Intonational phrasing in Romance: The role of syntactic and prosodic structure

Mariapaola D'Imperio, Gorka Elordieta, Sónia Frota, Pilar Prieto and Marina Vigário

1. Introduction

One of the main functions of prosody is to structure the message into "chunks" which have a definite size and internal structure. These prosodic constituents have been proposed to be hierarchically organized and non-isomorphic relative to syntactic constituents. Within the framework of Prosodic Phonology (Selkirk 1984; Nespor and Vogel 1986; *inter alia*), these constituents are the domain of realization/application of specific phonological rules. Moreover, what provides the information for the prosodic constituent to be built is surface syntactic structure (cf. Nespor and Vogel 1979, 1986; Selkirk 1980). However, early Prosodic Phonology work as well as more recent research have also accumulated a large body of evidence showing that the placement of prosodic boundaries for the purpose of structuring the utterance in various languages is often beyond the scope of syntax. In fact, factors such as information structure, constituent weight as well as speech rate appear to play a major role in phrasing decisions (Nespor and Vogel 1986; Steedman 1991; Truckenbrodt 1999; Selkirk 2000).

The concept of syntactic branchingness has also been employed in order to account for specific phrasing patterns that could not be predicted by the syntax-prosody algorithms postulated by Nespor and Vogel's version of Prosodic Phonology. Some of the restructuring rules postulated by that theory are in fact based on branchingness. One of those rules, for instance, predicts that in Italian a complement immediately following a head can stand on its own only if the complement is branching. For instance, an object noun is phrased together with the preceding verb only if the NP is non-branching. This peculiarity is illustrated in (1). Among the most uncontroversially assumed constituents above the foot, some versions of Prosodic Phonology posit the phonological phrase (ɸ) and the Intonational phrase (Ip).[1] While in (1a) the noun is non-branching, in (1b) it is branching, thus giving rise to a second, separate phonological phrase, whose evidence is the absence of *Raddoppiamento Sintattico* in [p] of *panini* 'sandwiches'. Examples are from Nespor (1993: 204).

(1) a. *(mangerò [p:]anini)*ɸ
         will eat   sandwiches
         'I will eat sandwiches'

     b. *(mangerò)*ɸ *([p]anini col salame)*ɸ
         will eat   salami sandwiches
         'I will eat salami sandwiches'

A similar notion of branchingness has been invoked in a number of other studies. On the other hand, work by Selkirk (1984) and, more recently, by Ghini (1993), substitute this notion of branchingness with constraints of different nature. For instance, Ghini argues that Nespor and Vogel's branching conditions can be successfully reanalyzed in terms of average weight and balance of phonological phrases. Among the constraints proposed by Ghini, one finds a binarity constraint that has also been employed in more recent work on English (Selkirk 2000) as well as on Brazilian Portuguese (Sandalo and Truckenbrodt 2002). In the last two works, a MAX-BIN constraint is employed in order to enforce binarity on the phonological phrase by stating that the average phonological phrase, at a normal speech rate, will consist of two prosodic words.

Another interesting fact concerns the possible interaction between the placement of phrase boundaries and

the length of the constituent. For instance, Jun (2003) finds that, in Korean, a syntactic phrase containing five or fewer syllables often forms one Accentual Phrase (AP)[2], while if containing six or more syllables it often forms two APs. However, the author also finds that this constraint interacts with a constraint of semantic origin, that is the recoverability of the meaning of some modifiers, as well as with speech rate. Prosodic phrases tend in fact to include more syllables at fast than at slower rates.

Speech rate has also been invoked in order to explain the difference in phrasing that we can notice in (2) (from Nespor 1993: 207). In Italian, the intonational phrase (Ip) is at the same time the domain of an intonational contour and, in some varieties, the domain of specific phonological sandhi rules, such as the spirantization of [k] in Tuscan. In (1) we can notice that the spirantization rule is applied within the domain of Ip, while it is blocked across an Ip boundary. Some of these boundaries can be erased when speaking fast. For instance, the presence of an Ip boundary between the third and the fourth Ip in (2a) blocks spirantization, hence [k] is produced for *con* 'with'. When the boundary is erased, as a result of fast speech rate, in (2b), then spirantization can apply.

(2) a. [[*La*[*h*]*amomilla*]$_\phi$]$_{Ip}$ [[[*k*]*ome è noto*]$_\phi$]$_{Ip}$ [[*k*]*alma*]$_\phi$ [*anche i bambini*]$_\phi$]$_{Ip}$ [[*k*]*on problemi*]$_\phi$ [*di insonnia*]$_\phi$]$_{Ip}$

   Chamomile   as known   calms   even children   with problems of insomnia

   'Chamomile, as it is known, calms even children having insomnia problems'

   b. [[*La*[*h*]*amomilla*]$_\phi$]$_{Ip}$ [[[*k*]*ome è noto*]$_\phi$]$_{Ip}$ [[*k*]*alma*]$_\phi$ [*anche i bambini*]$_\phi$ [[*h*]*on problemi*]$_\phi$ [*di insonnia*]$_\phi$]$_{Ip}$

More recently, a number of studies have pointed out language-particular preferences in prosodic groupings (Prieto 1997; Nibert 2000; Frota 2000; D'Imperio 2001, 2002). For instance, Frota (2000), in a detailed study of prosodic organization of Standard European Portuguese (SEP), proposes a prosodic weight condition that would account for Ip placement. This condition states that in SEP "long phrases tend to be divided; balanced phrases, or the longest phrase in the rightmost position are preferred" (Frota 2000: 327). Hence, this condition would favor the phrasing in (3c) over the phrasing in (3b), although (3a) is the most common phrasing.

(3) a . [O galã andava de porsche]$_I$
     b. *[O galã andava]$_I$ [de porsche]$_I$
     c. [O galã]$_I$ [andava de porsche]$_I$
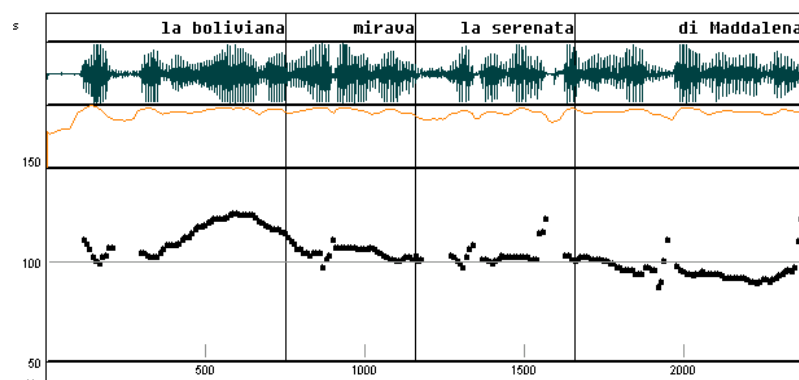   'The hero used to drive a Porsche'

Figure 1. F_0 curve and waveform for the Italian utterance *La boliviana mirava la serenata di Maddalena* 'The

Bolivian woman used to look at Maddalena's serenade', with broad focus and produced as a single (SVO) prosodic phrase.

A similar constraint has been employed for Italian by Ghini (1993), i.e. the principle of increasing units, stating that, in case of an asymmetric parse, phonological phrases in the recursive side (right side in Italian, as well as in Portuguese) are longer than in the non-recursive side (Ghini 1993: 59-60).

Among the tonal facts that are employed for the determination of phrasal domains, boundary tones often demarcate the end of an Ip. The vast majority of neutral declarative SVO utterances of Italian, for instance, are characterized by the presence of a L(ow) tone at the end of the utterance. In Neapolitan Italian (D'Imperio 2002), declarative utterances tend to be phrased as a single intonational phrase, characterized by internal downstep over successive H* preceding a final, nuclear H+L* (see Figure 1). These facts appear to be quite similar to Standard European Portuguese, where an initial H* followed by a final nuclear H+L* forms the neutral declarative tune (Frota 2000). In Catalan (Prieto 1997), instead, it appears that a prosodic break often occurs after S or even after V in a neutral declarative SVO utterance. Figure 2 intends to illustrate the second case, that is the placement of a break after the verb.
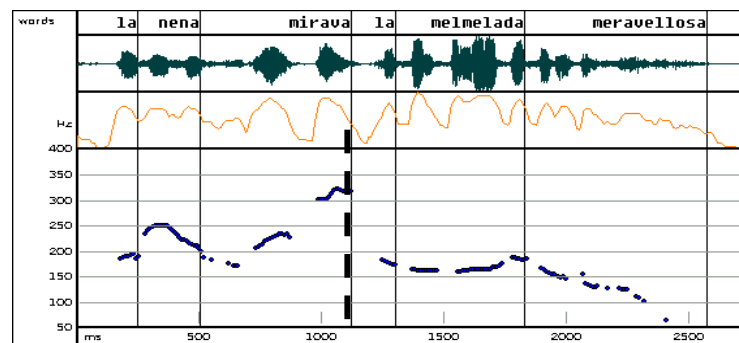

Figure 2. F_0 curve and waveform for the Catalan utterance *La nena mirava la melmelada meravellosa* 'The little girl was watching the wonderful marmalade', produced with two prosodic phrases, i.e. (SV)(O). The dashed line indicates the break location.

Although all the factors mentioned above (particularly length and syntactic branchingness) have been mentioned as having possible effects on prosodic boundary placement, these factors have not been examined in a systematic fashion crosslinguistically. Moreover, length effects have never been formalized in terms of a "number of syllable" constraint, but instead as a number of feet, prosodic word or ϕ condition (e.g. Ghini 1993, Guasti and Nespor 1999; Frota and Vigário 2002; Selkirk 2000). Hence, the first need that we identify is for both within-language and cross-linguistic systematic studies of the relevant factors.

Hence, this study sets out to investigate the influence of syntactic branching vs. constituent length (expressed in terms of number of syllables) on prosodic boundary placement in four Ibero-Romance languages: Catalan, European Portuguese (both Northern European Portuguese, NEP, and Standard European Portuguese, SEP) and Spanish. Another relevant issue concerns the disentangling of strictly prosodic from syntactic effects (prosodic vs. syntactic branching). Taking inspiration from weight claims such as Ghini's principle of uniformity and average weight ("average phonological phrases at a moderato tempo are formed by two phonological words (ω) [...]" (Ghini 1993: 52)), one can postulate that it is prosodic and not syntactic branchingness what influences phrase boundary placement. This is the second goal of our study.

2. Methods

A comparable database was created for the languages under study, consisting of sentences with a Subject-Verb-Object word order. The database was designed with exhaustive combinations of two constituent length conditions and seven syntactic branching conditions. The length variable applied to each of the words in the two constituents examined, the subject and the object. Short constituents contained 3 syllables and long constituents contained 5 syllables, including determiners and prepositions. The verb could also be short or long. Branching constituents were composed of a head noun and an adjective or a prepositional complement. Branchingness could be single (one level of branching) or double (two levels).

The seven syntactic branching combinations were: non-branching subject and object, non-branching subject and branching object, non-branching subject and double-branching object, branching subject and non-branching object, double-branching subject and non-branching object, branching subject and object, and double-branching subject and object. An example of some combinations of length conditions with syntactic branching conditions is shown in Table 1, only for one of the languages (Italian). The different combinations of length and branchingness produced a total of 76 utterances per language.

Table 1. Examples from the Italian corpus for the first experiment (syntactic branchingness) with branching Object conditions and both short and long subject (S).

	Non-branching O	**Short branching O**	**Long branching O**
Short S	*La bimba mirava la riga* 'The girl used to observe the line'	*La bimba mirava la riga di Vigo* 'The girl used to observe the line of Vigo'	*La bimba mirava la serenata meravigliosa* 'The girl used to observe the wonderful serenade'
Long S	*La boliviana mirava la riga* 'The Bolivian woman used to observe the line'	*La boliviana mirava la riga di Vigo* 'The Bolivian woman used to observe the line of Vigo'	*La boliviana mirava la serenata meravigliosa* 'The Bolivian woman used to observe the wonderful serenade'

Table 2. Examples from the Italian corpus for the second experiment (prosodic branchingness) with branching Object conditions and both short and long subject (S).

	Non-branching O	**Short branching O**	**Long branching O**
Short S	*Lorena Navona mirava la riga* 'Lorena Navona used to observe the line'	*La bimba meravigliava Morena Rinaldi* 'The girl used to astound Morena Rinaldi'	*La bimba meravigliava Maddalena Buonabitacolo* 'The girl used to astound Maddalena Buonabitacolo'
Long S	*Esmeraldina Secondigliano mirava la riga* 'Esmeraldina Secondigliano used to observe the line'	*La boliviana meravigliava Morena Rinaldi* 'The Bolivian woman used to admire Morena Rinaldi'	*La boliviana meravigliava Maddalena Buonabitacolo* 'The Bolivian woman used to astound Maddalena Buonabitacolo'

In a second experiment, in order to determine whether syntactic complexity or actual prosodic complexity determines prosodic phrasing, we substituted the syntactic branchingness condition of our materials with a prosodic branchingness condition (see Table 2). Prosodic words are usually defined as being above all the domain of primary stress (Vigário 2003b). Prosodic branching is defined at the prosodic word level and is given by structures with two prosodic words which are syntactically non-branching (proper names were used for all

languages; only for NEP and SEP compounds were also used - see Vigário 2003a for the prosodic analysis of the latter construction). Hence, in the second corpus, we substituted, for example, the branching object (or subject) *la serenata meravigliosa* 'the wonderful serenade' of Table 1 with a proper name consisting of two prosodic words, such as *Lorena Navona* (with primary stress on the penultimate syllable in both words[3]). We created thus a subset of 48 sentences.[4]

For each language, two speakers of the same region read the utterances of the two corpora three times in random order, with distractor sentences in between target sentences. A total of 744 (124 utterances per language: 76+48 from each experiment x 2 speakers x 3 repetitions) target utterances were obtained per language. The readers were instructed to read the sentences as if they were new information, i.e. as answers to a question such as 'Can you describe to me what happened?'. This allowed us to obtain broad focus utterances in reading style, avoiding possible effects of topic or focus intonation. The recordings took place in quiet rooms, using either DAT or analogical recorders. After digitization, the F_0 contours were examined using PitchWorks (Scicon), SpeechStation2 (Sensimetrics) and/or PRAAT. After digitization, the utterances were prosodically annotated using language-specific versions of the ToBI standard of transcription.

Note that, as it is well known, evidence for prosodic constituency can come from either tonal (Pierrehumbert 1980; Pierrehumbert and Beckman 1988; Jun 1993) or segmental phenomena (segmental sandhi, lenition/fortition, etc.; cf. Nespor and Vogel 1986), or alternatively from both (cf. Frota 2000, for an integrated approach). The initial approach taken here follows Ladd (1996: ch. 1), in that we do not start from a purely phonetic determination of constituent boundaries[5]. Instead, we adopt a more impressionistic approach in determining prosodic boundary location according to a perception-based, ToBI-like transcription (especially for its break indices proposals, cf. Price et al. 1991) of each of the languages/varieties included here. (Note that for some of the languages, such as Italian and EP, the intonational contours were also inspected. For EP, this was done in particular whenever the two transcribers did not agree.) Only in a following study shall we look at the actual phonetic correlates of the transcribed boundaries.[6]

3. Results

3.1 Catalan

Two speakers of Central Catalan (PG, a 36 year-old male from Barcelona and NM, a 34 year-old female from Arbúcies) randomly read the 76+48 Catalan utterances 3 times. In almost all cases, both speakers produced the sentences with two prosodic phrases and signalled the boundary between the two with a clear F0 continuation rise. Figure 3 illustrates the waveform and the F0 contour of the utterance (*La boliviana de Badalona*) (*rememorava la noia*) 'The Bolivian woman from Badalona remembered the girl', as produced by speaker NM. The F0 contour displays an H boundary tone after the verb, clearly indicating the (SV)(O) grouping. In some isolated instances, the transcriber also perceived a clear phrasing break with no clear tonal correlates.
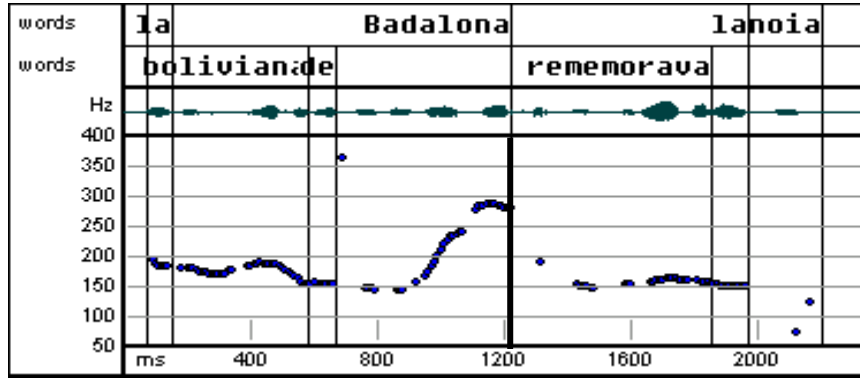
Figure 3. Waveform and F0 contour of the utterance (*La boliviana de Badalona*) (*rememorava la noia*) 'The Bolivian woman from Badalona remembered the girl', as produced by speaker NM. The dashed line indicates the break location.

Table 3 shows the percentages of (SVO), (S)(VO), and (SV)(O) phrasings in the non-branching subject/object conditions (e.g., *La nena mirava la noia* 'The little girl was watching the girl') in the conditions short/long subject and short/long object. In all conditions, the normal rendering is (S)(VO) for the two speakers, followed by (SV)(O) and, sporadically, by (SVO) groupings.[7] The results reveal that (SV)(O) phrasings increase both when subjects are short and when objects are long (between an 8%-24% increase in the short subject and long object conditions). [NB: numbers are percentage values].

Table 3. Percentages of phrasing patterns in the Non-branching S/O condition, with Short and Long S and O, for the two speakers NM and PG.

Length conditions	Non-branching S/O				
	PG			NM	
	(S)(VO)	(SV)(O)	(SVO)	(S)(VO)	(SV)(O)
Short S	76	16	8	84	16
Long S	84	8	8	92	8
Short O	84	16	0	100	0
Long O	76	24	0	76	24

When non-branching subjects are followed by branching objects, the number of (SV)(O) renderings tends to increase. Table 4 shows the phrasing patterns in two conditions: non-branching S/short branching O (e.g., *La nena mirava la noia del banc* 'The little girl was watching the bank's girl') and non-branching S/long branching O (e.g., *La nena mirava la melmelada de la Maria*) 'The little girl was watching Mary's marmalade'), with short and long subjects. The results indicate that the (S)(VO)-(SV)(O) distribution in the short branching O condition is quite similar to the long non-branching O condition shown in Table 1 (an average of 75% of (S)(VO) vs. 25% of (SV)(O) phrasings). Yet, this situation drastically changes when the object is long and branching: a prosodic phrase after the verb is placed 41-50% of the times for speaker PG and 75-83% of the times for speaker NM, a very substantial increase from the short branching O condition. We can also notice a tendency to produce more (SV)(O) phrasings when the subject is short rather than long, especially for speaker NM. Thus, phrase length computed in terms of number of syllables crucially affects phrasing in Catalan. It appears that in such cases a strategy is followed that balances somewhat the difference in length between the

constituents, i.e., a small prosodic break after V compensates for the difference in length between S and O.

Table 4. Percentages of phrasing patterns in the Non-branching S/Short or Long branching O conditions, with Short and Long S, for the two speakers NM and PG.

Length condit.	Non-branching S/Short branching O				Non-branching S/Long branching O			
	PG		NM		PG		NM	
	(S)(VO)	(SV)(O)	(S)(VO)	(SV)(O)	(S)(VO)	(SV)(O)	(S)(VO)	(SV)(O)
Short S	75	25	50	50	50	50	17	83
Long S	75	25	75	25	59	41	25	75

Let us now look at the phrasing patterns found in the proper name condition, i.e., cases which are prosodically but not syntactically branching. Table 5 shows the percentages of phrasing patterns in two conditions: non-branching S/short **prosodic** branching O (e.g., *La nena mirava l'Elena Miralles* 'The little girl was watching Elena Miralles') and non-branching S/long **prosodic** branching O (e.g., *La nena mirava la Margarida Vilajoiosa*) 'The little girl was watching Margarida Vilajoiosa'), with short and long subjects. The results show that the percentages of breaks with the proper name condition (Table 4) are as high as with the previous ambiguous branching condition (Table 3), indicating that prosodically branching proper name objects behave similarly to the previous ambiguous branching objects. Similarly, the length effect found with branching objects is replicated for speaker NM: when branching objects are longer in number of syllables, this speaker produces substantially more instances of (SV)(O).

Table 5. Percentages of phrasing patterns in the Non-branching S/Short or Long prosodic branching O conditions, with Short and Long S, for the two speakers MC and AG., for the two speakers NM and PG.

Length condit.	Non-branching S/Short branching O				Non-branching S/Long branching O			
	PG		NM		PG		NM	
	(S)(VO)	(SV)(O)	(S)(VO)	(SV)(O)	(S)(VO)	(SV)(O)	(S)(VO)	(SV)(O)
Short S	58	42	34	66	50	50	42	58
Long S	83	17	75	25	66	34	83	17

In the condition non-branching S/double-branched O, the phrasing patterns are similar to the ones displayed by long branching O, as prosodic boundaries are more usually placed after the verb than after the subject (Table 6).

Table 6. Percentages of phrasing patterns in the Non-branching S/Double branching O condition, with Short and Long S, for the two speakers NM and PG.

Length conditions	Non-branching S/Double-branching O			
	PG		NM	
	(S)(VO)	(SV)(O)	(S)(VO)	(SV)(O)
Short S	50	50	42	58
Long S	59	41	42	58

Systematically, sentences with branching (and double-branching) subjects were grouped as (S)(VO) by the 2 speakers, regardless of length conditions (e.g., *La boliviana de Badalona rememorava la noia* 'The Bolivian from Badalona remembered the girl', *La dona morena de Lugo rememorava la melodia* 'The dark-haired little

girl from Lugo remembered the melody'). The results in Table 6 reveal that (S)(VO) phrasings are completely systematic in utterances with branching and double-branching subjects: that is, the two speakers placed a prosodic break after the subject 100% of the time, regardless of constituent length.[8]

Table 7. Percentages of phrasing patterns in the Branching and Double-branching S conditions, with Short and Long S, for the two speakers NM and PG.

Length condit.	Branching S				Double-branching S			
	PG		NM		PG		NM	
	(S)(VO)	(SV)(O)	(S)(VO)	(SV)(O)	(S)(VO)	(SV)(O)	(S)(VO)	(SV)(O)
Short S	100	0	100	0	100	0	100	0
Long S	100	0	100	0	100	0	100	0

Table 8 shows that syntactically non-branching proper name subjects behave as the former syntactically branching subjects: that is, a prosodic break is found 100% of the time after a prosodically branched proper name subject.

Table 8. Percentages of phrasing patterns in the Syntactic Branching vs. Prosodic branching S conditions, with Short and Long S, for the two speakers NM and PG.

Length Condit.	Syntactic Branching S				Prosodic Branching S			
	PG		NM		PG		NM	
	(S)(VO)	(SV)(O)	(S)(VO)	(SV)(O)	(S)(VO)	(SV)(O)	(S)(VO)	(SV)(O)
Short S	100	0	100	0	100	0	100	0
Long S	100	0	100	0	100	0	100	0

The analysis of the Catalan data clearly demonstrates that although the most common phrasing pattern is (S)(VO), Catalan presents a tendency to divide utterances into phrases of similar syllabic lengths, often producing (SV)(O) phrasings. For example, when a short subject is followed by a long object speakers tend to group the S with the following V, revealing a strong tendency to balance the length of the prosodic constituents. Thus, a clear effect of length (in terms of number of syllables) is found in Catalan phrasing decisions. Finally, the branching effect in Catalan is clearly prosodic: by comparing the proper name condition (prosodically branched) with the ambiguous branching condition, the Catalan data shows that speakers produce similar patterns of phrasing in both conditions.[9] In fact, branching subjects, whether branchingness is prosodic or syntactic, trigger exclusively a (S)(VO) phrasing pattern.

3.2 European Portuguese

The two varieties of European Portuguese studied are the Standard variety spoken in Lisbon (SEP - for previous comparative work on intonational phrasing in SEP and other Romance languages see Elordieta *et al.* 2003, and Elordieta, Frota and Vigário 2003) and the Northern urban variety spoken in Braga (NEP - for previous comparative work on NEP and SEP see Vigário and Frota 2003). Examples of the speech materials are given in the Appendix.

Tables 9 to 12 show the percentages of (SVO) and (S)(VO) phrasings in the different conditions analyzed. (SVO) and (S)(VO) respectively indicate that no major intonational break is inserted in the first case whereas an intonational break is inserted between subject and verb in the second case. These were the only phrasing patterns attested.

Table 9. Percentages of phrasing patterns in the Non-branching S/O condition, with short and long Subject for SEP and NEP.

Condition	SEP		NEP	
Non-branching S and O	(S)(VO)	(SVO)	(S)(VO)	(SVO)
short S	0	100	50	50
long S	6	96	56	44

Table 10. Non-branching subject / syntactically (and prosodically) branching object condition with both short and long Subjects for SEP and NEP.

Condition	Non-branching S / Short branching O				Non-branching S / Long branching O			
	SEP		NEP		SEP		NEP	
	(S)(VO)	(SVO)	(S)(VO)	(SVO)	(S)(VO)	(SVO)	(S)(VO)	(SVO)
short S	0	100	56	44	4	96	69	31
long S	4	96	31	69	8	92	50	50

Table 11. Syntactically (and prosodically) branching and double branching subjects.

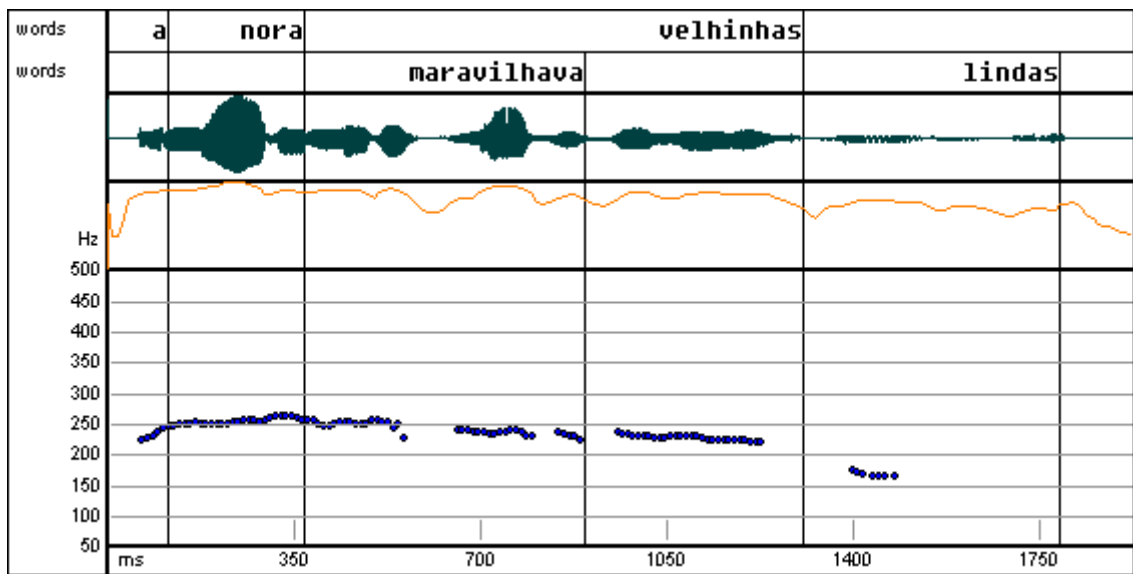| Condition	Branching S				Double branching S			
	SEP		NEP		SEP		NEP	
	(S)(VO)	(SVO)	(S)(VO)	(SVO)	(S)(VO)	(SVO)	(S)(VO)	(SVO)
short S	4	96	69	31	67	33	83	17
long S	40	60	63	37	94	6	71	29

Figure 4. Waveform and F0 contour of the utterance *(A nora maravilhava velhinhas lindas)* 'the daughter-in-law marveled beautiful old women' .

In SEP, when subjects are non-branching the predominant phrasing pattern is (SVO), i.e., everything is phrased together into one phrase, whether the subject is short or long and regardless of object complexity or length (Tables 9 and 10). This case is illustrated in Figure 4. Branching short subjects show a similar pattern (Table 12). Only for branching long subjects does the (S)(VO) phrasing pattern become relevant (40%). This phrasing pattern, with a major prosodic break after the subject, becomes the predominant one in double branching subject sentences, reaching 94% in the double branching long subject cases (Table 11). These results clearly show that subject length, and not syntactic or prosodic complexity, is playing a role in intonational phrasing in SEP. By contrast, length or syntactic/prosodic complexity of the object is not relevant.
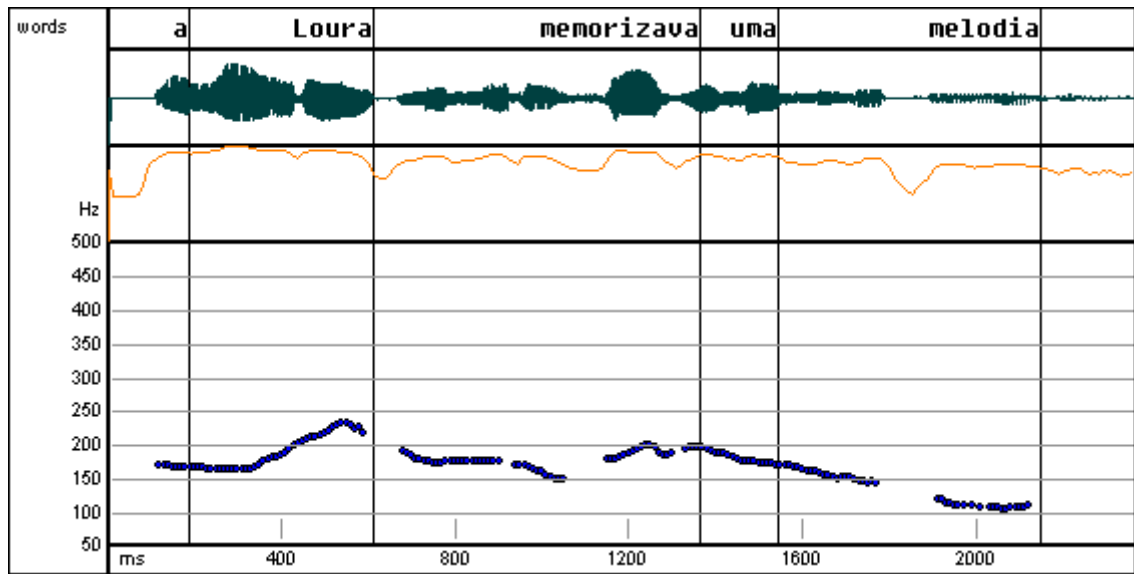

Figure 5. Waveform and F0 contour of the utterance *(A loura)(memorizava uma melodia)* 'the blond girl learned a song by heart'.

The results for NEP show a very different pattern. Here, (S)(VO) predominates in all conditions, even in the non-branching subject case (Tables 9-10). This phrasing pattern is illustrated in Figure 5. Constituent length plays a role (non-branching long subjects yield more (S)(VO) phrasings than non-branching short - 56% versus 50% - Table 8), as well as syntactic complexity ((S)(VO) increases with branching short subjects although they have the same length as non-branching long ones - 69% versus 56% - Tables 9 and 11), unlike in SEP. Also, in contrast with SEP, object length is relevant to intonational phrasing in NEP (branching long objects favor (S)(VO) more than branching short objects - 69% versus 56% - Table 10).

Overall, the differences found show that in SEP (S)(VO) phrasing is only triggered by length in number of syllables, whereas in NEP this pattern is present throughout and is mainly boosted by syntactic (and prosodic) branchingness, with length playing a relatively minor role.

An important question that has to be raised is whether the branching effect found in NEP is syntactic or prosodic. The data on the prosodic branching condition shown on Table 12 clarifies the issue. The percentages found for (S)(VO) in this condition are similar to those reported above for the syntactic complexity case (or higher). This clearly means that the effect is prosodic and not syntactic, that is it is the number of prosodic words that counts for favoring (S)(VO) in NEP.

Table 12. Prosodically branching subject / object for NEP.

Condition		NEP	
		(S)(VO)	(SVO)
Short non-branching S	Short branching O	58	42
	Long branching O	75	25
Long non-branching S	Short branching O	54	46
	Long branching O	54	46
Branching S	Short	80	20
	Long	98	2

The data on the prosodic branching condition generally replicate the results of the syntactic branching case. It also shows that length may play a role in intonational phrasing in NEP, both for the object (e.g. the contrast between a short and a long branching O when the subject is short non-branching) and the subject (e.g. the difference between short and long branching subjects). Moreover, it highlights that prosodic branchingness is more important than length, being the crucial factor at stake in this variety. Unlike the length factor which may (but does not have to) favor (S)(VO), the branching factor always boosts this phrasing pattern, as shown by the results of non-branching S contra branching S (see Table 11) or those of the non-branching long O (75% of (S)(VO)) and the short branching O conditions (86% of (S)(VO)) with a short branching S.

In short, the two varieties of EP intonationally phrase their utterances in different ways: in SEP, constituent length, defined in terms of number of syllables, is the decisive factor, whereas in NEP this role is taken by prosodic branching (i.e. number of prosodic words), with length possibly contributing to the effect.

3.3 Italian

Two speakers of the Neapolitan variety of Italian (LD and LC) produced the sentence from the two corpora (76+48 utterances) three times in random order. In Italian, the vast majority of the utterances was produced by both speakers as a single intonational phrase. This was especially true in non-branching sentences, as it is shown in Table 13. The (SVO) phrasing pattern is generally produced, as mentioned above, as a downstepping series of prenuclear H* accents, with a nuclear H+L* and a L-L% boundary at the end (see Figure 1). The only exceptions to this pattern in non-branching sentences were represented by a few (S)(VO) phrasings, produced by only one of the speakers (LD). On the other hand, (SV)(O) phrasings were never attested for this particular condition.

Table 13. Percentages of phrasing patterns in the Non-branching S/O condition, with Short and Long S and O, for the two speakers LC and LD.

Length	Non-branching S/O			
	LC		LD	
	(SVO)	(S)(VO)	(SVO)	(S)(VO)
short S	92	8	100	0
long S	100	0	100	0
Short O	92	8	100	0
Long O	92	8	100	0

Though (SVO) is still the preferred grouping, when a non-branching subject is followed by a long

branching object, the number of (SVO) renderings decreases sensibly only for speaker LC (while other phrasing strategies increase). This can be noticed in Table 14. Note that in this condition, (SV)(O) phrasings are found for speaker LC. Interestingly, though, in the double branching object condition a vast majority of (SVO) phrasings was found for both speakers (see Table 15), though LC retains a certain percentage of (SV)(O) phrasings. Note also that none of the subjects produced (S)(VO) phrasings in this condition.

Table 14. Percentages of phrasing patterns in the Non-branching S/Short or Long branching O conditions, with Short and Long S, for the two speakers LC and LD.

Length condit.	Non-branching S / Short branching O				Non-branching S / Long branching O				
	LC		LD		LC			LD	
	(SVO)	(S)(VO)	(SVO)	(S)(VO)	(SVO)	(S)(VO)	(SV)(O)	(SVO)	(S)(VO)
short S	92	8	92	8	75	8	17	100	0
long S	100	0	100	0	75	17	8	92	8

Table 15. Percentages of phrasing patterns in the Branching /Double branching O condition, with Short and Long O, for the two speakers LC and LD.

Leng. condit.	Branching O					Double branching O				
	LC			LD		LC			LD	
	(SVO)	(S)(VO)	(SV)(O)	(SVO)	(S)(VO)	(SVO)	(S)(VO)	(SV)(O)	(SVO)	(S)(VO)
Short O	96	4	0	96	4	100	0	0	100	0
Long O	75	12.5	12.5	96	4	67	0	33	100	0

Table 16 presents the results for sentences with short and long branching subjects, which, as one can notice, were mainly grouped as (S)(VO) by the 2 speakers. Here the situation drastically changes from a preference for (SVO) to a strong preference for (S)(VO) phrasings across conditions. Note that the percentages of (S)(VO) phrasings are above 60% in all cases. Interestingly, though, we can notice an effect of length of the object for both speakers, i.e. from 75% to 92% (S)(VO) in short subject sentences for LC and from 67% to 83% (S)(VO) in long subject sentences for LD. However, the effect of length of the subject is not homogeneous for the subjects. While for speaker LC (S)(VO) phrasings increase with subject length and short object (from 75% to 92%), LD presents the opposite trend for both short and long object sentences (from 83% to 67% of (S)(VO) in short object sentences, and from 92% to 83% for long object sentences). Despite the few (SV)(O) renditions of speaker LC, which can be accounted for in terms of some kind of 'balanced phrasing' constraint, a prosodic break, if present, is always placed after S in Italian. This suggests that the (SV)(O) phrasing patterns is not allowed or at least tends to be strongly avoided.

Table 16. Percentages of phrasing patterns in the Short and Long-branching S/non-branching O condition, with Short and Long O, for the two speakers LC and LD.

Length Condit.	Short branching S/non-branching O				Long branching S/non-branching O			
	LC		LD		LC		LD	
	(S)(VO)	(SVO)	(S)(VO)	(SVO)	(S)(VO)	(SVO)	(S)(VO)	(SVO)
Short O	75	25	83	17	92	8	67	33
Long O	92	8	92	8	92	8	83	17

Table 17. Percentages of phrasing patterns in the Branching and Double-branching S conditions, with Short and Long S, for the two speakers LC and LD.

Length condit	Branching S				Double Branching S			
	LC		LD		LC		LD	
	(SVO)	(S)(VO)	(SVO)	(S)(VO)	(SVO)	(S)(VO)	(SVO)	(S)(VO)
Short S	7	83	2.5	87.5	0	100	8	92
Long S	8	92	25	75	8	92	0	100

However, when looking at double branching subject condition, as shown in Table 17, (S)(VO) patterns are even more frequent. As shown in the table, we find 100% (S)(VO) cases for the short subject condition in LC as well as for the long subject condition in LD.
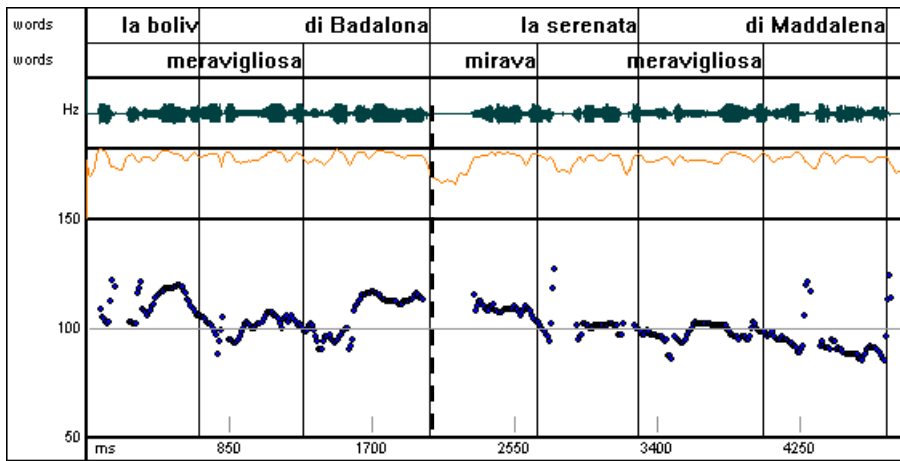

Figure 6. Waveform and F0 contour of the utterance (*La boliviana meravigliosa di Badalona*)(*mirava la serenata meravigliosa di Maddalena*) 'The wonderful Bolivian woman from Badalona used to observe Maddalena's wonderful serenade', as produced by speaker LD. The dashed line indicates the break location.

Table 18. Percentages of phrasing patterns in the Syntactic Branching vs. Prosodic branching S conditions, with Short and Long S, for the two speakers LD and LC.

Length condit.	Syntactic Branching S				Prosodic Branching S			
	LD		LC		LD		LC	
	(S)(VO)	(SVO)	(S)(VO)	(SVO)	(S)(VO)	(SVO)	(S)(VO)	(SVO)
Short S	88	12	89	21	100	0	96	4
Long S	75	25	100	0	92	8	100	0

Regarding the acoustic and perceptual correlates of the prosodic breaks marked in the tables, the end of a phrase was, in most cases, marked with a clear F_0 continuation rise or sustained pitch at the end of the Subject constituent, a prominent phrase-final H* accent and an optional pause. Figure 6 shows a case of (S)(VO) phrasing for a double branching Subject and Object condition, for the sentence *La boliviana meravigliosa di Badalona mirava la serenata meravigliosa di Maddalena* 'The wonderful Bolivian woman from Badalona used to observe Maddalena's wonderful serenade'. The prosodic break between S and the rest of the sentence is here

represented by the sustained pitch after the H* on *Badalona,*as well as by the marked pause and pitch reset after the same word. In general, the most prominent cue to a perceived break was the presence of sustained pitch (or a pitch rise), as well as the absence of downstep on the next accented word immediately following the perceived break.

When looking at phrasing patterns found prosodically branching subjects (the proper name condition), we notice that they reflect the situation found for syntactically branching subjects. Table 18 shows the percentages of phrasing patterns in two conditions: short vs. long **syntactic** branching S and short vs. long **prosodic** branching S (e.g., *Manolo Vallone mirava la riga* `Manolo Vallone used to observe the line' vs. *Esmeraldina Buonabitacolo mirava la riga* 'Esmeraldina Buonabitacolo used to observe the line'). Note that the behavior of syntactically and prosodically branching cases is very similar, in that a prosodic break is found almost around 100% of the cases after a prosodically branching proper name subject. Moreover, as for the syntactically branching cases, there appears to be an effect of length of the subject which goes in a direction opposite to the expected one (88% to 75% (S)(VO) phrasings for syntactic branching cases, and 100% to 92% for prosodically branching cases). We can only speculate at this point that long subjects might trigger a higher speech rate for this speaker, with a consequent tendency to group the sentence material into an (SVO) phrasing pattern. This issue needs of course more data as well as careful acoustic analysis in order to be answered.

Clearly, constituent branching is an important variable in Italian, but only as far as the Subject position is concerned. Also, it appears that the branching effect for the Subject is clearly prosodic in nature, since a similar predominance of the (S)(VO) phrasing patterns was found in both syntactically and prosodically branching cases. Here we found in fact a strong effect of branchingness, irrelevant of object length. As to subject length, there seems to be an effect only for speaker LD, which goes counter to the expectations, in that longer subjects trigger less (S)(VO) phrasings. However, the effect seems quite small and might be due to rate of speech changes.

3.4 Spanish

In Spanish the most common phrasing is (S)(VO). As we will see below, only when objects contain proper names (i.e., when objects are prosodically branching, not syntactically branching) do other phrasings occur with relatively high frequency. The breaks or boundaries in phrasing are realized by continuation rises, usually accompanied by final lengthening and in a few cases by pauses. The material following the boundary is pronounced with a lower pitch level.

Table 19. Percentages of phrasing patterns in the Non-branching S/O condition, with Short and Long S and O, for the two speakers MR and LM.

Length condition	Non-branching S/O			
	MR		LM	
	(S)(VO)	(S)(VO)/(SVO)	(S)(VO)	(S)(VO)/(SVO)
Short S	92	8	67	33
Long S	75	25	83	17
Short O	75	25	67	33
Long O	92	8	83	17

Table 19 presents the percentages for non-branching contexts. As we can see, the most common phrasing is (S)(VO), but there were some utterances which native listeners found very difficult to classify

unambiguosly as either (S)(VO) or (SVO), or for which non-coincident responses between (S)(VO) and (SVO) were obtained. Those are the cases that appear under (S)(VO)/(SVO) in Table 19.

Tables 20-21 present the results for non-branching subjects with syntactically branching objects. Table 20 corresponds to short branching objects and Table 21 corresponds to long branching objects. The comparison between the results from Table 19 with long non-branching objects and those from Table 20 with short branching objects is interesting, as in both cases the object contains the same number of syllables, i.e., 10. (S)(VO) phrasing percentages are higher in short branching object conditions than in long non-branching object conditions, especially for speaker LM, for whom there are more cases of utterances which could be classified as either (S)(VO) or (SVO) with long non-branching objects (cf. Table 19). The differences are not so evident when results from the two speakers are taken into account. However, once the results from the two speakers are averaged out, the differences in percentages of occurrence of (S)(VO) phrasings with long non-branching objects and short branching objects are clearer: whereas unambiguous (S)(VO) phrasings occur in 79% of cases with long non-branching objects, they increase to 96% with short branching objects. Thus, branchingness seems to play a role in intonational phrasing in Spanish, as utterances with branching objects are phrased more commonly as (S)(VO) than with non-branching objects. An illustrative example of an utterance phrased in this way is provided in Figure 7, corresponding to the sentence *La niña de Lugo miraba la mermelada* 'The girl from Lugo watched the marmalade'.

A few instances of other phrasings also arise, but they are not significant in number. Speaker LM produced three tokens of (SV)(O) phrasings with long branching objects, the only instances in the data for the first experiment. It is also important to note that constituent length does not affect intonational phrasing, as illustrated by the fact that the percentage of (S)(VO) phrasings does not increase with long branching objects, and by the fact that (S)(VO) is not more common with long subjects than with short subjects.

Table 20. Percentages of phrasing patterns in the Non-branching S/Short branching O condition, with Short and Long S, for the two speakers MR and LM.

| Length Condition | Non-branching S / Short branching O | | | |
| | MR | | LM | |
	(S)(VO)	(S)(V)(O)	(S)(VO)	(S)(VO) / (SVO)
Short S	100	0	92	8
Long S	92	8	100	0

Table 21. Percentages of phrasing patterns in the Non-branching S/Long branching O conditions, with Short and Long S, for the two speakers MR and LM.

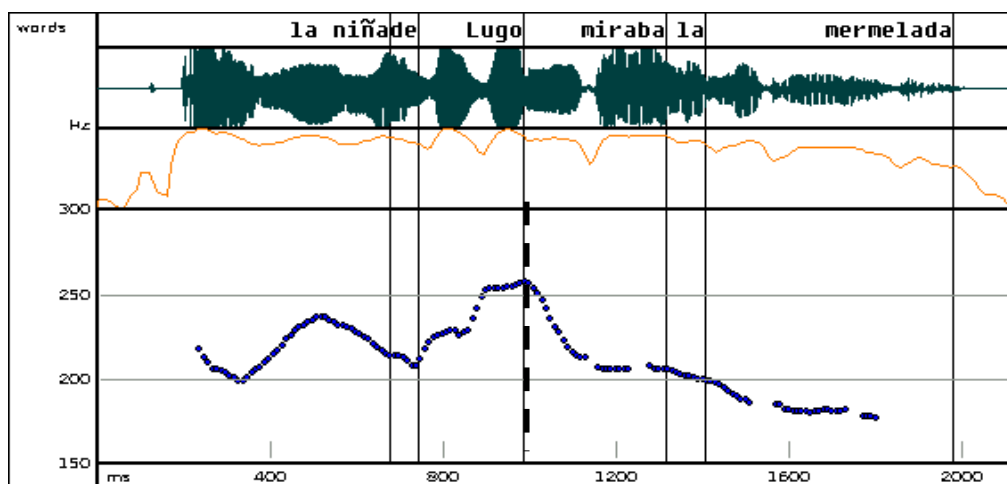| Length Condition | Non-branching S / Long branching O | | | | |
| | MR | | LM | | |
	(S)(VO)	(S)(V)(O)	(S)(VO)	(SV)(O)	(S)(V)(O)
Short S	92	8	92	8	0
Long S	92	8	84	8	8

Figure 7. Waveform and F0 contour of the utterance (*La niña de Lugo) (miraba la mermelada)* 'The girl from Lugo watched the marmalade'. The dashed line indicates the break location.

Like for the other languages studied, for Spanish a parallel database with complex objects which are not syntactically branching was created. These objects were composed of two prosodic words forming one noun (i.e., a proper name), so that the object was prosodically branching but syntactically non-branching[10]. This way it is possible to elucidate whether it is syntactic branchingness or prosodic branchingness that affects intonational phrasing in Spanish. Interestingly, the situation is different when prosodically branching objects are considered. In these cases, the percentage of occurrence of (S)(VO) phrasings decreases significantly, while other types of phrasing increase their frequency of occurrence. These percentages are presented in Tables 22 and 23, for short and long branching objects, respectively.

Table 22. Percentages of phrasing patterns in the Prosodic/Short branching O conditions, with Short and Long S, for the two speakers MR and LM.

Length Cond.	Non-branching S / Short branching O							
	MR				LM			
	(S)(VO)	(SVO)	(S)(V)(O)	(SV)(O)	(S)(VO)	(SVO)	(S)(V)(O)	(SV)(O)
Short S	8.3	83.4	0	8.3	59	6	6	29
Long S	42	25	33	0	50	50	0	0

Table 23. Percentages of phrasing patterns in the Prosodic/Long branching O conditions, with Short and Long S, for the two speakers MR and LM.

Length Cond.	Non-branching S / Long branching O							
	MR				LM			
	(S)(VO)	(SVO)	(S)(V)(O)	(SV)(O)	(S)(VO)	(SVO)	(S)(V)(O)	(SV)(O)
Short S	16.5	41.5	17	25	28.5	21.5	28.5	21.5
Long S	0	42	58	0	78	22	0	0

As can be seen in the tables, (SVO), (S)(V)(O) and (SV)(O) phrasings are quite common with proper names in object position. An explanation is in order as for why (S)(VO) phrasing is less frequent with proper names in object than with common nouns, which appeared in the corpus with syntactically branching objects.

That is, there appears to be a tendency to produce utterances in which the object is made intonationally salient, by way of phrasing or accentuation, possibly triggered by the processing (i.e., retrieval) difficulties of proper names compared to common nouns and by their cognitive and pragmatic saliency (see discussion in §4 below).

In utterances with syntactically double-branching objects the predominance of (S)(VO) phrasings is almost absolute, with a few cases by speaker LM with another break between the head noun and the adjective or between the adjective and the prepositional complement. No results for prosodically double-branching objects can be provided, as our hypothesis was that single-branching cases would elucidate the relative relevance of syntactic or prosodic branchingness. For Catalan, EP and Italian this was indeed the case, but for Spanish it now seems that analyzing double-branching objects could be interesting. We leave this for future research.

In branching and double-branching subject conditions, (S)(VO) occurred in almost 100% of the cases, both with syntactically and prosodically branching subjects. As for utterances in which both the subject and the object were branching, (S)(VO) occurred 96%-100% of the cases both with syntactically and prosodically branching constituents. Interestingly, in these cases there seems to be a combined effect of two forces or constraints favoring (S)(VO): the default strategy or general tendency in Spanish for this type of phrasing on the one hand, and the added saliency in the subject carried by the proper name, on the other.

In sum, in Spanish the most frequent phrasing type is (S)(VO), more clearly with branching subjects. In non-branching conditions there seem to arise possible (SVO) phrasings, probably triggered by the shortness of this type of utterances, and in branching objects composed of proper names (S)(V)(O), (SVO) and (SV)(O) phrasings arise with more frequency. As for why this pattern was not observed in Catalan, NEP or Italian, we can only hypothesize that its occurrence is language-dependent, as well as speaker-dependent. For future work, it would be interesting to design a unitary corpus in which common and proper names were combined, to falsify our hypothesis with the same corpus.

4. Discussion

The results presented show that the five Romance languages studied display different intonational phrasing patterns for broad focus SVO utterances. Controlling for phonological length and syntactic complexity of the constituents we observed that although in Catalan and Spanish the most common phrasing pattern is (S)(VO), Catalan presents a clear tendency to divide utterances into phrases of similar lengths, often producing (SV)(O) phrasings. In Spanish, (SV)(O) phrasings are much less common, while (S)(VO) is pervasive across conditions of length and syntactic complexity, although it is categorical in branching subject cases. Unlike Catalan and Spanish, in SEP (SVO) is the usual phrasing. Only a long branching subject often triggers the phrasing (S)(VO), not a short branching subject or a short or long branching object, showing that length rather than syntactic complexity plays a role in boundary placement in this language. (SV)(O) is unattested in EP. On the other hand, in NEP the pervasive phrasing pattern is (S)(VO), which is even more frequent with syntactically and prosodically branching constituents. Hence, if a length effect exists in NEP, it is not a very strong one. In Italian, like in SEP, (SVO) phrasings are the most common ones, but, more like Spanish, syntactic and prosodic branching Subjects seem to require an (S)(VO) phrasing pattern. A length effect seems to be at work when considering the syntactic branching cases, since, as noticed, double-branching subjects trigger even more (S)(VO) patterns. However, the length effect is inconsistent across subjects when considering the role of Object length in Subject branching cases. As for the other three languages, the branching effect appears to be clearly prosodic in nature.

Hence, Catalan and SEP show above all a clear length effect, beyond any effect of syntactic or prosodic complexity. Constituent length has recently become the subject of investigation on the part of psycholinguists. For instance, Watson and Gibson (2002), in a study intended to evaluate the performance of various algorithms

predicting intonational boundary placement, find that a quite simple incremental model based on constituent size scores better than complex algorithms in some cases. According to their hypothesis (Left/Right Constituent Boundary Hypothesis), the likelihood of boundary placement is a factor of the size of the recently completed syntactic constituent as well as the size of the upcoming syntactic constituent, at a specific word boundary. The hypothesis takes into account also syntactic argument relationship. Hence, one prediction of this hypothesis is that, as the size of an upcoming constituent increases, the likelihood of boundary placement before it increases as well. The claim behind such a generalization is that "speakers may need extra processing time to plan a longer syntactic constituent" (Watson and Gibson 2002: 16), with empirical support coming from work by Stenberg et al. (1978) and Ferreira (1991). In the first study, a positive correlation was found between the initiation time of lists of words and the number of stressed syllables included (number of "stress groups"), while Ferreira found a correlation between sentence initiation times and both length and syntactic complexity of the subject of the sentence. Moreover, Ferreira also found that the likelihood of placing an intonational boundary before the verb was a function of the syntactic complexity of the object.

Note that our results for Catalan represent a potential problem for Watson and Gibson's algorithm, since they require that a boundary be placed after a "completed" syntactic constituent, defined as a constituent that has no rightward dependents. In Catalan, as we saw above, an intonational boundary is often placed after a verb. This location is predicted by Watson and Gibson to have a very low likelihood for boundary placement for two reasons: the left hand constituent is not completed and the right hand constituent is an argument of the most recent head. Another problem for this kind of theories is that, though some processing requirement might be universal (such as memory requirements for planning a linguistic unit), different languages, even within a restricted family such as Romance, might behave quite differently. This finding suggests that language-specific factors play a role in intonational boundary placement, which can often override the requirements of the processing apparatus.

Moreover, even work that takes into account the effect of length, such as Watson and Gibson (2002), does not explicitly employ number of syllables as a variable. For instance, Watson and Gibson use phonological phrases as a measure of size, claiming that in most of their examples number of syllables correlates with number of feet as well as with number of phonological phrases. In other words, a "long" constituent will always contain a larger number of each of these constituents in their material. On the other hand, in our work we have tried to explicitly separate the contribution of number of syllables from number of constituents that are higher in the prosodic hierarchy. For instance, SEP utterances are pronounced almost categorically in a single (SVO) intonational phrase, unless the subject is at least 8 syllables long. In the latter case, an (S)(VO) phrasing pattern will appear. We can also say that this tendency is gradual, in that it gets stronger as subjects become longer, with the longest subjects of our corpus (15 syllables long) almost always showing an (S)(VO) pattern. Also, in Catalan, object length expressed in terms of number of syllables plays a strong role in determining unexpected (SV)(O) phrasing. Specifically, when a subject (either long or short) is followed by a long branching object (e.g., *La nena mirava la melmelada de la Maria* 'The little girl was watching Mary's marmalade') the percentages of (SV)(O) are substantially higher than when it is followed by a short branching object (e.g., *La nena mirava la noia del banc* 'The little girl was watching the girl from the bank'), namely, an average of 25% increase for the two speakers.

Hence, in SEP and in Catalan it is clearly constituent length (understood as number of syllables in a constituent) what is relevant for prosodic phrasing. In Spanish, on the other hand, our results show a strong tendency to separate subjects from the rest of the utterance material (i.e., verb and object) in independent prosodic phrases (i.e., in an (S)(VO) phrasing pattern). Moreover, this pattern is virtually categorical with branching subjects, which is true also for Italian. As to the branching issue, there are at least two possible

accounts for the patterns observed. Either it is number of prosodic words in a prosodic phrase that matters (e.g., a prosodic phrase can contain a maximum of two prosodic words) or it is syntactic complexity (i.e., branchingness) that plays a role.

Our prosodic branching corpus helped us to settle this issue in that in most of the languages under study the prosodic branching results paralleled the syntactic branching ones. The only exception was Spanish, in which a wealth of phrasing patterns was found for our prosodic branching corpus. The results were accounted for in terms of processing complexity of proper nouns vs. common nouns. Electrophysiological data (ERP) reveal that proper names and common nouns are processed differently (cf. Müller & Kutas 1996 and references therein), and results from computational models (Hyperspace Analogue to Language, cf. Burgess and Conley 1999) show that proper names are more susceptible to retrieval problems than common nouns. These properties of proper names are closely linked to the observation by Burgess and Conley (1999) that a failure to retrieve proper name information is more likely to result in a completely incorrect response, that is, a response that could hinder comprehension, whereas mistakes in common noun retrieval are not so serious for comprehension. Burgess and Conley (1999) mention the concrete example that if an individual is attempting to produce the word *chair* and instead produces *seat*, the result, although not precisely correct, would usually be sufficient for comprehension in normal conversation. However, uttering the name *Peter* to refer to *Bill* for instance leads to complete misunderstanding. That is, although common nouns possess features that belong to or are shared by two or more concepts, proper names have no conceptual meaning, and are only paired associates. They are cognitively and pragmatically more salient than common nouns. In a way, then, individuals may need more time, effort or care when processing a proper name, and in a reading task such as the one that occupies us in this paper it could very well be that such a processing property is reflected by producing breaks between the verb and the object more often than with common nouns. This could explain the higher frequency of (S)(V)(O) and (SV)(O) phrasings. As for the relative frequency of (SVO) phrasings, we can only speculate at this point that this type of phrasing may be motivated by a competition between the saliency of the proper name in object position and the general tendency observed in the language to phrase separately subjects and predicates (i.e., (S)(VO)), such that both the subject and the object are pronounced in the same phrase, with more or less equal prominence. In fact, almost all of speaker MR's and many of speaker LM's (SVO) phrasings were just variants of (S)(V)(O) uttered at a faster rate. Unfortunately, this property of proper names was not anticipated or foreseen, and hence our goal of elucidating which one of the two branching conditions favored (S)(VO) phrasings (i.e., syntactic or purely prosodic) could not be tested appropriately.

Regarding the acoustic correlates of the prosodic boundaries labelled in this study, while European Portuguese and Italian scale the first F_0 peak after a boundary approximately at the same height of the first peak in the utterance, Catalan and Spanish seem not to show pitch reset, while displaying very low initial peaks after the prosodic break. Future research will concentrate on these and other acoustic-phonetic correlates of prosodic boundaries in Romance languages.

5. Conclusion

The main goal of this work was to separately investigate phonological weight (in terms of number of syllables) as well as prosodic and syntactic parameters in order to discover the role of each in determining intonational phrasing. We found that constituent length is an active phrasing constraint in Catalan, SEP and NEP (though much stronger in the first two languages than in the latter). Only branchingness matters for Spanish, in that branching subjects and objects increase even more the frequency of (S)(VO) phrasings. This effect was also found in Italian, where a length effect was found but was inconsistent across subjects. In fact, the length effect

was present only in one speaker (and only for branching objects). Finally, only Catalan shows a strong tendency to produce (SV)(O), which might be due to a constraint on balancing the weight of each prosodic phrase. Above all, we showed that when a constituent branchingness effect was present, it was actually due to prosodic branchingness. The only exception was constituted by Spanish branching objects composed of proper names, where three phrasing patterns appeared with higher frequency, i.e. (S)(V)(O), (SVO) and (SV)(O). Future research will aim at investigating the acoustic correlates of the phrasing levels found in this study.

Acknowledgments

Appendix

DATABASE ON PHRASING (SYNTACTIC BRANCHINGNESS)

CONDITIONS
(Subject, Verb, Object)
S = short (3 syllables)
L = long (5 syllables)

1. NON-BRANCHING (S=3σ; L=5σ), 8 cases   -   CATALAN

SSS   La nena mirava la noia
           'The little girl was watching the girl'
LLL   La boliviana rememorava la melodia
           'The Bolivian woman remembered the melody'
SLL   La nena rememorava la melodia
           'The little girl remembered the melody'
SSL   La nena mirava la melmelada
           'The little girl was watching the marmalade'
LSS   La boliviana mirava la noia
           'The Bolivian woman was looking at the girl'
LLS   La boliviana rememorava la noia
           'The Bolivian woman remembered the girl'
SLS   La nena rememorava la noia
           'The little girl remembered the girl'
LSL   La boliviana mirava la melmelada
           'The Bolivian woman was watching the marmalade'

2. SHORT BRANCHING SUBJECT/OBJECT (N+PP=5σ; N+AP=5σ) (24 cases)

      BRANCHING SUBJECT   -   ITALIAN

N+AP SS   La bimba mora mirava la riga.
               'The dark-haired girl used to observe the line'
N+AP LL   La bimba mora memorizzava la melodia.
               'The dark-haired girl used to memorize the melody'
N+AP SL   La bimba mora mirava la serenata.
               'The dark-haired girl used to memorize the serenade'
N+AP LS   La bimba mora memorizzava la riga.
               'The dark-haired girl used to memorize the line'
N+PP SS   La bimba di Malmo mirava la riga.
               'The girl from Malmo used to observe the line'
N+PP LL   La bimba di Malmo memorizzava la melodia.

'The girl from Malmo used to memorize the melody'

N+PP SL   La bimba di Malmo mirava la serenata.

'The girl from Malmo used to observe the serenade'

N+PP LS   La bimba di Malmo memorizzava la riga.

'The girl from Malmo used to memorize the line'

BRANCHING OBJECT   -   NEP

SS N+AP   A nora mirava velhinhas lindas '

'The daughter-in-law looked at beautiful old women'

LL N+AP   A boliviana maravilhava velhinhas lindas

'The Bolivian girl marveled beautiful old women'

SL N+AP   A nora maravilhava velhinhas lindas

'The daughter-in-law marveled beautiful old women'

LS N+AP   A boliviana mirava velhinhas lindas

'The Bolivian girl looked at beautiful old women'

SS N+PP   A mulher levava liras na mala

'The woman had lirae in her bag'

LL N+PP   A boliviana manuseava liras na mala

'The Bolivian girl handled lirae inside her bag'

SL N+PP   A mulher manuseava liras na mala

'The woman handled lirae inside her bag'

LS N+PP   A boliviana levava liras na mala

'The Bolivian girl had lirae in her bag'

BRANCHING SUBJECT AND OBJECT   -   SEP

N+AP S N+AP   A nora loura mirava velhinhas lindas

'The blond daughter-in-law looked at beautiful old women'

N+AP S N+PP   A nora loura levava liras na mala

'The blond daughter-in-law had lirae in her bag'

N+PP S N+AP   A nora da mãe mirava velhinhas lindas

'The daughter-in-law of (my) mother looked at beautiful old ladies'

N+PP S N+PP   A nora da mãe levava liras na mala

'The daughter-in-law of (my) mother had lirae in her bag'

N+AP L N+AP   A mulher loura maravilhava velhinhas lindas

'The blond woman marveled beautiful old women'

N+AP L N+PP   A mulher loura manuseava liras na mala

'The blond woman handled lirae inside her bag'

N+PP L N+AP   A nora da mãe maravilhava velhinhas lindas

'The daughter-in-law of (my) mother marveled beautiful old women'

N+PP L N+PP   A nora da mãe manuseava liras na mala

'The daughter-in-law of (my) mother handled lirae inside her bag'

3. LONG BRANCHING SUBJECT/OBJECT (N+PP=10σ; N+AP=10σ) (24 cases)

    LONG BRANCHING SUBJECT   -   SPANISH

N+AP SS   La boliviana maravillosa miraba la noria
                 'The wonderful Bolivian woman watched the turning-wheel'
N+AP LL   La boliviana maravillosa rememoraba la melodia
                 'The wonderful Bolivian woman remembered the melody'
N+AP SL   La boliviana maravillosa miraba la mermelada
                 'The wonderful Bolivian woman watched the marmalade'
N+AP LS   La boliviana maravillosa rememoraba la noria
                 'The wonderful Bolivian woman remembered the turning-wheel'
N+PP SS   La boliviana de Badalona miraba la noria
                 'The Bolivian woman from Badalona watched the turning-wheel'
N+PP LL   La boliviana de Badalona rememoraba la melodia
                 'The Bolivian woman from Badalona remembered the melody'
N+PP SL   La boliviana de Badalona miraba la mermelada
                 'The Bolivian woman from Badalona watched the marmalade'
N+PP LS   La boliviana de Badalona rememoraba la noria
                 'The Bolivian woman from Badalona remembered the turning-wheel'

    LONG BRANCHING OBJECT   -   CATALAN

SS N+AP   La nena mirava la melmelada meravellosa
                 'The little girl was watching the marvelous marmalade'
LL N+AP   La boliviana rememorava la melmelada meravellosa
                 'The Bolivian woman remembered the marvelous marmalade'
SL N+AP   La nena rememorava la melmelada meravellosa
                 'The little girl remembered the marvelous marmalade'
LS N+AP   La boliviana mirava la melmelada meravellosa
                 'The bolivian woman was watching the marvelous marmalade'

SS N+PP   La nena mirava la melmelada de la Maria
                 'The little girl was watching Mary's marmalade'
LL N+PP   La boliviana rememorava la melmelada de la Maria
                 'The Bolivian woman remembered Mary's marmalade'
SL N+PP   La nena rememorava la melmelada de la Maria
                 'The little girl remembered Mary's marmalade'
LS N+PP   La boliviana mirava la melmelada de la Maria
                 'The Bolivian woman was watching Mary's marmalade'

4. BRANCHING SUBJECT AND OBJECT   -   ITALIAN

N+AP S N+AP   La boliviana meravigliosa mirava la serenata meravigliosa.
'The wonderful Bolivian woman used to observe the wonderful serenade'

N+AP S N+PP   La boliviana meravigliosa mirava la serenata di Maddalena.
'The wonderful Bolivian woman used to observe Maddalena's wonderful serenade'

N+PP S N+AP   La boliviana di Badalona mirava la serenata meravigliosa.
'The Bolivian woman from Badalona used to observe the wonderful serenade'

N+PP S N+PP   La boliviana di Badalona mirava la serenata di Maddalena.
'The Bolivian woman from Badalona used to observe Maddalena's serenade'

N+AP L N+AP   La boliviana meravigliosa memorizzava la riga meravigliosa.
'The wonderful Bolivian woman used to memorize the wonderful line'

N+AP L N+PP   La boliviana meravigliosa memorizzava la serenata di Maddalena.
'The wonderful Bolivian woman used to memorize Maddalena's serenade'

N+PP L N+AP   La boliviana di Badalona memorizzava la serenata meravigliosa.
'The wonderful Bolivian woman used to memorize the wonderful serenade'

N+PP L N+PP   La boliviana di Badalona memorizzava la serenata di Maddalena.
'The wonderful Bolivian woman used to memorize Maddalena's serenade'

SHORT DOUBLE BRANCH (3+3+3=9σ)

SUBJECT   -   NEP

N-AP-PP SS   A nora morena da velha levava marmelos
'The old lady dark-haired daughter-in-law carried quinces'

N-AP-PP LL   A nora morena da velha memorizava uma melodia
'The old lady dark-haired daughter-in-law learned a song by heart'

N-AP-PP SL   A nora morena da velha gravava uma melodia
'The old lady dark-haired daughter-in-law recorded a song'

N-AP-PP LS   A nora morena da velha maravilhava meninos
'The old lady dark-haired daughter-in-law marveled small children'

OBJECT   -   SEP

SS N-AP-PP   O homem mirava a nora morena da velha
'That man looked at the old lady dark-haired daughter-in-law'

LL N-AP-PP   O namorado maravilhava a nora morena da velha
The boyfriend marveled the old lady dark-haired daughter-in-law'

SL N-AP-PP   O homem maravilhava a nora morena da velha
'That man marveled the old lady dark-haired daughter-in-law'

LS N-AP-PP   O namorado mirava a nora morena da velha
'The boyfriend looked at the old lady dark-haired daughter-in-law'

5. DOUBLE-BRANCHING SUBJECT AND OBJECT (4 cases)

SHORT DOUBLE BRANCH   -   CATALAN

N-AP-PP S N-AP-PP   La dona morena de Lugo mirava la nena morena del banc
                     'The dark-haired woman from Lugo was watching the dark-haired little girl on the bench'
N-AP-PP L N-AP-PP   La dona morena de Lugo rememorava la nena morena del banc
                     'The dark-haired woman from Lugo remembered the dark-haired little girl on the bench'

LONG DOUBLE BRANCH   -   ITALIAN

N-AP-PP S N-AP-PP   La boliviana meravigliosa di Badalona mirava la serenata meravigliosa di Maddalena.
                     'The wonderful Bolivian woman from Badalona used to observe Maddalena's wonderful serenade'
N-AP-PP L N-AP-PP   La boliviana meravigliosa di Badalona memorizzava la serenata meravigliosa di Maddalena.
                     'The wonderful Bolivian woman from Badalona used to memorize Maddalena's wonderful serenade'

References

Beckman, M. and J. Pierrehumbert   1986   Intonational structure in English and Japanese. *Phonology* 3: 255–309.

Burgess, C. and Conley, P.   1999   Representing proper Names and objects in a common semantic space: a computational model.   Brain and Cognition 40: 67-70.

D'Imperio, M.   2001   Focus and tonal structure in Neapolitan Italian. *Speech Communication*, 33(4): 339-356.

2002   Italian intonation: An overview and some questions. *Probus*, 14(1): 37-69 (Special Edition on Romance Intonation).

Elordieta, G., Frota, S., Prieto, P. and Vigário, M.   2003   Effects of constituent weight and syntactic branching on intonational phrasing in Ibero-Romance. In *Proceedings of the 15th International Congress of Phonetic Sciences* (M.J. Solé, D. Recasens and J. Romero, editors), vol. 1, 487-490. Barcelona.

and Frota, S., and Vigário, M.   Forthcoming Subjects, objects and intonational phrasing in Spanish and Portuguese. Paper presented at the *GLOW Workshop Boundaries on Intonation*, Lund University, April 12. To appear in *Studia Linguistica*.

Ferreira, F.   1991   Effects of length and syntactic complexity on initiation times for prepared utterances. *Journal of Memory and Language*, 30: 210-233.

Frascarelli, M. 1997   The phonology of focus and topic in Italian. *The Linguistic Review* 14: 221–248.

Frota, Sónia   2000   *Prosody and Focus in European Portuguese*. New York: Garland.

and Marina Vigário   2002   Efeitos de peso no Português Europeu. In *Saberes no Tempo. Homenagem à Professora Henriqueta Costa Campos,* Maria Helena Mateus e Clara Nunes Correia (orgs.), Lisboa: Edições Colibri, 315-333.

Ghini, Mirco 1993 φ-formation in Italian: a new proposal. *Toronto Working Papers in Linguistics* 12: 41– 79.

Guasti, M. Teresa, and Marina Nespor   1999   Is syntax phonology-free? In *Phrasal Phonology*, René Kager and Wim Zonneveld (eds.), Nijmegen: Nijmegen University Press, 73-98.

Jun, Sun-Ah   1996   The phonetics and phonology of Korean prosody. New York: Garland.

2003   The effect of phrase length and speech rate on prosodic phrasing. In *Proceedings of the 15th International Congress of Phonetic Sciences*. (M.J. Solé, D. Recasens and J. Romero, editors), vol. 1, 483-486.

Ladd, Dwight Robert   1996. *Intonational Phonology.* Cambridge: Cambridge University Press.

Müller, H. and Kutas, M.   1996   What's in a name? Electrophysiological differences between spoken nouns, proper names, and one's own name. NeuroReport 8: 221-225.

Nespor, Marina, and Irene Vogel   1986   *Prosodic Phonology*. Dordrecht: Foris.

Nibert, Holly   2000   Phonetic and Phonological Evidence for Intermediate Phrasing in Spanish Intonation. PhD Thesis, University of Illinois.

Oliva, S. 1992   *La mètrica i el ritme de la prosa*. Barcelona: Quaderns Crema.

Pierrehumbert, Janet   1980   *Phonetics and phonology of English intonation*. PhD dissertation, MIT.

Pierrehumbert, Janet and Mary Beckman   1988   *Japanese tone structure*. Cambridge, MA: MIT Press.

Prieto, Pilar   1997   Prosodic manifestation of syntactic structure in Catalan. In *Issues in the Phonology and Morphology of the Major Iberian Languages*, F. Martínez-Gil and A. Morales-Front (eds.), 173-194. Washington, DC.: Georgetown Univ. Press.

Forthcoming   Syntactic and eurhythmic constraints on phrasing decisions in Catalan, Paper presented at the *GLOW Workshop Boundaries on Intonation*, Lund University, April 12. To appear in *Studia Linguistica*.

Price, P.J., Ostendorf, M., Shattuck-Hufnagel, S. And C. Fong 1991   The use of prosody in syntactic disambiguation. *Journal of the Acoustical Society of America*, 90: 2956-70.

Sandalo, F., and H. Truckenbrodt   2002   Some notes on phonological phrasing in Brazilian Portuguese. *MIT Working Papers in Linguistics* 42: 285-310.

Selkirk, Elizabeth 1978   On prosodic structure and its relation to syntactic structure. *Nordic Prosody II*, ed. T. Fretheim, 111–140. Trondheim: Tapir.

1984   *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.

1986   Derived domains in sentence phonology. *Phonology Yearbook* 3: 371–405.

2000   The interaction of Constraints on Prosodic Phrasing. *In Prosody: Theory and Experiment. Studies Presented to Gösta Bruce,* M. Horne (ed.), 231-261. Dordrecht: Kluwer Academic Publishers.

Steedman, Mark   1991   Structure and intonation. *Language*, 67: 260-296.

Sternberg, S., Monsell, S., Knoll, R. L., and C. E. Wright. 1978.   The latency and duration of rapid movement sequences: comparisons of speech and typewriting. In G. E. Stelmach (Ed.), *Information processing in motor control and learning.* New York: Academic Press, 117-152.

Truckenbrodt, H. 1999   On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry*, 30: 219-255.

Vigário, Marina   2003a Prosody and sentence disambiguation in European Portuguese. *Catalan Journal of Linguistics* 2 (Special Issue on Romance Intonation, edited by Pilar Prieto).

2003b   The Prosodic Word in European Portuguese. Interface Explorations, 6. Berlin/New York: Mouton de Gruyter.

and Sónia Frota   2003   The intonation of Standard and Northern European Portuguese. *Journal of Portuguese Linguistics* 2-2 (Special Issue on Portuguese Phonology edited by W. L. Wetzels), 115-137.

[1] A different intermediate domain, larger than ϕ, has been proposed in the framework of metrical and intonational phonology, according to which the level of phrasing that is above the prosodic word and below the intonational phrase is the intermediate phrase (Beckman and Pierrehumbert 1986; Pierrehumbert and Beckman 1988). This constituent is entirely defined in terms of specific intonational properties that are internal to the domain.

[2] An AP in Korean is a prosodic constituent marked by an initial and a final rising tone and can contain a word or more.

[3] For precision's sake, we must admit that, at some level of the structure, our proper names N (or N') branch internally into two Ns. Thus, by syntactic branchingness we mean a specific ramification at a given level (e.g., NP). Hence, we contrast a construction such as [DP el [NP *fill* [PP *de la Roser*]]] 'Roser's son' with constructions such as [DP la [NP [N [N *Marina*] [N *Lamuela*]]] 'Marina Lamuela'.

[4] Fewer sentences for the prosodic branching corpus relative to the syntactic branching one resulted from excluding the double branching cases, since prosodically single-branching cases are sufficient to

elucidate the question at hand.

which we considered superfluous for the issue we meant to explore.

[5] Ladd (1996, p. 10/11): "Moreover, I assume that constituent boundaries in prosodic structure are in the first instance abstractions, not actual phonetic events [...]. In some cases this means that conspicuous phonetic breaks occur at major constituent boundaries, but this is neither the essence of the boundary nor the only factor governing the distribution of the intonational features".

[6] In this work we will not specifically tackle the issue of whether our higher level prosodic breaks are intonational phrase breaks proper (Ip) or intermediate phrase breaks (ip), though some discussion will be offered in the individual language sections.

[7] Speaker PG sporadically produced the whole utterance in a single prosodic phrase, probably because he was speaking somewhat faster than speaker NM.

[8] When branching (or double-branching) subjects were followed by double-branching objects (e.g., *La boliviana de Badalona mirava la melmelada meravellosa de la Maria*) we observed two possible strategies to resolve the unbalanced phrase lengths. First, the speaker produced the verb as a plateau, separating it from both the subject and the object. Second, the speaker further separated the object constituent into two prosodic phrases, separating the last prepositional phrase (e.g., *(La boliviana de Badalona)(mirava la melmelada meravellosa)(de la Maria)*). That is, when unbalanced (S)(VO) groupings appear, other prosodic compensation strategies emerge such as further regrouping of the [VO] constituent.

[9] A contrast is found between the behavior of short branching S (*La nena mora*) and long non-branching S (*La boliviana*), which contain the same number of syllables: the latter cases have a higher tendency to have (SV)(O) renderings (around a 10% increase). This effect shows that the "phonological weight" of the phrase is computed not only by looking at number of syllables but also at the number of stresses and/or prosodic words.

[10] However, see note 3 for a more precise syntactic analysis.