

# The duration of filled pauses and prolongations in northern and southern dialects of Spanish<sup>1\*</sup>

Kata Baditzné Pálvölgyi  
bpkat79@gmail.com  
*Eötvös Loránd University (Hungria)*

## ABSTRACT

This research focuses on the prosodic patterns of hesitations contrasted in Northern and Southern dialects of European Spanish, more precisely, on their duration. A corpus of 200 spontaneous utterances has been compiled (including 100 utterances from the northern dialects and 100 utterances from the southern ones, produced by 16 male and 16 female informants, respectively). The analysis has been carried out following the standardization protocol offered by Cantero (2019), in which the representative values of duration (in seconds) are taken for each syllable, and then these values undergo a process of standardization, in order to be comparable objectively and speaker-independently. Due to difficulties in establishing exact syllable boundaries in Spanish, it is not the relative duration of the syllables, but rather the relative duration of the distances between intensity peaks which is compared. It is expected that certain “neutral” hesitations –lengthenings and filled pauses with no specific communicative function – show durational differences in the examined dialects, as Southern dialects are considered to be of higher speech rate due to frequent segment elision than northern ones. This would imply that the relative duration of the examined hesitation phenomena with respect to its context– supposedly of the same absolute duration as in the northern dialects – is longer in the southern dialects. According to the results, nevertheless, Southern dialects present shorter absolute duration also in case of hesitation phenomena, and thus relative duration of hesitation phenomena with respect to their context coincide in the two examined Spanish variants.

## KEYWORDS

hesitation, filled pauses, lengthenings, duration, standardization

## 1. Introduction

Hesitation is considered a disfluency phenomenon. There are several possible categorizations related to these phenomena (Neuberger 2014: 23), including, for instance, noises, repetitions, false starts, silent pauses, repairs,

---

<sup>1\*</sup>Supported by the ÚNKP-19-4 New National Excellence Program of the Ministry for Innovation and Technology.

truncations, filled pauses and prolongations (Eklund, 2004; Gósy, 2002; Lickley, 1994, 2015; Rodríguez et al., 2001; Shriberg, 1994), the last two phenomena being the two most common subtypes of hesitation (Deme & Markó, 2013).

Prolongations form part of phenomena that are applied to gain time without implying necessarily the interruption of elocution. Their aim is to slow down the velocity of speech without affecting communication (Rebollo Couto, 1997: 667). Filled pauses lack specific meaning and are employed to provide the speaker with time in order to plan and to reorganize his/her speech (Stepanova, 2007, cited by Machuca & Ríos, 2016). Filled pauses and lengthening can be considered two different acoustic disfluency phenomena (Rodríguez et al., 2001, 2015), but the present study is based on authors holding that prolongations are subsumed under filled pauses (Maclay & Osgood, 1959, cited by Machuca, 2018), as a special, “lexical” type (Blondet, 2001, Villa et al., 2017), and the two phenomena will be treated together, focusing on the durational aspect of prolongations and filled pauses in two well-defined dialects of European Spanish.

Spanish is spoken by more than 400 million speakers all over the world, and due to this fact, it presents a considerable dialectal variation. There are two main dialectal areas in Europe sharing several common characteristics in their pronunciation: those from the north (including also central varieties) and the dialects from the south (including the varieties spoken in the Canary Islands, cf. Hualde, 2014: 285-288). The objective of this study is to compare the strategies of hesitation (with no specific communicative functions other than turn-holding) applied by the dialects of Spanish from both the north and the south, from the point of view of the duration of prolonged segments and filled pauses. A research was conducted that compared 100 sentences provided by speakers of northern Spanish dialects with 100 sentences from informants from southern Spain. In both corpora the same methodology was applied, a three-phase prosodic analysis of speech proposed by Cantero (2019), in order to compare the data objectively. According to my prediction, there will be significant differences in duration, as southern dialects will present relatively longer utterance-internal hesitations. This supposed difference is due to the fact that southern dialects are generally considered to be of higher speech rate (cf. Toledo, 2010) because of the frequent elision

of certain sounds (such as the syllable-final *-s* or the intervocalic *-d-*), and this phenomenon causes that the absolute duration of the syllables is shorter. This also means that lengthenings and filled pauses in the southern dialects, of the same assumed absolute duration as in the northern dialects, are relatively longer as compared to the length of previous segments.

Based on what has been revealed so far, this study focuses on hesitation phenomena from the point of view of duration, by formulating the following hypothesis: in southern Spanish dialects, as they are regarded to be of higher speech rate, the relative duration of lengthenings and filled pauses with no specific communicative functions is longer as compared to their context than in northern Spanish dialects.

## 2. Corpus and informants

The corpus was obtained from two sources: the ‘Map Task’ activities in the interactive Atlas of Romance intonation compiled by Prieto et al. (2010-2014), and of spontaneous interviews uploaded to YouTube. This way only spontaneous speech samples were obtained. 32 speakers were selected, 16 informants from the north (8 men and 8 women), and 16 from the south of Spain (also 8 men and 8 women), from recordings of 291 minutes and 30 seconds in total. In the northern corpus, 146 hesitation phenomena have been detected, compared to the 136 cases in the southern corpus. The following table sums up the data related to the informants.

Table 1 - The informants’ data.

Northern Spanish corpus							
origin	speakers	monograph and code	job	sex	age	nº of utterances selected	duration
Gijón (Map Task)	Speaker 1	GP; EGI-1	student	f	24	5	5:49'
	Speaker 2	AL; EGI-2	student	f	22	7	
Oviedo (Map Task)	Speaker 1	GP; EOv-1	student	f	20	9	4:52'
	Speaker 2	MR; EOv-2	student	f	25	6	
Cabezón de la Sal (Map Task)	Speaker 1	NEO; ECA-1	teacher	f	31	12	11:50'
	Speaker 2	AOZ; ECA-2	teacher	f	31	2	

Madrid (Map Task)	Speaker 1	CTN; EMA-1	none	f	33	3	14:00'
	Speaker 2	SBL; EMA-2	none	f	37	5	
Salamanca (interviews)	Speaker 1	FI; ESA-1	politician	m	49	7	4:39'
	Speaker 2	CG; ESA-2	politician	m	57	6	21:36'
Burgos (interviews)	Speaker 1	JV; EBU-1	politician	m	59	6	15:53'
	Speaker 2	LT; EBU-2	politician	m	36	6	16:16'
Ávila (interviews)	Speaker 1	ÁA; EAV-1	politician	m	60	6	22:31'
	Speaker 2	M Á G N ; EAV-2	politician	m	59	7	25:23'
León (interviews)	Speaker 1	AS; ELE-1	politician	m	56	6	26:30'
	Speaker 2	JADD; ELE-2	politician	m	51	6	27:24'
age (years; mean)	40,63						
utterances (total)	100						

Southern Spanish corpus							
origin	speakers	monograph and code	job	sex	age	nº of utterances selected	duration
Canary Islands (Map Task)	Speaker 1	AAG; ECAN-1	teacher	f	38	6	4:41'
	Speaker 2	AAH; ECAN-2	teacher	m	38	8	
Jaén (Map Task)	Speaker 1	SB; EJA-1	student	f	22	10	4:14'
	Speaker 2	DG; EJA-2	student	m	21	1	
Constantina (Map Task)	Speaker 1	LSN; ECO-1	student	f	23	8	2:42'
	Speaker 2	MSC; ECO-2	student	f	22	8	
Jerez de la Frontera (Map Task)	Speaker 1	MGA; EJE-1	entrepreneur	f	41	3	3:15'
	Speaker 2	PM; EJE-2	entrepreneur	m	46	6	
Málaga (interviews)	Speaker 1	MÁJ; EMAL-1	politician	m	49	6	14:05'
	Speaker 2	EB; EMAL-2	politician	m	45	6	18:14'
Sevilla (interviews)	Speaker 1	PP; ESE-1	politician	f	50	7	15:19'
	Speaker 2	JE; ESE-2	politician	m	51	6	
Badajoz (interviews)	Speaker 1	FJF; EBA-1	politician	m	47	6	12:18'
	Speaker 2	GFV; EBA-2	politician	m	59	6	6:20'
Granada (interviews)	Speaker 1	MJLG; EGR-1	politician	f	55	6	8:37'
	Speaker 2	LGCH; EGR-2	politician	f	47	7	5:03'
age (years; mean)	40,88						
utterances (total)	100						

The following map presents the regions of the corpora. Only monolingual areas were chosen for the analysis (leaving apart thus, territories such as Catalonia, Valencia or the Balearic Islands (partly Catalan-speaking zones), Galicia (partly a Galician-speaking zone), the Basque Country and La Rioja (partly Basque-speaking zones), because these regions could have shown influences by other peninsular languages).

Figure 1 - The origin of the informants.



### 3. Method

The theoretical background used in this work is based on the protocol for Prosodic Analysis of Speech, (abbreviated as PAS, Cantero,2019). Pitch, duration and intensity are considered to be suprasegmental features, and due to this reason, they are relatively difficult to interpret. First, because we must disregard speaker-dependent characteristics that carry no linguistic significance, and second, because prosodic units must be understood as

bearing relative prominence only with respect to adjacent units, so they have no information alone.

A solution to overcome these difficulties is offered by Cantero's Melodic Analysis of Speech (MAS) (2009) and his latter implementation to the theory, Prosodic Analysis of Speech (PAS) (2019). As for the analysis, there is an acoustic phase, assisted by an acoustic analysis software such as Praat (Boersma & Weenink, 2019). The second step is the prosodic representation: in order to concentrate only on the prosodically relevant features, it is necessary to ignore irrelevant variations and reduce the data in case of each syllable to a characteristic prosodic value (in case of intonation, in Hz; in case of intensity, in dB, and in the case of duration, in seconds). The third step is the data standardization: the prosodic data curves are represented taking into account not the absolute values, but the relative ones, as each syllable is given a percentage based on its prosodic rise/fall experienced with respect to the previous syllable.

This analysis permits us to describe more objectively the prosodic features of a given language, and compare prosodically, for instance, dialects. The process is illustrated in the next section, using examples from my corpora.

### 3.1 The standardization of prosodic data in the PAS model

The standardization of prosodic data is not a recent idea in linguistic research. Standardization of melodic contours was first done using semitones in the 'Dutch School', also known as the IPO model. The most emblematic work of this approach is 't Hart et al. (1990), followed later by various researches in different languages (Adriaens, 1991; Beaugendre, 1994; Odé & van Heuven, 1994). In Spanish, Garrido (1991, 1996) and Estruch et al. (2007) worked with similar automatic stylization methods (Baditzné, 2012).

The difference between the standardized curves in the PAS model and the ones in the Dutch School is that the PAS model uses percentages for the standard values, which is a system easier to handle than the one with semitones. The standardized contour is represented by a line which starts at an arbitrary value of 100% and anchors in each syllable, which is itself characterized by a percentage based on its tonal position as compared to the previous syllable. If the syllable is located lower, it is a negative percentage,

and if it is higher than the previous syllable, it is a positive one.

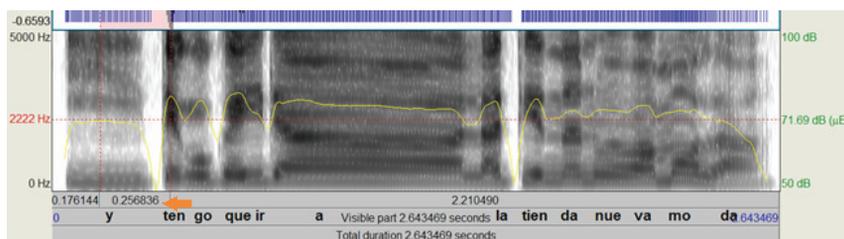
Though first applied to describe Spanish intonation (Cantero et al., 2005; Cantero & Font-Rotchés, 2007, 2020; Font-Rotchés & Mateo, 2011), it has been extended to the study of intonation in other languages as well, such as Catalan (Font-Rotchés, 2007; 2008; 2009), Chinese (Kao, 2011), and also in the description of the intonation of interlanguages, for example the Spanish spoken by Brazilians (Fonseca & Cantero, 2011), Italians (Devís, 2011), Swedes (Martorell, 2011) or Hungarians (Baditzné, 2011; 2012; 2018). For a partial Spanish application, cf. Patiño (2008). In Hungarian, a similar analysis was carried out in Olaszy & Koutny's investigation, also based on percentages and standardized contours (Olaszy & Koutny, 2001: 182-183).

In the following sections I will explain how standardization works on durational data in the PAS model.

### 3.2 The steps of durational standardization

The first phase of the analysis is to identify syllables, and search for the intensity peak of each syllable. The following screenshot shows (with yellow continuous line) intensity values that accompany the utterance (the text is my addition). Every syllable usually has its intensity peak in the middle of its nuclear vowel, so these are the values we are searching for. The absolute duration of each syllable is calculated based on the distance between its intensity peak and the intensity peak of next syllable.

Figure 2 - The distance between the intensity peaks of the first two syllables of the utterance *y tengo que ir a la Tienda Nueva Moda* 'and I have to go to the shop called New Fashion', from Madrid (in seconds, indicated by the red arrow).

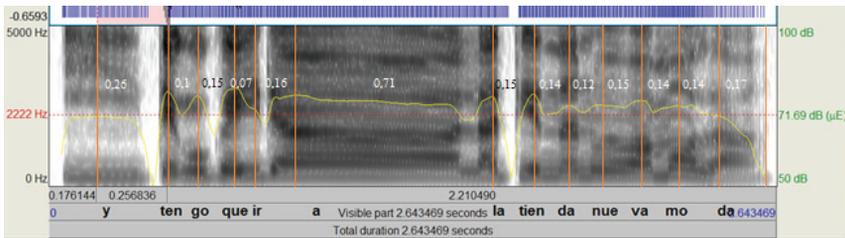


We can see that the first value we need is 0,26s, whereas the next distance, between syllables *ten-* y *-go*, is of 0,1s, which means that the duration of *ten-* is -61,54% shorter than the duration of the previous syllable *y*.

We could simply compare the duration of each syllable instead of comparing the duration of distances between intensity peaks; the argument supporting this latter view is that in Spanish, it is definitely difficult to separate syllables clearly, as the exact location of syllable boundaries is not obvious (Cantero, p.c., 2019). This is why it is preferred to measure distances between intensity peaks, as such peaks are easily recognizable acoustically as compared to the exact extension of a syllable.

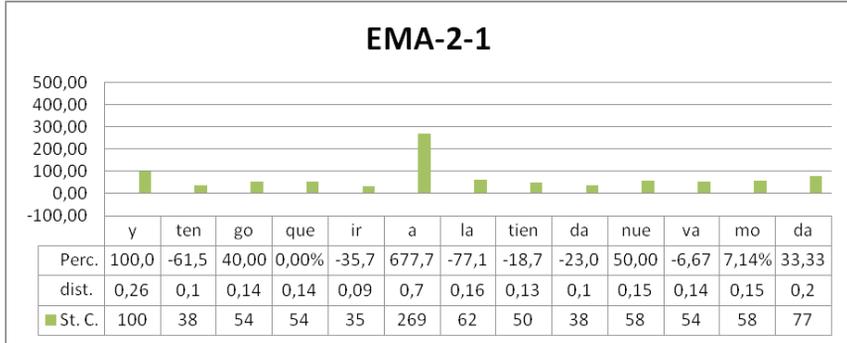
The distance between the intensity peaks of the following syllables is calculated the same way, and as we can see in Figure 3, where all the intervals are displayed, there is a strikingly high value of 0,71s of the preposition *a* ('to'), which is by far the longest segment of the utterance. The last value is calculated from the last intensity peak of the last syllable *-da* till the end of the utterance.

Figure 3 - Absolute duration intervals of the syllables in the utterance *y tengo que ir a la Tienda Nueva Moda* 'and I have to go to the shop called New Fashion', from Madrid



The next step in the PAS methodology is to standardize the measured absolute values by converting them to percentages with respect to the value of the previous syllable. In case of utterance-first syllables, this distance is added the arbitrary value of 100%, but the next values are calculated as compared to the previous one, so the percentage of the second syllable *ten-* is of -61,54%, as the duration of 0,1s is a decrease of 61,54% as compared to the previous duration value 0,26s. Figure 4 shows the standardized values of all the syllables of the utterance in a plot.

Figure 4 - Relative duration intervals of the syllables in the utterance *y tengo que ir a la Tienda Nueva Moda* ‘and I have to go to the shop called New Fashion’, from Madrid



The standardized curve thus ensures that the described durations are objectively comparable to each other, regardless of the individual durational characteristics of the speakers (e.g. if an informant speaks much slower than others, especially regarding the articulatory speech rate; what would matter are the proportions of the durations and not the absolute values of each syllable).

### 3.3 The analysis of durational data in the present research

In accordance with my objectives, segments affected by hesitation phenomena (lengthenings and filled pauses) will be examined from a durational point of view, assuming that in southern dialects, relative durational data will be longer, as southern dialects are considered to be of higher speech rate due to the higher proportion of elision.

As in this study the focus is on prolongations and filled pauses as disfluency phenomena, first we must define which segments were considered as part of this group.

The first criterion to fulfil was, in case of lengthenings, that the segment should be prolonged. There are several ways to determine if a segment should be considered lengthened. Lengthening is detectable by ear, so Deme & Markó (2013) for instance based their judgement on perceptive tests: if a segment was recognized as lengthened by 6 out of 10 listeners (all linguists),

it was considered lengthened. My choice was to base this judgement on durational data: as the minimum duration of filled pauses is considered to be 0,2s by Goldman-Eisler, 1973 and Guaitella, 1996, cited by Blondet, 2001: 8, and filled pauses are generally longer in Spanish than lexical vowel lengthenings (Villa et al., 2017: 167), I took 0,2s as the minimum value for both lengthenings and filled pauses. The other criterion was that the segment in question should be longer than the previous one.

As the study focuses on lengthenings as disfluency phenomena, we must also take a look at two cases in which lengthening is a natural by-product of certain conditions but does not serve as a tool to gain time, thus, is definitely not a disfluency phenomenon, and exclude those samples from our analysis. First, we must bear in mind that phrase-final lengthening is a natural process in numerous languages, including Spanish (Gósy & Krepesz, 2018). Second, stressed syllabic position can also result in longer duration in Spanish (Ortega-Llebaria, 2006), so stressed syllables were also discarded from the analysis (lengthenings in Spanish mostly affect unstressed segments anyway, cf. Rebollo, 1997).

In my analysis, I have worked with two values in case of each hesitation phenomena: the relative duration data of segments affected by hesitation, and also their absolute duration data. As we are analyzing relative prosodic values, we cannot take into consideration either utterance-initial lengthening or utterance-initial filled pauses when we measure the proportion of duration change to the affected segment, as these values cannot be contrasted with any previous value, so they cannot even be relativized. Thus, these hesitation phenomena were discarded from the analysis.

Based on what has been just said, only the following hesitation phenomena were analysed in this study:

- lengthenings on unstressed syllables with duration superior to 0,2s, which are neither utterance-initial nor utterance-final;
- filled pauses with duration superior to 0,2s, which are neither utterance-initial nor utterance-final.

#### 4. Results

In the following section we will focus on the analysis of the received data in order to validate my hypothesis, i.e. that the absolute duration of the examined hesitation phenomena (lengthenings and filled pauses with no specific communicative meaning) is coincident in northern and southern Spanish dialects, but the relative duration of the examined hesitation phenomena is longer in southern Spanish dialects than in the northern ones, due to the shorter syllabic intervals in the southern dialects.

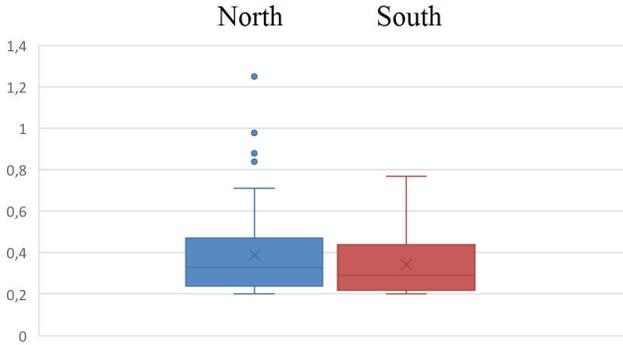
Table 2 sums up the durational data related to the two hesitation phenomena attested in both corpora. As can be seen, in the northern corpus 146 cases were detected as compared to the 136 cases in the southern corpus. However, as utterance-initial and utterance-final hesitation phenomena, as well as lengthenings on stressed syllables, were excluded from the analysis, this number was reduced to 106 northern cases and 117 southern ones.

Table 2 - Data related to the tonal movements of hesitation phenomena

	North	South
nº of hesitation phenomena attested	146	136
nº of non-initial hesitation phenomena	106	117
lowest value of absolute duration (s)	0,2	0,2
highest value of absolute duration (s)	1,25	0,77
mean value of absolute duration (s)	0,39	0,34
results of two-sample t-test (unequal variances)	t Stat: 2,04 > t Critical two-tail: 1,97; different means	
lowest value of relative duration (%)	4,76	5,26
highest value of relative duration (%)	675	540
mean value of relative duration (%)	161	153
results of two-sample t-test (equal variances)	t Stat: -0,5 < t Critical two-tail: 1,97; same means	

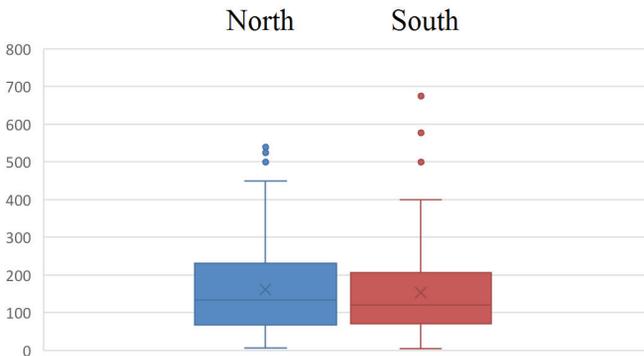
As attested in the results, the average value of the absolute duration data, in case of the northern dialects, is 0,39s, while in the case of the southern ones, it is 0,34s (cf. Figure 6), and according to the two-sample two-tailed t-test of unequal variances, this is a statistically significant difference, we cannot consider them the same.

Figure 6 - Absolute duration data (in sec) in the two corpora (lengthening and filled pauses; boxplots and statistical analysis are generated by Excel 365 pro plus).



In the case of the proportion of relative duration of the affected segments, the mean value is 161% in case of the northern dialects and 153% in case of the southern ones (cf. Figure 7), but according to the two-sample two-tailed t-test of equal variances, it is not a statistically significant difference.

Figure 7 - Relative duration data in % in the two corpora (lengthening and filled pauses; boxplots and statistical analysis are generated by Excel 365 pro plus).



## 5. Discussion and conclusion

In this study the two main dialectal areas of European Spanish were examined from the point of view of the durational aspect of two hesitation phenomena, lengthenings and filled pauses, when they lack any specific communicative function. A corpus of 200 utterances (100 northern Spanish and 100 southern Spanish ones, taken from Map Task activities and spontaneous interviews downloaded from YouTube videos) were contrasted, in order to verify the following hypothesis:

In both dialects, the absolute duration of segments affected by neutral hesitation phenomena (lengthenings and filled pauses with no specific communicative meaning) is the same, but in southern dialects, which are assumed to be of higher speech rate due to elision, relative values of hesitated segments will be longer, as they are compared to shorter previous intervals.

There is an argument to suppose that southern dialects present relatively longer hesitation phenomena in utterance-internal segments than their northern counterparts. This could be so because the southern variants are characterized by elision more frequently than the northern ones, thus, even if segments affected by hesitation are of the same absolute duration in both dialects, they are perceived relatively longer in the southern dialects where syllables are realized shorter due to elision than in the northern dialectal zone.

We have seen based on our results that the average absolute duration of hesitation segments, contrarily to the prediction, is lower in southern dialects: 0,34s as compared to the northern mean of 0,39s, which is a statistically significant difference. This means that the first part of our hypothesis –i.e. that absolute duration data of hesitation segments in both dialects will coincide – is not verified. As for the second part of the hypothesis, according to which relative duration data of hesitation segments will be longer in southern dialects, it again has proved to be false: the 153% of the southern mean as compared to the northern mean of 161% after statistical testing has turned out to be not significantly different.

This means that, contrarily to what has been predicted, southern dialects are characterized by higher velocity also in case of hesitation phenomena (lengthenings and filled pauses), and this is why we find no differences in the relative durational data. Northern Spanish speakers are characterized by

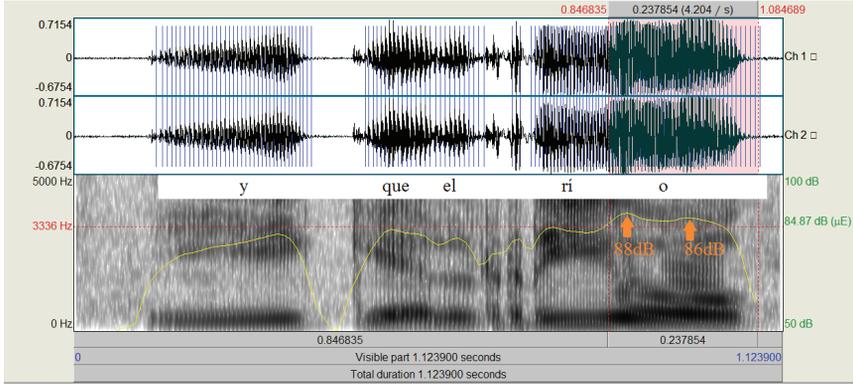
slower speech rate, and slower hesitation phenomena, whereas Southern Spanish speakers do not only present higher speech rate, but also the velocity of their hesitation phenomena is higher. In other words, the speakers of both dialects of Spanish tend to adjust the velocity of hesitation phenomena to their velocity of speech; the velocity of hesitation phenomena is also determined by the dialect of the speaker.

In this paper we have only seen the durational aspect of utterance-internal hesitation phenomena in European Spanish dialects, but by Cantero's (2019) extended Prosodic Analysis of Speech (PAS) model we could define prosodic features other than duration in case of hesitation phenomena. The PAS model offers a standardization protocol for intonation and intensity similar to the procedure we have seen in the case of duration; as for intensity, the standardized intensity curve is generated by reducing every syllable to its prominent intensity peak value, and these values are standardized in terms of proportion always compared to the previous value. Melodic standardization calculates the tonal difference in terms of percentages for each segment, always comparing syllables to the previous ones, and thus the standardized curve is generated over these values. The perception threshold is already established for intonation: the Spanish ear is sensitive to at least 10% of tonal difference (Font-Rotchés & Mateo, 2011). These values are not yet discovered for neither intensity nor duration in case of Spanish, and, especially in duration, they would be essential for objective data interpretation.

Perceptive thresholds are important in duration, but it is also a question whether the human mind finds a segment lengthened comparing it to the duration of the adjacent previous segment; if this is the case, Cantero's (2019) duration standardization model would be useful, but if our mind stores "average durations" of various previous syllables and compares the duration of every new segment to this stored average value, the PAS model would be insufficient to describe relative durations.

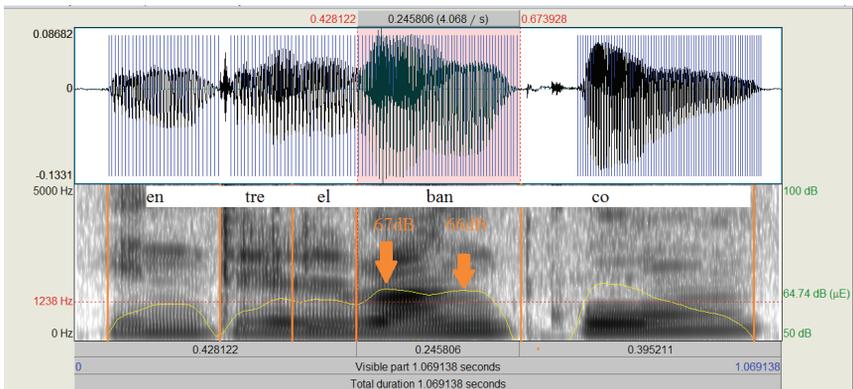
There are some special cases in which the PAS analysis may not calculate correctly the relative duration values if we do take into consideration some elemental rules. The position of intensity peaks is a crucial point in this analysis. When there are two peaks in a syllable, as in *-o* in Figure 8, it is the higher one which counts.

Figure 8 - Two intensity peaks in one syllable in *y que el río* 'and that the river', an utterance from Badajoz.



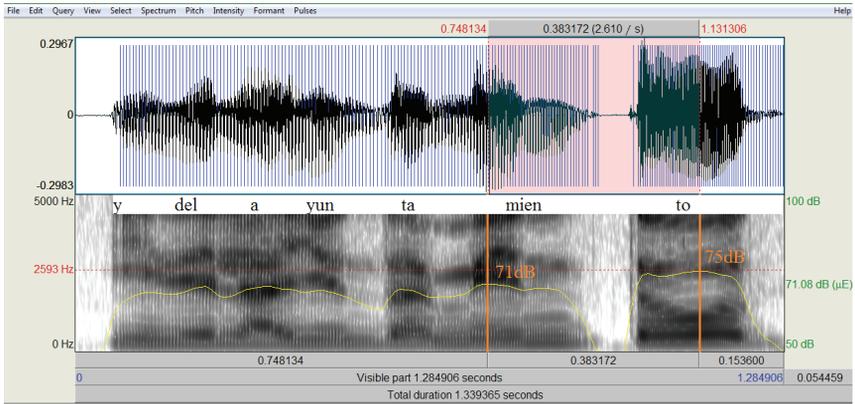
We must bear in mind that it is only an intensity peak which is located within a vowel which must be taken into account. Consonantal peaks (as in Figure 9) should not be taken into consideration. There are apparently two peaks in the same syllable, *ban-* in Figure 8, but the higher one should be taken, in this case, 67dB, as this peak is located on the vowel.

Figure 9 - Two intensity peaks in one syllable in the utterance *entre el banco* (...) 'between the bank', from Madrid, the second peaks is on a consonant, [n].



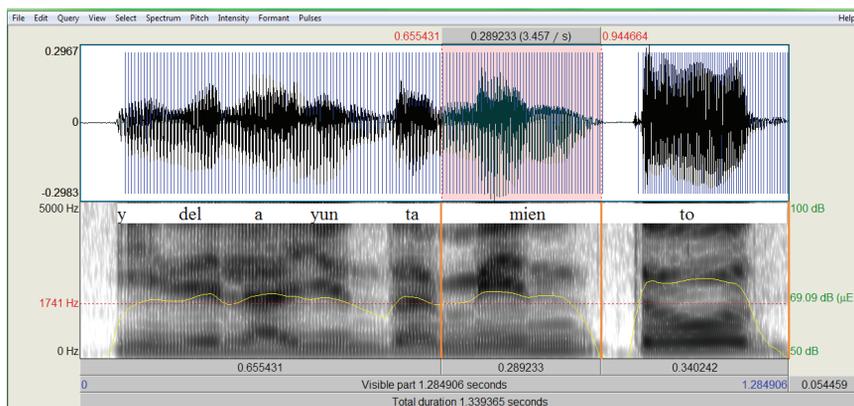
In some cases, the intensity peak is not located in the middle of the nuclear vowel. These cases suppose a challenge in the method, as they may cause that the relative duration of a syllable is predicted incorrectly, as in Figures 10 and 11. In Figure 10, it is clear that the two peaks that we must take into consideration are the ones with 71dB and 75dB, the first one is located before the middle point of the syllable *-mien-*, whereas the second is located towards the end of syllable *-to*, not in the middle.

Figure 10 - Some intensity peaks are not located in the middle of the nuclear vowel in *y delayuntamiento* 'and from the city council', an utterance from Madrid.



This is problematic because this way we have the following two absolute distances between intensity peaks: 0,38s and 0,15s, respectively. After standardization, it would turn out that the relative duration corresponding to the final syllable *-to* is 60,53% less than the previous value. This is definitely not the case if we look at the absolute duration of the syllables (not the distances), illustrated on Figure 11: the syllable *-mien-* has the duration of 0,29s as compared to the following syllable *-to*, with 0,34s. The relative duration data in this case would be of a 17,24%, a positive value instead of the negative one we got by the PAS analysis.

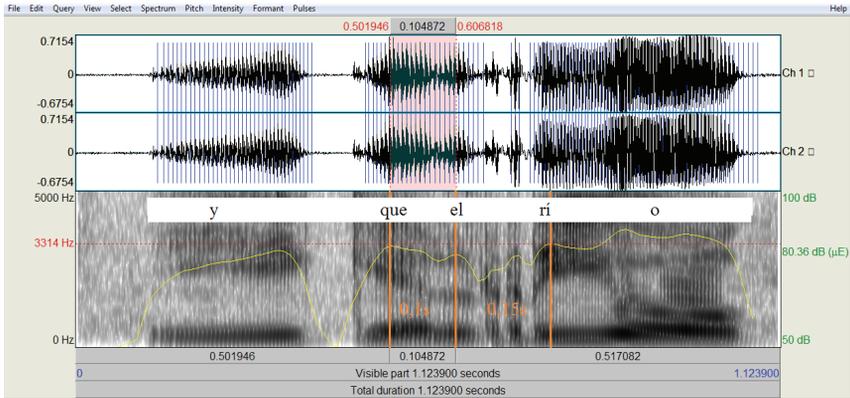
Figure 11 - The absolute duration of the last two syllables in *y delayuntamiento* 'and from the city council', an utterance from Madrid.



This means that if intensity peaks are not located exactly where they should (i.e. in nucleus-central position), the PAS model could predict relative duration data incorrectly. The lack of an easily discernible intensity peak would also result in difficulties (such as in case of intensity plateaus instead of peaks), as in this case taking the intensity value at the middle point of the syllable for example would be an arbitrary solution. The opposite case – the presence of multiple intensity peaks of the same or of almost the same value within the vowel would also raise the question of correct interpretation: which peak should be measured? Taking into consideration all these difficulties related to uncommon intensity peaks, measuring durational data between syllable boundaries instead of the distance between intensity peaks still might be a solution with less complications.

Another problem is related to the case when we have two, very short syllables, the second one being even the double of the duration of the first one, but still very short (cf. Figure 12).

Figure 12 - Short syllabic intervals in *y que el río* 'and that the river', an utterance from Badajoz.



In these cases, the model predicts that there is a lengthening (as there is a +50% of lengthening between the two syllables, but the distances are very short, only 0,1s and 0,15s). For this reason, based on perceptive tests a threshold value for minimal lengthening should be established (which might turn to be coincident with the value we applied in this study based on previous literature) and this threshold value could be a complementary solution to the model.

Finally, as for future research, there are numerous aspects that could be investigated apart from carrying out the same investigation on a more extended corpus, such as the comparison of durational data in hesitations between European Spanish and LatinAmerican dialects (the latter ones considered to be slower, for example the Mexican or Argentinian ones, which might again trigger the prediction of different relative durational values in hesitation segments).

## REFERENCES

- Adriaens, L. M. H. 1991. *Ein Modell deutscher Intonation. Eine experimentell-phonetische Untersuchung nach den perzeptiv relevanten Grundfrequenzänderungen in vorgelesenem Text*. PhD thesis, Technological University of Eindhoven.
- Baditzné Pálvölgyi, K. 2018. La presencia de rasgos prelingüísticos en la entonación de la interlengua húngaro-española. *Colindancias: Revista de la Red de Hispanistas de Europa Central* 2018: 9, 237-248.
- Baditzné Pálvölgyi, K. 2017. El castellano hablado por húngaros: rasgos prelingüísticos de transferencia negativa en la entonación. In: Victoria Marrero Aguiar, Eva Estebas Vilaplana (coords.) *Tendencias actuales en fonética experimental: Cruce de disciplinas en el centenario del Manual de Pronunciación Española (Tomás Navarro Tomás)* (pp. 397-398). Madrid: Universidad Nacional de Educación a Distancia (UNED), 2017.
- Baditzné Pálvölgyi, K. 2012. *Spanish Intonation of Hungarian Learners of Spanish: yes-or no questions*. Tesis Doctoral. Budapest: Eötvös Loránd University. *Biblioteca Phonica* 15 (2012), from the World Wide Web: <http://www.publicacions.ub.edu/revistes/phonica-biblioteca/>. [last accessed 12/02/2020]
- Baditzné Pálvölgyi, K. 2011. The intonational patterns used in Hungarian students' Spanish yes-no questions, *Phonica* 7 (2011): 80-99., from the World Wide Web: <http://www.publicacions.ub.edu/revistes/phonica7/>. [last accessed 12/02/2020]
- Beaugendre, F. 1994. *Une étude perceptive de l'intonation du français*, PhD Thesis, University of Paris XI, Orsay, France.
- Blondet, M. A. 2001. Las pausas llenas: marcas de duda e identidad lingüística. *Lingua Americana* V No 8 (2001) 5-15.
- Blondet, M. 2006. *Variaciones de la velocidad de habla en español: patrones fonéticos y estrategias fonológicas*. Un estudio desde la producción. Tesis doctoral. Mérida: Universidad de Los Andes.
- Boersma, P. & Weenink, D. 2019. *Praat: Doing phonetics by computer* (version 6.0.49). <http://www.praat.org/> (accessed 22 March 2019).
- Cantero Serena, F. J., Alfonso, R., Bartolí, M., Corrales, A. & Vidal, M. 2005. Rasgos melódicos de énfasis en español. Laboratori de Fonètica Aplicada – LFA, *Phonica* vol.1., from the World Wide Web: [http://www.publicacions.ub.edu/revistes/phonica1/PDF/articulo\\_03.pdf](http://www.publicacions.ub.edu/revistes/phonica1/PDF/articulo_03.pdf) [last accessed 12/02/2020]

- Cantero Serena, F. 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* (pp. 20-47). Universidad del País Vasco.20-47.
- Cantero Serena, F. J. & Font-Rotchés, D. 2009. Protocolo para el análisis melódico del habla, *Estudios de Fonética Experimental 18*: 17-32.
- Cantero Serena, F. J. & Font-Rotchés, D. 2007. Entonación del español peninsular en habla espontánea: patrones melódicos y márgenes de dispersion. *Moenia13*: 69-92.
- Cantero Serena, F. J. 2019. Análisis prosódico del habla: más allá de la melodía. In: Álvarez Silva, M. R.; A. Muñoz Alvarado & L. Ruiz Miyares (eds.): *Comunicación Social: Lingüística, Medios Masivos, Arte, Etnología, Folclor y otras ciencias afines*. Volumen II (pp. 485–498). Santiago de Cuba: Ediciones Centro de Lingüística Aplicada.
- Cantero Serena, F. J. 2002. *Teoría y análisis de la entonación*, Barcelona: Ed. Universitat de Barcelona.
- Deme, A. & Markó, A. 2013. Lengthenings and filled pauses in Hungarian adults' and children's speech. In: Eklund, R. (Ed.) *Proceedings of DiSS 2013, The 6th Workshop on Disfluency in Spontaneous Speech KTH Royal Institute of Technology* (pp. 21-24). Stockholm, Sweden: Department of Speech Communication and Music Acoustics, Royal Institute of Technology.
- Devís, E. 2011. La entonación del español hablado por italianos. *Didáctica (Lengua y Literatura) 23* (2011): 35-58.
- Di Cristo, A. 1982. *Prolegomènes à l'étude de l'intonation*. Micromelodie. Paris: CNRS.
- Eklund, R. 2004. *Disfluency in Swedish human-human and human-machine travel booking dialogues*. PhD thesis, Linköping Studies in Science and Technology, Dissertation No. 882, Department of Computer and Information Science, Linköping University, Sweden.
- Estruch, M., Garrido, J. M., Llisterri, J., & Riera, M. 2007. Técnicas y procedimientos para la representación de las curvas melódicas. *RLA, Revista de Lingüística Teórica y Aplicada*45: 59-87, from the World Wide Web: [http://liceu.uab.es/~joaquim/publicacions/Estruch\\_Garrido\\_Llisterri\\_Riera\\_Metodos\\_Entonacion\\_07.pdf](http://liceu.uab.es/~joaquim/publicacions/Estruch_Garrido_Llisterri_Riera_Metodos_Entonacion_07.pdf)[last accessed 12/02/2020]
- Fonseca, A. & Cantero Serena, F. J. 2011. Características da entonação do espanhol falado por brasileiros. *Anais do VII Congresso Internacional Abralin*. Ed. Abralin(pp. 84-98). Associação Brasileira de Lingüística, Curitiba (Brasil).

- Font-Rotchés, D. 2007. *L'entonació del català*. Biblioteca Milà i Fontanals 53. Barcelona: Publicacions de l'Abadia de Montserrat.
- Font-Rotchés, D. 2008. Els patrons entonatius de les interrogatives absolutes del català central. *Llengua i Literatura* 19: 299-329.
- Font-Rotchés, D. 2009. Les interrogatives pronominals del català central. Anàlisi melòdica i patrons entonatius. *Els Marges. Revista de llengua i literatura* 87: 41-64.
- Font-Rotchés, D. & Mateo Ruiz, M. 2011. Absolute interrogatives in Spanish: a new melodic pattern. *Actas do VII congresso internacional da ABRALIN* (pp. 1111-1125). Curitiba (Brasil).
- Garrido, J. M. 1991. *Modelización de patrones melódicos del español para la síntesis y el reconocimiento*. Bellaterra: Departament de Filologia Espanyola, Universitat Autònoma de Barcelona.
- Garrido, J. M. 1996. *Modelling Spanish Intonation for Text-to-Speech Applications*. PhD dissertation. Departament de Filologia Espanyola, Universitat Autònoma de Barcelona.
- Goldman-Eisler, F. 1973. *Psycholinguistics. Experiments in Spontaneous Speech*. New York: Academic Press.
- Gósy, M. 2002. A megakadásjelenségek eredete a spontán beszéd tervezési folyamatában. [The origin of disfluency phenomena in the spontaneous speech planning process]. *Magyar Nyelvőr* 126: 192-204.
- Gósy, M. y Kresz, V. 2018. Phrase-final Lengthening of Phonemically Short and Long Vowels in Hungarian Speech across Ages. In: Gósy, M. y Grácz, T. E. (Eds.) *Challenges in analysis and processing of spontaneous speech*. Research Institute for Linguistics, Hungarian Academy of Sciences. 99-126.
- Guaitella, I. 1996. Analyse prosodique des hésitations vocales: propositions pour un modèle rythmique. *R.P.A.* 118-119, 113-145.
- Hualde, J. I. 2014. *Los sonidos del español*. CUP.
- Kao, W. 2011. *La entonación de enunciados declarativos e interrogativos en chino mandarín hablado por taiwaneses*. Trabajo de investigación final de Máster. Laboratori de Fonètica Aplicada de la UB.
- Lickley, R. J. 1994. Detecting disfluency in spontaneous speech. PhD dissertation, University of Edinburgh.
- Lickley, R. J. 2015. Fluency and disfluency. In: Redford, M. (Ed.). *The Handbook of Speech Production*. John Wiley Blackwell: 445-469.
- Machuca Ayuso, M. J. 2018. Pausas sonoras y bilingüismo. *Estudios de Fonética Experimental XXVII*, 2018, 75-95.

- Machuca Ayuso, M. J. & Ríos Mestre, A. 2016. Estructura formántica de las pausas sonoras en español. In: Fernández Planas, A. Ma. (Ed.) (2016): *53 reflexiones sobre aspectos de la fonética y otros temas de lingüística*, Barcelona, pp. 67-76.
- Maclay, H. & Osgood, C. 1959. Hesitation phenomena in spontaneous English speech. *Word* 15, pp. 19-44.
- Martorell, L. 2011. *Les interrogatives absolutes de l'espanyol parlat pels suecs*. Trabajo Final de Máster. Barcelona: Universitat de Barcelona, Facultat de Formació del Profesorado, 2011.
- Neuberger, T. 2014. *A spontán beszéd sajátosságai gyermekkorban*. [Characteristics of spontaneous speech in childhood]. *Beszéd. Kutatás. Alkalmazás*. Budapest: ELTE Eötvös Kiadó.
- Odé, C. & van Heuven, V. J. 1994. *Experimental studies of Indonesian prosody*. Dep. of Languages and Cultures of Southeast Asia and Oceania, University of Leiden.
- Olaszy, G. & Koutny, I. 2001. Intonation of Hungarian Questions and their prediction from text. In: Puppel, S. & Demenko, G. (Eds.) *Prosody 2000, Speech recognition and synthesis*, Poznań 2001.
- Ortega-Llebaria, M. 2006. Phonetic Cues to Stress and Accent in Spanish. In: *Selected Proceedings of the 2nd Conference on Laboratory Approaches to Spanish Phonetics and Phonology*, ed. Manuel Díaz-Campos, Somerville, MA: Cascadilla Proceedings Project, 104-118.
- Patiño, E. 2008. *Prosodic Comparative Study of Mexico City and Madrid Spanish*. Freie Universität Berlin, Germany. Escuela Nacional de Antropología e Historia, Mexico.
- Prieto, P., Borràs-Comes, J. & Roseano, P. (coords.). 2010-2014. *Interactive Atlas of Romance Intonation*. Web page: <http://prosodia.upf.edu/iari/> [last accessed 12/02/2020]
- Rebollo Couto, L. 1997. Pausas y ritmo en la lengua oral. Didáctica de la pronunciación. *ASELE Actas VIII* (Centro Virtual Cervantes), pp. 667-676.
- Rodríguez et al. 2001. Annotation and analysis of disfluencies in a spontaneous speech corpus in Spanish. In: *ITRW on Disfluency in Spontaneous Speech (Diss '01)*, Edinburgh, Scotland, UK, August 29-31. ISCA Archive, [https://www.isca-speech.org/archive/open/diss\\_01/dis1\\_001.html](https://www.isca-speech.org/archive/open/diss_01/dis1_001.html), pp. 1-4.[last accessed 12/02/2020]
- Rodríguez et al. 2015. Las pausas en el discurso de individuos con demencia tipo Alzheimer. Estudio de casos. *Revista de Investigación en Logopedia* 1 (2015) 40-59.
- Shriberg, E. 1994. *Preliminaries to the theory of speech disfluencies*. PhD dissertation, University of California, Berkeley.

- Stepanova, S. 2007. Some features of filled hesitation pauses in spontaneous Russian. In: J. Trouvain y W. J. Barry (Eds.): *Proceedings of ICPhS 2007. 16th International Congress of Phonetic Sciences, Saarbrücken*, pp. 1325-1328.
- t'Hart, J. T., Collier, R. & Cohen, A. 1990. *A perceptual study of intonation. An experimental-phonetic approach to speech melody*. Cambridge: Cambridge University Press
- Toledo, G. 2010. Métricas rítmicas en tres dialectos Amper-España Rhythmic metrics in three dialects of Amper-Spain. *Estudios Filológicos* 45: 93-110. [https://scielo.conicyt.cl/scielo.php?pid=S0071-17132010000100008&script=sci\\_arttext](https://scielo.conicyt.cl/scielo.php?pid=S0071-17132010000100008&script=sci_arttext) [last accessed 12/02/2020]
- Villa Villa, J. et al. 2017. Las vocales de relleno en español: nuevos datos y algunas reflexiones. In: Ruiz Minares et al. (Eds): *Nuevos estudios sobre Comunicación Social, Vol. I.*, Centro de Linguística Aplicada, Santiago de Cuba. pp. 165-169.
- The source of the image in Figure 1: from the World Wide Web: [https://www.freepik.com/free-vector/map-spain\\_2454242.htm#page=1&query=spain%20maps&position=4](https://www.freepik.com/free-vector/map-spain_2454242.htm#page=1&query=spain%20maps&position=4) [last accessed 12/02/2020]