REVISTA DA **ABRALIN**

RESEARCH REPORT

Prosodic perceptive thresholds in Spanish: intensity and duration

Kata Baditzné PÁLVÖLGYI
Eötvös Loránd University (ELTE)

ABSTRACT

The following study aims at examining prosodic thresholds of perception for Spanish, parting from the algorithm of analysis offered by the Prosodic Analysis of Speech (PAS) model by Cantero (2019). The approach is based on defining the prosodic values for each syllable after a process of standardization, by relativizing the prosodic value associated with each syllable to that of the preceding one. As the perception threshold in the case of melody was defined as 10% for Spanish (FONT-ROTCHÉS; MATEO RUIZ, 2011), this research investigates, by the help of an anonymous perceptive test submitted to 30 native Spanish speakers, what prosodic relative values of intensity and duration change with respect to the previous syllable make a syllable perceived as prosodically prominent. The informants of the test had to decide whether in an invented Spanish word syllables with different range of increase in intensity or duration result in perceptibility as far as prosodic prominence is concerned, and whether these prosodic cues contribute to stress identification. As for the results, a 30% of increase in duration with respect to the previous syllable is already perceivable, but intensity growth of 5% is not, and neither of these thresholds trigger automatically stress identification. The results of the research can be used as a point of reference in speech synthesis programs.

RESUMO

O seguinte estudo visa examinar os limiares prosódicos de percepção para o espanhol, a partir do algoritmo de análise oferecido pelo modelo de Prosodic Analysis of Speech (PAS) de Cantero (2019). A abordagem baseia-se na definição dos valores prosódicos de cada sílaba após um processo de padronização, relativizando o valor prosódico associado a cada sílaba ao da anterior. Como o

OPEN ACCESS

EDITED BY
- Monique Araújo (UnB)
- Sara Recio (UB)

EVALUATED BY
- Carolina Jorge Trujillo (ULL)
- Maria Amparo Devis Herraiz (UB)
- Juan María Garrido Almiñana (UNED)

DATES
- Received: 26/09/2022
- Acepted: 20/12/2022
- Published: 31/12/2022

HOW TO CITE
Pálvölgyi, Kata Baditzné (2022). Prosodic perceptive thresholds in Spanish: intensity and duration. *Revista da Abralin*, v. 21, n. 2, p. 193-208, 2022.

limiar de percepção no caso da melodia foi definido como 10% para o espanhol (Font-Rotchés; Mateo Ruiz, 2011), esta pesquisa investiga, por meio de um teste perceptivo anônimo submetido a 30 falantes nativos de espanhol, quais os valores relativos prosódicos de intensidade e a mudança de duração em relação à sílaba anterior tornam uma sílaba percebida como prosodicamente proeminente. Os informantes do teste tiveram que decidir se em uma palavra inventada em espanhol sílabas com diferentes amplitudes de aumento de intensidade ou duração resultam em perceptibilidade respeito à proeminência prosódica e se essas pistas prosódicas contribuem para a identificação do acento. Quanto aos resultados, já é perceptível um aumento de 30% na duração em relação à sílaba anterior, mas um aumento de intensidade de 5% não, e nenhum desses limiares aciona automaticamente a identificação do acento. Os resultados da pesquisa podem ser usados como ponto de referência em programas de síntese de fala.

KEYWORDS
Prosody. Intensity. Duration. Perception. Spanish.

PALAVRAS-CHAVE
Prosódia. Intensidade. Duração. Percepção. Espanhol.

Introduction

Prosody as a suprasegmental feature can only be interpreted in terms of relative differences: a syllable can be louder, higher in fundamental frequency (F0) or longer only with respect to its adjacent context. Relative differences can be expressed through standardized data, which is a common strategy in prosodic analysis in order to get rid of individual speaker-dependent characteristics. Melodic standardization was first done by using semitones in the 'Dutch School', also known as the IPO model, cf. e.g. 'T HART, COLLIER et al. (1990), followed later by various researches (ADRIAENS, 1991, BEAUGENDRE, 1994, ODÉ; VAN HEUVEN 1994). In Spanish, Garrido Almiñana (1991, 1996) and Estruch et al. (2007) worked with automatic stylization methods.

In this study, the Melodic Analysis of Speech (MAS) method is applied as a starting point, which represents prosodic change in terms of percentages. The model was first applied to the intonation of Spanish (CANTERO SERENA et al., 2005, CANTERO SERENA; FONT-ROTCHÉS, 2007, 2020, FONT-ROTCHÉS; MATEO RUIZ, 2011), and later it has been extended to the study of intonation in other languages as well, such as Catalan (FONT-ROTCHÉS, 2005, 2007, 2009) and Chinese (WEI-LI, 2011). The model has also been used for the description of the interlanguage intonation, for instance the Spanish spoken by Brazilians (FONSECA; CANTERO SERENA, 2011), Italians (DEVÍS, 2011), Swedes

(MARTORELL, 2011), Polish (URBANIK-PĘK, 2021) or Hungarians (BADITZNÉ, 2011, 2012, 2018, 2019, 2021, 2022a,b,c). The extended version of the model (later known as the Prosodic Analysis of Speech method, abbreviated as PAS, cf. CANTERO, 2019) does not only analyse melodic traits, but relative intensity and duration differences as well, by representing these values in terms of percentages, too.

Regarding the analysis, the first phase is acoustic, assisted by an acoustic analysis software such as Praat (BOERSMA; WEENINK, 2020): by the help of the program, the prosodic value for each syllable is detected (in the case of intonation, measured in Hz; in the case of intensity, in dB, and in the case of duration, in seconds). The second step is the prosodic representation: in order to concentrate only on the prosodically relevant features, it is necessary to ignore the irrelevant variations and reduce the data in the case of each syllable to a characteristic prosodic value. The next step is the standardization of the data: the graphs of prosodic data do not represent the absolute values, but rather the relative ones, since each syllable receives a percentage based on its prosodic rise / fall with respect to the previous syllable.

In the case of the melodic aspect, the process outlined in Cantero (2019, p. 489-491) is followed. The utterances are visualized using the Praat acoustic analysis software (BOERSMA; WEENINK, 2020), and then we proceed to the phase of obtaining the F0 in the case of each syllable. First, we get rid of irrelevant micromelodic variations by reducing each syllable to a characteristic tonal value. In the case of pitch instability within a syllable, the extreme values of F0 are taken. Each absolute value (measured in Hertz) is then converted to a relative value, depending on the value directly preceding it: the first value in the utterance is assigned an arbitrary value '100', and the following values represent the tonal distance measured in % from to the previous syllable.

The standardization of intensity and duration data follows the same pattern as the analysis: the values measured for each syllable are taken into account. In the case of intensity, the value measured at the intensity peak of each vowel, and in the case of duration, Cantero (2019) recommends using the values of the tonal distance between intensity peaks; Weeber (2021), however, suggests comparing the duration of each syllable to the average duration of the syllables of individual speakers as well. In this study, the relative values of duration were calculated based on the measured duration of each syllable and not the pitch distance between peaks, and the relativization was carried out always with respect to the previous syllable.

In former studies (cf. e. g. RIETVEL; GUSSENHOVEN, 1985), the perceptive threshold for melody in Spanish has been established as 1,5 semitones; for intensity, Dorta et al. (2017) set up 4 dB. As for duration, Rossi (1972) considers 27% as an already perceivable difference, whereas for Fernández Planas & Celdrán (2003), the threshold of minimally perceivable differences is of 33,33%. For Toledo (1988), this value is of 30-40 ms for the “barely perceivable differences”, these defined as somewhat longer, 40-60 ms by Marrero (2008).

A 10% of melodic change is considered to be perceivable (FONT-ROTCHÉS; MATEO RUIZ, 2011, p. 1113) within the PAS framework (CANTERO, 2019), but no such thresholds had been discovered for intensity or duration based on the same approach. This study aims to fill this gap, by establishing

prosodic thresholds of perceivability in the case of intensity and duration applied to Spanish, as well as their relation to stress identification.

Word stress is the result of the prominence given to a syllable compared to the other syllables in the word (HUALDE et al., 2010, p. 103) by changes in the fundamental frequency, intensity or duration with respect to its context (QUILIS, 1999, p. 385). Thus, the three prosodic characteristics that can have a prominent role in stress perception are tone, intensity and duration, but until this day there is no complete unanimity in the literature on whether stressed syllables in Spanish are pronounced in a higher tone, with longer duration or with greater intensity compared to their adjacent context. According to Navarro Tomás (1964), stressed syllables are characterized by higher intensity, according to Llisterri et al. (2003), by higher F0, and the latter is accompanied by longer duration according to Ortega-Llebaria (2006). The second purpose of this study hence is to contribute to these investigations by finding out what type of prosodic prominence besides melody (i.e., intensity or duration) indicates for listeners that the syllable in question is stressed.

The research questions, thus, are the following:

(1) What range of intensity and duration changes from the previous syllable can be considered as perceptive thresholds in Spanish?

(2) Is it prominence in intensity or rather prominence in duration that contributes to the syllable's perception as stressed by native Spanish speakers?

In order to answer these questions, an anonymous survey submitted to 30 native speakers of Spanish was carried out, in which 19, prosodically different but segmentally identical Spanish audio samples produced by a robot were listened to and judged according to the perceived prosodic prominence.

1. Method and informants

As a first step in order to establish the perceptive thresholds of prosodic prominence, a non-word was chosen, which, based on Spanish phonotactic rules, could as well be a loanword adapted to Spanish, and the spelling of which does not clearly indicate which syllable is stressed: "Katasah". Basically, words ending in a consonant in Spanish (except for "n" and "s") are stressed on the last syllable (based on this, the assumed pronunciation would be [kata'sah], since the otherwise silent "h" is pronounced in the case of loanwords, cf. e. g. Spanish "hámster" ['xamster] < English "hamster"), however, if the last consonant is realized as a silent "h", the word actually would end in a vowel and as such would be stressed on the penultimate syllable, thus [ka'tasa]. This "insecurity" was necessary so that the listeners would not assume with certainty in advance which syllable they should hear the accent on. Native informants heard the non-word inserted in a sentence ("A Katasah no hemos ido", 'We didn't go to Katasah') 19 times, but each sample contained exactly one difference compared to the others.

The original sentence was read by a robot with Castilian Spanish pronunciation in a male voice (using an application at http://texttospeechrobot.com/ with default pitch and speed settings, version 4 of 2022) and the first four syllables ("A ka ta sah") were reproduced all at the same volume (77 decibels) and

duration (0 .16 sec) and with an F0 value between 107-118 Hz (i.e. a pitch difference of no more than 10%, so as to guarantee tonal stability perceived as level tone).

The use of a synthesized speech sample read by a robot was chosen instead of a natural speech sample because in the later phase of the research, when in each sample one segment needed to be manipulated, the manipulated segment could have been easily recognizable in its natural context. In a completely synthetized voice, on the other hand, a manipulated segment would more naturally remain unnoticed and fit in the context.

Although intonation itself is not dealt with in this paper, stable tonal values within syllables were also important in order to avoid possible differences in perception, as, according to Barnes et al. (2012), not only the duration of an F0 movement, but global contour shape and the localization of F0 turning points also do influence hearers' perception about how they identify the exact alignment of a certain tonal movement. The stability of melodic values besides intensity and duration thus was a requirement to contribute to the objectivity of the research, so as not to influence stress perception.

Although the robot had not originally produced these values in the original audio, the sound recording was manipulated in Praat (BOERSMA; WEENINK, 2020) in such a way that these values were produced, so that in the prototype version there should be no objectively detectable prosodic prominence between the first four syllables. After that, manipulated files were created, based on the prototype, in which one syllable per each sample (more precisely, in which one syllable of the word "Katasah") is 1<2%, 2<5%, or 5<% higher in terms of intensity compared to the previous syllable, and in terms of duration, it was 10<20%, 20<30%, or 30<% longer. These values of proportional growth were chosen based on previous research, according to which stressed syllables are accompanied by a relative increase within this interval of prosodic change (typically less than 5% in intensity and 30% in duration in stressed syllables as an average, and only superior associated with first melodic peaks in the case of duration, and in both values when aligned to the nuclear (i.e. utterance-final) accent, cf. BADITZNÉ, 2022b,d, 2022e: 79). The manipulation had to be carried out in Praat (and not in the original text-to-speech synthesis (TTS) program), because, as opposed to the TTS program used, Praat is a software that permits acoustic analysis, by visualizing prosodic values, this way allowing their identification. Based on Cantero's (2019) PAS model, it could be easily calculated on the basis of the identified prosodic values in dB and ms, how loud a segment should be of over 5% of intensity growth, or how long a 30% longer segment should last with respect to the previous segment, for example.

In each case, only one segment was affected by the change, in the case of intensity, the syllable was replaced by a syllable pronounced with the necessary volume extracted from a recording played several times at different volumes (1<2%, 2<5% and 5<% louder versions of the same audio file) and then copied in Praat, and in the case of the duration, the syllables were lengthened to the necessary duration (10<20%, 20<30% and 30<% longer) by partial segment copying in Praat. This means that, when e. g. in the word "Katasah", the syllable "Ka" was 1<2% higher in intensity compared to the previous preposition "a", there was no difference in the duration of the first four syllables ("A ka ta sah"), all of them were exactly 0.16 seconds long. Similarly, in the sample containing the syllable "Ka" lengthened by 10<20% compared to the duration of the previous syllable, each of the first four syllables ("A ka ta sah") were exactly of 77 decibels.

Figure 1 shows the spectrogram image of the prototype sample sentence, in which it can be seen that the first four syllables share the prosodic values of intensity and duration, and syllables are also melodically identical. In Figure 2, an utterance with over 5% of intensity growth on the syllable "Ka-" with respect to the previous syllable is observable, but the rest of the prosodic values are identical in the case of the first four syllables ("A Ka ta sah"), whereas in Figure 3 there is an increase in duration on the syllable "Ka-" as compared to the previous syllable, with no prosodic change otherwise on the first four syllables.

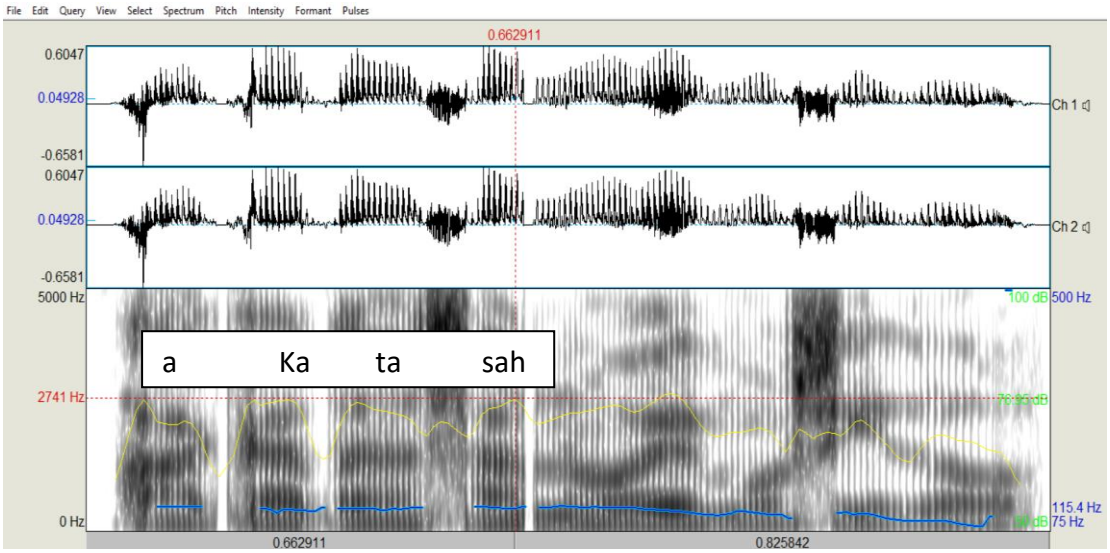

FIGURE 1 - Praat image of the utterance prototype "A Katasah no hemos ido", showing no prosodic prominence on the first four syllables.
Source: the author.

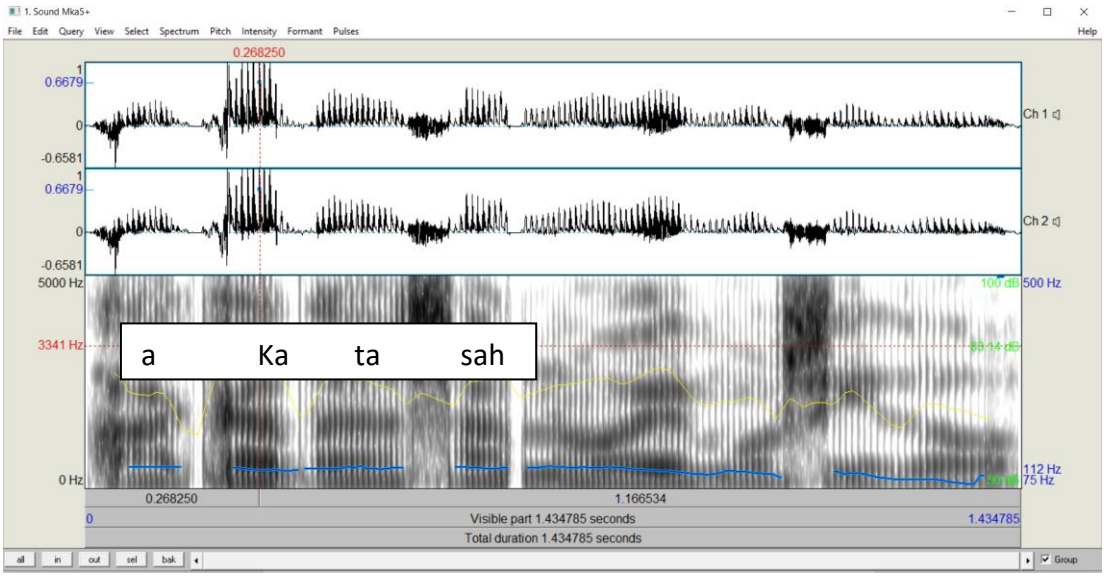

FIGURE 2 - Praat image of the utterance "A Katasah no hemos ido", with 5%< increase in intensity on the syllable "Ka-" with respect to the previous syllable
Source: the author

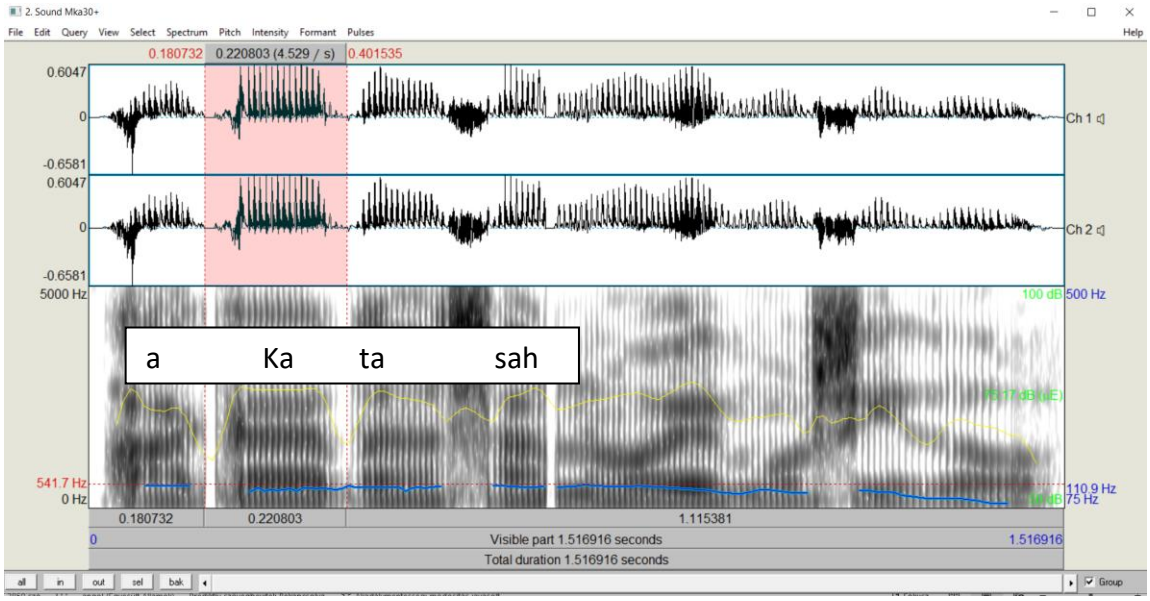

FIGURE 3 - Praat image of the utterance "A Katasah no hemos ido", with 30%< increase in duration on the syllable "Ka-" with respect to the previous syllable.
Source: the author

The following table illustrates the properties of the manipulated audio files. It can also be seen that the samples appeared in an aleatory order (but usually a sample manipulated on one of the syllables in terms of intensity was followed by a sample manipulated on one of the syllables in terms of duration).

Code	properties	order
Meq	intensity: 77 dB, duration 0.16 s on the first 4 syllables	5
Mka1+	duration 0.16 s on the first 4 syllables, ka- 1<2% higher in intensity than the previous syllable	11
Mka2+	duration 0.16 s on the first 4 syllables, ka- 2<5% higher in intensity than the previous syllable	1
Mka5+	duration 0.16 s on the first 4 syllables, ka- 5<% higher in intensity than the previous syllable	13
Mta1+	duration 0.16 s on the first 4 syllables, -ta- 1<2% higher in intensity than the previous syllable	3
Mta2+	duration 0.16 s on the first 4 syllables, -ta- 2<5% higher in intensity than the previous syllable	15
Mta5+	duration 0.16 s on the first 4 syllables, -ta- 5<% higher in intensity than the previous syllable	7
Msa1+	duration 0.16 s on the first 4 syllables, -sah 1<2% higher in intensity than the previous syllable	17

Msa2+	duration 0.16 s on the first 4 syllables, -sah 2<5% higher in intensity than the previous syllable	19
Msa5+	duration 0.16 s on the first 4 syllables, -sah 5<% higher in intensity than the previous syllable	9
Mka10+	intensity 77 dB on the first 4 syllables, ka- 10<20% longer than the previous syllable	18
Mka20+	intensity 77 dB on the first 4 syllables, ka- 20<30% longer than the previous syllable	8
Mka30+	intensity 77 dB on the first 4 syllables, ka- 30<% longer than the previous syllable	16
Mta10+	intensity 77 dB on the first 4 syllables, -ta- 10<20% longer than the previous syllable	6
Mta20+	intensity 77 dB on the first 4 syllables, -ta- 20<30% longer than the previous syllable	14
Mta30+	intensity 77 dB on the first 4 syllables, -ta- 30<% longer than the previous syllable	4
Msa10+	intensity 77 dB on the first 4 syllables, -sah 10<20% longer than the previous syllable	12
Msa20+	intensity 77 dB on the first 4 syllables, -sah 20<30% longer than the previous syllable	2
Msa30+	intensity 77 dB on the first 4 syllables, -sah 30<% longer than the previous syllable	10

TABLE 1 - the prosodic traits of the sample sentences.
Source: the auhor

In the questionnaire (provided by Google Forms), 30 native speakers heard 19 identical sentences from a segmental point of view, in which only one prosodic difference could be discerned per sentence at the suprasegmental level. As for the informants' data, two of the native speakers were also native speakers of Catalan, and the proportion of European and Latin-American informants was 53% and 47%, respectively, their age ranging between 27 and 75 years (the average being 43,4), with 14 women and 16 men. For each sentence, they had to decide which of the syllables in "Katasah" is the loudest, the longest, or is the stressed syllable. The audios were accessible in the questionnaire through links. After each sample, the received questions were the following:

1) In the word "Katasah", which syllable is the longest?
2) In the word "Katasah", which syllable is the loudest?
3) In the word "Katasah", which syllable is stressed?

The possible answers for each question were "Ka", "ta", "sah", "none" and "I don't know".

2. Results

The following table shows the percentage of the informants who guessed correctly the prosodically prominent syllable based on the perception test. Correctly in this case means that they identified the exact syllable which was, after the manipulation, the prominent one (i. e. bore higher values of intensity or duration than the adjacent syllables), or, in the case of the prototype (sample 5), chose the option "none" for all the three questions (as they were not supposed to hear any prosodic differences between the first four syllables of the utterance). The gray fields indicate the data where this percentage

represented the majority of the votes, i.e. when the majority of those completing the test chose correctly the prosodic prominence, and in the case of stress, the percentage expresses the extent to which the real prosodic prominence of the sample was connected to the assumed stressed nature of the syllable (here the underlined value indicates the majority of votes).

Code	properties	% of correct answers	stressed nature of prosodically prominent syllable
Meq	intensity: 77 dB, duration 0.16 s on the first 4 syllables	26,7/0	3,3
Mka1+	duration 0.16 s on the first 4 syllables, ka- 1<2% higher in intensity than the previous syll.	13,3	3,3
Mka2+	duration 0.16 s on the first 4 syllables, ka- 2<5% higher in intensity than the previous syll.	23,3	6,7
Mka5+	duration 0.16 s on the first 4 syllables, ka- 5<% higher in intensity than the previous syll.	16,7	0
Mta1+	duration 0.16 s on the first 4 syllables, -ta- 1<2% higher in intensity than the previous syll.	43,3	<u>83,3</u>
Mta2+	duration 0.16 s on the first 4 syllables, -ta- 2<5% higher in intensity than the previous syll.	53,3	<u>76,7</u>
Mta5+	duration 0.16 s on the first 4 syllables, -ta- 5<% higher in intensity than the previous syll.	46,7	<u>86,7</u>
Msa1+	duration 0.16 s on the first 4 syllables, -sah 1<2% higher in intensity than the previous syll.	16,7	13,3
Msa2+	duration 0.16 s on the first 4 syllables, -sah 2<5% higher in intensity than the previous syll.	16,7	20
Msa5+	duration 0.16 s on the first 4 syllables, -sah 5<% higher in intensity than the previous syll.	20	16,7
Mka10+	intensity 77 dB on the first 4 syllables, ka- 10<20% longer than the previous syllable	23,3	10
Mka20+	intensity 77 dB on the first 4 syllables, ka- 20<30% longer than the previous syllable	26,7	20
Mka30+	intensity 77 dB on the first 4 syllables, ka- 30<% longer than the previous syllable	63,3	20
Mta10+	intensity 77 dB on the first 4 syllables, -ta- 10<20% longer than the previous syllable	60	<u>83,3</u>
Mta20+	intensity 77 dB on the first 4 syllables, -ta- 20<30% longer than the previous syllable	43,3	<u>70</u>

Mta30+	intensity 77 dB on the first 4 syllables, -ta- 30<% longer than the previous syllable	66,7	76,7
Msa10+	intensity 77 dB on the first 4 syllables, -sah 10<20% longer than the previous syllable	26,7	20
Msa20+	intensity 77 dB on the first 4 syllables, -sah 20<30% longer than the previous syllable	30	20
Msa30+	intensity 77 dB on the first 4 syllables, -sah 30<% longer than the previous syllable	36,7	16,7

TABLE 2 - The results of votes of the questionnaire on Spanish prosodic prominence.
Source: the authors.

It can be seen, that in the case of intensity, the majority of listeners guessed correctly the prosodic prominence only when it fell on the syllable "-ta-", which was the penultimate syllable in the pseudoword. Since in Spanish the penultimate syllable is the one usually stressed (DELATTRE, 1965), it is possible that listeners subconsciously thought to identify prosodic prominence there, so the results regarding intensity cannot be considered completely objective. In the case of duration, it also appears that the prosodic prominence detected on the syllable "-ta-" (already above 10% of lengthening) is merely the subconsciously supposed prominence on the penultimate syllable, so it is likely that only the lengthening above 30% is actually perceptible (in this case, the lengthenings of the first and the last syllables were also noticed by the informants, not only on the penult).

It is also visible in the case of the question related to stress, that even if informants perceived the place of prosodic prominence correctly, they did not connect it to the perception of stress, only in the case of the penultimate syllable. This again can be explained by the underlying assumption of native speakers to identify word stress on the penultimate syllable anyway. This inconsistency of the informants regarding correct prosodic prominence identification associated with stress identification when it is not the penultimate syllable prevents us from generalizing and accepting that this perceptive threshold for duration – 30% – can trigger stress perception.

3. Discussion

The results of the study concerning a minimal threshold as 30% for duration are in accordance with the findings of earlier literature: 27% for Rossi (1972) and 33,33% for Planas & Celdrán (2003). As informants recognized the segment lengthened by >30% even if it was not the penultimate syllable – which they could subconsciously regard as the prosodically prominent syllable because of the predominant word stress pattern of Spanish on the penult – the conclusions regarding durational threshold can be considered objective.

As for intensity, we can see that informants only identified prosodic prominence on syllables with at least 5% of intensity growth, but only if this syllable was the penult. On syllables other than the penult this prominence was not perceived, and although Dorta et al. (2017) establish 4 decibels as the perceptive threshold, which, converted into percentages, would be around 6%, as normal conversation intensity is between 60 and 70 decibels (Abramovicz et al., 2012), we cannot consider this over 5% as a perceptive threshold justified in this paper, as independently of the penultimate syllable – the assumed prevalent stress position for words - it is not identified.

The perception of stress again may have been influenced by the dominant Spanish stress pattern on the penultimate syllable, this is why it is possible that the respondents found this syllable to be stressed in all cases, even if the prosodic prominence was correctly detected by them on another syllable. This means that based on this perception test, it is not obvious that even in the case of duration, when informants found the prosodic prominence independently of syllable position correctly, a 30% of perceptive threshold is enough to identify word stress, as even in those cases the penultimate syllable was chosen almost unanimously by listeners as the stressed one.

Regarding the perception test, although it only involved Spanish-speaking informants, it could be further diversified based on the dialect of the informants and the research could be amplified in this aspect in the future. Also, due to the possibility that the sentence with the same segmental content heard many times in a row could cause difficulty in judging the consecutive samples, it would be worth repeating the experiment in such a way that distractors are placed between sentences with the same segmental content.

4. Conclusions

Prosodic phenomena are only interpretable in relative terms, this is why in the last decades various prosodic analysis approaches have emerged that are based on standardized data. Cantero's Prosodic Analysis of Speech (PAS) (2019) method represents prosodic values assigned to syllables as a succession of relative values in percentages, always expressing the percentual growth/decrease of a syllable with respect to the previous syllable. Within this analysis, the melodic perceptive threshold is 10% in Spanish: this is a minimum value to be identified as a melodic movement by the native ear. The aim of this research is to discover perceptive thresholds for intensity and duration between syllables as well, within the PAS framework.

In this study, a 19-item questionnaire was submitted to 30 native Spanish speakers so as to test two prosodic features – intensity and duration – from the point of view of perceptibility thresholds in terms of percentages. Based on the perception test conducted with native speakers, in which informants had to judge a Spanish audio sample manipulated in terms of intensity and duration, the difference in intensity below 5% is not perceptible, or at least we cannot rely on the data received, as they might be the result of a biased judgement influenced by the dominant Spanish stress pattern of penultimate stressed syllables, because only in the case of penultimate syllables did most of the listeners identify

prosodic prominence on syllables with this proportion of intensity growth as compared to the previous syllable. Lengthening above 30% on a syllable with respect to the previous syllable, however, should be considered as a perceptive threshold, as independently of syllable position this proportion of lengthening was recognized as prosodic prominence by the majority of the informants.

Informants only connected prosodic prominence to stress perception in the case of penultimate syllables, and even when they identified prosodic prominence correctly, if it was not on the penultimate syllable, the syllable was not considered to be stressed by them. This allows us to declare that based on our results, neither of the prosodic thresholds set previously in the research (the highest being 5%< of intensity growth and 30%< of duration growth) are enough by themselves to indicate the stressed nature of a syllable to listeners.

As a conclusion, the results obtained in the perceptive test suggest that for Spanish, over 30% of duration growth is already perceived as a prosodic prominence, but for intensity, not even an over 5% of growth has proved to be enough to be considered as prominent. Even if in former research (e. g. Baditzné 2022: 79) stressed syllables in Spanish tended to be associated with over 2% of intensity change and over 30% of duration change, these thresholds separately were not enough to indicate the stressed nature of a syllable, and in the case of intensity, neither over 5% of growth was sufficient by itself. This means that in the future the interplay of various factors of prosodic increase could be investigated, as the presence of more than one prosodic cue might indicate more objectively stress perception.

The outcome of the present study, with the main finding that underlines the existence of a 30% as perceptive threshold for duration in Spanish, can be applied as a reference in for instance designing speech synthesis output data, in which there is a desired prosodic prominence of duration. As for setting up a perceptive threshold for intensity and defining the role of intensity and duration thresholds in stress perception, further research is required.

Additional information

Evaluation: https://doi.org/10.25189/rabralin.v21i2.2077.R

Conflict of interest
The author has no competing interests to declare.

Research protocol and pre-registration
The conducted research was not preregistered in any independent, institutional registry.

Link to preprint
Available from: https://osf.io/9n7cu/?view_only=9fe5a50edb244ceda75212e20b4608fa
DOI 10.17605/OSF.IO/9N7CU

Data accessibility
https://docs.google.com/spreadsheets/d/1pSxhKwY1qraBodx9y1pEyquoCO0YWAVBR-jSqoumvQrs/edit?usp=sharing

Acknowledgements

Supported by the ÚNKP-21-5 New National Excellence Program of the Ministry for Innovation and Technology and by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

REFERENCES

ABRAMOWICZ, Jacques; KREMKAU, Frederick William; MERZ, Eberhard. Obstetrical ultrasound: can the fetus hear the wave and feel the heat? *Ultraschall Med.* 33(3), p. 215–217. 2012. doi: 10.1055/s-0032-1312759.

ADRIAENS, Léon Marie Henri. Ein Modell deutscher Intonation. Eine experimentell-phonetische Untersuchung nach den perzeptiv relevanten Grundfrequenzänderungen in vorgelesenem Text. Doctoral dissertation. Technological University of Eindhoven, 1991.

BADITZNÉ PÁLVÖLGYI, Kata. The intonational patterns used in Hungarian students' Spanish yes-no questions. *Phonica* n. 7, p. 80-99, 2011.

BADITZNÉ PÁLVÖLGYI, Kata. Spanish Intonation of Hungarian Learners of Spanish; Yes or No Questions. Biblioteca *Phonica* v. 15. Barcelona: Laboratori de Fonètica Aplicada. 2012.

BADITZNÉ PÁLVÖLGYI, Kata. La presencia de rasgos prelingüísticos en la entonación de la interlengua húngaro-española. Colindancias: *Revista de la Red de Hispanistas de Europa Central* n. 9, p. 237-248, 2018.

BADITZNÉ PÁLVÖLGYI, Kata. ¿Debería importarnos la pronunciación en la enseñanza del español con fines específicos? In: LINGUA. CORVINUS NYELVI NAPOK TANULMÁNYKÖTET. Budapest: Budapesti Corvinus Egyetem Corvinus Idegennyelvi Oktató- és Kutatóközpont. 2019. p. 196-207,

BADITZNÉ PÁLVÖLGYI, Kata. Rasgos del perfil melódico del español hablado por húngaros. In: DOLORS, FONT-ROTCHÉS; FRANCISCO, JOSÉ CANTERO SERENA *Entonaciones del español*. Barcelona: Ediciones Octaedro. 2021. p. 149-166.

BADITZNÉ PÁLVÖLGYI, Kata. Utterance-initial F0 movements in the spontaneous speech of Hungarian learners of Spanish. In: PROCEEDINGS OF THE 3RD INTERNATIONAL SYMPOSIUM ON APPLIED PHONETICS, Tarragona, Univ. Rovira i Virgili, 2022ª. p. 42-43.

BADITZNÉ PÁLVÖLGYI, Kata. Las características prosódicas del acento en la interlengua húngaro-española. *Verbum Analecta Neolatina* n. XXIII, 2022/1, p. 69-88, 2022b.

BADITZNÉ PÁLVÖLGYI, Kata. Tonal Peaks in the Spontaneous Speech of Vantage Level Hungarian Learners of Spanish. *Acta Linguistica Academica* n. 69(1), p. 59-73, 2022c. https://doi.org/10.1556/2062.2021.00436

BADITZNÉ PÁLVÖLGYI, Kata. The prosody of word stress in the spontaneous speech of Hungarian learners of Spanish. In: BOOK OF ABSTRACTS: ADVANCES IN SECOND/FOREIGN LANGUAGE ACQUISITION

CONFERENCE (ASEFOLA). Deptartment of Languages and Literature. School of Humanities and Social Sciences. University of Nicosia. 2022d. p. 42-43.

BADITZNÉ PÁLVÖLGYI, Kata. La prosodia del tramo prenuclear en dialectos europeos y americanos del español. *Acta Hispanica* 27, p. 25-39. 2022e.

BARNES, Jonathan, VEILLEUX, Nanette; BRUGOS, Alejna; SHATTUCK-HUFNAGEL, Stefanie.Tonal Center of Gravity: A global approach to tonal implementation in a level-based intonational phonology. *Laboratory Phonology* v. 3(2), p. 337–383, 2012.

BEAUGENDRE, Frédéric. Une étude perceptive de l'intonation du français. Doctoral dissertation. University of Paris XI, Orsay. 1994.

BOERSMA, Paul.; WEENINK, David. Praat: doing phonetics by computer [Computer program]. 2020.

CANTERO SERENA, Francisco José. Análisis prosódico del habla: más allá de la melodía. In: ÁLVAREZ SILVA et al. *Comunicación Social: Lingüística, Medios Masivos, Arte, Etnología, Folclor y otras ciencias afines*. Volumen II, Santiago de Cuba: Ediciones Centro de Lingüística Aplicada, 2019, p. 485–498.

CANTERO SERENA, Francisco José; RAÚL, Alfonso; BARTOLÍ, Marta; CORRALES, Anna; VIDAL, Maribel. Rasgos melódicos de énfasis en español. *Phonica* 1 (2005), p. 1-40, 2005.

CANTERO SERENA, Francisco José; FONT-ROTCHÉS. Dolors Entonación del español peninsular en habla espontánea: patrones melódicos y márgenes de dispersión. *Moenia*, n. 13., p. 69-92, 2007.

CANTERO SERENA, Francisco José; FONT-ROTCHÉS, Dolors Protocolo para el análisis melódico del habla. *Estudios de Fonética Experimental* n. 18., p. 17-32, 2009.

CANTERO SERENA, Francisco José; FONT-ROTCHÉS, Dolors. Melodic Analysis of Speech (MAS). Phonetics of Intonation. In: Abasolo, J.; de Pablo, I.; Ensunza, A. *Contributions on education*. Bilbao: Universidad del País Vasco, 2020, p. 20-47,

CANTERO SERENA, Francisco José. *Teoría y análisis de la entonación*. Barcelona: Universitat de Barcelona. 2002.

DELATTRE, Pierre. *Comparing the Phonetic Features of English, German, Spanish and French*. Heidelberg: Julius Groos Verlag. 1965.

DEVÍS, Empar. La entonación del español hablado por italianos. *Didáctica (Lengua y Literatura)* n. 23, p. 35-58, 2011.

DORTA, Josefa; MARTÍN GÓMEZ, José Antonio; JORGE TRUJILLO, Carolina. Intensity threshold: beyond pure tones. *Estudios de Fonética Experimental*, n. 26, p. 133-163, 2017.

ESTRUCH, Mónica; GARRIDO ALMIÑANA, Juan María; LLISTERRI, Joaquim; RIERA, Montserrat. Técnicas y procedimientos para la representación de las curvas melódicas. *Revista de Lingüística Teórica y Aplicada* n. 45, p. 59-87, 2007.

FERNÁNDEZ-PLANAS, Ana María; MARTÍNEZ CELDRÁN, Eugenio. El tono fundamental y la duración: dos aspectos de la taxonomía prosódica en dos modalidades de habla (enunciativa e interrogativa) del español. *Estudios de Fonética Experimental*, n. 12, p. 166-200, 2003.
FONSECA, Aline; CANTERO SERENA, Francisco José Características da entonação do espanhol falado por brasileiros. ACTAS DO VII CONGRESSO INTERNACIONAL DA ABRALIN. Curitiba: Abralin. Associação Brasileira de Lingüística. 2011, p. 84-98.

FONT-ROTCHÉS, Dolors; MATEO RUIZ, Miguel. Absolute interrogatives in Spanish: a new melodic pattern. ACTAS DO VII CONGRESSO INTERNACIONAL DA ABRALIN. Curitiba, 2011. p. 1111-1125.

FONT-ROTCHÉS, Dolors. L'entonació del català. Biblioteca Milà i Fontanals 53. Barcelona: Publicacions de l'Abadia de Montserrat. 2007.

FONT-ROTCHÉS, Dolors. Els patrons entonatius de les interrogatives absolutes del català central. *Llengua i Literatura* n.19, p. 299-329, 2008.

FONT-ROTCHÉS, Dolors. Les interrogatives pronominals del català central. Anàlisi melòdica i patrons entonatius. Els Marges. *Revista de llengua i literatura* n. 87, p. 41-64, 2009.

GARRIDO ALMIÑANA, Juan María. *Modelización de patrones melódicos del español para la síntesis y el reconocimiento*. Barcelona: Servei de Publicacions de la Universitat Autònoma de Barcelona. 1991.

GARRIDO ALMIÑANA, Juan María. Modelling Spanish Intonation for Text-to-Speech Applications. Doctoral dissertation. Departament de Filologia Espanyola, Universitat Autònoma de Barcelona. 1996.

HUALDE, José Ignacio; ESCOBAR, Anna María; OLARREA, Antxon; E. TRAVIS, Catherine. *Introducción a la lingüística hispánica*. Cambridge: Cambridge University Press. 2010.

ILLÉSNÉ WEEBER, Ágnes. Mexikói és spanyolországi hezitációs mintázatok a spontán beszédben. In: MENCZEL, G.; SPISÁK, J.: *Lazarillo fiatal hispanisták tanulmányai* v. 8, 2021, p. 91-108.

WEI-LI, Kao. La entonación de enunciados declarativos e interrogativos en chino mandarín hablado por taiwaneses. Master's Thesis. Laboratori de Fonètica Aplicada de la Universitat de Barcelona. 2011.

LLISTERRI, Joaquim; DE-LA-MOTA, Carme; MACHUCA, María; RIERA, Montserrat; WELLINGTON SALES RIOS, Antonio. The perception of lexical stress in Spanish. In: SOLÉ, M. J. et al. PROCEEDINGS OF THE XV INTERNATIONAL CONGRESS OF PHONETIC SCIENCES. 2003.

MARRERO, Victoria. La fonética perceptiva: trascendencia lingüística de mecanismos neurofisiológicos. *Estudios de Fonética Experimental*, n. 17, p. 207-245, 2008.

MARTORELL, Laura. Les interrogatives absolutes de l'espanyol parlat pels suecs. Master's Thesis: Universitat de Barcelona, Facultad de Formación del Profesorado. 2011.

NAVARRO TOMÁS, Tomás. La medida de la intensidad. *Boletín del Instituto de Filología de la Universidad de Chile* n. 16, p. 231-235, 1964.

ODÉ, Cecilia; J. van HEUVEN, Vincent. Experimental studies of Indonesian prosody. Leiden: Dep. of Languages and Cultures of Southeast Asia and Oceania, University of Leiden. 1994.

ORTEGA-LLEBARIA, Marta. Phonetic Cues to Stress and Accent in Spanish. In: DÍAZ-CAMPOS, M. SELECTED PROCEEDINGS OF THE 2ND CONFERENCE ON LABORATORY APPROACHES TO SPANISH PHONETICS AND PHONOLOGY. Somerville, MA: Cascadilla Proceedings Project. 2006, p. 104-118.

QUILIS, Antonio. *Tratado de fonología y fonética españolas*. Madrid: Ed. Gredos. 1999.
RIETVELD, Toni; GUSSENHOVEN, Carlos. On the relation between pitch excursion size and prominence. *Journal of Phonetics*, N. 13, p. 299-308, 1985.

ROSSI, Mario. Le seuil différentiel de durée. In: VALDMAN, Albert. *Papers in Linguistics and Phonetics to the Memory of P. Delattre*. The Hague-Paris: Mouton, 1972, p. 64-94.

T'HART, Johan; COLLIER, René; COHEN, Abraham. *A perceptual study of intonation*. An experimental-phonetic approach to speech melody. Cambridge: Cambridge University Press. 1990.

TOLEDO, Guillermo Andrés. *El ritmo en el español. Estudio fonético con base computacional*. Madrid: Gredos, 1988.

URBANIK-PĘK, Weronika. El Análisis Melódico del Habla. Un método para abordar el análisis de la entonación del español hablado por polacos. *Studia Iberystyczne: Universos de lingua portuguesa em debate*: Swiaty jezyka portugalskiego n. 19, p. 253-284, 2021.