formació Professional i Ensenyament d'Idiomes

TREBALL FINAL DE MÀSTER

Curs 2019-2020

Does L1 Catalan rhythm transfer to L2 English?

ESTUDIANT: Patrícia Marsà Morales

DIRECCIÓ: Laia Aixalà Pàmies

ABSTRACT

Languages have been classified according to their rhythmic properties into two main groups: stress-timed languages (such as English, German or Dutch) and syllable-timed languages (such as Spanish, French or Italian). Due to a lack of rhythm studies regarding Catalan language, the aim of this study is to analyse if there is rhythmic transference from L1 Catalan to L2 English. In order to carry out this task, eighteen native Catalan speakers with different English levels from A1 to C2 have been recorded reading a Catalan and an English phonetically balanced text, they were subsequently classified into three groups (A, B, C)Then, the recordings were segmented with *Praat* software and analysed through *Correlatore* using the parameters suggested from recognized literature. The results obtained in this study suggest that rhythmic transference is produced from L1 Central Catalan to L2 English spoken by Catalan speakers.

Keywords: rhythm, transference, L1 Catalan, EFL, L2 acquisition

RESUM

Les llengües han estat classificades segons les seves propietats rítmiques en, principalment, dos grups: les llengües de ritme accentual (com l'anglès, l'alemany o l'holandès) i les llengües de ritme sil·làbic (com l'espanyol, el francès o l'italià). A causa d'una manca d'estudis sobre el ritme del català, aquest treball té com a objectiu analitzar si es produeix transferència del ritme del català L1 a l'anglès L2. Per dur a terme la tasca proposada, s'han enregistrat divuit parlants nadius de català amb diferents nivells d'anglès, des de l'A1 fins el C2, llegint dos textos fonèticament equilibrats, un en anglès i un altre en català, per separar-los posteriorment en tres grups (A, B, C). Les gravacions es van segmentar mitjançant el programa *Praat* i es van analitzar amb *Correlatore* fent servir les combinacions de paràmetres proposades per estudis reconeguts. Els resultats obtinguts en aquest estudi suggereixen que es produeix transferència rítmica del català a l'anglès.

Paraules clau: ritme, transferència, Català L1, EFL, adquisició d'L2

RESUMEN

Las lenguas han sido clasificadas según sus propiedades rítmicas en, principalmente, dos grupos: lenguas de ritmo acentual (como el inglés, el alemán y el holandés) y lenguas de ritmo silábico (como el español, el francés y el italiano). Debido a una falta de estudios sobre el ritmo del catalán, este trabajo pretende analizar si se produce transferencia del ritmo del catalán central L1 al inglés L2. Para llevar a cabo la tarea propuesta, se han grabado dieciocho hablantes nativos de catalán con distintos niveles de inglés, desde A1 hasta C2, leyendo dos textos fonéticamente equilibrados, uno en inglés y otro en catalán, para posteriormente separarlos en tres grupos (A, B, C). Las grabaciones se segmentaron mediante el programa *Praat* y se analizaron con *Correlatore* mediante el uso de parámetros propuestos por estudios reconocidos. Los resultados obtenidos en este estudio muestran que se produce transferencia rítmica del ritmo del catalán al inglés

Palabras clave: ritmo, transferencia, L1 catalán, EFL, adquisición de L2

TABLE OF CONTENTS

1. INTRODUCTION

Rhythm is a property of oral languages described as "the perceived regularity of prominent units in speech". These regularities may be stated in terms of patterns of stressed vs. unstressed syllables, syllable length (long vs. short) or pitch (high vs. low) - or some combination of these variables (Crystal, 2011). According to rhythmic properties, languages have been classified into stress-timed and syllable-timed. Stress-timed languages are most of the Germanic languages such as English, German, Dutch or Swedish, but it also includes Thai, Russian or Persian, on the other hand, syllable-timed languages are many Romance languages such as Spanish, French, Italian and others such as Turkish, Korean or Mandarin Chinese.

English, as a stress-timed language, was thought to present isochronous periods between stressed syllables, regardless of the number of unstressed syllables the utterance presented, so that according to Abercrombie (1967), the duration of syllables would rather depend on the context: they will compress or expand to maintain the rhythm recurrence. On the contrary, syllable-timed languages present an interval regularity that will vary depending on the unstressed syllables, since both present a stable duration. Catalan, on the other hand, has been a controversial language because of the properties it shares with both types of rhythmic classifications. In the 80s, the theory of isochronic intervals was no longer supported of the two rhythmic classifications. Dauer (1987) suggested that linguistic rhythm depends on two language properties: syllabic structure and vowel reduction. Then, Ramus et al. (1999), Grabe and Low (2002), and Dellwo and Wagner (2003) proposed a series of parameters to measure vowel and consonantal intervals from utterances. %V, ΔV, ΔC, CrPVI, VnPVI and varcoΔC are the ones chosen to carry out this study.

This dissertation is divided into five main sections. The first section will include the hypotheses and objectives raised for this study (see section 2). Then, the theoretical framework section will provide a definition for linguistic rhythm (3.1), a description of the rhythm typology (3.2) and will put an insight on linguistic transference theories (3.3). Section 4 will focus on the methodology of this study starting from an explanation on the corpus used (4.1) and how the data was collected (4.2). The following point will provide general data about the informants who participated in the study (4.3). Then, the instruments and data segmentation sections (see 4.4 and 4.5) will be used to explain two software programs used to carry out the segmentation and the data analysis for the

recordings. Finally, section 4.6 will provide a detailed description about the acoustic parameters employed. Section 5 will present and discuss the results obtained with numeral data and graphic representations, and section 6 will state the conclusions obtained from this research.

2. HYPOTHESES AND OBJECTIVES

This study will try to find an answer to the following research questions:

- Is there rhythm transference from Catalan (L1) to English (L2)?
- If so, is there a progressive loss of Catalan rhythm as the speakers' English level (from A1 to C2) increases?
- Is there a significant difference between each language stage?
- Are phonetically balanced texts a good tool for evaluating such range of English levels?

The following objectives are the ones aimed to achieve to answer the previous questions:

- To record and process, using the appropriate instruments, the data from eighteen Central Catalan speakers with varying levels of L2 English.
- To explore and apply the different acoustic parameter combinations that scholars proposed for the last two decades.
- To detect the most accurate parameters to investigate rhythmic transference from Catalan to English.

3. THEORETICAL FRAMEWORK

This section will explore past developed theories, ideas, and models from researchers on linguistic rhythm. A description of the study of rhythm, its history, and its typology will be explained subsequently.

3.1 Linguistic rhythm

Rhythm refers to the periodicity in which certain events occur. It is used to describe recurring events that happen in a wide variety of fields: when it comes to nature, the very first examples that come to our mind are the transition between the day and night, heartbeats, or a galloping horse. Rhythm is also present in the field of human

communication, in fact, Ancient Greek philosophers first identified it as "a measured movement" (Partridge 1961), at a later date, Aristotle described rhythm in *The art of rhetoric* as "a pattern of recurrence imposed on speech or other sounds" (Adams 1979: 10) and used it to describe the citizens' speech who used iambic patterns, in contrast to Rhetoricians, the ones who had a major language knowledge, that made use of a trochaic rhythm.

Any rhythmical pattern is characterized by two parameters: structure and succession (Mairano, 2011). A structure consists of a set of permanent and basic elements that are not subject to circumstantial changes. On the other side, succession will take place when an event occurs after another in a period of space or time. In music, the stave bars indicate succession, and its structure is given by the notes. In poetry, lines are made of stressed and unstressed syllables, and strong syllables usually occupy fixed recurring positions in every line. More in detail, Allen (1975), outlines that a hearer would always perceive rhythmic patterns in any regular succession of sonorous events if the sequence lasts between 0.1 and 3.0 seconds.

Linguistic rhythm has been described from different perspectives: first studies from Pike (1945) and Abercrombie (1967) describe it as a feature of human languages: the characteristics that justify these occurrences in a periodic succession of time are stress signs and syllables, the first ones are the main characteristic of stress-timed languages and the second ones of syllable-timed languages.

Pinker and Jackendoff (2005) state that the ability to perceive and produce rhythm would purely be a human trait, although monkeys can perceive formants in their own species' vocalizations and, more specifically, tamarin monkeys can be taught to discriminate the gross rhythms of different languages. Even though, monkeys cannot learn how to move to a rhythm, whereas it is one of the elementary characteristics that can be observed in young adults. According to Cumming (2010), speech rhythm can be defined as "the perceived regularity in an utterance, induced by the acoustic multi-dimensionality of the speech signal, which is the result of a realisation of phonological structure, and influenced by the listener's native language".

3.2 Rhythmic typology

The very first classification of oral languages according to its rhythm was carried out by James Lloyd (1940), who observed that languages such as English, German or

Dutch have a rhythm that resembles Morse-code messages (stress-timed languages). On the other side, he stated that languages such as Spanish, French, or Turkish sounded alike a machine-gun (syllable-timed languages). This theory is taken as a point of departure of the "isochrony hypothesis" (Pike 1945, Abercrombie, 1967), a position which distinguishes languages between the two rhythm types and has two basic claims: every existing language belongs to one particular rhythm type, and rhythm types are based on a timing unit (such as the syllable or the foot), which will occur in sequences of regular intervals of similar durations.

Years later, the so-called classical "isochrony hypothesis" started to lose strength due to an augment of studies of acoustic measurements on a wider amount of languages: it was proved that the duration of syllables in syllable-timed languages varies according to the number of segments. In the case of stress-timed languages, the duration of feet depends on the number of syllables. Subsequently, the debate on the isochrony hypothesis was divided into two different paths: one followed a phonological framework, and the other one had a phonetical interpretation. The outcome for the "phonological turn" was a renewal of the perspective phonetics had on analysing speech rhythm. The most influential study was Ramus et al. (1999), who proposed a series of new acoustic measures for the two rhythm classes. Instead of looking for equal lengths at the syllable or foot level, they formulated acoustic correlates that should help give a better description of the two rhythm types (see 4.5).

From a phonological perspective, Dauer (1987) and Dasher & Bolinger (1982) proposed that rhythmic patterns rely on other linguistic properties of language such as lexicon and syntax, and to their phonological and phonetical attributes. Dauer (1987) pointed out the three main features that influenced speech rhythm: syllable structure, stress patterning, and the presence or absence of the vowel reduction phenomenon. She suggested that stress-timed languages tend to have more complex syllabic structures, and that syllable-timed languages do not tend to make strong distinctions between full and reduced vowels. Dasher and Bolinger (1982) added that syllable-timed languages tended not to have vowel length distinctions. Dauer (1983) claimed that rhythm types could not be found in the speech signal itself, they would rather derive from a bundle of properties of the phonological system where the complexity of syllable structure and the reduction of unstressed vowels gained more importance. Secondly, rhythm types were not absolute categories, but they would rather constitute poles of a typological continuum, allowing

the presence of mixed or intermediate types. Since a strict dichotomy was no longer tenable, Nespor (1990) supported the existence of intermediate languages, which would include the case of Catalan or Polish.

Studies on Catalan rhythm did not agree on the state of the language. Nespor (1990) classified it as an intermediate language. Her findings indicate that the syllabic structures of Catalan are more complex than Italian or Spanish, which are prototypical syllable-timed languages and their syllabic structures. Catalan, then, would be positioned closer to stress-timed languages. In addition to this, Catalan presents a stress-timed feature which is the vowel reduction process, by which vowels are weakened and then, centralized to an unstressed position. Ramus et al. (1999) study, according to the results for %V, ΔV and ΔC, classify Catalan as a syllable-timed language. In their following study (Ramus et al, 2003), Catalan and Spanish rhythms are considered undistinguishable due to ΔV results, which confirm their similarity despite their differences. In opposition, Grabe & Low (2002) analysis supports Nespor's (1990) conclusions by using a different set of parameters. Still, there is no conclusion on the state of Catalan rhythm.

English, on the other hand, is a stressed-timed language because its syllables tend to last different amounts of time: there is a perceived amount of time between consecutive stressed syllables, and consequently, unstressed syllables tend to be "compressed" to adapt into the time segment.

3.3 Linguistic transference

Linguistic transference, also known as crosslinguistic influence, is a phenomenon caused in contexts where language contact is produced. Language contact is, in its simplest definition, the use of more than one language in the same place at the same time by the same social group (Thomason, S., 2001). It can lead to situations where a momentaneous linguistic contact is produced, which does not imply any significant effect on any of the languages involved, or to an eventual bilingualization of its following generations (Roseano et al., 2015).

During the last decade, the study of the effect of language contact on prosody has grown significantly. Although its underrepresentation, most studies on L2 acquisition suggest that language contact does not only influence lexical, syntactical, or morphological elements but also on prosody, and more precisely, on intonation and rhythm (Roseano et al., 2015).

In the case of studies involving these fields, it is argued that L1 and L2 prosodic system differences limit the acquisition process and that it could lead to interferences between them. In other words, learners' native language will have an influence on their second language acquisition because of the crosslinguistic influence that depends on the structural relationship between the two languages. L1 influence initially occurs when transferring its structures to L2 (Shatz, I., 2017). When transfer facilitates L2 acquisition is referred to as "positive", and it is found in cases where L1 and L2 structures are similar or identical. Conversely, transfer which interferes on acquisition is referred to as "negative", and it is found in cases where languages share fewer similarities. This group would include the languages this paper is based on. "Adult diglossia" is a phenomenon where L2 students acquire the language imperfectly because of its usage takes place in different contexts and for different aims than the speakers' L1. The prosodic properties of their L1 will tend to be transferred involuntarily to the language learnt, which is a common case in foreign language self-learning (Roseano et al. 2015).

Regarding the present state of research on rhythm in linguistic transference, there is a lack of L2-based corpus when it comes to prosody. Most of the conducted research over the past forty years in L2 phonetics and phonology relies on the limited use of laboratory settings (Gut, U., 2009), and it is only recently when L2 oral corpora have been created to study segmental and supra-segmental fields of phonetics and phonology.

4. METHODOLOGY

This section will define the methodology followed to achieve the given objectives. An explanation on every selected criterion will be detailed hereunder, along with the reason why they were chosen among other approaches from previous rhythm studies.

4.1 Corpus

As explained further in section 3.5, the corpora used in every study varied according to the scholars' criteria. This paper takes as reference studies that employed two different types of texts that can be classified into two main groups: controlled syllabic structure utterances (CV or CVC) as in Ramus et al. (1999) study, where the speakers of every language from the study had to read a series of declarative sentences containing between 15 and 19 syllables. The second group would contain texts that followed no strict

requirements, such as the corpus from Dellwo and Wagner (2003) which consisted of a literary text which was translated into the different languages investigated. Depending on the scholars' own criteria, they preferred to choose one type of corpora or another.

The corpora used in this work are two phonetically balanced texts that were created exclusively for research studies on linguistic rhythm. The English text is *The North Wind and the Sun* (Handbook of the IPA 1999: 39) and the Catalan text, *El vent del nord i el sol* (Marsà Morales and Roseano 2019) (see Annex 1), consists on an adaptation using progressive adjustments for up to a 1% difference according to the Catalan phonological unit frequency of use from Rafel (1980, 1996). According to the concept stated in section 3.1 (Lloyd James, 1940; Pike, 1945; Abercrombie, 1967), rhythm was a phenomenon firstly described on spontaneous speech, which means that the usage of a corpus that followed no requirements (from the second group) would be the closest way to simulate spontaneous speech in order to obtain the optimal results. The most significant feature of using phonetically balanced texts is that, apart from resembling spontaneous speech, it is also controlled speech, which assures that the data obtained will correspond to the real use Catalan and L2 English speakers make of the language.

4.2 Data collection

The data gathering process was carried out online via voice recording applications. Informants were asked to read the Catalan text in order to familiarize with it. Then, they were requested to record themselves while reading it on a speech rhythm they considered as normal. Pronunciation mistakes or not respecting the punctuation marks would result in an invalidation of the recording attempt. It was necessary to pay special attention to punctuation marks due to the strategical placing of the majority of phonemes in the phonetically balanced texts. For example, when recording the Catalan text, not respecting the pause in "*Un dia, de sobte"* (see Annex) would lead to the articulation of a voiced dental fricative [ð], instead of the voiced dental plosive [d], affecting to the expected length the consonant would have and resulting on an inadequate segmentation of the sound. After obtaining a valid recording for the Catalan text, informants would read the English text with the objective of familiarizing with it, and then record themselves until obtaining a valid sample, which would usually take about four to five attempts. The output format of the audio recordings was not compatible with the data segmentation program *Praat* (Boersma, P., and Weenink, D., 1995) thus audios were converted from .OGG stereo format to .WAV mono channel. Then, informants were tagged according to their

English proficiency: the first informant with an A1 English level was tagged as *Eng_A1_1*, and the third most proficient informant was *Eng_C2_3*, resulting in a total number of 18 informants, classified in groups of three according to the 6 levels of English of the Common European Framework of Reference (CEFR).

4.3 Informants

In order to gather the data to carry out the recordings and its subsequent segmentation process, it was required to find a total number of 18 L1 Catalan speakers who also had competence on L2 English. Informants speak the central Catalan dialect, known for being the most commonly spoken variety of the language. It is spoken in the Barcelona province, in the eastern half of the Tarragona province, and mostly in Girona. A strict requirement for informants from Barcelona was not to present the most characteristic traits from the variant of central Catalan spoken in Barcelona, the *barceloní*. These are the substitution of the schwa /ə/ for the near-open central vowel [ɐ] and the substitution of the voiced palatal lateral approximant /ʎ/ for an iodized sound /j/.

Informant	Age	Birthplace	Finished studies
A1_1	63	Barcelona	ESO
A1_2	58	Terrassa	Master
A1_3	30	Barcelona	CFGS
A2_1	26	Barcelona	ESO
A2_2	56	Terrassa	CFGS
A2_3	18	Vilafranca del Penedès	Batxillerat
B1_1	42	Terrassa	Degree
B1_2	27	Barcelona	ESO
B1_3	27	Terrassa	CFGS
B2_1	25	Barcelona	Degree
B2_2	26	Terrassa	Degree
B2_3	28	Barcelona	Master
C1_1	29	Font-rubí	Degree
C1_2	27	Terrassa	Degree
C1_3	23	Quart	Degree
C2_1	35	Terrassa	Master

C2_2	27	Terrassa	Master
C2_3	29	Sabadell	Degree

Table 1. Informants' age, birthplace, and finished studies

Table 1 shows the data of all the informants who participated in the research: adult learners of English with different levels of education, from secondary obligatory school to university studies. Informants are native speakers of Catalan thus, table names correspond to the English CEFR level each participant has, which has been kept as a classification code.

4.4 Instruments

In order to process the gathered data for this research, it was necessary to select valid and reliable tools. These instruments are *Praat*, and *Correlatore*, and they will be explained in the following sections, along with its functionality.

4.3.1 *Praat*

Praat (Boersma, P., Weenink, D., 1996) (*Figure* 1) is a free computer software used for recording and analysing speech, it allows its users to load or record sounds with a microphone connected to the computer and it reads and shows on the upper half of the "sound" window the corresponding *wave* form for the recording and its spectrogram on the lower half. Users can navigate the sound file by zooming for more detail or scrolling to the fragment of interest.

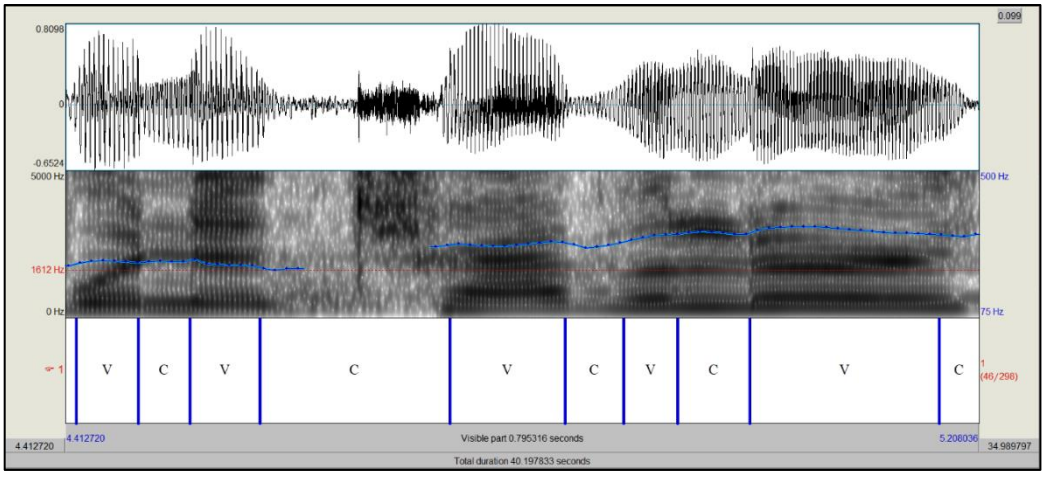


Figure 1. Vocalic and consonantal segments displayed on *Praat*

The most relevant feature of *Praat* for the elaboration of this research is the annotation section where users can segment their speech recordings and label them on multiple levels. These annotations can be exported into text files called "TextGrid", that other software can process.

4.3.2 *Correlatore*

Rhythmic analysis was carried out using *Correlatore* (Mairano, P., Romano, A., 2010). *Correlatore* offers a wide variety of options: the most relevant for this study is used to calculate the most used rhythmic parameters (ΔV, ΔC, %V, varcoΔC, varcoΔV, and PVI) from TextGrid files created with *Praat.* The second next remarkable feature is its module for graphic representations that follow the chosen rhythmic parameters. Users can create any kind of graphic by configurating the visuals and choosing which data they want to display on each axis.

4.5 Data segmentation

Once converting the audio files in a format *Praat* could read, namely .wav, a segmentation and annotation process into .TextGrid text files was carried out by segmenting vocalic (V) and consonantal (C) intervals. Some sound clusters needed an adaptation in order to obtain a faithful representation of the segments, these were the following:

- Semivowels /w/, /j/ were considered vocalic sounds.
- Vowel-consonant boundaries were placed when formant structures presented a break.
- Plosives and fricatives preceded by a pause were set to a fixed duration of 50ms with the purpose of homogenizing sentence and word onsets

4.6 Acoustic parameters

During the past two decades there have been various rhythm studies that applied different parameter criteria to analyse the data segmented. The parameters featured in the section below are classified according to the authors that proposed them.

4.6.1 ΔV, ΔC and %V

Ramus et al. (1999) was the first rhythm study that suggested segmenting the data into vocalic and consonantal intervals. Therefore, their paper was meant to be an

implementation of the phonological account of rhythm perception (Ramus et al., 1999), as they state, ΔC and %V appeared to be directly related to syllabic structure. The first two (ΔV and ΔC) are used to calculate the standard deviation, which is known to be used to measure the amount of variation of a set of values which, in this case, are the vocalic and consonantal intervals within each sentence. %V determines the proportion of vocalic intervals within the sentence.

The results for this study demonstrated that the optimum combination to discriminate the rhythm properties of each language was ΔC/%V. This is because ΔC is an indicator for syllabic complexity: the higher a value is, the more complex onsets and codas in that language in particular will be. Languages with a very simple syllabic structure (identified as syllable-timed languages) have simple consonantal clusters, while languages with a complex syllabic structure (identified as stress-timed languages) have both simple and complex consonantal clusters, consequently, the ΔC value for stress-timed languages will be higher.

On the other hand, vocalic percentage (%V) is a variable that brings data about the presence or absence of vowel reduction through syllable complexity: the higher the %V is, the lower level of vowel reduction the given language will have.

Similarly, the standard deviation of vocalic intervals (ΔV) indicate the presence or absence of a high level of vowel reduction. Once analysing all the data, Ramus et al. found out that ΔV reflected results for other factors that could influence rhythm measurements such as vowel lengthening in specific contexts (Italian), contrastive vowel length (Japanese), long vowels (tense vowels and diphthongs in English and Dutch and nasal vowels in French) and vowel reduction (English, Dutch, Catalan). In fact, ΔV scale could not be interpreted as transparently as ΔC, since the previous phenomena combined with each other and influenced the variability of vocalic intervals. For this reason, and despite %V and ΔV were expected to give similar data, ΔV was put aside since %V results were much more stable. The results for Ramus et al. (1999) ΔC/%V parameter combination are shown Figure 2.

According to stress-timed languages (English [EN], Dutch DU] and Polish [PO], the value for ΔC is much higher than syllable-timed languages (Spanish [ES], Italian [IT], French [FR] and Catalan [CA]). Conversely, the value for %V lowers as a result of having

a high ΔC. As stated before, syllable-timed languages have a higher %V value than stress-timed languages.

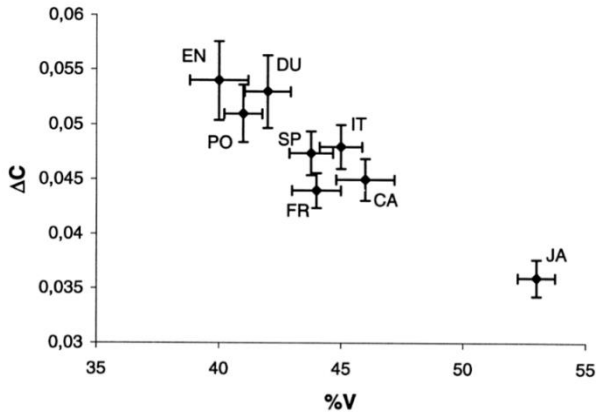


Figure 2. Distribution of languages over the (%V, ΔC) plane.

4.6.2 Pairwise Variability Index

Grabe and Low (2002) studied the relationship between speech timing and rhythmic classifications of languages based on the Pairwise Variability Index (PVI). This study was based on a previous one that focused on the differences between British English (stress-timed) and Singapore English (syllable-timed) (Low, Grabe and Nolan, 2000). In that study, instead of measuring interstress intervals or syllable duration (which are listed as phonological units), it focused on measuring vowels and consonants and the duration of intervals between them in a passage of speech. Taking it as a reference, Grabe and Low (2002) computed a PVI for each type of measurement. For their study, the Pairwise Variability Index expresses the level of variability between successive vocalic or consonantal intervals: in contrast to the parameters from Ramus et al., the data is obtained by sequencing its segments from a sequence of time.

Grabe and Low (2002) used two PVI variations for this index. The raw Pairwise Variability Index (rPVI) was used for consonantal intervals and the normalised Pairwise Variability Index (nPVI) was used for vocalic intervals. Henceforth, both parameters will be noted as CrPVI and VnPVI.

$$n\,PVI = 100 \times \left[\sum_{k=1}^{m-1} \left| \frac{d_k - d_{k+1}}{(d_k + d_{k+1})/2} \right| / (m-1) \right],$$

Figure 3. Normalized Pairwise Variability Index formula

Figure 3 shows that nPVI is obtained by calculating the durational difference between each pair of successive measurements by taking the absolute value of the difference and dividing it by the mean duration of the pair. rPVI is calculated by omitting both the third step ($(d_k + d_{k+1})/2$) and the output multiplication.

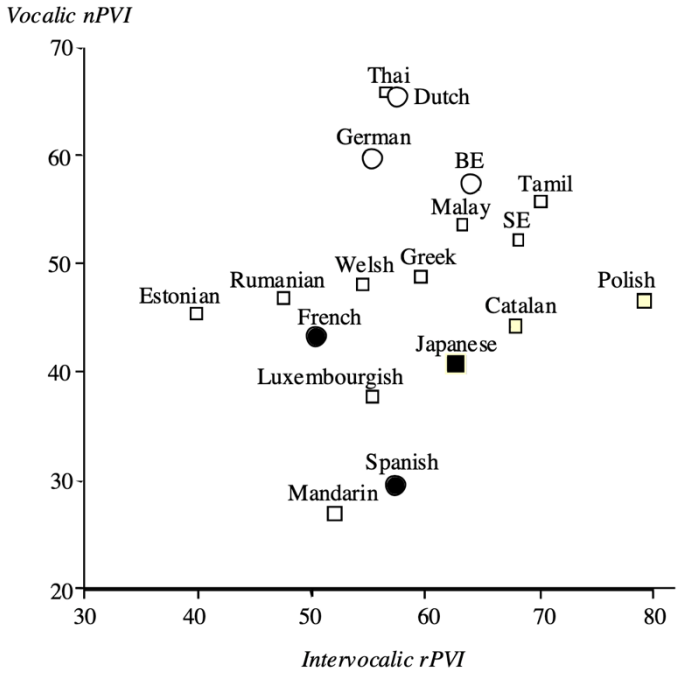


Figure 4. PVI profiles for data from eighteen languages (Grabe and Low, 2002)
○ = stress-timed, ● = syllable-timed, ■ = mora-timed, □ = mixed

As noted in *Figure 4* stress-timed languages such as British English [BE], German, and Thai show a higher VnPVI and CrPVI index than syllable-timed languages (Spanish and French). This result is mainly because stress-timed languages have a wider variety of complex syllabic structures and sometimes present vowel reduction. It is important to emphasise the position Catalan occupies in the table, and because of this, it was classified as an intermediate language (along with Polish) since it has complex syllable structures and presents vowel reduction.

4.6.3 varcoΔC/varcoΔV

In 2003, a new point of view of the language rhythm research is proposed: speech rate and rhythm as a whole were a topic that had hardly been studied and applied for both types of rhythm. It was necessary to put a focus on them since it was found that both

parameters were dependent on speech timing and, at that time, were suspected to interact between them to a great degree (Dellwo and Wagner, 2003). Taking into account that previous studies proved that ΔC was extremely speech rate dependant while %V remained stable across different speech rates, Dellwo and Wagner presented an alternative enhanced model for a cross-language rate control based on the parameters from Ramus et al. (1999) where speech rate would be much more stable. Their proposal was based on a study where native speakers read the same short text in German (original), French, and English, being the two last ones translated by philologically educated native speakers (Dellwo and Wagner, 2003). The speakers had to read the text in a way they considered "normal reading". After that, they read it two more times: the first one they were asked to read it "slower" and the second time "even slower".

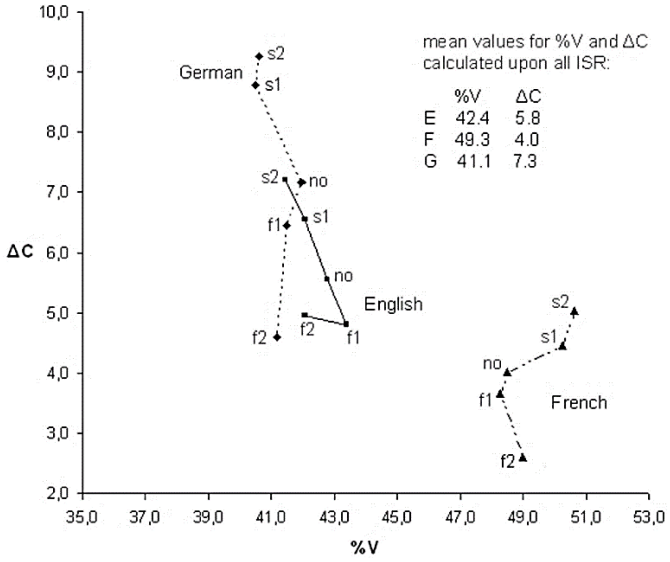


Figure 5. German, English and French ΔC and %V distribution (Dellwo and Wagner 2003)

According to *Figure 5*, stress-timed languages (German, English) show higher ΔC values than syllable-timed languages (French), just as in Rammus et al. (1999), the data remains distributed following the same pattern. In the case of %V in both studies, syllable-timed languages show higher values than the stress-timed ones. According to Dellwo and Wagner (2003), the data obtained with the new parameters (varcoΔC and varcoΔV) revealed that the tendencies supported the hypothesis that stress-timed and syllable-timed languages are distinguishable by %V and ΔC, but the new parameters offered much more accurate and detailed results.

5. RESULTS AND DISCUSSION

This section will show and discuss the results obtained for the rhythmic analysis using the parameters suggested by scholars.

Once the segmentation process for the Catalan and English recordings was completed, TextGrid files were processed with *Correlatore* according to the languages and the levels of English investigated. The following parameter combinations were chosen in order to obtain the most accurate data interpretations regarding the hypotheses and objectives raised:

- %V, ΔV and ΔC from Ramus et al. (1999) (see section 4.6.1)
- VnPVI and CrPVI from Grabe and Low (2002) (see section 4.6.2)
- varcoΔC and V% from Dellwo and Wagner (2003) (See section 4.6.3)

Table 2 shows the mean scores for the Catalan rhythm metrics from processed the TextGrid files. In case of *Table 3*, results are displayed according to the three stages of English from the Common European Framework of Reference (CEFR). Basic includes levels A1 and A2, Independent includes B1 and B2, and Proficient shows the mean for C1 and C2.

	ΔV	ΔC	%V	VnPVI	CrPVI	varcoΔV	varcoΔC
Mean_Cat	36.74	56.24	42.53	45.13	62.81	50.82	57.46

Table 2. Acoustic parameter values for Central Catalan

	ΔV	ΔC	%V	VnPVI	CrPVI	varcoΔV	varcoΔC
Mean_Eng_A	44.63	75.04	39.26	44.38	79.19	44.85	51.21
Mean_Eng_B	46.55	75.20	39.34	51.96	82.48	48.15	53.53
Mean_Eng_C	49.97	73.65	39.62	55.14	82.31	56.73	57.78

Table 3. Acoustic parameter values for English spoken by Catalan speakers

Table 4 Shows the prototypical data of British English (Received Pronunciation), which will be exclusively used as a reference for English as a stress-timed language in graphs.

	ΔV	ΔC	%V	VnPVI	CrPVI	varcoΔV	varcoΔC
Eng_Eng	66.79	67.76	47.24	63.45	79.46	59.07	56.15

Table 4. Acoustic parameter values for British English (RP)

Regarding Ramus et al. (1999) parameter combinations, %V, ΔV and ΔC were chosen because of the relation they present regarding syllabic structures. Stress-timed languages, which have a high onset and coda variability, are expected to obtain higher ΔC scores and a lower %V than syllable-timed languages. The analysis results show higher values for L1 British English (67.76) and also for every L2 English stage. In fact, A and B levels show very similar results (75.04, 75.20) and levels C (73.65) have obtained a closer value to the prototypical L1 BE. %V is a parameter that shows the proportion of vocalic intervals within the text. The value obtained for BE is 47.24, and it is 39,26, 39,34 and 39,62 for respectively, A, B and C levels which, for this parameter, demonstrate a homogeneous usage of vowel intervals within each language stage. When it comes to ΔV, which shows standard deviation of vocalic intervals, we can see a progression between A, B and C levels, where the higher levels have higher scores. Mean_Eng_A score is 44.65, which is closer to Mean_Cat value (36.74) and Mean_Eng_C result is 49.97, which approaches the British English (RP) score (66.79).

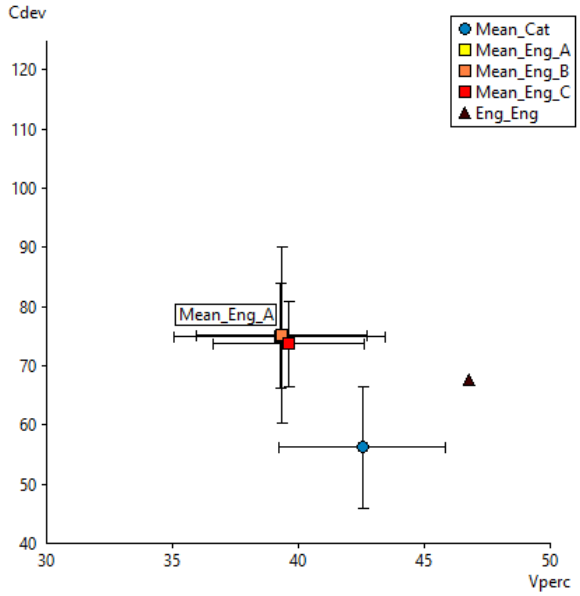


Figure 6. Mean %V and ΔC values for L1 Central Catalan (Mean_Cat), L2 English spoken by Catalan speakers (Mean_Eng) and L1 English (RP) (Eng_Eng)

Figure 6 shows the results on an axis which displays the relationship between %V and ΔC. Taking into account the placing that the mean English values occupy in the plane, we can state that, in order to measure rhythm transference, this parameter combination does not provide concluding results for this study. In the same way, the reason why the mean values appear stacked in the *x-y* plane could be because this combination seems to be the option for defining the current rhythmic position of a language rhythmic classification rather than rhythm transference, which, in this case it could be labelled as a so-called "L2 English spoken by Central Catalan speakers", which was the initial aim of Ramus et al. (1999) study. In other words, this parameter does give information about the position English spoken by Catalan speakers occupies in comparison with L1 British English and L1 Catalan. However, Mean_Eng_C variable displays that there seems to be a lower dispersion on higher levels of language, which, according to the results, level C speakers tend to cluster on a more delimited area than A and B levels.

Figure 7 shows combinations for the second parameters suggested from Ramus et al. (1999). The use of ΔV seems to clarify the positioning of the variables along the plane, but with minimum distance between them, even though, there seems to be a progression from stage A to stage C. It is also visible how L1 Catalan and L1 English are in opposed poles, whilst Mean_Eng cluster is situated in a range between their means.

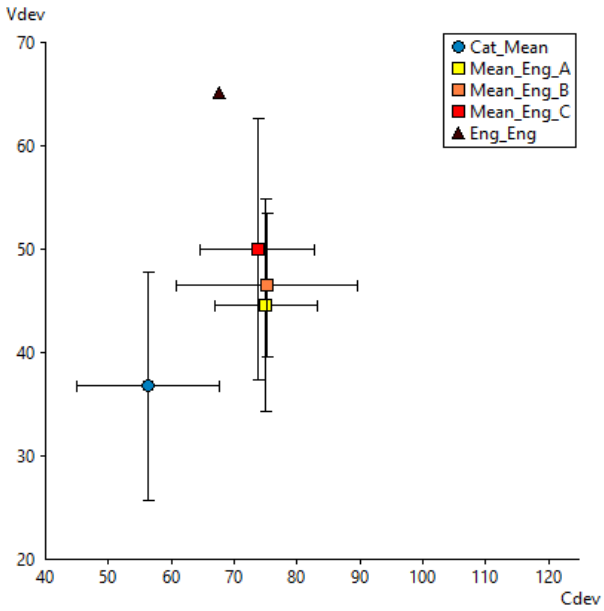


Figure 7. Mean ΔV and ΔC values for L1 Central Catalan (Mean_Cat), L2 English spoken by Catalan speakers (Mean_Eng) and L1 English (RP) (Eng_Eng)

Grabe and Low (2002) used the Pairwise Variability Index, a parameter that was created in order to measure consonantal or vocalic intervals and shows how much each subsequent segment varies in relation to their durations in the text. Their findings suggest that VnPVI shows a relation with vowel reduction, and CrPVI is related with syllabic structures. These parameters will have greater values with stress-timed languages because these languages have a wider variety of complex syllabic structures. Then, the authors concluded that syllable-timed languages should show lower scores for both values because they have both complex and simple syllabic structures. The results for these values in the present study meet the hypothesis of Grabe and Low: scores for CrPVI and VnPVI are both the lowest of the five variables in Catalan (45.13 and 62.81) and L2 English results show a meaningful progression between each language stage.

The results shown in *Figure 8* reflect the mentioned progression regarding vocalic intervals (VnPVI): values for Mean_Eng_A are very close to the L1 Catalan mean scores, which indicates that the durational difference between Catalan L1 and English L2 "A" speakers is very similar. Prototypical British English (RP) and L1 Catalan spoken by informants are in opposite poles in the plane, and so, as VnPVI increases in the *y* axis scale, the mean values for L2 English speakers gradually increase along with each CEFR level.

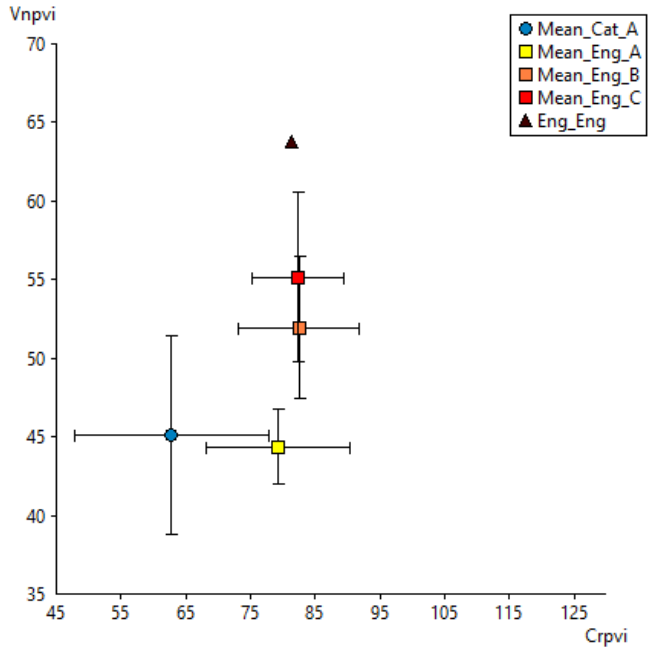


Figure 8. Mean CrPVI and VnPVI values for L1 Central Catalan (Mean_Cat), L2 English spoken by Catalan speakers (Mean_Eng) and L1 English (RP) (Eng_Eng)

Regarding the values obtained for CrPVI parameters, results suggest that there is no difference when it comes to consonantal intervals, in other words, the occurrence of L2 English intervals is very similar to English (RP). This finding may indicate that L2 English informants, regardless of their level, easily adapt to the syllabic structures of the target language.

Dellwo and Wagner (2003) presented an alternative parameter for the classic standard deviation for consonantal intervals, varcoΔC. Because of ΔC was said to be sensitive to speech rate, they suggested to apply the coefficient of variation in order to normalize it. This new adjustment eliminates the speech rate differences from the variables (English as a stress-timed language and Catalan as an intermediate or syllable-timed language) and correlates its speech rhythm, this is the reason why the values obtained between each rhythm type are more even.

In this case, Figure 9 shows that the parameter %V demonstrates the differences between each type of variable, namely, L1 Catalan, L2 English (A, B, C levels) and L1 English. However, varcoΔC scores for L1 Catalan (57.46), L1 English (56.15) and L2 English spoken by level C Central Catalan speakers (57.78) show that the value for L2 English level C share a similar variability of consonantal intervals as speakers with a native language proficiency. On the other hand, lower level L2 English speakers (A and B) tend to have a lower variability of consonantal intervals compared with proficient speakers. This could mean that, depending on how proficient are speakers in their language (L1 Catalan, Ls2 English (C) and L1 English speakers), the accuracy these speakers have will be much higher to the corresponding consonantal segments. When it comes to both A and B levels, although there seems to be a progression, it could suggest that speakers' consonantal intervals are less similar to the prototypical L1 English, which is expected to be the desired language.

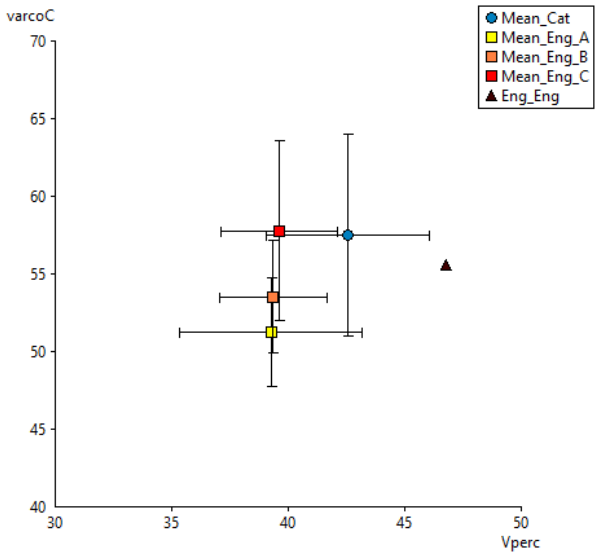


Figure 8. Mean %V and varcoΔC values for L1 Central Catalan (Mean_Cat), L2 English spoken by Catalan speakers (Mean_Eng) and L1 English (RP) (Eng_Eng)

6. CONCLUSIONS

The present study has attempted to explore if there is rhythmic transference from native Catalan speakers to L2 English from a range of 6 different levels (from A1 to C2) from the Common European Framework of Reference (CEFR). In order to comply with the first objective, informants have been searched in order to fulfil the eighteen slots, and, when it came to graphic representations, they have been classified into three groups: basic, which includes levels A1 and A2, independent, which includes B1, defined as the stage where speakers have mastered practical language skills for everyday use, and B2, which is the qualification that proves the speaker has language skills to live and work independently in an English-speaking country, and proficient, which includes C1, an in-depth higher-level qualification for B2, and C2, a qualification that proves holders have mastered English and can communicate as fluently and sophisticated as a highly competent English speaker (Cambridge English, 2020). Then, thirty-six audio files were obtained from recording the informants, plus a prototypical L1 British English speaker (RP) (IPA Handbook, 1999).

According to the acoustic parameters used, the values obtained along with the graphic representations show that L1 Catalan and L1 English are displayed in opposite poles on the graphs for the most used acoustic parameter combinations (ΔC and ΔV,

CrPVI and VnPVI). In order to discriminate between rhythm types, the previous parameters can be used because they displayed the results for the variables in a more precise way than the rest of the charts presented (ΔC and %V, varcoΔC, %V). On the other hand, the parameters that seem to provide better numerical results between the three variables (L1 Catalan, L2 English and L1 English) are those which measure consonantal intervals and do not present speech normalization, which are ΔC and CrPVI. Regarding the results for the parameters that analyse vocalic intervals, %V data reveals very similar scores between L2 English from all levels, which seems to be determining in order to discriminate between language rhythms. ΔV and VnPVI show a progression between language levels, and clearer different scores, consequently, they seem to offer good results for investigating rhythmic transference, apart from rhythmic typology.

On the whole, the results from the present study demonstrate that there is rhythm transference from L1 Catalan to L2 English, but, even if some parameter combinations display a progression within each language level, the results should not be considered conclusive enough to state that the rhythmic transference from Catalan decreases as English speakers gain a major language proficiency.

6.1 Limitations

It is worth mentioning that some burdens were encountered during this research. Initially, the recording process was meant to be done face-to-face with a high-quality microphone, but due to the context this study has been carried out, informants had to be taught how to record themselves. The vast majority of participants created the audios with proper microphones and in a noise-free environment, but there were a few cases where it was necessary to use a cleaning software in order to reduce the amount of noise that affected the audio. It is important to highlight that audio quality does not directly affect the segmentation process, but a cleaner spectrogram is much better to differentiate formants in *Praat.* Apart from that, the use of phonetically balanced texts could have not been an adequate corpus for beginner English levels because it was necessary to provide them with some help with the vocabulary. On the other hand, since the segmentation process has been done manually, interval boundaries may have not been applied under all the possible requirements that a script could have, and there could be minimum segmentation mistakes.

6.2 Further research

Considering everything mentioned above, further research is still required in the field, firstly, in order to define the rhythmic typology of Catalan, namely, if it is syllable-timed or an intermediate language, this "label" would help better identify rhythmic transference compared with other languages, and secondly, since there is still a lack of rhythm transference studies on this language, it would be interesting to see different approaches using different methodologies and corpora. A diachronic study starting from basic level informants that share similar ages, birthplaces, and studies (such as primary or secondary school students) would provide an interesting approach to this field.

Thus, considering the conclusions of this research, it can be seen that the use of phonetically balanced texts is a good tool for evaluating rhythm transference, but perhaps it is not so adequate with initial and intermediate levels of L2 English speakers. It would be interesting to carry out a similar study adapting the corpus used.

REFERENCES

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

Allen, G. D. (1975). *Speech rhythm: its relation to performance universals and articulatory timing*. Journal of Phonetics, 3, 75-86.

Boersma, P., Weenink, D. (1995). *Praat: doing phonetics by computer* [Computer program]. Version 6.1.15.

Cambridge English. (2020). *Cambridge English Qualifications*. Retrieved from https://www.cambridgeenglish.org/exams-and-tests/qualifications/

Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment.* Cambridge, U.K: Press Syndicate of the University of Cambridge.

Crystal, D. (2011). *A dictionary of linguistics and phonetics*. Hoboken, NJ: John Wiley & Sons.

Dauer, R. (1983) *Stress-timing and syllable-timing reanalysed*. Journal of Phonetics 11: 51–62

Dauer, R. M. (1987*). Phonetic and phonological components of language rhythm*. T. Gamkrelidze (Ed.), Proceedings XIth ICPhS: The eleventh International Congress of Phonetic Sciences (p. 447-450). Talin, Estònia: Academy of Sciences of the Estonian S.R.R

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

Fitch, W., Hauser, M., Chomsky, N. (2005). *The evolution of the language faculty: clarifications and implications.* Cognition 97(2): 179-210.

Grabe, E., Low, E. (2002). *Durational variability in speech and the rhythm class hypothesis.*

Gut, U. (2009). *Non-native speech*: *A corpus-based analysis of phonological and phonetic properties of L2 English and German* (Vol. 9). Peter Lang.

International Phonetics Association, Cambridge University Press. (1999) *Handbook of the International Phonetic Association: A guide to the Use of the International Phonetic Alphabet* (Cambridge University Press, 1999)

Ling, L., Grabe, E & Nolan, F. (2000). *Quantitative Characterizations of Speech Rhythm: Syllable-Timing in Singapore English*. Language and speech. 43. 377-401.

Lloyd, J. A. (1940*). Speech Signals in Telephony*. London: Sir Isaan Pitman & Sons.

Mairano, P., Romano, A. (2010) *Un confronto tra diverse metriche ritmiche usando Correlatore.* In: Schmid, S., Schwarzenbach, M. & Studer, D. (eds.) La dimensione temporale del parlato, (Proc. of the V National AISV Congress, University of Zurich, Collegiengebaude, 4-6 February 2009), Torriana (RN): EDK, 79-100.

Marsà Morales, P., Roseano, P. (2019). *El ritme del català: anàlisi a partir de textos fonèticament equilibrats*. Estudios de Fonética Experimental XXVIII.

Nespor, M. (1990). *On the rhythm parameter in phonology. In Logical Issues in Language Acquisition* (p. 157-175). Dordrecht: Fori.

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

Ramus, F., Nespor, M., Mehler, J. (1999). *Correlates of Linguistic Rhythm in the Speech Signal*. Cognition. 73. 265-292. 10.1016/S0010-0277(00)00101-3.

Roseano, P., Planas, A. M., Elvira-García, W., Celdrán, E. (2015*). Contacto lingüístico y transferencia prosódica bajo una perspectiva diacrónica: El caso del alguerés*. Dialectologia et Geolinguistica. 23. 95-123. 10.1515/dialect-2015-0006.

Shatz, I. (2017). *Native language influence during second language acquisition: A large-scale learner corpus analysis. Proceedings of the Pacific Second Language Research Forum* (PacSLRF 2016) (pp. 175–180). Hiroshima, Japan: Japan Second Language Association.

Thomason, S. (2001) *Language Contact - An Introduction* (Edinburgh University Press 2001).

ANNEX

Phonetically balanced texts

Catalan

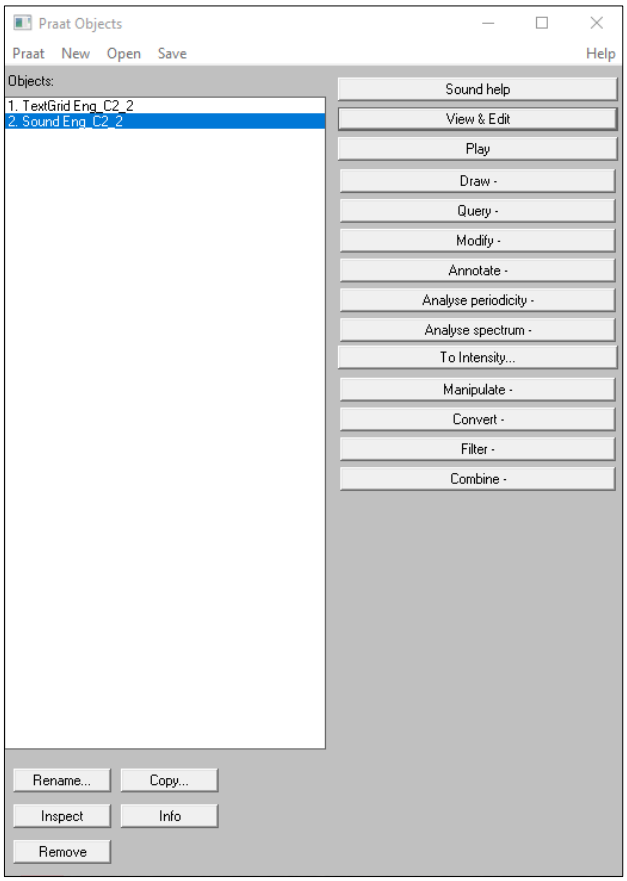
Un dia, de sobte, el vent de la tramuntana i el sol discutien fortament, per decidir el millor dels dos. Per atzar, veuen un viatger sense calçat i amb un abric roig passejant tranquil per un terreny. Decidiren que qui aconseguís que el senyor es tragués una capa es quedaria el cel aquella temporada. La tramuntana bufà amb moltíssimes ganes que la notessin, no obstant, com més bufava, l'home més s'abrigava, i decidí deixar-ho estar. El sol sortí al cap de res, de sobte, l'home es tragué la capa perquè s'acalorava molt, i doncs, la tramuntana confessà que era una fluixa.

English

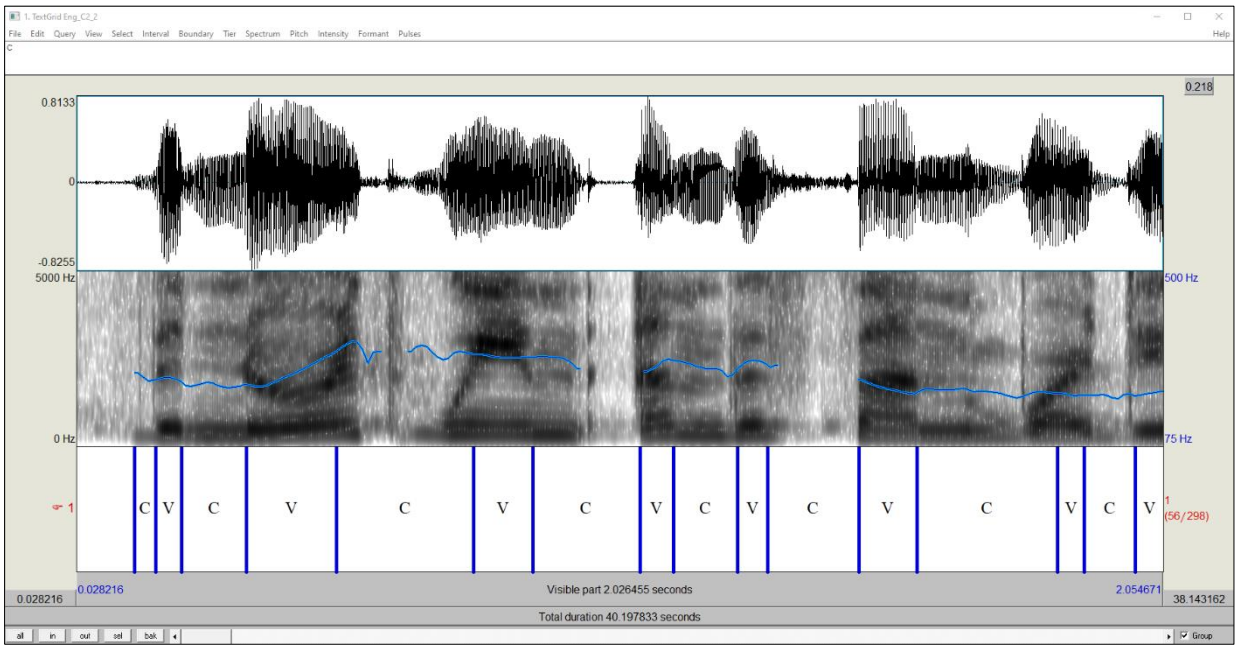
The north wind and the sun were disputing which was the stronger, when a traveller came along wrapped in a warm cloak. They agreed that the one who first succeeded in making the traveller take his cloak off should be considered stronger than the other. Then the north wind blew as hard as he could, but the more he blew the more closely did the traveller fold his cloak around him. And at last the north wind gave up the attempt. Then the sun shone out warmly, and immediately the traveller took off his cloak. And so the north wind was obligated to confess that the sun was the stronger of the two.

Praat

Main Menu

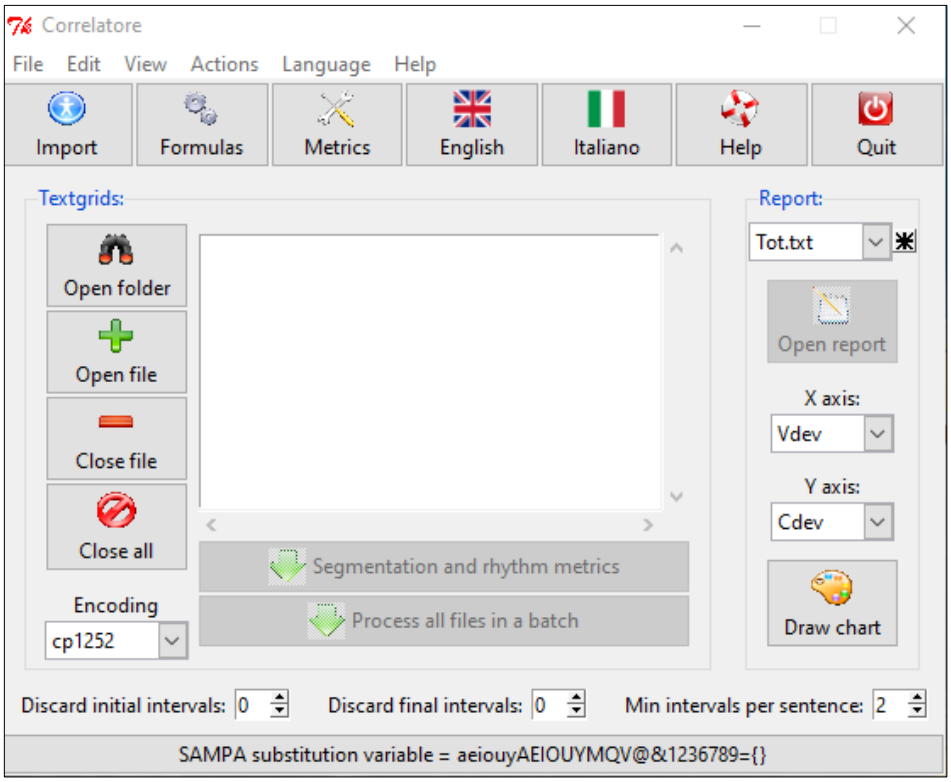


Praat "View & Edit" screen displaying consonantal (C) and vocalic (V) segmentation for informant Eng_C2_3

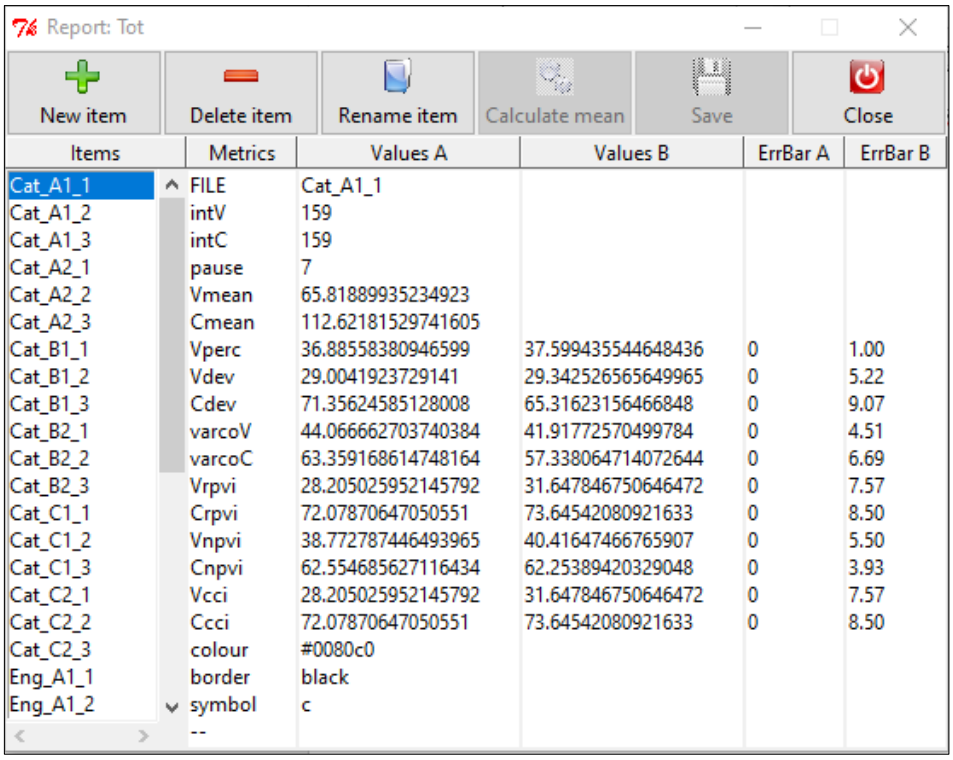


Correlatore

Main menu



Data report for informant Cat_A1_1



Items	Metrics	Values A	Values B	ErrBar A	ErrBar B
Cat_A1_1	FILE	Cat_A1_1			
Cat_A1_2	intV	159			
Cat_A1_3	intC	159			
Cat_A2_1	pause	7			
Cat_A2_2	Vmean	65.81889935234923			
Cat_A2_3	Cmean	112.62181529741605			
Cat_B1_1	Vperc	36.88558380946599	37.599435544648436	0	1.00
Cat_B1_2	Vdev	29.0041923729141	29.342526565649965	0	5.22
Cat_B1_3	Cdev	71.35624585128008	65.31623156466848	0	9.07
Cat_B2_1	varcoV	44.066662703740384	41.91772570499784	0	4.51
Cat_B2_2	varcoC	63.359168614748164	57.338064714072644	0	6.69
Cat_B2_3	Vrpvi	28.205025952145792	31.647846750646472	0	7.57
Cat_C1_1	Crpvi	72.07870647050551	73.64542080921633	0	8.50
Cat_C1_2	Vnpvi	38.772787446493965	40.41647466765907	0	5.50
Cat_C1_3	Cnpvi	62.554685627116434	62.25389420329048	0	3.93
Cat_C2_1	Vcci	28.205025952145792	31.647846750646472	0	7.57
Cat_C2_2	Ccci	72.07870647050551	73.64542080921633	0	8.50
Cat_C2_3	colour	#0080c0			
Eng_A1_1	border	black			
Eng_A1_2	symbol	c			
	--				

Segmentation and rhythm metrics for informant Eng_C2_3