Issues in the Phonology and Morphology of the Major Iberian Languages

Edited by
Fernando Martínez-Gil and Alfonso Morales-Front

Georgetown University Press  ♦  Washington, D. C.
Prosodic manifestation of syntactic structure in Catalan

Pilar Prieto

1 Introduction

In the last few years experimental research has shown that a variety of small differences in the scaling of $F_0$ peaks might be attributed to syntactic structure (Ladd & Johnson 1987, Ladd 1988, 1990, 1993, 1994, Kubozono 1989, 1992). For example, Ladd & Johnson (1987) argue that branching structure affects the height of the first peak of phrase initial constituents in English, and Kubozono (1992) reports that right-branching utterances in Japanese exhibit a downtrend pattern which is more distinct from that of left-branching utterances: initial accents of right-branching constituents undergo a process of pitch or metrical boosting, allowing listeners to identify syntactic structure through pitch scaling differences.

Recent results on the investigation of such phenomena are affecting the views about the mapping from syntactic to prosodic structure. In the last decade, a large body of evidence showing that phonological rules do not refer directly to a rich syntactic structure has lead researchers to postulate an intermediate and impoverished prosodic structure, which defines the domains to which phonological rules can refer. Even though there are differences between the various syntax-prosody mapping algorithms (Nespor & Vogel 1986, Selkirk 1984), the so-called Strict Layer Hypothesis (SLH) has been widely accepted. In short, SLH defends that prosodic structure is strictly layered, that is, that higher nodes must always contain lower nodes in the hierarchy.

This view has been inherited by what we might call the standard theory of intonational phonology (Pierrehumbert 1980, Beckman & Pierrehumbert 1986, Pierrehumbert & Beckman 1988). This framework proposes two levels of phrasing in English: intermediate phrase (which consists of one or more pitch accents plus a phrase accent), and intonational phrase (which consists of one or
more intermediate phrases plus a boundary tone). In accordance with SLH, intonational phrases cannot contain any other prosodic domain except intermediate phrases. Regarding the scaling of peaks, Pierrehumbert and her colleagues state that two possible manipulations of pitch height are predicted by their framework. First, a strictly local downstep, or accent-by-accent compression of the pitch range, whose application is bound to the domains of a given prosodic constituent (generally, intermediate or intonational phrases). And second, pitch reset or a pitch height increase attested at the beginning of a prosodic constituent. This pitch reset is described as a nonphonological manipulation of pitch height, a continuous variable whose changes are attributed to a free choice on the part of the speaker and constrained by the overall pragmatics of the situation.

Ladd (1988, 1990, 1993, 1994) and Kubozono (1992) argue that the Strict Layer Hypothesis is too strict to account for the pitch downtrend differences sensitive to branching structure mentioned before, and that a richer recursive prosodic structure is needed. (1) represents two Japanese sentences with left vs. right-branching structures using the recursive prosodic constituency proposed by Kubozono (1992: 385). A strictly layered prosodic structure would not be able to account for the pitch scaling differences exemplified by these two sentences, given that upstep or pitch boosting exclusively affects initial accents belonging to a right-branched constituent. Instead, if we allow for a richer recursive prosodic structure, we obtain a mechanism to single out the accents undergoingmetrical upstep:

(1)  

<table>
<thead>
<tr>
<th>a. Left-branching</th>
<th>b. Right-branching</th>
</tr>
</thead>
<tbody>
<tr>
<td>Utterance</td>
<td>Utterance</td>
</tr>
<tr>
<td>MP</td>
<td>MP</td>
</tr>
<tr>
<td>mp</td>
<td>mp</td>
</tr>
<tr>
<td>mp mp mp</td>
<td>mp mp mp</td>
</tr>
<tr>
<td>ao'yma-in a'ru daigaku</td>
<td>a'ni-no me'n-no eri'maki</td>
</tr>
<tr>
<td>Aoyama-in exist university</td>
<td>brother-GEN cotton-GEN muffler</td>
</tr>
<tr>
<td>‘a university in Ayoama’</td>
<td>‘(my) brother’s cotton muffler’</td>
</tr>
</tbody>
</table>
In light of the present controversy, there is a need to explore the relevance of metrical and syntactic structure in the scaling of $F_0$ peaks in more depth. In this regard, Catalan seems to be an appropriate case to study. In her description of Central Catalan intonation, Bonet (1984) mentions that the presence of constituent boundaries is signaled through pitch scaling differences, namely, by the presence of a low tone on the constituent-final accented syllable. According to her, the scaling of $F_0$ peaks (in particular, the pitch level attained by the constituent-final stressed syllable) is crucial to disambiguate right- vs. left-branching utterances like the ones shown in (2). In (2a) the stressed syllable in the word llança receives a low tone, and in (2b) it is vella that receives the low tone:

\[(2) \quad \text{a. Left-branching} \quad \text{b. Right-branching} \]

\[
\begin{align*}
\text{Utterance} & \\
\text{NP} & \\
\text{AP} & \quad \text{N} & \quad \text{VP} \\
| & | & | \\
\text{La vella} & \quad \text{llança} & \quad \text{l’amenaça} \\
\text{‘The old lance threatens him/her’} & \quad & \\
\end{align*}
\]

\[
\begin{align*}
\text{Utterance} & \\
\text{NP} & \quad \text{VP} \\
\text{NP} & \quad \text{V} & \quad \text{NP} \\
| & | & | \\
\text{La vella} & \quad \text{llança} & \quad \text{l’amenaça} \\
\text{‘The old lady threatens him/her’} & & \\
\end{align*}
\]

The aim of this study is twofold. First, to investigate how Catalan speakers disambiguate structurally ambiguous sentences of the type shown above and whether listeners are able to recognize their corresponding meaning. And second, to further explore Bonet’s hypothesis about the relationship between $F_0$ scaling and constituent boundaries in Catalan and, more generally, to seek evidence regarding the influence of syntactic/prosodic structure on pitch downtrend patterns. If pitch scaling were the only cue serving to disambiguate the structures in (2) (with no accompanying phrasing differences), this would indicate the need to revise the Strict Layer Hypothesis and advocate for a recursive type of prosodic structure.
2 Method

2.1 Production test

For many languages, it has been reported that prosodic features such as duration, amplitude and F₀ changes reflect syntactic structure and that speakers and listeners are able to use this prosodic information to disambiguate structurally ambiguous sentences (Klatt 1973, Lehiste 1973, Terken & Collier 1992, Venditti 1994). The main goal of the production part of the experiment is to analyze duration and fundamental frequency differences (concentrating on the scaling of F₀ peaks) between potentially ambiguous sentences in Catalan.

Methodologically, it is important to test the effects of syntactic structure on F₀ and duration patterns of potentially ambiguous sentences, that is, sentences that differ exclusively in constituent structure. It is well-known that F₀ and segmental durations are influenced by a variety of factors like the phonetic structure of neighboring segments, degree of stress, within-word position, sentence position, etc. Thus, by using sentences containing the same string of words, we are neutralizing the potential confounding effects of other factors, and more accurate comparisons are possible. (3) lists the four potentially ambiguous utterances used in the experiment, which involve right vs. left-branching structures of the type [NP][VP][NP] vs. [[AP][NP]][VP]. The list provides the two possible meanings of the utterances and separates the syntactic subject from the VP by a comma:

(3)
La vella, llança l'amaença 'The old lady threatens him/her'
La vella llança, l'amaença 'The old lance threatens him/her'
La vella, escolta la veu 'The old lady LISTENS to the voice'
La vella escolta, la veu 'The old girl scout sees her'
La poderosa, crema la casa 'The powerful lady sets the house on fire'
La poderosa crema, la casa 'The powerful cream helps her get married'
El vell, guarda la porta 'The old man guards the door'
El vell guarda, la porta 'The old guard is wearing it (the scarf)'

Several experimental studies have found that degree of embedding or constituency length significantly affects the height of the first F₀ peak at the start of each constituent (Ladd and Johnson 1987, Kubozono 1992, Venditti 1994).
The second set of sentences in (4) was designed to analyze the potential effects of degree of embedding on pitch scaling. They were obtained by adding a complement to the utterance-final constituent of each of the sentences in (3). By comparing each of the new sentences in (4) with its corresponding utterance in (3) we can test whether or not the height of the constituent-initial peak is affected by the depth of the constituent:

(4)  

La vella, llança l'amenaça terrible
'The old lady threatens him/her terribly'

La vella llança, l'amenaça terriblement
'The old lance threatens him/her terribly'

La vella, escolta la veu especial
'The old lady listens to the special voice'

La vella escolta, la veu amb ençís
'The old girl scout sees her with delight'

La poderosa, crema la casa de Vilamalla
'The powerful lady sets the Vilamalla house on fire'

La poderosa crema, la casa ben casada
'The powerful cream helps her get married for good'

El vell, guarda la porta del carrer
'The old man guards the main door'

El vell guarda, la porta ben content
'The old gard is happily wearing it (the scarf)'

Each utterance in (3) and (4) was placed in a disambiguating paragraph (see Appendix I). Then, three speakers of Central Catalan (PG, JS, and PP) read each paragraph from six to eight times with an effort to differentiate the two meanings of each sentence. In order to obtain two different pronunciations, speakers made a conscious attempt at differentiating between the two possible meanings of each sentence. To avoid the use of pauses to disambiguate the test sentences, subjects were asked to read the paragraphs at a relatively fast tempo.
The final recordings totalled 336 target utterances/paragraphs (16 paragraphs x 7 repetitions x 3 speakers = 336).

After completing the recordings, the target sentences were separated from their paragraphs and prosodically labelled using the TOBI annotation conventions (Pitrelli et al. 1994). The following $F_0$ and duration measures were taken, using the waveform, spectrogram and corresponding $F_0$ contour of each file: 1) phrase-initial $F_0$ value; 2) maximum $F_0$ peak for each pitch accent (H* or L*); 3) maximum $F_0$ boundary tone (H'); 4) phrase-final $F_0$ value; and 5) duration of each of the syllables in the sentence (only for speaker PP).

2.2 Perceptual test

The second part of the experiment consists of a perceptual test. Its main goal is to investigate whether listeners can easily identify the meaning corresponding to the pronunciations of the left- vs. right-branching utterances. The stimuli consisted of the sentences in (3) read by speakers PP and PG, for a total of 16 sentences. A five-second interval appeared between each stimulus, presented in random order within sentence-types. Eighteen university students and staff having Catalan as a first language participated in the experiment. The stimuli sentences were presented over a portable mini-disc recorder and the listeners were asked to decide on the meaning of each of the sentences (see Appendix 2). In each session, listeners were allowed to pause to make their judgment, but they were not permitted to hear a given sentence twice.

3 Analysis and discussion

3.1 Results of the production experiment

This section describes the intonational and durational differences found between the two productions of each sentence in (3) and (4), focusing on the potential differences in pitch downtrend patterns between right-branching and left-branching structures. In her study of Catalan intonation patterns, Bonet (1984: 15) claims that declarative sentences with more than one syntactic constituent have the following tonal characteristics: 1. All accented syllables belonging to the last constituent in the utterance have a low tone; 2. The last
accent in a constituent has a low tone; 3. The rest of the accents have a high tone. According to Bonet, the tonal distribution predicted by the patterns above is crucial to disambiguate the two meanings of ambiguous sentences such as La vella llança l’amenaça. When vella is the syntactic subject, the phonological constituent ends right after this word and the accent in ve is low. In contrast, when vella llança is the subject, it is the accent on llan that is low. She also notes that an optional high boundary tone can appear at the end of this constituent. The sentences in (5) below show Bonet’s intonational transcription of a left-branching utterance (5a) Una jove veu, l’amenaça ‘a young voice threatens him/her”and a right-branching utterance (5b) Una jove, veu l’amenaça ‘a young girl sees the threat’ (Bonet 1984: 20). The main difference between the two is the pitch level attained by the word jove: in (5a), being non-final in the constituent, it is high; instead, in (5b) it is low ( constituent-final accent):

\[
\begin{array}{c|c|c}
(5) & a. & Una jove veu & l’amenaça \\
& & \uparrow \uparrow & \uparrow \\
& & \uparrow \downarrow \uparrow & \downarrow \downarrow \uparrow \\
\end{array}
\]

\[
\begin{array}{c|c|c}
& b. & Una jove & veu & l’amenaça \\
& & \uparrow \downarrow \uparrow & \uparrow & \downarrow \downarrow \uparrow \\
\end{array}
\]

We employed a recent version of Pierrehumbert’s accent inventory (Beckman & Hirschberg 1993) to transcribe the examples. This theory of intonational transcription is based on the binary classification of local F0 minima or maxima located around accented syllables (pitch accents) and phrase boundaries (boundary tones). The claim is that we can obtain a phonologically adequate tonal transcription by using exclusively H and L tones. The decision between Hs and Ls is based on adjacent tonal values and the speaker’s register; for example, L tones are generally manifested as local minima, generally coinciding with the baseline of the speaker. There are two levels of phrasing.
An intermediate phrase, which consists of one or more pitch accents plus a phrase accent (or a simple H- and L- tone that controls the pitch from the last pitch accent to the end of the phrase). The other is an intonational phrase, which consists of one or more intermediate phrases plus a final boundary tone, which may also be high (H%) or low (L%), occurring at the end of the phrase.

(6) shows the intonational transcription of our speaker’s pronunciation of right-branching (6a) vs. left-branching utterances (6b).\textsuperscript{8} Below each transcription we display the actual F\textsubscript{0} contour of speaker JS producing the right-branching utterance La vella, llança l’amença in (6a) or the left-branching utterance La vella llança, l’amença in (6b). Intonational phrasing is the most salient difference between the two realizations of the sentence: an intermediate phrase boundary (phonetically marked by a high phrase accent H\textsuperscript{*}) is placed in two different positions depending on the constituent structure of the sentence (after the syntactic subject). The presence of a high boundary tone H\textsuperscript{-} is perceptually clear. Also, it is phonetically identified by a steady increase of the pitch level from the pitch accent on vella and on llança to the phrase boundary. Usually, the three stressed syllables are accented. With the exception of the last accents in the utterance, pitch accents are transcribed as high (H\textsuperscript{*}), as they all attain levels well above the baseline of the particular speaker. Utterance-final pitch accents are transcribed as L\textsuperscript{*}, since F\textsubscript{0} levels at this point generally approach the baseline of each speaker.\textsuperscript{9}

(6) Intonational transcription of right vs. left-branching utterances

\begin{quote}
a. H\textsuperscript{*} H\textsuperscript{-} H\textsuperscript{*} L\textsuperscript{*} L\textsuperscript{-} L\textsuperscript{\%}
La vella, llança l’amença
\end{quote}

\begin{center}
\includegraphics[width=\textwidth]{JS_contour.png}
\end{center}

JS contour of La vella, llança l’amença
The $F_0$ contours in (6) exhibit different pitch downtrend patterns. The first accent, aligned with the syllable *ve*, attains more or less the same pitch level in both sentences (the phrase accent, $H^-$, is present only in (6b), and aligned with the postonic syllable *lla*). The pitch level of the $H^*$ accent on the word *llança* is higher when *llança* starts a new intermediate phrase (after $H^-$) than when it is the last word in the constituent. Thus, the particular pronunciation in (6) shows that $F_0$ resets at the beginning of a new intermediate phrase. To test whether the two types of branching structure exhibit consistent differences in pitch downtrend patterns, we need to analyze the measures of the production test for the three speakers (JS, PP, PG). Table 1 shows the mean $F_0$ values of four target values corresponding to the following positions in the $F_0$ contour, for the two types of branching structure: first pitch accent in the sentence ($H^*$), phrase accent ($H$), utterance-final accent ($L^*$), and second pitch accent of the sentence (X or Y). The second $H^*$ pitch accent is named X or Y depending on its position within the prosodic constituent: X when the accent is at the beginning of a new intermediate phrase, and Y when it is constituent-final. Means were obtained by averaging all of the values of the sentences in (3) (8 sentences x 7 repetitions = 56 examples in each cell). Since the pitch range (or minimum and maximum $F_0$ values each speaker reaches) is different for each speaker, values in the $F_0$ contour are comparable only within speakers. The results in Table 1 show that speakers JS and PP exhibit a consistent difference between the pitch
level attained by X and Y: when the accent is the last in the constituent (Y), its peak is around 10 Hz lower than when it is constituent-initial (X). Speaker PG, though, does not exhibit a distinct pitch downturn pattern between the two sentences: the height of the second accent is practically the same in constituent-initial (X) or constituent-final (Y) position. Probably this is an indication that differences in pitch downturn are not crucial but they can help to disambiguate right-branching and left-branching structures. If that is the case, the perceptual score obtained using PG as stimuli sentences might be lower than using JS or PP sentences.

Table 1. Mean $F_0$ values (in Hz) of sentences in (3)

<table>
<thead>
<tr>
<th>Right-branching utterances (La vella, llança l’amenaça):</th>
<th>H*</th>
<th>H^-</th>
<th>X</th>
<th>L*</th>
</tr>
</thead>
<tbody>
<tr>
<td>JS</td>
<td>115</td>
<td>133</td>
<td>107</td>
<td>80</td>
</tr>
<tr>
<td>PP</td>
<td>190</td>
<td>213</td>
<td>184</td>
<td>149</td>
</tr>
<tr>
<td>PG</td>
<td>147</td>
<td>161</td>
<td>132</td>
<td>105</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Left-branching utterances (La vella llança, l’amenaça):</th>
<th>H*</th>
<th>Y</th>
<th>H^-</th>
<th>L*</th>
</tr>
</thead>
<tbody>
<tr>
<td>JS</td>
<td>115</td>
<td>98</td>
<td>119</td>
<td>80</td>
</tr>
<tr>
<td>PP</td>
<td>195</td>
<td>172</td>
<td>188</td>
<td>154</td>
</tr>
<tr>
<td>PG</td>
<td>148</td>
<td>134</td>
<td>148</td>
<td>106</td>
</tr>
</tbody>
</table>

The results in Table 1 also confirm our previous observations. We find that variation in intonational phrasing is a consistent strategy to disambiguate right-vs. left-branching utterances: to distinguish the two meanings of the sentence, Catalan speakers produced an intermediate phrase boundary after the syntactic subject. A consistent and clear phonetic correlate of phrasing is the presence of a H^- boundary tone and, optionally, a short pause at the end of the intermediate phrase. As Table 1 shows, there is a salient distance between the pitch height
attained by a Y and a H', namely, between 14 and 20 Hz difference between a Y and a following H'.

In sum, Bonet's hypothesis is partially confirmed. When Catalan speakers try to disambiguate sentences with left- vs. right-branching structures, there is a unanimous tendency to have phrasing differences. Against Bonet's claim, the actual phonetic correlates of phrasing structure are found to be a consistent presence of a H’ boundary tone, and not pitch lowering of constituent-final accents. Using Pierrehumbert's framework to interpret the data, we find that the height of the first peak H* (i.e., the one corresponding to the word vella) has the same pitch level in both types of branching structure for the three speakers (see Table 1). Again, Table 2 shows the mean F0 peak (in Hz) of the first H* for right- vs. left-branching structures, as well as utterance-initial and utterance-final F0 values. The results in Table 2 confirm that the three measures are rather constant in the two pronunciations of the sentence for a given speaker (they also display very low standard deviation values). For each speaker, the differences between the two populations were found to be nonsignificant (at p < 0.01).

Table 2. Mean F0 (in Hz) of different locations in the F0 contour

<table>
<thead>
<tr>
<th></th>
<th>Initial</th>
<th>H*</th>
<th>Final</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Right-branching utterances</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(La vella, llança l’amenaça)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JS</td>
<td>100</td>
<td>115</td>
<td>77</td>
</tr>
<tr>
<td>PP</td>
<td>176</td>
<td>190</td>
<td>136</td>
</tr>
<tr>
<td>PG</td>
<td>116</td>
<td>147</td>
<td>92</td>
</tr>
</tbody>
</table>

| **Left-branching utterances**   |         |     |       |
| (La vella llança, l’amenaça)    |         |     |       |
| JS                              | 102     | 115 | 79    |
| PP                              | 174     | 195 | 134   |
| PG                              | 118     | 148 | 98    |
In general, we conclude that Catalan displays a tendency to reset the pitch of a constituent-initial accent, even though it seems that such a reset is not an obligatory phonetic cue of phrasing in this language (cf. speaker PG in Table 3). We find an optional pitch reset at the beginning of a new intermediate phrase (after a high boundary tone). The difference in pitch scaling is one of the expected variations in pitch height, and follow from standard assumptions made by Pierrrehumbert and her colleagues. Thus, a strictly layered prosodic structure can account for the pitch downtrend differences found in the two types of branching structures, and the Strict Layer Hypothesis is thus not challenged by this data.

With respect to the role of duration, several studies have defended that preboundary lengthening is also a phonetic cue that serves to signal phrasing differences (Klatt 1973, Cooper & Paccia-Cooper 1980). In our data, we expected a significant amount of syllable lengthening before intermediate phrase-boundaries (i.e., syllables with a H). Table 3 compares the mean duration (in ms) of the word-final and penultimate syllables of the same word in preboundary and phrase-medial positions for speaker PP. The table shows very inconsistent results: out of the eight cases analyzed, only three have longer syllables in preboundary position. Thus, the duration patterns of the two branching structures reveal no significant differences, suggesting that variations in pitch play a more crucial role in distinguishing the two sentences.

<table>
<thead>
<tr>
<th>Target word</th>
<th>Pre-boundary Penult</th>
<th>Final</th>
<th>Constituent-medial Penult</th>
<th>Final</th>
</tr>
</thead>
<tbody>
<tr>
<td>vella</td>
<td>12</td>
<td>15</td>
<td>11</td>
<td>15</td>
</tr>
<tr>
<td>vella</td>
<td>8.5</td>
<td>8</td>
<td>7.3</td>
<td>8</td>
</tr>
<tr>
<td>poderosa</td>
<td>11</td>
<td>13</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>guarda</td>
<td>21</td>
<td>9.3</td>
<td>23</td>
<td>9.4</td>
</tr>
</tbody>
</table>
3.2 Degree of embedding

With the goal of exploring the potential effect of degree of embedding on the height of the constituent-initial peak, we created the new set of sentences shown in (4). We expected that longer constituents would trigger a pitch boost of the constituent-initial accent. For example, assuming that Catalan speakers would read the right-branching utterances in (7) with a phrase break after the syntactic subject *vella*, we expected the peak of the H* in *escolta* to be higher in (7b) than in (7a):\(^10\)

(7)  
a. \textit{La vella, escolta la veu}  
   ‘The old lady listens to the voice’

b. \textit{La vella, escolta la veu amb encís}  
   ‘The old lady listens to the voice with delight’

Yet, the 3 speakers produced unexpected phrase breaks in the case of (7b), which was pronounced as (8). The actual phrase boundary, which was expected to be after the syntactic subject (as in 7a) was moved to the right and located after the verb:

(8) \textit{La vella escolta, la veu amb encís}

The fact that speakers placed intonational boundaries in different locations made it impossible to compare the pitch levels attained by the constituent-initial peak in a word like *escolta*, and thus we were unable to test the effects of degree of embedding. Why did speakers consistently change the expected phrasing in sentences like (8)? Clearly, a sentence like *La vella, escolta la veu amb encís*, with a phrase break after the syntactic subject *vella*, is a very marked pronunciation of the sentence. This fact suggests that there is a certain degree of optionality in phrasing decisions and that the choice of prosodic domains is not exclusively dictated by syntax. Other factors, in this case constituency length, seem to play an active role in ruling out a pronunciation like (7b).
3.3 Results of the perceptual test

The goal of the perceptual test is to determine whether the production of the two types of branching structures can be reliably identified by Catalan listeners. To our surprise, the rate of identification was only moderate: each of the 18 listeners correctly classified between 60-80% of the stimuli. In general, the male stimuli (PG) presented more difficulties than the female (PP). This fact might be due to the following factors: 1., the more drastic F0 movements attested in the female contours might help in the perceptual task; 2., the pitch downtrend differences between the two types of branching structure shown by the female contours might help in identifying phrase boundaries.

The perceptual results reveal an interesting phenomenon regarding the influence of syntactic structure on phrasing decisions. For some reason, a phrasing like La vella llança, l’amenança (with a H’ phrase accent after llança) leads to ambiguity, especially when this pronunciation corresponds to the first presentation of the pair of sentences to be disambiguated. Such left-branching utterances reached a mean identification rate of around 40%, indicating that the listener consistently doubted in assigning an interpretation to these sentences. In contrast, the identification rate for right-branching utterances (cf. La vella, llança l’amenança) was 100% for the female speaker and 90% for the male speaker. Thus, we find a sharp contrast between the two phrasing options: while a phrasing such as La vella, llança l’amenança has just one interpretation, La vella llança, l’amenaça can be considered ambiguous. Why is this the case? In our opinion, such a contrast can be explained if we interpret phrasing decisions as partially constrained by syntax: while a phrase break between vella and llança immediately rules out an interpretation where vella llança is the syntactic subject, the non-presence of such a break leads to ambiguity. Generally, a phrase break cannot separate two words that belong to the same inner constituent; thus, if a phrase break is present between two words this is generally an indication that they belong to different constituents. Instead, if no phrase breaks are present, a structurally ambiguous sentence will probably be left with various grammatical interpretations.

4 Conclusion

We have shown that variation in intonational phrasing is the common strategy Catalan speakers use to disambiguate otherwise ambiguous sentences
involving right-branching vs. left-branching structures. Specifically, an utterance like *la vella llança l’amenaca* is generally disambiguated by means of a distinct intonational phrasing: an intermediate phrase break is placed at the end of the first constituent (i.e., the syntactic subject). Phonetically, a constituency break is marked with an obligatory high boundary tone (H’, in Pierrehumbert’s terms) and an optional short pause. Also, pitch downtrend differences between the two types of branching structures seem to be an optional cue: only two of the speakers exhibited a pitch reset in constituent-initial accents.

Thus, the analysis of intonational contours in our production data partially confirm Bonet’s observations. Like Bonet, we find that the main difference between the two types of branching structures is phrasing. Yet, our results on the manifestation of such phrasing differences do not support Bonet’s claim that pitch lowering of constituent-final accents is crucial. Interpreting the data within Pierrehumbert’s framework, pitch lowering is not attested in constituent-final accents, given the equal pitch level attained by the first pitch accent in the two structures. Instead, we attest an optional pitch reset at the beginning of intermediate phrases.

The intonational facts differentiating the two ambiguous sentences can be easily explained by a strictly layered prosodic structure and the standard assumption that pitch tends to be reset at the beginning of a new prosodic domain. The current controversy about the property of SLH, then, is not challenged by our data.

The perceptual results of the test revealed some interesting facts about phrasing differences and their relationship to syntactic structure. We find the following contrast in Catalan listeners’ intuitions: a sentence such as *la vella llança, l’amenaca* (pronounced with an intermediate phrase boundary after the *llança*) is considered ambiguous, in contrast with *la vella, llança l’amenaca* (pronounced with an intermediate phrase break after *vella*), which has only one meaning. Our interpretation of this contrast is that phrasing in natural speech is partially guided by syntax, in the sense that certain groupings cannot be separated by phrase breaks. This would explain why *la vella, llança l’amenaca* has only one possible interpretation: the interpretation where noun and adjective together would form the syntactic subject would violate the above-mentioned constraint on phrasing. Instead, *la vella llança, l’amenaca* does not violate any constraint and thus permits two possible interpretations. Thus, we assume that linguistic context is the key disambiguating factor in natural speech, and that
phrasing decisions, subject to certain constraints, are highly optional. Still, the status of syntactic constraints and other factors like constituency size and their influence on phrasing decisions deserves further investigation.\textsuperscript{14}

We hope to have started the necessary groundwork in the way to predict the $F_0$ patterns in different prosodic conditions. The description of the effects of syntactic/prosodic structure on downstep patterns and variations in pitch scaling needs to be completed.\textsuperscript{15} In order to provide a full description of the patterns of downstep in Catalan it would be important, among other things, to test possible interactions between syntax and downstep and to develop $F_0$ assignment algorithms as a function of syntactic/prosodic domains.

\textbf{Appendix I}

1: La Maria visita un museu d’armes antigues. És diumenge i el museu és pràcticament buit. En una sala hi ha una llança vella que penja del sostre, amenaçant, amb el filó apuntant cap a terra. La Maria entra en la sala i queda parada. La vella llança l’amenaça. S’espanta i marxa.

2: És un poblet del sud d’Itàlia on ja quasi no hi viu ningú. Uns lladregots estan a punt d’entrar a robar en una de les cases abandonades del poble, i, de sobre, surt una vella del balcó. Va tota vestida de negre i sembla una bruixa. Ells s’espanten. La vella llança l’amenaça. Si tornen diu que els maten.

3. Cada dia la vella Maria escolta la novel·la radiofònica després de dinar: \textit{Camila}. Aquell dia comença a sentir-la i nota una cosa estranya en la veu de la protagonista. La vella escolta la veu. No és la de sempre.

4. Fa vint anys el seu grup d’escoltes va escalar l’Aneto i, amb els anys, van a muntar una pel·lícula. La vella escolta la veu. Ho recorda tot amb nostàlgia.

5. Era una dona no gaire agraciada que es va quedar per vestir sants. Veient-la tan infeliç, una de les seves amigues li regala una crema d’Avon i li diu que es casarà enguany. Després de posar-se la crema va a provar sort a un ball de gent gran organitzat pel casino del poble. Ara que està a punt de casar-se tothom diu: La poderosa crema la casa.

6. Aquesta notícia que he llegit al diari és increïble. Explicava fins on poden arribar els matrimonis amb problemes i peles. En aquest cas, la dona, pubilla que havia heretat fortuna de casa seva, sabia que si es divorciava del seu marit hauria de deixar-li la casa on vivien. Pel títol de l’article endevinareu què va fer abans de demanar el divorci: \textit{La poderosa crema la casa}.  

7. Uns amics meus volien llogar un pis al barri vell de Barcelona. El van anar a veure. Mentre esperaven que arribés l'amo, van estar parland amb un vellet que s'estava a la porta. Quan va arribar, van pujar a dalt. Els va agradar el pis, però la veritat, no el trobaven gaire segur. L'amo els va dir: "No us preocupeu. El vell guarda la porta".

8. Aquella nit no vam poder dormir. Estàvem preocupats per la bufanda vermella!! Segur que me l'havia deixada al museu!! I m'havia costat 100 pessetes al mercat!! De bon matí, vaig tornar al museu. L'encarregat va dir-me: Sé que ve a buscar la bufanda. El vell guarda la porta.

Appendix II

Quiestionari

Sentiràs 8 frases primer llegides per un parlant femení (part I) i després per un parlant masculí. Has d'endevenir el signifcat de les frases (que podrien ser ambigües si no fos per la pronúncia).

Nom: ____________________________

PART I. Parlant femení

1. *La vella llança l'amença.* Quin és el verb de la frase?
   a. llançar
   b. amenaçar

2. *La vella llança l'amença.* Quin és el verb de la frase?
   a. llançar
   b. amenaçar

3. *La vella escolta la veu.* Quin és el verb de la frase?
   a. escoltar
   b. veure

4. *La vella escolta la veu.* Quin és el verb de la frase?
   a. escoltar
   b. veure

5. *La poderosa crema la casa.* Quin és el verb de la frase?
   a. cremar
   b. casar-se

6. *La poderosa crema la casa.* Quin és el verb de la frase?
   a. cremar
   b. casar-se
7. *El vell guarda la porta.* Quin és el verb de la frase? a. guardar  
b. portar

8. *El vell guarda la porta.* Quin és el verb de la frase? a. guardar  
b. portar

**PART II. Parlant masculí**

1. *La vella lança l’amenaça.* Quin és el verb de la frase? a. llançar  
b. amenaçar

2. *La vella lança l’amenaça.* Quin és el verb de la frase? a. llançar  
b. amenaçar

3. *La vella escolta la veu.* Quin és el verb de la frase? a. escollir  
b. veure

4. *La vella escolta la veu.* Quin és el verb de la frase? a. escollir  
b. veure

5. *La poderosa crema la casa.* Quin és el verb de la frase? a. cremar  
b. casar-se

6. *La poderosa crema la casa.* Quin és el verb de la frase? a. cremar  
b. casar-se

7. *El vell guarda la porta.* Quin és el verb de la frase? a. guardar  
b. portar

8. *El vell guarda la porta.* Quin és el verb de la frase? a. guardar  
b. portar

**Pregunta:** Quina és la característica que creus que t’ha pogut ajudar a desambuguar les frases? __________________________________________________________________________

__________________________________________________________________________

__________________________________________________________________________
Notes

* I am grateful to Eulàlia Bonet, Corinne Cortés, Gorka Elordieta, Sonia Frota, Joan Mascaró, and Jaume Solà, for useful comments and criticism. I would also like to thank my Catalan informants, Joan Salavedra and Pere Gifra, who were always ready to step into the sound booth and start reading. Also, I am grateful to my university colleagues for answering the perceptual test and providing valuable feedback and impressions: Anna Bartra, Eulàlia Bonet, Albert Branchadell, Teresa Cabré, Antonio Chacón, Matilde Delgado, Mariona Ferrer, Mª Teresa Inglés, Teresa Iribarren, Lídia Lluís, Núria Martí, Joan Mascaró, Jaume Mateu, Mª Carme Picallo, Nil Santiáñez, Dolores Sitjà, Jaume Solà, Teresa Vallverdú. I also want to thank Fernando Martínez-Gil for his help in editorial matters. This work was funded by grants awarded by the Comissionat per a Universitats i Recerca de la Generalitat de Catalunya and by the DGICYT from the Ministerio de Educación y Ciencia (research projects 1995-GR00486 and PB-93-0893-C04-04).

1 In some languages, a time-dependent declination function can also lower the peaks in a given utterance.

2 According to Kubozono (1992: 383), MP represents a major phrase and is defined as the domain of downstep in Japanese (it corresponds to the intermediate phrase posited by Beckman & Pierrehumbert). mp corresponds to a minor phrase.

3 Kubozono (1992: 382), for example, defends that in Japanese "the magnitude of pitch rises in each syntactic boundary can largely be predicted by the depth of right-branching structure."

4 The informants were speakers of Central Catalan, all postdoctoral students in their late twenties: PG (from Barcelona), JS (from Olot) and PP (from Figueres).

5 In order to test the contribution of pitch scaling differences to the meaning identification, the PP and PG productions of the sentences were selected. As we will see, while speaker PG does not exhibit differences in pitch scaling between the two types of branching structures, speaker PP does (shows a consistent pitch reset at the beginning of a new prosodic constituent).

6 Utterances containing only one accent constitute an exception to these rules.

7 The meaningful tonal movements of a sentence have generally been located around accented syllables.

8 Our subjects consistently used the same disambiguating strategy.

9 For more details on the specifics of intonational transcription in Catalan, see Prieto (1996).
The effect of degree of embedding cannot be analyzed in the left-branching utterances either (see 2 below). The reason is that the last pitch accent of the sentence in (ia) has a L* in the word veu and thus cannot be compared with the H* pitch accent of veu in (ib):

(i) a. La vella escolta, la veu
    b. La vella escolta, la veu especial

Catalan speakers can also use phrasing variation to indicate differences in PP- or AP-attachment in the same way different types of branching structures are resolved. A sentence like Un home cansat de París has two meanings, depending on whether the PP is attached to the NP or the AP: ‘a tired man from Paris’, and ‘a man sick of Paris’. In this case, different phrasing differences unequivocally lead to two different interpretations: Un home, cansat de París leads to ‘a man sick of Paris’, and Un home cansat, de París leads to ‘a tired man from Paris’. Furthermore, variation in intonational phrasing is a common disambiguation strategy in languages like Spanish, Italian and English (Avesani et al. 1995).

Catalan, then, seems to have a smaller reset effect than Japanese. In this language, some phoneticians report a strong pitch reset at the beginning of a constituent: the average difference between peaks in constituent-initial and medial positions is around 40 or 45 Hz (Azuma & Tsukuma 1991, Kubozono 1992: 374).

It would be interesting to test the relative importance of acoustic cues like the presence of a H- boundary tone and pauses, differences in the pitch downtrend pattern, etc., in phrase boundary identifications.

If speakers did not make a conscious effort to disambiguate the two types of sentences under examination, phrase breaks tended to be the same for both structures.

The effect of degree of embedding could not be tested in our data because the informants did not place prosodic boundaries as initially expected; so, it remains to be seen whether lengthy constituents do trigger a stronger pitch reset (Cooper and Paccia-Cooper 1980).

References


Beckman, Mary, and Julia Hirschberg. 1993. The ToBI annotation conventions. Ohio State University and AT&T Bell Laboratories.


