BOUNDARY TONES IN SPANISH DECLARATIVES: MODELLING SUSTAINED PITCH

Eva Estebas-Vilaplana[1], Yurena M. Gutiérrez[2], Francisco Vizcaíno[3], Mercedes Cabrera[3]

[1]Universidad Nacional de Educación a Distancia, [2]Universitat Autònoma de Barcelona, [3]Universidad de las Palmas de Gran Canaria (Spain)

eestebas@flog.uned.es, yurenagg@gmail.com, francisco.vizcaino@ulpgc.es, mercedes.cabrera@ulpgc.es

ABSTRACT

The aim of this paper is to present a phonological description of the boundary tones in final and non-final declarative sentences in Spanish, drawn from a read news corpus and a dialogue corpus. The final clauses tend to finish with L*L% and sometimes L+H*L%. Four different pitch configurations can be found for non-final patterns: a rise (L*H%), a fall-to-mid (H*!H%), a fall-rise (H*LH%) and a sustained tone, which presents different phonetic manifestations depending on the pitch level of the previous accent (H*, !H* or L*). These findings question the validity of the traditional Sp_ToBI convention (HL%) to describe a sustained tone since it cannot account for a level pitch after !H* or L*. A new boundary tone, =%, is proposed whose feature for pitch height is underspecified. For this reason, it can adopt the values of H, !H or L, according to the pitch height of the last accent.

Keywords: sustained pitch, boundary tones, Sp_ToBI, underspecification, declarative intonation.

1. INTRODUCTION

The first descriptions of Spanish intonation by Navarro Tomás [9], based on a read corpus, state that declarative sentences are usually divided into two parts, namely, the *protasis,* which indicates that the sentence has not yet finished and there is more information to follow, and the *apodosis* where the information is final or complete. Each part is usually associated with an intonation unit, as in *La casa de Juan* (*protasis) / está a la venta (apodosis)* ("John's house / is on sale"). However, there might be cases in which the *protasis* includes more than one intonation unit as in *La casa de Juan / comprada en 2012 (protasis* with two intonation units*) / está a la venta* ("John's house / bought in 2012/ is on sale"). The *apodosis* is always made up of one intonation phrase.

The tone inventory proposed by [9] to describe the pitch patterns of declarative sentences is closely linked to the part of the utterance they are associated with. Thus, whereas the *apodosis* always ends with a *cadencia* or fall, the intonation units in the *protasis* can have different endings: (i) *anticadencia* (rise),

(ii) *semicadencia* (fall from high-to-mid), (iii) *semianticadencia* (rise from low to mid) and (iv) *suspensión* (sustained tone). The equivalences of these tones with the latest Sp_ToBI conventions ([5], [8]) are as follows.

Table 1: Equivalences between Navarro Tomás's tone inventory and the Sp_ToBI conventions.

Navarro Tomás tone inventory	Sp_ToBI
Cadencia	L* L%
Anticadencia	L* H%
Semicadencia	H* !H%
Semianticadencia	L* !H%
Suspensión	H* HL%

Navarro Tomás also mentions the existence in Spanish of a rise-fall or circumflex pattern (L+H*L%) and a fall-rise inflexion (H*LH%) but he does not treat them as being phonologically relevant.

Studies on the intonation of read news ([4]) as well as dialogues ([10]) have shown that both types of speech acts tend to divide sentences into small fragments, in the case of broadcasters to maintain the audience's attention and, in the case of dialogues, as part of the ongoing conversational process. Thus, the declarative sentences used in these speech acts tend to have a long *protasis* made up of several intonation phrases.

In this paper, we examine the boundary tones of Spanish declarative sentences obtained in a read news corpus and in a dialogue corpus, following the tenets of the Sp_ToBI system. In particular, we analyse the tonal configurations that contribute to express a final vs. no final information.

2. CORPUS

The data used in this study belongs to the *Glissando* corpus ([7]). This corpus includes two different sets of data: a news sub-corpus and a dialogue sub-corpus. Over 25 hours of speech are recorded by 28 speakers in two languages, Catalan and Spanish. 8 of the recorders are professional radio news broadcasters or advertising actors. The corpus is transcribed, aligned with the acoustic signal, and prosodically annotated.

2.1. The read news sub-corpus

The Spanish news sub-corpus includes a selection of news from *Cadena SER* radio. For the present study, 24 pieces of news have been annotated: 12 read by a female radio news broadcaster, and 12 read by a male professional reader with an 'advertising' profile. Each piece of news lasts approximately one minute. A total of 30 minutes of speech is analysed. The news data set is essentially composed of long declarative sentences made up of several intonation phrases.

2.2. The dialogue sub-corpus

The dialogue sub-corpus includes two kinds of interactions: 1) oriented dialogues (inspired in the *Map Task* [1]), and 2) informal conversations. For the present study, only one oriented dialogue is analysed. It was recorded by two advertising actors. One of the interlocutors had to find information about a business trip. The dialogue lasts 10 minutes. Even though the dialogue contains various sentence types, in this study only declaratives are analysed.

3. DATA ANALYSIS

The annotation conventions used for the analysis of the data follow the Sp_ToBI system ([2]) with its further revisions ([5], [6], [8]). Sp_ToBI describes intonation by means of two tones: (H)igh and (L)ow. *Pitch accents* are associated with stressed syllables and *boundary tones* with the right edge of the intonation phrase. Sp_ToBI includes six pitch accents which can be monotonal (L*, H*) or bitonal (L+H*, L+>H*, L*+H, H+L*). In the first version of the Sp_ToBI system, boundary tones were only monotonal, such as L% and H%. A mid boundary tone (M%) was also incorporated to account for those final pitch movements where the f_0 rises or falls into a mid pitch (as in *semicadencia* or *semianticadencia*). More recent revisions of the system ([6]) also included bitonal boundary tones to describe final complex pitch movements (LH%, HL%, HH%). The last versions of the Sp_ToBI system ([5], [8]) substitute the notation M% by !H%. This notation stands for mid pitch.

Apart from the tonal movements, Sp_ToBI also indicates the break indices (BI) at the end of the different levels of prosodic phrasing. In this study we use BI0 to show syllable reduction and BI3 and BI4 to indicate a minor and a major prosodic domain respectively. Following the same procedure as in [5], BI3 and BI4, apart from signalling prosodic structure, also indicate the status of the information contained in the prosodic unit. Thus, BI3 is annotated in those cases in which the information is

interpreted as *non-final*, i.e., there is more information to follow; and BI4 when the information is interpreted as *final* or complete, and consequently, its IP is normally followed by a break (silence or pause).

The annotation process was performed by four trained transcribers using *Praat* [3]. Each transcriber visualized a display of the signal (f_0 curve and waveform) and relied on auditive and visual information to annotate the intonation patterns. The annotations of the four transcribers were contrasted in order to reach a consensus on the final labelling.

4. RESULTS

The results show that both in the read news corpus and the dialogue corpus, declarative sentences were divided into rather short intonation phrases (IP) which can be classified into two groups: 1) final and 2) non-final. A final IP is usually followed by a break and the information is interpreted as complete. It corresponds to the *apodosis* section proposed by [9] and it is signalled by a BI4. Non-final IPs indicate that the information has not reached the end. They belong to the *protasis* and they have a BI3. In our data, the number of IPs in the *protasis* ranges from 1 till 5.

Both in the read news and in the dialogue, the nuclear pitch configuration presented at the end of the *apodosis* tends to be L*L%, as illustrated in Figure 1. In a few cases (12% in the read news corpus and 10% in the dialogue), the L+H*L% pitch configuration was used instead, as exemplified in Figure 6. The use of L+H*L% seems to prompt a more involved nuance. For the IPs in the *protasis*, four different pitch configurations were found: a rise (L*H%), as illustrated in Figure 1, a fall-to-mid (H*!H%), as in Figure 2, a fall-rise (H*LH%), as in Figure 3, and a sustained tone, as in Figure 4. The sustained pitch is transcribed as =% (see section 5 for more details on this notation).

Figure 1: Example of rise at the end of the *protasis* (L)+H*H% and a fall at the *apodosis* L* L% for the sentence *Being at the vivac, they could hear a howling wolf* (news corpus).
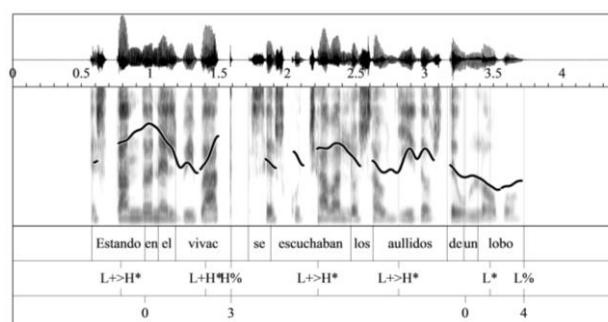
Figure 2: Examples of a fall-to-mid pitch (L+(!)H* !H%) at the *protasis* for the sentence *This Festival, Máxima Arte* (news corpus).
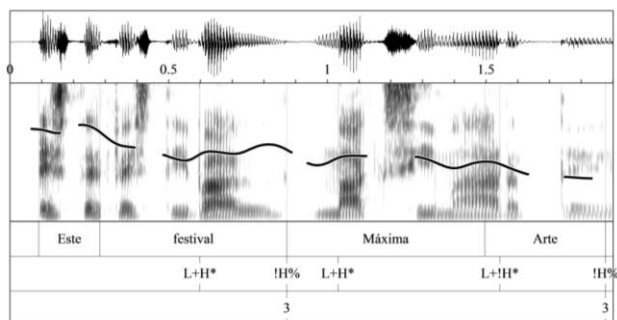
Figure 3: Examples of a fall-rise pitch (L+H* LH%) at the *protasis* for the sentence *which probabilities there are during the next year of suffering a depression* (news corpus).
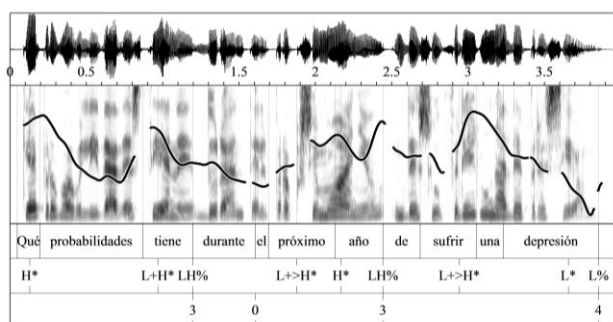
Figure 4: Examples of sustained boundary tones after a (L+)H* pitch accent at the *protasis* for the sentence *If it will be near the train station or the bus station* (dialogue).
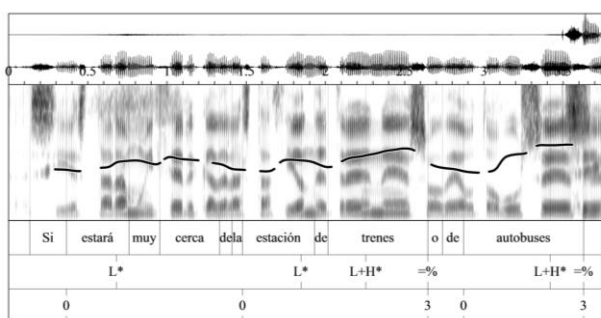

Table 2 shows the percentage of occurrence of the non-final pitch configurations both in the read news corpus and in the dialogue. The results presented in Table 2 show that the most recurrent pitch configuration at the end of non-final IPs in Spanish declaratives is sustained pitch, both in the news corpus and in the dialogue (32% and 43% respectively). Sustained pitch is, therefore, used in both speech acts as a means to indicate that the information has not finished and to maintain the interlocutor's expectations about forthcoming information. A more thorough analysis of the sustained pitch in our data shows that this tone may have different realizations depending on the pitch of the previous accent. If the last pitch accent is H*, sustained pitch remains high, as in Figure 4. If it is !H*, sustained pitch maintains the final mid pitch target, as in Figure 5. Finally, if the last pitch accent is L*, sustained pitch is now manifested as low pitch level, as in Figure 6. These findings prompt a revision of the traditional convention to indicate a sustained tone in Sp_ToBI (L+H* HL%) since the HL% boundary tone can in principle account for a sustained pitch after a H* pitch accent but not after !H* or L*.

**Table 2**: Pitch configurations at the end of non-final IPs and percentage of occurrence.

Non-final pitch configurations	% of occurrence	
	Read news	**Dialogue**
L*H%	28	23
H*!H%	26	21
H*LH%	14	13
=%	32	43

Figure 5: Example of a sustained boundary tone after a !H* pitch accent at the *protasis* for the sentence *The students' hall of residence has presented today a new audio book of its series "Poetry at the residence"* (news corpus).
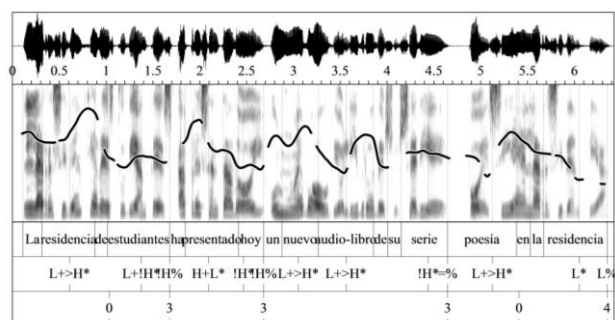
Figure 6: Example of a sustained boundary tone after a L* pitch accent at the *protasis* and L+H*L% at the *apodosis* for the sentence *and the shortest amount of time also possible* (dialogue).
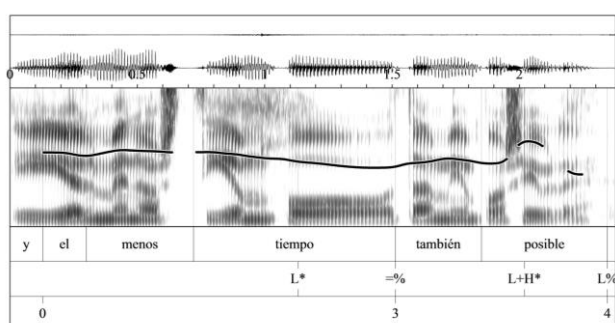
5. DISCUSSION

The intonational analysis of the declarative sentences in a news corpus as well as in a dialogue shows that one of the most recurrent pitch patterns at the end of non-final IPs is sustained pitch level. The data show that such final tune can have three different manifestations depending on the pitch level of the previous accent. Thus, sustained pitch can remain high if the previous accent is H*, mid if it follows !H* and low after L*. The height of this final sustained pitch, therefore, is contextually determined, and it adopts the pitch level of the preceding pitch accent.

The annotation convention for final sustained pitch within the Sp_ToBI system consists in an (L+)+H* pitch accent followed by an HL% boundary tone ([6], [8]). This tonal sequence has been attested in vocatives and calling contours. The HL% convention in principle accounts for sustained pitch from a preceding H* pitch accent, but it fails to describe sustained pitch from a mid or low tone, since the expected sequences !H*HL% and L*HL% would both stand for a rise-fall f_0 movement. As stated in [5], the HL% convention presents further problems. From a phonological point of view, L is used to prevent the previous H from rising and therefore it indicates that the pitch remains level. This interpretation, however, is at odds with other uses of the HL% boundary tone where the L is a real target, as in the sequence L*HL%, attested in contrastive sentences in Spanish. In this case, the L does not encode the information "maintain the H level" but it is an intended low target. On the other hand, a detailed phonetic analysis of the vocatives and calling contours show that, even though the perception of the final movement is sustained, the f_0 trace presents a final f_0 lowering, which does not attain the f_0 level of a target L% but still depicts a lowering trajectory. This lowering movement indicates that the L is not actually intended but responds to a physiological cause. Vocatives and calling contours are usually produced with a long duration, and a high amount of air to reach a long distance, all of which involves an extra laryngeal effort. There is a point in which this laryngeal effort cannot be maintained any longer, and hence it has to be relaxed due to physiological reasons. The falling pitch is, therefore, not a phonological target but the result of the relaxation of the laryngeal adjustments. Thus, the usage of the HL% boundary tone to describe a sustained pitch fails even in the cases where the final pitch accent is H*.

In this paper, we propose a new annotation convention to describe sustained pitch, namely, =%. It represents a phonological prime whose feature for pitch height is underspecified and whose phonological nature could be described as 'remain sustained'. As stated earlier, its height is contextually determined by the last tone in the final pitch accent (either H*, !H* or L*). Thus, =% can adopt the pitch values of H, !H or L. In this way, the phonological entity =% can have different phonetic realizations depending on the shape of the last tone in the final pitch accent.

6. CONCLUSIONS

In this paper, the boundary tones in final and non-final declarative sentences in Spanish extracted from a read news corpus and a dialogue corpus have been examined. The results show two pitch configurations for final clauses (L* L% and L+H* L%) and four patterns for non-final clauses: a rise (L*H%), a fall-to-mid (H*!H%), a fall-rise (H*LH%) and a sustained tone. This tone has different phonetic manifestations depending on the pitch of the previous accent (H*, !H* or L*). A new Sp_ToBI annotation convention has been proposed to model sustained pitch, =%. This boundary tone represents a phonological prime with an underspecified pitch height and with the encoding 'remain sustained'.

7. REFERENCES

[1] Anderson, A. et al. 1991. The HCRC Map Task Corpus. *Lanuage. and Speech* 34, 351-366.

[2] Beckman, M. Díaz-Campos, M., McGory, J. T., Morgan, T. A. 2002. Intonation across Spanish in the Tones and Break Indices framework. *Probus* 14, 9-36.

[3] Boersma, P., Weenik, D. 2013. *Praat: doing phonetics by computer.* http://www.praat.org.

[4] De la Mota, C., Rodero, E. 2011. La entonación en la información radiofónica. *Anejo de Quaderns de Filologia. La entonación hispánica*, Univ. València.

[5] Cabrera, M., Gutiérrez, Y., Vizcaíno, F., Estebas-Vilaplana, E. Submitted. Relevance and boundary tone choice in Castilian Spanish. Evidence from the read news Glissando corpus.

[6] Estebas-Vilaplana, E, Prieto, P. 2008. La notación prosódica del español: una revisión de Sp_ToBI. *Estudios de Fonética Experimental* 17, 265-283.

[7] Garrido, J. M., et al. 2013. Glissando: a corpus for multidisciplinary prosodic studies in Spanish and Catalan. *Lang. Resources and Evaluation* 47, 945-971.

[8] Hualde, J. I., Prieto, P. 2015. Intonational variation in Spanish: European and American varieties. In: Frota, S., Prieto, P. (eds.) *Intonational variation in Romance*. Oxford: OUP.

[9] Navarro Tomás, T. 1974 [1944]. *Manual de entonación española.* New York: Spanish Institute in the United States.

[10] Shriberg, E., et al. 1998. Can prosody aid the automatic classification of dialog acts in conversational speech? *Language and Speech* 41, 439-487.