

# Hezkuntzari buruzko ekarpenak

## Contributions on education (EUDIA-8)



**EDITOREAK:**  
**Juan Abasolo**  
**Irati de Pablo**  
**Ariane Ensunza**

Reference: Cantero, Francisco J., & Font-Rotchés, D. (2020). Melodic Analysis of Speech (MAS). Phonetics of Intonation. In Abasolo, J.; de Pablo, I.; Ensunza, A.(eds): *Contributions on education*. Universidad del País Vasco, 20-47.

eman ta zabal zazu



Universidad  
del País Vasco

Euskal Herriko  
Unibertsitatea

**Hezkuntzari** buruzko ekarpenak (EUDIA-8) [Recurso electrónico] = Contributions on education : Leioan, 2019ko apirilaren 5ean /[editores] Juan Abasolo Isasa, Irati de Pablo Delgado, Ariane Ensunza Aldamizetxebarria. – Datos. - Bilbao : Universidad del País Vasco / Euskal Herriko Unibertsitatea, Argitalpen Zerbitzua = Servicio Editorial, [2020]

1 recurso en línea : PDF (137 p.)

Modo de acceso: World Wide Web.

Textos en euskara e inglés.

ISBN: 978-84-1319-098-3.

1. Euskara (Lengua) - Dialectos. 2. Lenguaje y lenguas - Variación. 3. Prosodia (Lingüística). I. Abasolo Isasa, Juan, coed. II. Pablo Delgado, Irati de, coed. III. Ensunza Aldamizetxebarria, Ariane, coed. IV. Título: Contributions on education.

(0.034)809.169-087

(0.034) 800.87



UPV/EHUko EUDIA Ikerketa-taldea

Euskal Herriko Unibertsitateko ikerketa taldea (GIU 22/16)

Euskal Unibertsitate Sistemako ikerketa taldea (IT 1028/16)

# MELODIC ANALYSIS OF SPEECH (MAS): PHONETICS OF INTONATION

Francisco José Cantero Serena, Dolors Font-Rotchés

Universitat de Barcelona

[cantero@ub.edu](mailto:cantero@ub.edu); [dolorsfont@ub.edu](mailto:dolorsfont@ub.edu)

## Abstract

The acoustic-perception-based Melodic Analysis of Speech method (MAS) is a formal, objective and complete method that encompasses the most relevant aspects to be considered when analysing intonation. It provides researchers who wish to study intonation with criteria for establishing a corpus, the identification of melodic units, the extraction and relevance of acoustic data, its standardisation and the representation and interpretation of graphs, as well as the execution of perception tests and validation of the results obtained. It also offers a way of interpreting the melodic data based on three levels of intonation: one which provides an explanation of the structure of speech, the prelinguistic level, the linguistic level and the paralinguistic level, giving it significance with regard to the speaker's intentions or the communicative context. In addition, a phonological description based on the Autosegmental-Metrical model (AM) and its labelling convention known as Tones and Break Indices (ToBI), which is complementary and totally compatible with the phonetic description of MAS, can be added. The method is constantly advancing, being reformulated and added to as new requirements arise resulting from its use by researchers in the description of different languages, interlanguages and dialects. It is these developments, which have occurred over more than a decade and particularly in the two years, together with the basic pillars of the method, which are presented in this article.

Keywords: Melodic Analysis of Speech method (MAS), Intonation, Acoustic analysis, Perception tests, semi-automatic analysis.

## 1 Introduction

This paper presents the most relevant characteristics of the Melodic Analysis of Speech (MAS) method, together with the developments that have taken place to date<sup>1</sup>. It was proposed by Cantero in 1999, published in 2002, and since then has been used by various researchers, initially to describe languages and interlanguages and later dialect variants and aspects of paralinguistic intonation (in relation to the speaker's intentions and the communicative context). Currently, there are open lines of research on description and analysis relating to different aspects of intonation in different languages.

Traditionally, intonation has been one of the aspects of speech characterisation that has been greatly ignored by linguists. The first people to show an interest in describing intonation were authors from the *British school*. Authors such as Jones (1909), Palmer (1922) and Armstrong

---

<sup>1</sup> This paper is part of the research project financed by the Ministry of Economy and Competitiveness *Analysis of Speech and Teaching Models*, ref. FFI2013-41915-P.

& Ward (1926) aimed to write manuals designed for the correct teaching of English. In their work, they provided melodic curves, their context and meaning, so that learners could imitate them. In later decades, other linguists followed this tradition, including O'Connor & Arnold (1961), Crystal (1969), Cruttenden (1986) and the American, Bolinger (1986, 1989). Close to this tradition, although with original contributions that have greatly influenced intonation studies for the Spanish language, Navarro Tomás published his manual for Spanish intonation in 1944, designed as a practical language manual and a theoretical treatise on Spanish intonation.

Also based on the phonetic characterisation of intonation, authors of the Dutch school, t'Hart, Collier & Cohen (1990) and their *IPO model* (Institute for Perception Research) are of note, as they provide a stylisation of the intonation curves through the selection of relevant pitch movements in the melodic curve, eliminating other pitch variations.

At the same time as these schools, from a structural perspective of the phenomenon and a phonological viewpoint, linguists of the *American tradition* emerged with their level analysis proposal; authors such as Pike (1945) and Trager & Smith (1951), among others, described the curve based on four pitch levels, four stress levels (primary, secondary, etc.) and junctures (marking the beginning and ending of the phrase).

Along this baseline of phonology, and from the theoretical approach of generative phonology, stems the *Autosegmental Metrical model* (Pierrehumbert, 1980, 1987), which, based on the metrical structure of the phrase, suggests the representation of its pitch variations using two levels: H (high) and L (low). For this model, during the 1990s, the ToBI (Tone and Break Indices) transcription system was developed for English (Silverman et al., 1992; Beckman & Hirschberg, 1994), as well as systems for other languages, including Italian (ToBI<sub>It</sub>, Avesani, 1995), German (GToBI, Grice et al., 1996) and Spanish (Sp\_ToBI, Beckman et al., 2002; Prieto & Roseano (eds.), 2010), among others, and a transcription variant for Dutch (ToDI, Gussenhoven et al., 1999).

Later, and influenced by the Autosegmental Metrical model, the Aix-en-Provence model emerged, also generative, which suggested obtaining phonological forms of the prosodic system of different languages. To do so, it follows an automatic styling process using the MOMEL (MOdelling MELody) algorithm and the INTSINT notation model (Hirst et al., 2000).

In this context, the *Melodic Analysis of Speech* (MAS) model aims to provide a robust approach for the acoustic analysis of intonation, regardless of its phonological interpretation. It is a formal and objective method that provides the researcher with a way of analysing intonation, described in detail in the protocol established in 2009 by Cantero & Font-Rotchés<sup>2</sup> and with a way of interpreting the data from three different levels of intonation: prelinguistic, linguistic and paralinguistic (Cantero, 2002).

---

<sup>2</sup> There is an English version in Font-Rotchés & Cantero (2009).

This protocol is divided into three sections: a) the identification of melodic units in a corpus of spontaneous speech; b) the development of an acoustic phase, including the determination of relevant frequency values and standardisation; and lastly, c) a perception phase, during which the analyses and interpretation of data is validated. It remains the same for the procedure described below, although with some new contributions resulting from research, such as matters regarding labelling or the relevance of a number of values in the analysis that have proven to be important for some levels of intonation. Other significant developments include the creation of the semi-automated script by Mateo (2010a), which has also evolved over the years, adapting to the requirements of the group's research.

With regard to levels of intonation in the prelinguistic level (melodic features that structure the discourse and delimit it in comprehensible units), motivated by the development of several studies on dialect variants and interlanguages, the new term *melodic profile* has been coined and defined, as a series of melodic features that characterise an *accent* (such as dialect or foreign accent). On a paralinguistic level, three types of intonation have been defined: emotional, focus and (im)politeness. And finally, given that some melodic features simultaneously inform us of different characteristics (dialect or interlanguage, linguistic or pertaining to the speaker's intentions or the communicative context), the term *melodic syncretism* has been coined, which will be developed further in future research.

Therefore, the following sections explain the steps to be followed for applying the method, which include all of the developments made to date and the interpretation of melodic data, as well as a final discussion chapter where current research being made is described and some possible additions to the method in the near future are set forth.

## **2. Identification of melodic units**

### *2.1. Phonic hierarchy*

Sounds of speech do not occur indiscriminately but are organised in phonic blocks around accents, which are the true nuclei of discourse and have a hierarchical relationship, called *phonic hierarchy* (Cantero, 2002), shown schematically in Figure 1.

According to this hierarchical relationship, consonant sounds are organised around vowel sounds creating syllables and these join to form words where stress is the most relevant component (e.g. the syllable *par-* in *parque* 'park'). This stress is not only combined with the unstressed syllables of the word but also by adjacent unstressed words, such as prepositions, articles, conjunctions, unstressed pronouns, etc. to create *phonic words* or *rhythmic groups* (e.g. *por el parque* 'through the park', where *par-* is the nucleus).

Phonic words or rhythmic groups also form blocks around the accent, which stands out and is higher in the hierarchy. These create the *phonic group* or "the major phonic unit that segments the discourse and the only one that has its own identity and is not conditional upon other units" (Cantero, 2002: 83). The accent or nucleus of the phonic group, which is the most

relevant, is normally found on the last phonic word (e.g. the syllable *par-* of the phonic group *Ayer paseamos por el parque* ‘Yesterday we walked through the park’).

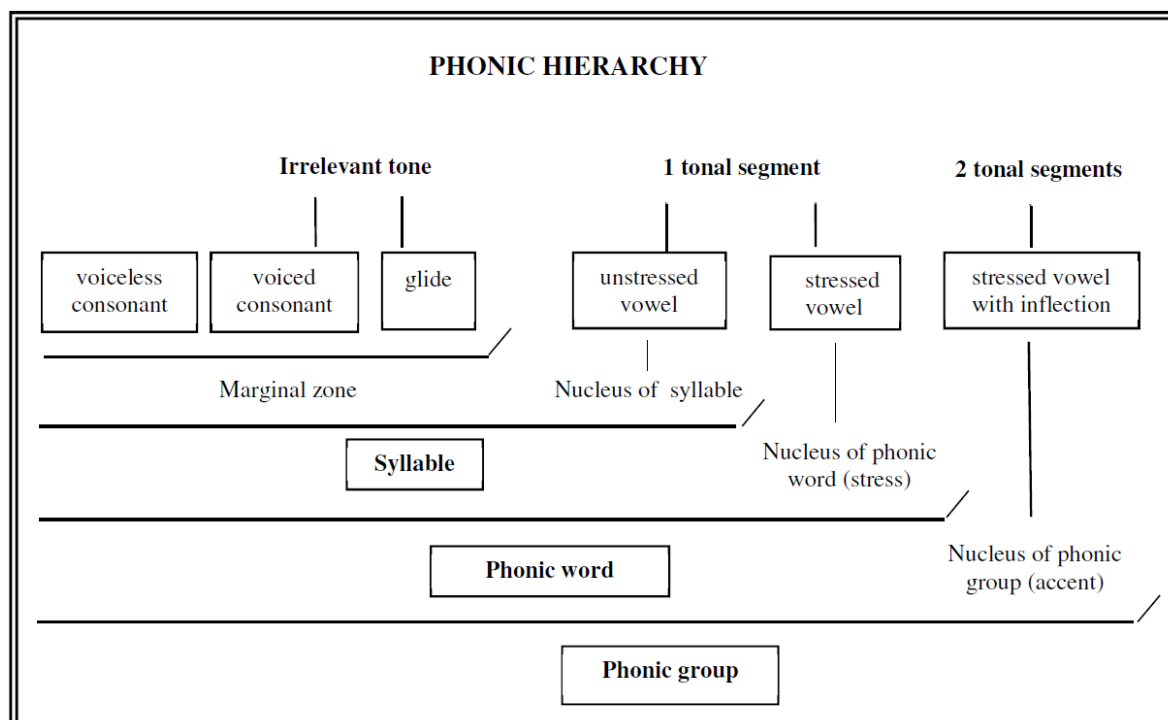


Figure 1. Diagram of phonic hierarchy.

The way of organising speech or phonic hierarchy comprises the personality of each language beyond the mere pronunciation of sounds and has two functions. On the one hand, it integrates the sounds organised hierarchically and forming phonic blocks (syllables, rhythmic groups and phonic groups) by the speaker and, on the other, it delimits these phonic blocks and allows recognition of the words in each block facilitating comprehension of the discourse by the listener.

## 2.2. Melodic units

The phonic group, which we have defined as the major phonic unit of discourse, where sounds of the language are organised hierarchically, contains a coherent lexical-grammatical structure together with a melody. This melody, which represents a significant melodic unit in the intonation of speech, is called the *intonation contour*.

The smallest melodic unit of the intonation contour is the *tonal segment*, in other words, each one of the more or less stable and perceptibly clear phases, which tend to coincide with a *mora* (equivalent to the duration of a short vowel). In this regard, a vowel typically lasts a mora and therefore contains a tonal segment, whilst a long vowel, such as the tonic vowel where the nucleus of the phonic group occurs, tends to last for two moras and as a result contains two tonal segments with different tonal values. This succession of two adjacent and different tonal segments is what we call *tonal inflection*.

The different values of the tonal segments that comprise the melody of the intonation contour follow the hierarchy of the phonic group. Therefore, the only tonally relevant segments are the vowels (these comprise an identifiable tonal segment in the melody of the contour) and, particularly, the stressed vowels (nuclei of rhythmic groups), which stand out among those are unstressed because they present a tonal contrast. Unlike vowels, voiceless consonants are not sounds but noises because they lack phonation. Voiced consonants or glides, despite being considered sounds, tend to be marginal because they always occur associated with a vowel.

Among the different nuclei of rhythmic groups that form a phonic group there is one that stands out for containing the most relevant accent or *nucleus*, melodically characterised for including a tonal inflection with two tonal segments and usually found in the last rhythmic group. Similarly, the first tonic vowel of the contour or *first peak*, is particularly relevant.

Following this hierarchy of the phonic group, which contains the intonation contour, we can see, in Figure 2, that its structure is divided into three parts: the *anacrusis*, the *body* and the *final inflection (FI)*, limited by two accents, the *first peak (1st peak)* and the *nucleus*.

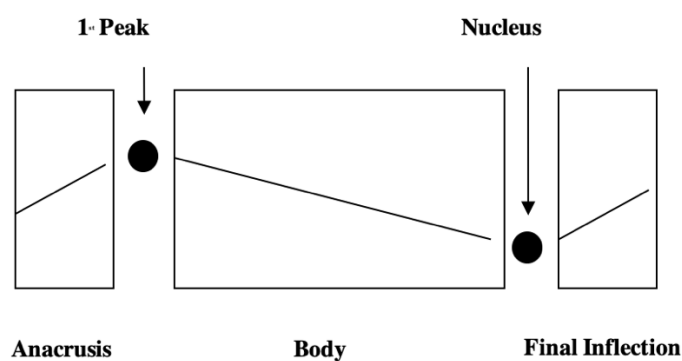


Figure 2. Diagram of the three parts of the contour.

The *anacrusis* is formed by the rising tonal segments prior to the first stressed vowel in the contour, called the *first peak*. Not all of the contours have anacrusis as on occasion it is non-existent. The *body* is the syllables from the first peak to the last stressed vowel in the contour, called the *nucleus*, which consists of a gradual tonal fall of the different segments (although in some emphatic utterances it can be rising) and the point from which the *final inflection* begins. The description of these elements allows us to define the contour melody (especially the description of the *final inflection* and its tonal movements, falling, rising, flat, rising-falling, among others).

To obtain the melodic units or intonation contours of a spontaneous speech corpus, initially it is advisable to choose the utterances that coincide with a turn to speak in the dialogue (normally short and easy to identify) until the researcher has acquired enough confidence in identifying phonic groups and in defining speech melodies. However, it is not always possible to select a turn from the dialogue. For this reason, to identify phonic groups we use the nucleus of the phonic group as a formal criterion for segmenting speech. This nucleus is the

accent that has a tonal inflection and is the most significant part of the contour, usually occurring on the last phonic word.

### *2.3. Establishing the corpus*

To establish a corpus for the purpose of describing a language or a dialect variant, we must obtain models of spontaneous speech that provide an intonation reality, in other words, fragments of speech, uttered by speakers who produce them in a real context. Therefore, the corpus must be obtained outside the laboratory and without any prepared or induced written or spoken support from the researcher. It must also include a significant number of speakers, who vary in age, gender and social-cultural background, who have no speech impediments and who are native speakers of the language.

In certain types of TV programmes such as talk shows, street interviews, debates, discussions and game shows, among others, we have found a resource that meets these basic requirements: multiple native and anonymous speakers of a language who spontaneously express themselves in a context of dialogue. It is worth mentioning that when selecting them we avoid professional speakers and all TV show hosts who follow a script or who are reading. This ensures that the utterances of the corpus are produced by speakers in a spontaneous context.

Another point to be considered in the selection of the utterances is that, in addition to being produced naturally in a real conversation, there must be no overlapping of voices or background noise or music that impede analysis. First interventions by speakers are also disregarded as these could be prepared, selecting those expressed in an uninhibited way and with no form of prior preparation.

Each utterance chosen must be coded, which involves noting down the transcription, its location on the recording, identifying data of the television programme (as well as the time and day it was broadcasted) and, where appropriate, the type of recording, a description of the conversational context in which the utterance was produced and information about the speaker. Once the utterances have been selected and coded, the audio is extracted from the video using suitable software. It is then segmented in Praat (Boersma & Weenink, 2019) and given a code. A maximum of ten utterances per speaker is established so that the corpus is as varied as possible.

The corpus of Peninsular Spanish created by Alfonso (2010), which does not distinguish the origin of the speakers and includes a total of 90 utterances produced by 37 speakers, meets these requirements. Later, corpora were created by Ballesteros (2011) and Mateo (2014), each including samples from five different regions: Asturias, Navarre, the Basque Country, Castile-Leon and Madrid in the case of Ballesteros and Andalusia, the Canary Islands, Castile-La Mancha, Extremadura and Murcia in the case of Mateo. In total the two corpora include 373 hours of recordings, 2941 utterances and 814 speakers (Cantero, 2016). These corpora are the ones used to characterise the linguistic intonation of spontaneous speech in



Cantero et al. (2005), Cantero & Font Rotchés (2007) and Font-Rotchés & Mateo (2013), as well as the dialect intonation in Ballesteros (2011) and in Mateo (2014).

Also of note is the Catalan corpus by Font-Rotchés (2007), which includes 46 hours of recordings, 580 utterances and 160 speakers. This corpus, apart from being useful for describing the linguistic intonation of Catalan and emphatic intonation (Font-Rotchés, 2011), has also been used to describe the intonation of (im)politeness (Devís & Cantero, 2014). Following this same protocol, corpora have also been created of spontaneous speech of German (Torregrosa, 2015), Brazilian Portuguese (Sena, 2013; Araújo, 2014) and Chinese (Kao, 2011).

Other types of corpus also exist that have different objectives, such as the description of an interlanguage, which implies creating a corpus following different criteria to those indicated for spontaneous speech. In studies carried out by our group on interlanguage, corpora have been created using other strategies. In research on Spanish spoken by Taiwanese (Liu, 2005), Hungarians (Baditzné, 2012), Brazilian Portuguese (Fonseca, 2013) and Swedish (Martorell, 2010), conversations between interlanguage speakers have been recorded using microphones outside a laboratory environment. This involved Spanish learning classroom activities and other different situations, focusing the interest on the content: personal experiences, opinions of their relationship with the Spanish language and culture or hot topics in current affairs, among others, with the aim of obtaining conversations that were as spontaneous as possible.

Lastly, with regard to describing models of speech by television news readers, advertising actors and current affairs programme presenters, who produce a very formal and media-oriented model of speech from a written text, a corpus has been created taken from recordings of television broadcasts of each one of these genres (see Font-Rotchés & Paloma, 2012; and Torrent, 2015 for Catalan; and Font-Rotchés & Machuca, 2010 for Spanish).

### **3. Acoustic phase**

#### *3.1. Determination of relevant frequency values*

Once the utterances have been separated, we establish the  $F_0$  value of each vowel, of each tonal segment. This involves differentiating the relevant frequency values (the tonal segments) from the irrelevant values (the  $F_0$  of the sonorous consonants and of the glides). This is done using Praat analysis and synthesis software (Boersma & Weenink, 2019), which is a reliable acoustic analysis instrument.

We identify the vowels (guided by a sonogram) and establish their value (the average of the  $F_0$  values of the vowel). Figure 3 shows an example of the extraction of frequency values of the utterance *Tiene usted hijos?* ‘Do you have children?’ obtained by using the Praat application. The dark areas of the spectrogram correspond to each one of the vowel segments (or diphthongs) and for each one of them the  $F_0$  value in Hz is obtained from the average value of the vowel’s  $F_0$ .

When a stressed vowel contains a tonal inflection with a tonal distance of at least 10%<sup>3</sup> (100% is equivalent to an octave of the musical scale), a value of two tonal segments must be established (see syllable *-hi-* in Figure 3 and values in Table 5). This constitutes the inflection or three segments if it is a circumflex inflection (see syllable *-xò* in Table 2), which can be found in any part of the contour. These values are calculated from the initial and final stable values or from the extreme values of the inflection (if there is no tonal stability).

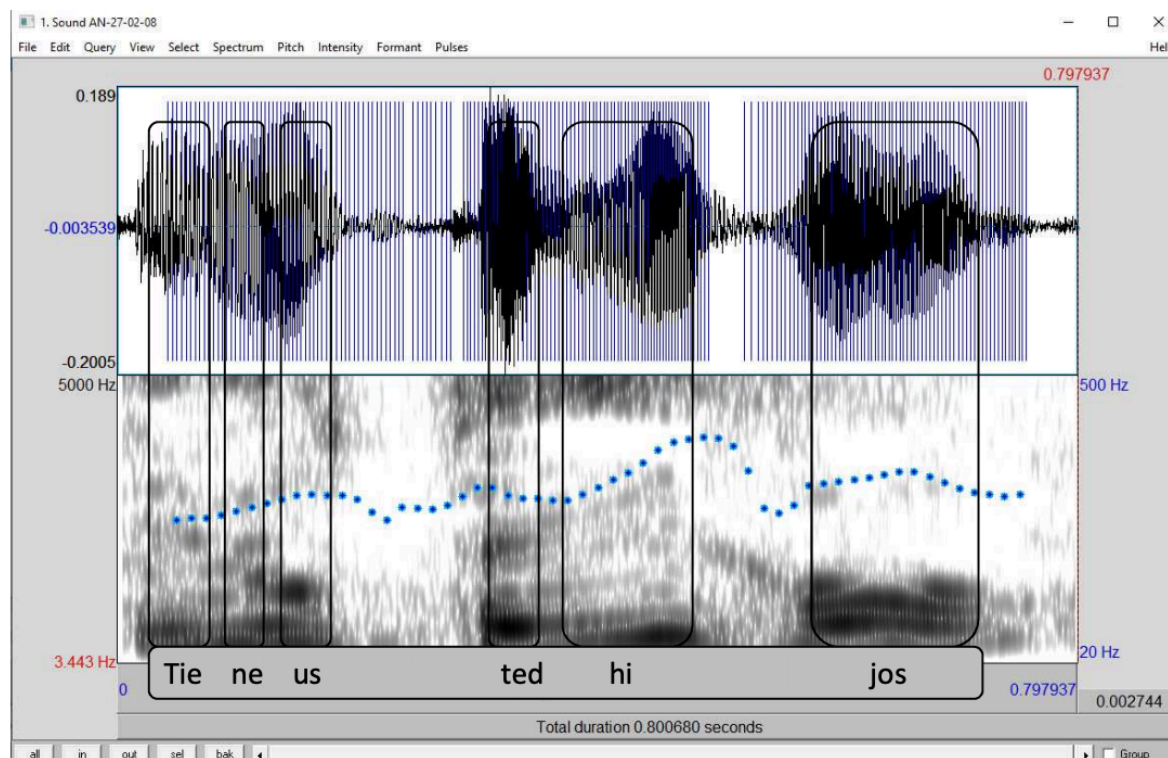


Figure 3. Extraction of frequency values of the utterance *¿Tiene usted hijos?* ‘Do you have children?’

Furthermore, when the inflection ends with a vowel followed by a voiced consonant (specially a nasal, a lateral or a vibrant) with a tonal distance of at least 10% (see an example in Table 6), this consonant usually represents the last tonal segment of the inflection (see perception tests in Font-Rotchés, 2007).

Initially, to describe linguistic intonation we only used very marked tonal inflection data, with more than 20% of movement. However, for research on dialects and interlanguages the need to use inflection data with a tonal distance of 10% for intonation relevance became evident. It is in these small features where some of the characteristics lie that allow us to describe and distinguish one dialect from another (see Ballesteros, 2011; Mateo, 2014) or to describe the transfer features of the interlanguage from the language of origin, as is the case of Brazilian Portuguese (Fonseca, 2013) and Italians (Devís, 2011b) speaking Spanish.

With regard to the extraction of tonal values for the final inflection, in other words, the inflection that occurs in the last tonic syllable of the contour, we can see that there are small

<sup>3</sup> This tonal distance has been verified in different perception tests carried out by the Applied Phonetics Lab of the University of Barcelona and applied in MAS research since 2011.

variations according to whether the word is oxytonic, paroxytonic or proparoxytonic. If the word is oxytonic, it may be that the accented syllable has a tonal inflection with two values — we distinguish the second with one asterisk(\*)—; this is the case of the syllable *Stadt* 'city' in German, found in the utterance *Schon viel Spass gehabt in der Stadt?* 'Did you have a good time in the city?' (see Table 1). When this syllable has only one value, the final inflection would begin on the preceding syllable.

Utterance	<i>Schon</i>	<i>viel</i>	<i>Spass</i>	<i>(ge)habt</i>	<i>in</i>	<i>der</i>	<i>Stadt</i>	<i>Stadt*</i>
Hertz	149	172	190	176	164	146	<b>169</b>	<b>195</b>

Table 1. Example of German whose last stressed vowel has two values(4).

Table 2 shows the Catalan example of *Bé, crisi és això!* 'Well, crisis is that!' where the last stressed vowel, -ò, has a tonal inflection with three values distinguished by asterisks (\*) —one if it is the second value and two if it is the third.

Utterance	<i>Bé</i>	<i>cri</i>	<i>si</i>	<i>és</i>	<i>ai</i>	<i>xò</i>	<i>xò*</i>	<i>xò**</i>
Hertz	328	347	400	365	340	<b>316</b>	<b>426</b>	<b>292</b>

Table 2. Example of Catalan whose last stressed vowel has three values.

On other occasions, the final inflection begins with a stressed vowel and ends with an unstressed vowel, for example, in final inflections that coincide with a word with the stress on the penultimate or antepenultimate syllable. See Table 3, the Catalan example of *Portes vambes?* 'Do you wear sneakers?', where the stress is on the penultimate syllable, and Table 4, the German example of *Kommt das aus Amerika?* 'Is this from America?', where the stress is on the antepenultimate syllable.

Utterance	<i>Por</i>	<i>tes</i>	<i>vam</i>	<i>bes</i>
Hertz	103	97	<b>110</b>	<b>208</b>

Table 3. Example of Catalan in which the final inflection takes place in a paroxytonic word.

---

4 Certain elements of the spoken chain are shown in brackets, such as *(ge)* of *(ge)habt*, to point out that these elements are elided or assimilated and that they do not have an acoustic correspondence in spectrographic analysis.

Utterance	<i>Kommt</i>	<i>das</i>	<i>aus</i>	<i>A</i>	<i>me</i>	<i>ri</i>	<i>ka</i>
Hertz	259	323	269	234	<b>217</b>	<b>296</b>	<b>443</b>

Table 4. Example of German in which the final inflection takes place in a proparoxytonic word.

It is also possible for the accent to be found on the penultimate syllable and that this syllable or the following unstressed syllable could have two values, therefore giving us a circumflex final tonal inflection with three values. This is the case of the Spanish example in Table 5 of *¿Tiene usted hijos?* 'Do you have children?' in which the final inflection has three values: the first and the second in the stressed vowel, *hi*- (293 Hz and 376 Hz), and the third in the next unstressed syllable: *-jos* (305Hz).

Utterance	<i>Tie</i>	<i>ne</i>	<i>us</i>	<i>ted</i>	<i>hi</i>	<i>hi*</i>	<i>jos</i>
Hertz	256	270	289	295	<b>293</b>	<b>376</b>	<b>305</b>

Table 5. Example of Spanish of a circumflex rising-falling final tonal inflection.

There is another Spanish example in Table 6 *¿Tú tienes internet, Carmen?* 'Do you have internet, Carmen?' in which the final inflection also has three values: the first in the stressed vowel, *Car*-, and the second and the third in the next unstressed syllables: *-men* (123 Hz) from the vowel and *-men\** (159 Hz) from the voiced consonant *-n*.

Utterance	<i>Tú</i>	<i>tie</i>	<i>nes</i>	<i>in</i>	<i>ter</i>	<i>net</i>	<i>Car</i>	<i>men</i>	<i>men*</i>
Hertz	234	244	210	198	189	178	<b>171</b>	<b>123</b>	<b>159</b>

Table 6. Example of Spanish of a circumflex falling-rising final tonal inflection.

### 3.2. Standardisation of the frequency data and graphic representation

The absolute values obtained in this analysis phase (extraction and determination of  $F_0$  of the tonal segments, measured in Hertz) do not yet constitute the contour melody because they still have the individual phonic characteristics of each speaker, and therefore comparable generalizations can not be extracted. In this sense, the melody uttered by a man that constitutes the succession of absolute values 100Hz–200Hz is not the same as one of values 200Hz–300Hz, uttered by a woman, although the difference is identical: 100Hz in both cases. In relative terms, the tonal interval is different. In the first case, there is a difference of 100% but in the second case, there is a difference of 50%.

Therefore, this data must be relativised, by transforming these frequency values into percentage data that reflect the tonal relationship existing between one absolute value and the

one immediately after it. The formula for the relativisation of the frequency values<sup>5</sup> is as follows:

$$(1) \text{ Relativisation formula: } REL_{v_{n+1}} = [(ABS_{v_{n+1}} - ABS_{v_n}) / ABS_{v_n}] * 100$$

The variable  $ABS_{v_n}$  corresponds to the absolute value of the vowel segment of reference (in Table 7, 259 is the value in Hertz from the first syllable, *Kommt*).  $ABS_{v_{n+1}}$  is the frequency value of the next vowel segment at the reference value and from which the relative difference in % with respect to the preceding vowel segment is determined (in Table 7, 323 is the value in Hertz of the next vowel segment).  $REL_{v_{n+1}}$  or the relative value expresses the melodic difference as a percentage between the absolute value of reference and  $ABS_{v_{n+1}}$  or the value immediately after.

Starting with the arbitrary value 100%, we obtain the relative values of each contour that make up the algorithm, which is the expression of the melody. Table 7 shows the percentages obtained: 100%, 25%, -17%, -13%, -7%, 36%, 50% (the rise is expressed as a positive percentage and the fall as a negative percentage). Some authors express this interval using the semitone (st) as a unit, but we consider the use of percentages to be more intuitive.

Utterance	Hertz	Percentage	Standard Curve
<i>Kommt</i>	259	100%	100
<i>das</i>	323	25%	125
<i>aus</i>	269	-17%	104
<i>A</i>	234	-13%	90
<i>me</i>	217	-7%	84
<i>ri</i>	296	36%	114
<i>ka</i>	443	50%	171

Table 7. Absolute (Hertz), Relative (Percentage) and Standard values (Standard Curve) of the contour.

To draw the graphic representation of these melodies, we convert the percentages obtained into standard values: starting, for example, with the value 100 (an arbitrary value). By applying (see Standard Curve on Table 7) the rise percentage seen in the second segment, 25%, we will obtain a second value, 125; to this we apply the fall percentage of the third segment, -17%, to obtain the third value, 104; and so on until the last segment. The formula for standardising the relative values is as follows:

$$(2) \text{ Standardisation formula: } ST_{v_{n+1}} = ST_{v_n} + [(REL_{v_{n+1}} * ST_{v_n}) / 100]$$

<sup>5</sup>Drawn from Torregrosa and Font-Rotchés (2017:1318).

$ST_{v_{n+1}}$  corresponds to the standardised value of  $REL_{v_{n+1}}$  with respect to  $ST_{v_n}$ . This last value is the standard value that acts as reference value. Using the standard values of each utterance, we can plot the graphic representation of each contour in Excel to compare them.

In Figure 4, showing an utterance produced by a female speaker, the dotted line represents the absolute values in Hertz, with a very marked rise at the end, whilst the lower solid line represents the standard values with a very gentle rise. In contrast, in Figure 5, uttered by a male speaker, once the standardisation has been done, the absolute and standard values are very similar.

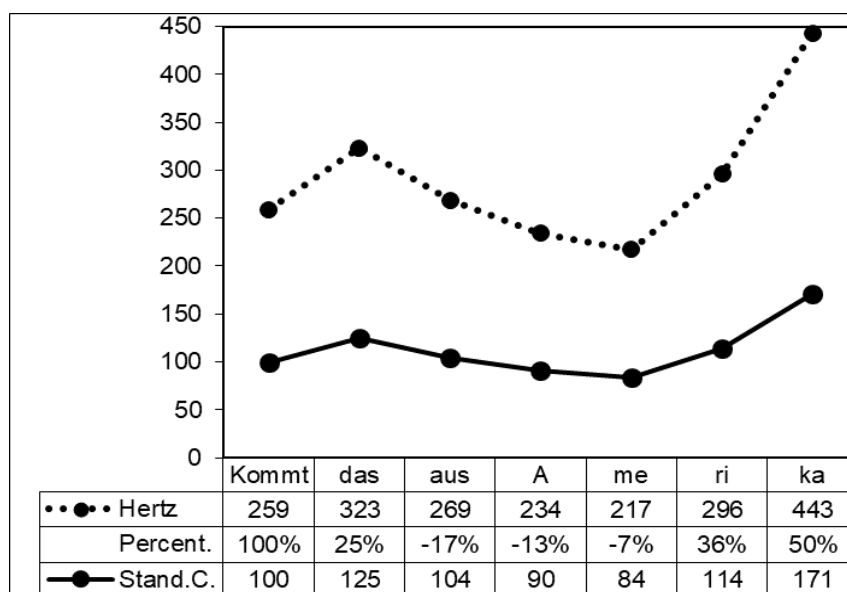


Figure 4. Absolute and standard values of the melody *Kommt das aus Amerika?* 'Is this from America?'

If we compare the absolute values of both graphs (dotted lines) we can conclude that they have a very similar final inflection; the rise that begins on the last tonic syllable, *-me* and *-bo(ren)*, respectively, and that this is the most significant part of the contour. In contrast, if we compare the standard values (solid lines) we can see that both questions have a final inflection with a very different rise: 104% (from the distance between *-me-*, 217, and *-ka*, 443) in Figure 4 and 148% in Figure 5.

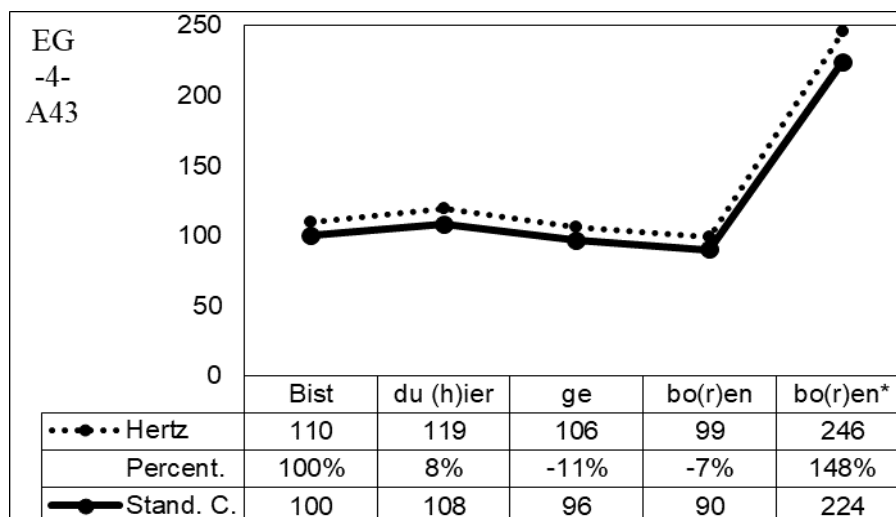


Figure 5. Absolute and standard values of the melody *Bist du hier geboren?* ‘Were you born here?’

With the application of this first phase of the method, or acoustic phase, we obtain the contour standardisation, which can now be compared and classified, regardless of the age, gender or any other characteristic of the speaker, as all the micromelodic variations have been removed and the values have been standardised. This process does not standardise the length of the different contours but isolates and replaces them with the number of tonal segments and their relative value, which is the only relevant data in a melodic analysis. The differences between two contours of different length should be understood from a melodic perspective, regardless of whether the melody contains a word or many words, whether it has a longer or shorter utterance, etc.

### 3.3. Praat Script

This entire acoustic phase process discussed in sections 3.1 and 3.2, which includes the extraction and standardisation of the data and the creation of graphs, is done manually. To speed up the process, Mateo (2010a, 2010b) designed a script for Praat so that this could be done semi-automatically.

To apply the script for Praat, first the corpus must be established, which can be in any language, dialect variety or register, taken from recordings of short and coded utterances.

Once the script has been installed in Praat (refer to Mateo, 2010b for instructions), the tonal segments for each audio file are established and labelled based on the creation of a TextGrid in Praat. This is a key point in the process as it requires precision and criteria in acoustic analysis. The work must be done using the expanded sonogram (selecting the *Sel* option) and in *Pitch settings* the option of speckles as drawing method must be enabled so that the limits of each tonal segment can be established with the greatest possible precision. At the same time the sound wave of the oscillogram must be considered, revealing the moments of maximum intensity that correspond to the tonal segments, in other words, to the vowels.

To create the labelling in TextGrid (see Figure 6), before the first segment of the utterance we must establish if it is a male (m) or female (f) voice. We can then begin to segment the first vowel of the utterance, transcribing the first syllable. The aim is to search for the stable values of the vowel and mark the limits, which must be located after the first speckle of the vowel and before the last one, to ensure that the values are correct. To mark the limits of a syllable when it is a tonal inflection, either because the vowel has two or three values or because the syllable ends in a voiced consonant, we use the same criteria because the script will use the two, or where appropriate, three values.

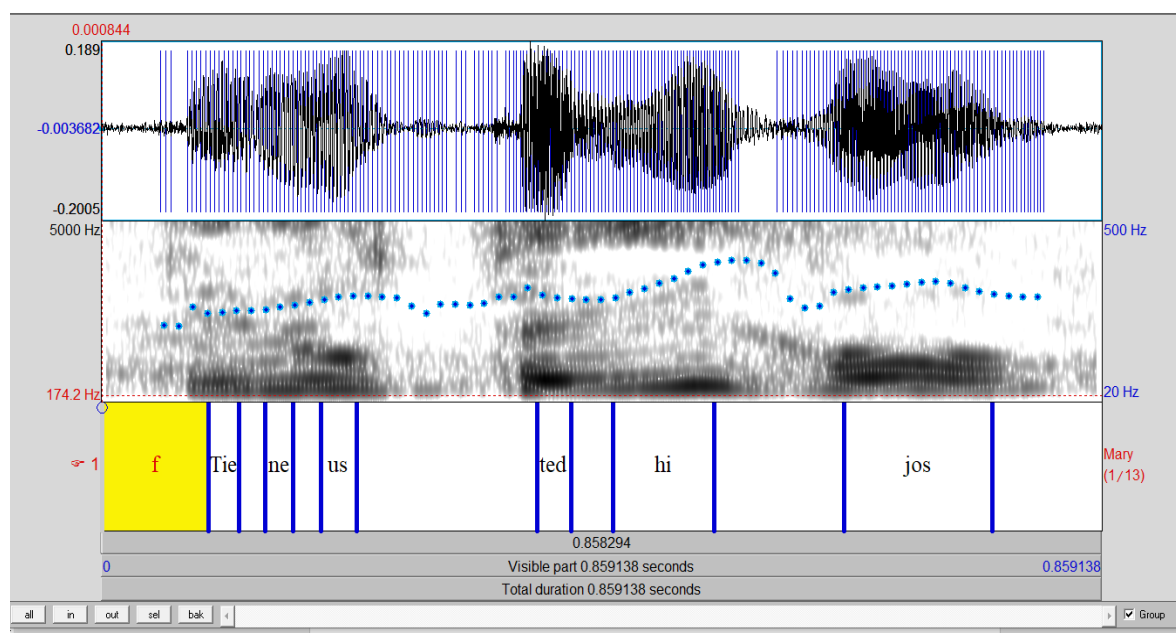


Figure 6. Sonogram and TextGrid of the utterance in Spanish *¿Tiene usted hijos?* 'Do you have children?'

Thus, in Figure 6 we can see that the utterance is spoken by a woman (f) and that the first syllable is *Tie-*. We have marked the limits taking into account the stable values of the diphthong *-ie-* and the point of intensity on the oscillogram, avoiding using values belonging to the next voiced consonant, *n-*, which is marginal. The second segment is the *-e-* of *-ne*. The voiced nasal *-n-* prior to the vowel is not a relevant value, as previously mentioned. Hence, the limits of the vowel are established respecting the same criteria. The same process is carried out for the other segments. The penultimate vowel *-i* has a tonal inflection with two values; similarly, we establish the limits at the beginning and the end of the vowel and the script will extract the two values.

The next step involves extracting the data, using the audio and the TextGrid of each one of the utterances. This is done by enabling the first of the scripts. This step must include a revision of warnings, such as values for segments that Praat cannot obtain or segments with three values that must be verified, among others, all of which will be flagged by the script. The subsequent modifications to the data, which may be necessary, must also be revised.

Once the correct data has been guaranteed, the standard curve of the utterance is created by enabling the second of the scripts. The graphs of the standard curve are generated using a



macro in an Excel document provided by the script. Once the graphs have been created, comparisons can be made and the intonation models and patterns established<sup>6</sup>. See Figure 7, showing a graph obtained by applying the script.

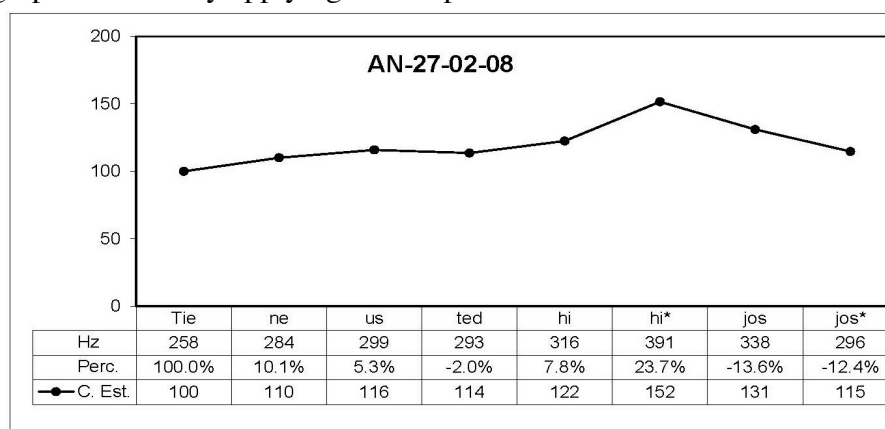


Figure 7. Intonation curve of the Spanish utterance ¿Tiene usted hijos? ‘Do you have children?’

In short, the Praat script is a reliable and simple tool that can be used to automate the data collection process for the analysis of intonation in any language, saving researchers a considerable amount of time. However, before using the tool, acoustic analysis should be practised manually on a significant number of utterances to obtain criteria and to understand its characterisation.

After analysing all the contours and presenting them graphically, they must be classified into different groups according to the path of the final inflection (rising, falling, rising-falling, falling-rising, among others). For example, the representation of the two German contours (Figure 4 and Figure 5) with a rising final inflection, shows that both belong to the same melodic class: contours with rising final inflection. In some cases, the melodic features of the body and the first peak must be taken into account, such as the Spanish interrogative pattern, the rising body and final inflection pattern (Font-Rotchés & Mateo, 2013; 2017). Subsequently each group of contours must be organised according to the percentage of tonal movement in the final inflection, from greater to lesser, in order to apply the perception phase.

#### 4. Perception phase

##### 4.1. Analysis validation

In the perception phase, first the standard values of an intonation contour are validated in the voice synthesizer by ensuring that it has the same melody as the original, but with a different voice tessitura. We use the option "To manipulate" to open each contour in Praat to obtain this synthesised copy. After opening a contour ready for manipulation (See Figure 9), we erase all

<sup>6</sup> For more detailed information about the scripts, see Mateo (2010a).

original data of the contour and replace it with our standardised data (See Figure 10)<sup>7</sup>. We can check the validation of our results with a series of perception tests submitted to the listener's judgement from an exact copy (thanks to voice synthesis) of the analysed sentences.

This validation of the melodies obtained during the perceptive phase of the analysis allows us to be completely sure that the analysis has been done correctly and that the obtained data are objective.

It also allows us to check that the validated melodies have been reduced to their most basic expression. Standardisation of the melodic curve provides us with the melody itself, regardless of the speaker's tessitura and clean of any superfluous micromelodic variations. This method is very similar to *close copy* by authors from the Dutch school ('t Hart, Collier & Cohen, 1990).

Following is an example, taken from research that defined the melodic patterns of interrogative intonation in Brazilian Portuguese (Cantero & Font-Rotchés, 2013). The utterance *Duzentos cinquenta os dois?* 'Two hundred and fifty for the two?', where the speaker is giving the price for two calves, revealed the following analysis (see Figure 8).

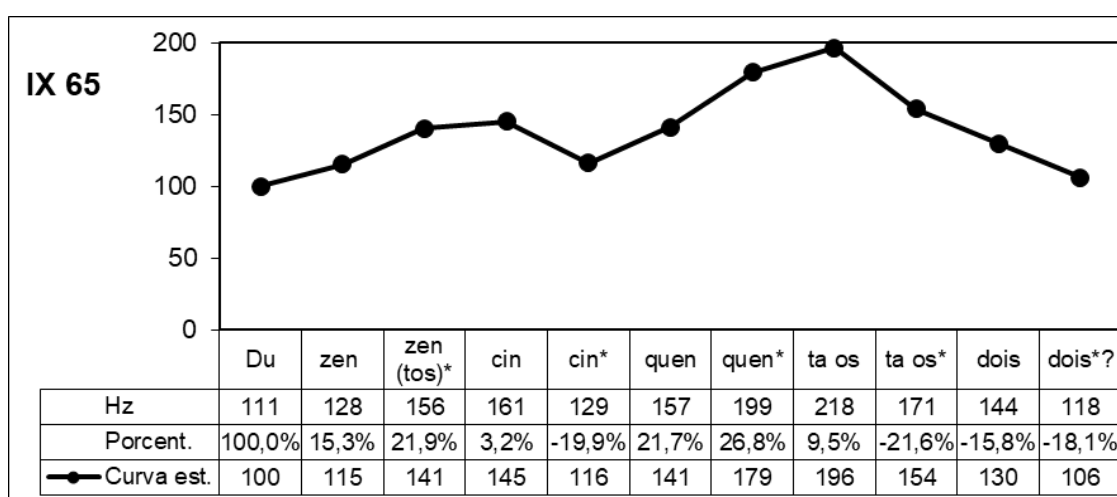
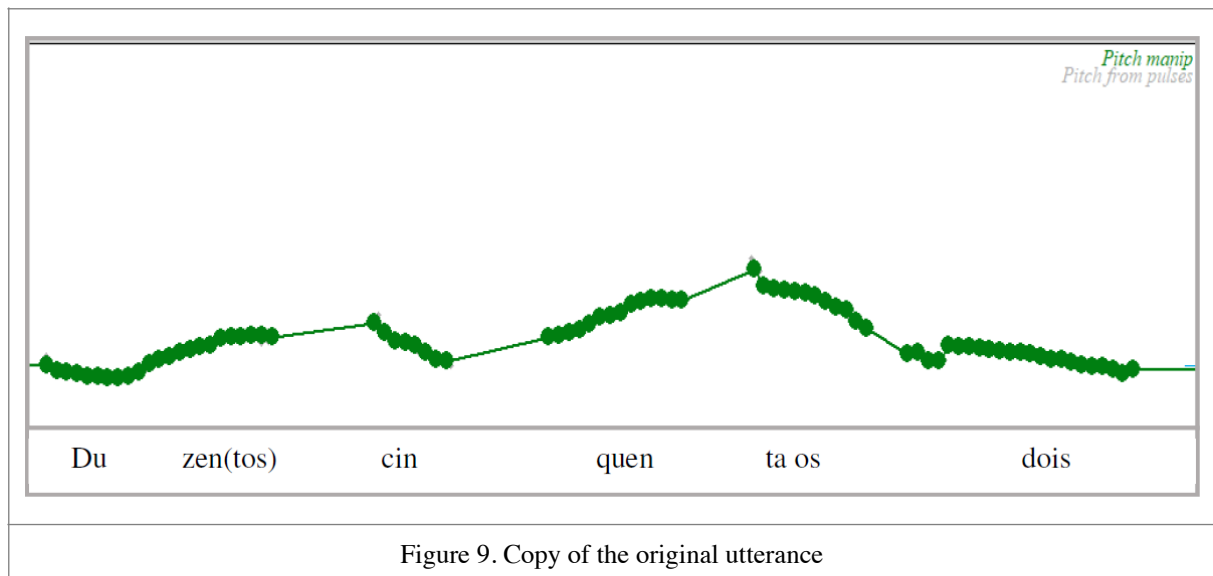


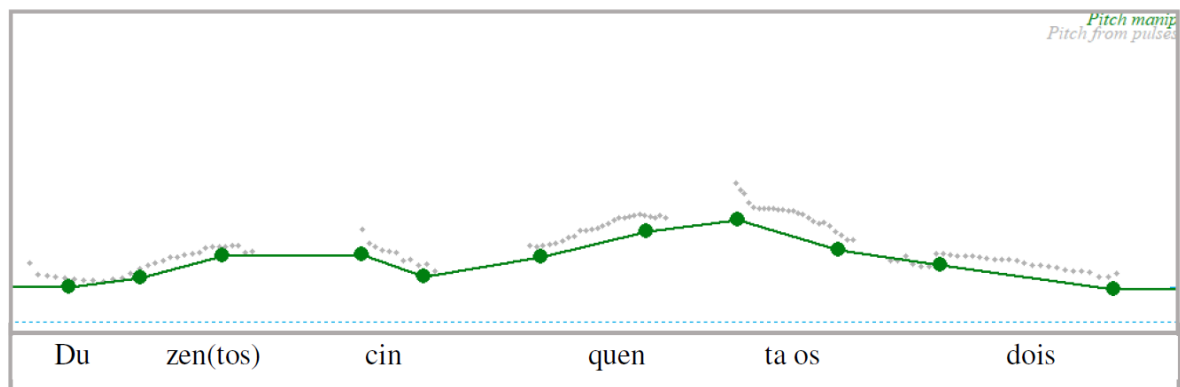
Figure 8. Intonation curve of the Brazilian Portuguese utterance *Duzentos cinquenta os dois?* 'Two hundred and fifty for the two?'

During the Manipulation process in Praat, the copy of the values is obtained (see Figure 9):

<sup>7</sup> To erase the original data, we must select it, click on 'pitch' and then 'remove pitch points' at the top of the screen. We can also replace it with our standardised data adding a dot where each vowel is located by clicking 'pitch' and 'add pitch point at cursor'.



To this copy the standardised values provided by the analysis are transferred (see Figure 10).



We can now record the synthesis of the result and perceptively validate the melody.

#### 4.2. Identification of the melodic features

Once the melody of the intonation contours, obtained during the acoustic phase, has been validated, we can begin to identify its melodic features: the first peak, body or declination and the final inflection.

Here it is important to distinguish between the different function that the contour's melodic features have in the three levels of intonation analysis:

- Prelinguistic intonation
- Linguistic intonation
- Paralinguistic intonation

*Prelinguistic intonation* refers to the melodic features that allow the discourse to be structured and integrated in comprehensible units. The *melodic profile* (Cantero & Devís, 2011) of a

language and of a dialect consists in this set of melodic features that characterise an *accent* (such as the dialect or foreign accent). The melodic profile of standard Peninsular Spanish, for example, coincides with the pattern of the intonation contour shown in Figure 2:

- a high first peak, whose centre is normally the first tonic vowel of the group (or the first post-tonic vowel);
- a body or declination that falls evenly, with gradual inflections that systematically begin on a tonic vowel;
- a final inflection that starts on the last tonic vowel of the group and can be falling (with values often greater than -15%) or rising (with values that can reach more than +120%).

*Linguistic intonation* refers to the level of intonation analysis that allows units that are different on a phonological level to be defined. For example, neutral intonation of interrogative, emphatic or suspended intonations, whose specific manifestations are the *melodic patterns*. In other words, the melodies commonly used by native speakers of a language in their communicative interchanges and which can become models of production and perception.

*Paralinguistic intonation* refers to the melodic features that provide a meaning related to the speaker's intentions or to the communicative context, but which do not form linguistically coded units (although they can form semi-stable codes: see Cantero (2014)). We consider paralinguistic intonation to have a role in at least three well-defined areas:

- *Emotional* intonation, including the expression of speaker's personal characteristics, his or her emotions and affection.
- *Focus* intonation, which centres on the actual melody of the utterance drawing attention to it, focusing on it or on a part of it to underline its relevance in the discourse. A good example is the intonation of irony.
- *Politeness* intonation, which is focussed on the speaker and calls upon him or her to intensify or attenuate the effect of the message and, in general, to look for confrontation or cooperation.

In any event, there are very few melodic features of an intonation contour, therefore these functions are assumed by the same melodic features simultaneously. This is what we call *melodic syncretism*. One same final inflection, which marks the end of the utterance and its *linguistic* form (for example, an interrogation), can at the same time mark the speaker's dialect accent (*prelinguistic*), express his or her concern (*emotional*), focus on the question so that in this context it is clearly an ironic question (*focus*) and reveal impoliteness to the listener (*politeness*).

### 4.3. Forming the hypotheses

To establish the generalisations that allow us to understand how speech melodies work, we use those obtained during the acoustic phase, which are validated perceptively, identifying and subsequently isolating the melodic features we wish to test. Only a perception test can guarantee which specific melodic feature is responsible for which prelinguistic, linguistic or paralinguistic, communication or discourse function.

For example, to determine the feature of the melody that is responsible for giving the utterance an interrogative value, we can isolate each one of the melodic values found in the analysis of corpus utterances that have been validated as melodies that confer an interrogative value. If we confirm that the melodic feature common to all of them is a rising final inflection, we can determine that this feature is the one we must test in the experiment. Therefore, based on the values that the melodic analysis provides for the questions analysed, we can formulate a hypothesis, such as: *a rising final inflection conveys an interrogative value as of a specific percentage rise*. This hypothesis must be counter-verified in the perception test.

### 4.4. Perception tests

Once the melodic feature to be tested has been identified, it is modified using Praat's synthesis routine (*manipulation*). Using the previous example, we generate various identical utterances where only the value of the rise in the final inflection is modified.

We can modify an utterance in the corpus directly, as long as there are no grammatical, lexical, contextual parameters, etc. that allow the sense or function of the melodic feature hypothesised. For example, we would not use phrases with an interrogative particle, however, we could choose a neutral utterance or create one ourselves by copying the standardised data and modifying only the feature to be tested (in this case, the final inflection). Therefore, the modified melodic feature becomes the variable we wish to counter-verify in the perception test.

This test involves asking a significant group of listeners to react to the manipulated utterances: various unmanipulated utterances, mixed with the manipulated ones (for example, with rises of 20%, 30%, 40%, 50%, etc. up to 120%) and with other non-experimental utterances, in different perception sessions, with various groups of listeners. The results allow us to determine the percentage of rise that led listeners to the conclusion that the utterance is interrogative.

This procedure allowed the melodic patterns of Peninsular Spanish (Cantero & Font-Rotchés, 2007) and Catalan (Font-Rotchés, 2007) to be established for spontaneous speech.

In Figure 11 we can see how the utterance used in the example in section 4.1 (in Figure 8, 9 and 10) is modified:

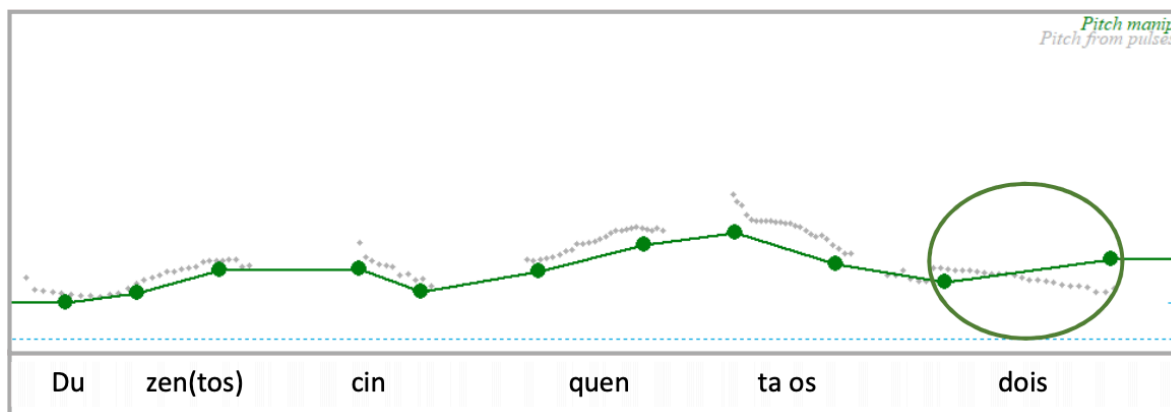


Figure 11. Synthetized copy obtained after modifying the last value (in the circle).

In the original final inflection of the utterance there is a fall of 18.1% (Figure 8). In the standardised copy (Figure 10) we manipulate the last value, *dois*\* and move from 106 to 158, creating a rise of 21.5% (Figure 11). In this example, the final rise by 21.5% did not allow us to identify the utterance as interrogative (as expected) but did as of a rise of 30% (and up to 50%). In Brazilian Portuguese, therefore, we can determine final rises with very slight interrogative values, which in Spanish or in Catalan are never understood as interrogatives: in Spanish, only as of 70% and in Catalan as of 80% (Cantero & Font-Rotchés, 2013).

## 5. Discussion and concluding thoughts

The theoretical assumptions of our method are not phonological, but exclusively phonetic, acoustic and perceptive. We determine the tonal segments that form the contour, we standardise their frequency values to express the melody and we validate them perceptively; we identify the melodic features that give the utterance different values (linguistic, pragmatic, discourse...) and we subject them to tests to determine them; we work with tonal intervals expressed as percentages (equivalent to semitones, but more easily and more intuitively handled).

Therefore, it is an independent method, which is at the same time compatible with any phonological focus (structural, notional, autosegmental, etc.), particularly with methods of phonological labelling such as ToBI. The melodies resulting from the analysis have the advantage of allowing any speaker to be compared, of handling large quantities of data and of establishing reliable statistics and precise dispersion margins thanks to the exact quantification of the intervals.

This makes it a solid method: it allows the analysis of spontaneous speech, it requires no preliminary alignments and works well with all types of corpus. It is also objective: the perception test ensures that all of the standardised values and resulting melodies are reliable.

Compared to other intonation analysis models, MAS allows:

- Study of intonation from exclusively phonetic levels of analysis, which do not require the use of other levels of linguistic or grammatical analysis, nor do they depend on them.
- Work with a formal analysis method, based on the acoustic and perceptive analysis of speech. This analysis provides a very detailed description of the intonation from a strictly phonetic viewpoint, regardless of its phonological interpretation.
- Characterisation not only of standard melodies of a language or dialect (whose phonetic generalisation is possible in the perceptive, experimental phase of the method) but also of the individual variants of a single speaker and even the characteristics of pathological speech.
- Work with spontaneous speech: although any type of corpus can be used in the analysis, the solidity of the method allows large volumes of data and a wide range of anonymous speakers to be handled.
- Use of the open phonetic analysis method, allowing any type of notation, analysis or subsequent transcription: pragmatic, grammatical, lexical, as well as any type of phonological labelling. For example, the Autosegmental-Metrical model (AM), which has a labelling convention known as *Tones and Break Indices* (ToBI), is a phonological labelling method for intonation, not a phonetic analysis method. In this sense, it is completely compatible with the MAS method. See two examples in Figures 12 and 13 which demonstrate compatibility.

According to the MAS method, the melodic contour of Spanish *¿Tiene usted hijos?* (see Figure 12) is characterized by a flat body without anacrusis and first peak and a rising-falling final inflection. This flat body is phonologically transcribed as L\* in the AM model because it is phonetically realized as a low plateau at the minimum of the speaker's range.

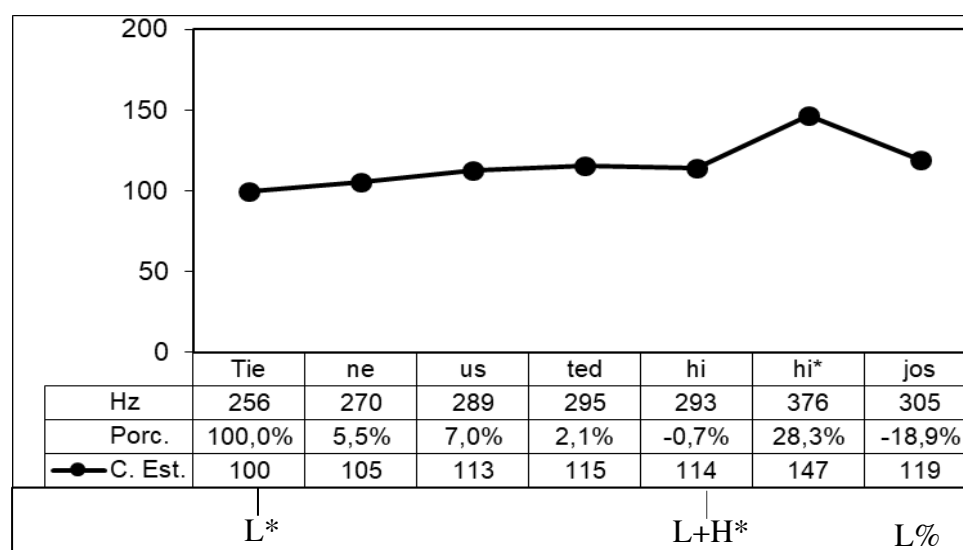


Figure 12. Contour phonetically analysed with the MAS method and labelled with the Sp ToBI system.

We can see another example in Figure 13. The melodic contour of the question *¿Usted es de Gran Canaria?* ('Are you from Gran Canaria?') starts with an anacrusis with a tonal rise of 68% until the first peak, *es*, which is displaced in the unstressed segment after the first stressed segment of the contour, *-ted*. After that, the melodic line falls abruptly until the nucleus, *-na-* and ends with a falling final inflection of 14.6%. These phonetic features can be complemented with a phonological description in the AM model: an L+ <H\* prenuclear pitch accent, which involves a rise with a displaced F<sub>0</sub> peak, an HL\* nuclear accent, realized as a F<sub>0</sub> fall within the accented syllable, and an L% boundary tone.



In fact, carrying out a detailed acoustic analysis of the melodic features of speech is a laborious task and this is undoubtedly the main disadvantage of the method. For this reason, we suggest simplifying the analysis process using the Praat script described in section 3.3, which we continue to optimise to facilitate the task. Many authors have preferred to limit themselves to an intuitive approach to the melodic curve, expressing only the pitch levels that may have a phonological relevance, the assumed stresses and the limits of the utterance. This



is a reasonable method when there is no need for a detailed analysis of the melodies but simply their phonological expression. However, just as the acoustic analysis of the segments cannot be replaced by their simple transcription, on occasion there is a need for a detailed melodic analysis. For example, to characterise the intonation of foreign speakers of a language, to describe dialectal intonations, to distinguish personal characteristics of a speaker, or to specify the melodic features that have pragmatic or discourse functions, etc. In all of these cases, where there is a need to go beyond the previously deduced phonological characterisation, a detailed melodic analysis must be made. This is also the case when the researcher wishes to establish linguistic generalisations based on the mass analysis of new acoustic data.

Since it was proposed (Cantero, 1999) and to date, the method has been used with success in different studies, by our research team<sup>8</sup> and by different international research groups (particularly South American).

At our research centre, the University of Barcelona's Applied Phonetics Laboratory<sup>9</sup>, we have created a wide corpus of spontaneous speech for Peninsular Spanish and Catalan, among other languages (see section 2.3).

Research on the following areas has been carried out:

*Prelinguistic intonation*: dialect accents of Peninsular Spanish (Ballesteros, 2011; Ballesteros & Font-Rotchés, 2019; Mateo, 2014) and Catalan (Pons, 2007); as well as for various interlanguages (foreign accents): the intonation of Spanish spoken by Swedish (Martorell, 2010), Hungarian (Baditzné, 2012), Italian (Devís, 2011b), Brazilian (Fonseca, 2013) and Chinese (Cortés, 2004; Liu, 2005) speakers; and the study of speech demarcation units for Spanish (Cabedo, 2013; Martínez, 2015).

*Linguistic intonation*: melodic patterns and dispersion margins for Peninsular Spanish (Cantero & Font-Rotchés, 2007), Cuban Spanish (García Riverón et al. 2010), Catalan (Font-Rotchés, 2007), Brazilian Portuguese (Araújo, 2014; Cantero & Font-Rotchés, 2013; Sena, 2013; Silva, 2009), German (Torregrosa, 2015; Torregrosa & Font-Rotchés, 2017) and Chinese (Kao, 2011).

*Paralinguistic intonation*: focusing on politeness and impoliteness for Spanish (Devís, 2011a) and Catalan (Devís & Cantero, 2014); mechanisms of focus in Peninsular Spanish (Cantero et al., 2005), South American Spanish (Pereira, 2014; Ruiz, 2011), Catalan (Font-Rotchés, 2011, 2014) and of professional speakers (Font-Rotchés & Machuca, 2010; Font-Rotchés & Paloma, 2012; Torrent, 2015); and an initial approach to emotional intonation (Cantero, 2014).

---

<sup>8</sup> See a compendium of the first ten years of work in Various Authors (2009) and a more recent revision in *Phonic* magazine, vol. 9-10 (2013-2014). <http://revistes.ub.edu/index.php/phonica/index>

<sup>9</sup> <http://www.ub.edu/lfa/>

Much of this research is still on-going. Currently, work is being made on melodic patterns for Portuguese, German and English, as well as different studies on dialects of Catalan and American Spanish. We have also begun working on the educational applications of our intonation research, particularly on paralinguistic intonation (Devís, 2015; Devís & Bartolí, 2015).

Among the urgent research tasks we propose is the implementation of the script for Praat in a new version that goes beyond melodic analysis and allows the standardisation of intensity (dynamic analysis) and duration, permitting a complete *prosodic analysis* of speech.

In effect, our method only allows the analysis of melodies. However, as we look further into the analysis of spontaneous speech and the prelinguistic and paralinguistic levels of intonation, we find differences that go beyond the strictly melodic, related to the dynamics and the relationships of duration between the melodic units, which doubtless play a very important role in speech.

Therefore, the next step we will take in the analysis of intonation will be the Prosodic Analysis of Speech (Cantero, 2019).

## 6. References

- Alfonso, R. (2010). *El vocalismo del español en el habla espontanea*. PhD thesis. Barcelona: Universitat de Barcelona.
- Araújo, M. (2014). *Entonação das interrogativas e das declarativas do português brasileiro falado em Minas Gerais: Modelos para o Ensino de Línguas*. (Masters Dissert.). Universidade de Brasília.
- Armstrong, L. E. & Ward, I. C. (1926). *Handbook of English Intonation*. Leipzig/Berlin: Teubner.
- Avesani, C. (1995). ToBIIt. Un sistema di trascrizione per l'intonazione italiana. In Lazzari, G. (Ed.), *Atti delle V Giornate di studio del gruppo di fonetica sperimentale* (pp. 85–98). Trento: Servizio editoria ITC.
- Baditzné, K. (2012). *Spanish Intonation of Hungarian Learners of Spanish; Yes or No Questions*. Biblioteca phonica, 15. Barcelona: Laboratori de Fonètica Aplicada UB. Retrieved from [http://www.publicacions.ub.edu/revistes/phonica-biblioteca/15\\_Kata.pdf](http://www.publicacions.ub.edu/revistes/phonica-biblioteca/15_Kata.pdf) (accessed 9/4/2019)
- Ballesteros, M. P. (2011). *La entonación del español del norte*. PhD thesis. Universitat de Barcelona. Retrieved from <https://www.tdx.cat/handle/10803/109963> (accessed 9/4/2019).
- Ballesteros, M. P. & Font-Rotchés, D. (2019). Acoustic analysis of intonation: Comparison between two dialects of Spanish from the north of the peninsula. *Lingua*, 221, 49-71.
- Beckman, M., Díaz-Campos, M., Tevis McGory, J. & Morgan, T. A. (2002). Intonation across Spanish, in the Tones and Break Indices framework. *Probus*, 14, 9–36.
- Beckman, M. & Hirschberg, J. (1994). *ToBI annotation conventions*. The Ohio State University Research Foundation.
- Boersma, P. & Weenink, D. (2019). *Praat: Doing Phonetics by Computer* (Computer program). Version 6.1.02. Retrieved from <http://www.praat.org>.
- Bolinger, D. L. (1986). *Intonation and its parts*. Stanford: Stanford University Press.
- Bolinger, D. L. (1989). *Intonation and its uses*. Stanford: Stanford University Press.

- Cabedo, A. (2013). Sobre prosodia, marcadores del discurso y unidades del discurso en español: evidencias de un corpus oral espontáneo. *Onomázein*, 28, 201–213.
- Cantero, F. J. (1999). Análisis Melódico del Habla: principios teóricos y procedimiento. *Actas del I Congreso de Fonética Experimental* (pp.127-133). Tarragona: Universitat Rovira i Virgili.
- Cantero, F. J. (2002). Teoría y análisis de la entonación. Barcelona: Edicions Universitat Barcelona.
- Cantero, F. J. (2014). Códigos de la entonación y entonación emocional. In *Comunicación, Cognición, Cibernetic@*. Actas del 31.er Congreso Internacional de la Asociación Española de Lingüística Aplicada (pp. 618–629). Tenerife: Universidad de la Laguna.
- Cantero, F. J. (2016). Corpus de habla espontánea para el estudio de la entonación. In: 53 reflexiones sobre aspectos de la fonética y otros temas de lingüística (pp. 151–160). Barcelona: Laboratori de Fonètica de la Universitat de Barcelona.
- Cantero, F.J. (2019): Análisis prosódico del habla: más allá de la melodía. In Álvarez, M.R.; Muñoz Alvarado, A. & Ruiz, L. (eds.): *Comunicación Social: Lingüística, Medios Masivos, Arte, Etnología, Folclor y otras ciencias afines*. Vol. II. Santiago de Cuba: Ediciones Centro de Lingüística Aplicada.
- Cantero, F. J., Alfonso, R., Bartolí, M., Corrales, A. & Vidal, M. (2005). Rasgos melódicos de énfasis en español. *Phonica*, 1, 1–40.
- Cantero, F. J. & Devís, E. (2011). Análisis melódico de la interlengua. In Hidalgo, A.; Congosto, Y.; Quilis, M. (eds.): *El estudio de la prosodia en España en el siglo XXI: perspectivas y ámbitos*. Anejo nº 75 de la Revista Quaderns de Filologia (pp. 285–300). València: Universitat de València.
- Cantero, F. J. & Font-Rotchés, D. (2007). Entonación del español peninsular en habla espontánea: patrones melódicos y márgenes de dispersión. *Moenia. Revista Lucense de Lingüística y Literatura*, 13, 69–92.
- Cantero, F. J. & Font-Rotchés, D. (2009). Protocolo para el análisis melódico del habla. *Estudios de Fonética Experimental*, 18, 17–32.
- Cantero, F. J. & Font-Rotchés, D. (2013). The intonation of absolute questions of Brazilian Portuguese. *Linguistics and Literature Studies*, 3, 142–149.
- Cortés, M. (2004). Análisis acústico de la producción de la entonación española por parte de sinohablantes. *Estudios de Fonética Experimental*, XIII, 142–149.
- Cruttenden, A. (1986). *Intonation*. Cambridge: Cambridge University Press.
- Crystal, D. (1969). *Prosodic systems and intonation in English*. Cambridge: Cambridge University Press.
- Devís, E. (2011a). La entonación de (des)cortesía en el español coloquial. *Phonica*, 7, 36–79.
- Devís, E. (2011b). La entonación del español hablado por italianos. *Didáctica. Lengua y Literatura*, 23, 35–58.
- Devís, E. (2015). Contextos para una aplicación didáctica de la entonación atenuadora en español. *Rilce. Revista de Filología Hispánica*, 31(1), 154–170.
- Devís, E. & Bartolí, M. (2015). Propostes didàctiques d'entonació de cortesia atenuant per al nivell B1 de català com a llengua estrangera. In Anejo 7 de la Revista Normas. *Revista de Estudios Lingüísticos Hispánicos*. Adrián Cabedo Nebot (ed.): *Perspectivas actuales en el análisis fónico del habla: tradición y avances en la fonética experimental* (pp. 199–210).
- Devís, E. & Cantero, F. J. (2014). The intonation of mitigating politeness in Catalan. *The Intonation of Mitigating Politeness in Catalan*, 10(1), 127–149.
- Estebas-Vilaplana, E. & Prieto, P. (2008). La notación prosódica en español. Una revisión del Sp\_ToBI. *Estudios de Fonética Experimental*, XVII, 263–283.

- Face, T. & Prieto, P. (2007). Using accents in Castilian Spanish: a revision of Sp\_ToBI. *Journal of Portuguese Linguistics*, 6.1, 117-146.
- Fonseca, A. (2013). Caracterización de la entonación del español hablado por brasileños. PhD Thesis. Universitat de Barcelona. Retrieved from [https://www.tesisenred.net/bitstream/handle/10803/134929/01.AFdO\\_1de3.pdf?sequence=1&isAllowed=y](https://www.tesisenred.net/bitstream/handle/10803/134929/01.AFdO_1de3.pdf?sequence=1&isAllowed=y) (accessed 9/4/2019).
- Font-Rotchés, D. (2007). *L'entonació del català*. Barcelona: Publicacions de l'Abadia de Montserrat.
- Font-Rotchés, D. (2011). Èmfasis bàsics de l'entonació en els contorns del català central. In Lloret, M.R.; Pons, C.: *Noves aproximacions a la fonologia i la morfologia del català* (pp. 193–216). València: Institut Interuniversitari de Filologia Valenciana.
- Font-Rotchés, D. (2014). Rasgos melódicos de la entonación paralingüística de foco. In Díaz Galán et al. (eds): *Comunicación, Cognición, Cibernética*. Actas del 31 Congreso Internacional de la Asociación Española de Lingüística Aplicada. Tenerife: Universidad de la Laguna.
- Font-Rotchés, D. & Cantero, F.J. (2009). Melodic Analysis of Speech Method applied to Spanish and Catalan. *Phonica*, 5, 33-47.
- Font-Rotchés, D. & Machuca, M. (2010). Énfasis en los eslóganes publicitarios de radio y televisión. *Icono 14. Revista de Comunicación y Tecnologías Emergentes*, Actas n. A4, 591–605.
- Font-Rotchés, D. & Mateo, M. (2013). Entonación de las interrogativas absolutas del español peninsular del sur en habla espontánea. *Onomázein*, 28, 256–275.
- Font-Rotchés, D. & Mateo, M. (2017). Melodías para confirmar, preguntar, sugerir o pedir en español. *Phonica*, 13, 49–67.
- Font-Rotchés, D. & Paloma, D. (2012). Tendencias entonativas en los titulares televisivos en catalán. *Círculo de Lingüística Aplicada a La Comunicación*, 51, 50–81.
- García Riverón, R., Bermúdez, M., Pedrosa, A. & Marrero, A. (2010). El sistema de entonación del español de Cuba a la luz del modelo de análisis melódico del habla. *Phonica*, 6, 3–25.
- Grice, M., Reyelt, M., Benz Müller, R., Mayer, J. & Batliner, M. (1996). Consistency in transcription and labelling of German intonation with GToBI. In Bunnell, T.; Idsardi, W. (Eds.), *Proceedings of the 4th International Conference on Spoken Language Processing* (pp. 1716–1719). New York: Institute of Electrical and Electronic Engineers.
- Gussenhoven, C., Rietveld, T. & Terken, J. (1999). ToDI: Transcription of Dutch Intonation. Retrieved January 5, 2019, from <http://gep.ruhosting.nl/carlos/todisun-ah.pdf> (accessed 9/4/2019).
- Hirst, D. J., Di Cristo, A. & Espesser, R. (2000). Levels of representation and levels of analysis for the description of intonation systems. In Home, M. (Ed.), *Prosody: Theory and Experiment* (pp. 51–87). Dordrecht: Kluwer Academic Press.
- Jones, D. (1909). *Intonation Curves*. Leipzig/Berlin: Teubner.
- Kao, W. L. (2011). La entonación de enunciados declarativos e interrogativos en chino mandarín hablado por taiwaneses. (Masters Dissert.). Universitat de Barcelona, Barcelona.
- Liu, Y. H. (2005). La entonación del español hablado por taiwaneses. *Biblioteca Phonica*, 2. Laboratori de Fonètica Aplicada (UB). Retrieved from <http://www.publicacions.ub.edu/revistes/phonica-biblioteca/> (accessed 9/4/2019).
- Martínez, D. (2015). La importancia del factor prosódico en el estudio de los marcadores del discurso: algunos problemas de su análisis acústico-melódico. *Círculo de Lingüística Aplicada a La Comunicación*, 62, 105–124.

- Martorell, L. (2010). Les interrogatives absolutes de l'espanyol parlat per suecs. (Masters Dissert.). Universitat de Barcelona. Retrieved from <http://diposit.ub.edu/dspace/bitstream/2445/32302/1/Treball%20fi%20m%C3%A0ster%20Laura%20Martorell.pdf> (accessed 9/4/2019).
- Mateo, M. (2010a). Protocolo para la extracción de los datos tonales y curva estándar en análisis melódico del habla (AMH). *Phonica*, 6, 49–90.
- Mateo, M. (2010b). Scripts en Praat para la extracción de datos tonales y curva estándar. *Phonica*, 6, 91–111.
- Mateo, M. (2014). La entonación del español meridional. PhD thesis. Universitat de Barcelona. Retrieved from <http://diposit.ub.edu/dspace/handle/2445/53156> (accessed 9/4/2019).
- Navarro Tomás, T. (1944). *Manual de entonación española*. New York: Hispanic Institute.
- O'Connor, J. D., & Arnold, G. F. (1961). *Intonation of colloquial English*. London: Longman.
- Palmer, H. E. (1922). *English Intonation, with systematic exercises*. Cambridge: Heffer.
- Pereira, D. I. (2014). Patrones y mecanismos prosódicos de cuantificación e intensificación en construcciones exclamativas. PhD thesis. Universidad de Concepción (Chile).
- Pierrehumbert, J. B. (1980). *The Phonology and Phonetics of English Intonation*. PhD thesis. MIT.
- Pierrehumbert, J. B. (1987). *The Phonology and Phonetics of English Intonation*. Bloomington: Indiana University Linguistics Club.
- Pike, K. L. (1945). *The intonation of American English*. Ann Arbor: University of Michigan Press.
- Pons, J. (2007). La ràdio en menorquí i l'entonació de les preguntes parcials. *Phonica*, 7, 100–117.
- Prieto, P. & Roseano, P. (eds.). (2010). *Transcription of Intonation of the Spanish Language*. München: Lincoln Europa.
- Ruiz Mella, M. (2011). Determinación de la incidencia de la entonación en la producción y percepción de oraciones declarativas con focalización de tipo contrastiva en hablantes de español de Concepción. PhD thesis. Universidad de Concepción (Chile).
- Sena, R. (2013). A entonação no processo de ensino-aprendizagem de PLE. Proposta didática para o ensino de modelos de entonação interrogativa do português do Brasil-Estado de São Paulo. (Masters Dissert.). Universidade de Brasília.
- Silva, C. (2009). Análisis melódico de declarativas e interrogativas absolutas en español/LE. *Phonica*, 6, 92–113.
- Silverman, K., Beckman, M., Pitrelli, J., Ostendorf, M., Wighman, C., Price, P. & Hirschberg, J. (1992). ToBI. A standard for labelling English prosody. In Ohala, J. et. al. (Eds.), *Proceedings of the 1992 International Conference on Spoken Language Processing* (Vol. 2). Edmonton: University of Alberta.
- t'Hart, J., Collier, R. & Cohen, A. (1990). *A perceptual study of intonation. An experimental-phonetic approach to speech melody*. Cambridge: Cambridge University Press.
- Torregrosa, J. (2015). Una propuesta metodológica para el análisis melódico de enunciados orales espontáneos con partículas modales del alemán. *Revista de Filología Alemana*, 23, 193–212.
- Torregrosa, J. & Font-Rotchés, D. (2017). Melodic patterns of absolute interrogative utterances in northern German in spontaneous speech. *Linguistics*, 55(6), 1311–1335.
- Torrent, N. (2015). Barberà i Cuní: dos models entonatius per a programes informatius. *Phonica*, 15, 87–102.
- Trager, G. L. & Smith, H. L. (1951). *An outline of English Structure*. Norman, Okla: Battenburg Press.

Various Authors (2009): Análisis Melódico del Habla (AMH): 1999-2009. Biblioteca Phonica, 10. [http://www.publicacions.ub.edu/revistes/phonica-biblioteca/10\\_AMH.pdf](http://www.publicacions.ub.edu/revistes/phonica-biblioteca/10_AMH.pdf) (accessed 9/4/2019).