1. Introduction

Syntactic structure is known to influence prosodic structure above the level of the prosodic word (cf. Truckenbrodt, 2007, and references cited there). The relation between syntactic and prosodic constituency is therefore typically reflected by (Optimality Theory) constraints that define prosodic constituent domains based on syntactic structure (e.g. Selkirk’s 1986, 1996 Alignment theory or Truckenbrodt’s 1995, 1999 Wrap theory). However, research on the syntax-phonology interface shows that non-syntactic factors such as constituent weight and speech rate (e.g. Nespor and Vogel, 1986, 2007; Ghini, 1993; Selkirk, 2000; D’Imperio et al., 2005; Prieto, this volume) are also of great importance for phrasing decisions. Catalan is a language that nicely shows the interaction between syntactic and eurhythmic constraints. In this Romance language, SVO (Subject-Verb-Object) structures typically show a prosodic boundary immediately after the subject, resulting in an (S)(VO) grouping, (1a). (Parentheses ‘( )’ represent prosodic grouping.) This grouping can be accounted for by using alignment constraints that make reference to the right edge of the subject DP (Prieto, 2005; Feldhausen, 2010). However, if the object is prosodically heavy (i.e. if it consists of more than one prosodic word (ω)), as in (1b), a different phrasing pattern occurs: (SV)(O) (D’Imperio et al., 2005; Prieto, 2005). In this case, eurhythmic constraints that make reference to prosodic weight have to account for the grouping (Prieto, 2005). Neither alignment nor wrap constraints are able to do so, since there is no right XP edge between V and O, only the left edge of the DP. As for what determines prosodic weight, D’Imperio et al. (2005:71) show that the number of prosodic words (ω) is the
decisive factor in Catalan. They obtained their results by varying the length of the subject and object constituents. I follow their assumptions on \( \omega \) throughout the present study.\(^1\)

1. As one reviewer points out, D’Imperio et al. (2005) argue that prosodic weight can be determined by two different things: levels of syntactic branching/number of prosodic words or length, measured in terms of number of syllables. Elordieta et al. (2003) and D’Imperio et al. (2005) show that length is not a factor creating prosodic weight in Spanish, while in the Catalan data, length shows an effect for one speaker (D’Imperio et al., 2005:68f.). With regard to syntactic vs. prosodic complexity, in Elordieta et al. (2003) it is argued that syntactic complexity (i.e. branching vs. non-branching constituents) is the relevant factor, while D’Imperio et al. (2005) argue that it is the number of prosodic words. Since the study by D’Imperio et al. (2005) must be seen as a sequel to that of Elordieta et al. (2003), I follow the newer study here.

2. Whenever I use the term DP object, I mean simple DP object (the kind of DP used in D’Imperio et al., 2005; Prieto, 2005). Simple DPs consist of a (null) determiner and a NP, while complex DPs contain CPs, e.g. relative clauses. A reviewer wonders whether hypothesis 1 (cf. section 2) would also be valid for complex DPs. I will leave this question for further research.

3. In their cross-linguistic study of prosodic phrasing in Romance languages – Catalan, (Peninsular) Spanish, Italian, and European Portuguese – D’Imperio et al. (2005) show that a tendency toward the \( \text{SV(O)} \) pattern occurs only in Catalan. In Spanish, for example, the \( \text{SV(O)} \) phrasing is almost non-existent (Elordieta et al., 2003, 2005:130; Feldhausen et al., 2010). Even though in D’Imperio et al. (2005:81) and Prieto (2006:55) some cases of Spanish \( \text{SV(O)} \) were found in the condition with a long branching object, it is much less common than in Catalan and not significant. As one reviewer points out, the occurrences of \( \text{SV(O)} \) in Prieto (2006) might be due to the fact that the study was based on Barcelona Spanish: there is the possibility that the speakers were influenced by Catalan, given the fluency in both Catalan and Spanish of citizens of Catalonia.

4. The term embedded clause refers to any CP contained in another CP. The term root clause will refer to any CP that is not contained in a higher CP (cf. also Downing, 1970:20ff.). A simple sentence such as [Peter sleeps] is a root clause; it is not contained in a higher CP. In a complex root clause such as [Mary supposes [that Peter sleeps]] the sentence [that Peter sleeps] is an embedded clause because it is contained in a higher CP. Furthermore, the term matrix clause refers to the root clause minus its embedded clause, i.e. [Mary supposes [...]]. cf. Truckenbrodt (2005:274) for a similar explanation of terms. Throughout my work, the term simple SVO structure refers to a root clause with a transitive verb that does not dominate another CP (i.e. which does not have embedded clauses). The term complex SVO structure refers to root clauses that contain an embedded clause.

The interesting and atypical finding of D’Imperio et al. (2005) and of Prieto (2005) that the subject can phrase together with the verb in Catalan if the DP object is heavy raises a natural question.\(^2,3\) It is common knowledge that objects can also be sentential. How are phrasing decisions made by Catalans when the object is a complete clause? Do speakers still produce a \( \text{SV(O)} \) pattern? For this reason, in the present study the DP object of structures like those in (1) is replaced by a CP object.

To anticipate the results of the study: the typical phrasing pattern of SVO structures in which the object is sentential is \( \text{SV(O)} \). This suggests that the presence of a CP object strongly influences the prosodic structure – irrespective of the actual prosodic weight of the object clause. In the Optimality Theory formalization of these results, the alignment constraint ALIGN-CP1 is introduced, which aligns the left edge of a CP with the left edge of a prosodic phrase and thus accounts for the

\[(1) \quad \text{SVO phrasing in simple clauses (i.e. with DP object):}\]

\[
\begin{align*}
(S) & \quad (V) & \quad (O) \\
\omega & \quad \omega & \quad \omega
\end{align*}
\]

\[\text{Prosodic grouping} \quad \text{Prosodic words}\]

\[
\begin{align*}
&[L\text{’aguila robà el ratolí.} & \text{The eagle stole the mouse.}\]
&[L\text{’aguila robà el ratolí del meu germà.} & \text{The eagle stole my brother’s mouse.}\]
\end{align*}
\]

\[\text{\textit{The eagle stole my brother’s mouse.}}\]

\[\text{\textit{The eagle stole the mouse.}}\]

\[\text{The interesting and atypical finding of D’Imperio et al. (2005) and of Prieto (2005) that the subject can phrase together with the verb in Catalan if the DP object is heavy raises a natural question.}^{2,3}\text{ It is common knowledge that objects can also be sentential. How are phrasing decisions made by Catalans when the object is a complete clause? Do speakers still produce a (SV)(O) pattern? For this reason, in the present study the DP object of structures like those in (1) is replaced by a CP object. Just like the DP objects in (1), the embedded object clauses differ in their prosodic weight, consisting of one \( \omega \) or three \( \omega \), as shown in (2) and (3) respectively (here and elsewhere, \( [\ldots]_{CP1} \) marks the complex root clause and \( [\ldots]_{CP2} \) the embedded clause).}^{4}\text{ This allows for comparable structures, sharing the property of having an XP edge between V and O: the left edge of DP or CP.}\]

\[\text{To anticipate the results of the study: the typical phrasing pattern of SVO structures in which the object is sentential is (SV)(O). This suggests that the presence of a CP object strongly influences the prosodic structure – irrespective of the actual prosodic weight of the object clause. In the Optimality Theory formalization of these results, the alignment constraint ALIGN-CP1 is introduced, which aligns the left edge of a CP with the left edge of a prosodic phrase and thus accounts for the}\]

\[\text{The interesting and atypical finding of D’Imperio et al. (2005) and of Prieto (2005) that the subject can phrase together with the verb in Catalan if the DP object is heavy raises a natural question.}^{2,3}\text{ It is common knowledge that objects can also be sentential. How are phrasing decisions made by Catalans when the object is a complete clause? Do speakers still produce a (SV)(O) pattern? For this reason, in the present study the DP object of structures like those in (1) is replaced by a CP object. Just like the DP objects in (1), the embedded object clauses differ in their prosodic weight, consisting of one \( \omega \) or three \( \omega \), as shown in (2) and (3) respectively (here and elsewhere, \( [\ldots]_{CP1} \) marks the complex root clause and \( [\ldots]_{CP2} \) the embedded clause).}^{4}\text{ This allows for comparable structures, sharing the property of having an XP edge between V and O: the left edge of DP or CP.}\]

\[\text{To anticipate the results of the study: the typical phrasing pattern of SVO structures in which the object is sentential is (SV)(O). This suggests that the presence of a CP object strongly influences the prosodic structure – irrespective of the actual prosodic weight of the object clause. In the Optimality Theory formalization of these results, the alignment constraint ALIGN-CP1 is introduced, which aligns the left edge of a CP with the left edge of a prosodic phrase and thus accounts for the}\]
boundary between V and sentential O. Eurhythmic constraints based on constituent weight cannot fulfill this task: the object clause in (2) merely consists of one \( v \), just like the DP object in (1a), yet they have different phrasings. The prosodic analysis of the data is carried out within the Autosegmental-Metrical (AM) model of intonation (see Ladd, 1996, 2008). To be more precise, the Cat_ToBI transcription system (Aguilar et al., 2009) is taken as the basic model for transcription (see section 3 for details).

The paper is organized as follows. In section 2 the research hypotheses are presented. In section 3 the Cat_ToBI system and the relevant boundary cues are introduced. Section 4 details the methodology of the two production experiments that were conducted. Section 5 is devoted to the presentation of the results. In section 6 the theoretical approach to the phrasing of embedded SVO is given. It is shown that a stochastic-Optimality Theory approach can easily account for the (variation in the) empirical results. Section 7 summarizes and concludes the paper.

2. The hypotheses

Two hypotheses are adopted in the present study. First, I assume that sentential objects have the same effect on phrasing as the heavy objects of D’Imperio et al. (2005): it is hypothesized that CP objects – compared to DP objects – increase the probability of matrix S and V phrasing together (see (4)). Second, I assume that even an increasing number of (SV) groupings does not lead to an obligatory prosodic separation of the embedded object from the preceding matrix clause (see (5)).

(4) Hypothesis 1 (H1)

\textit{Compared to heavy DP objects, sentential objects increase the single–group phrasing of (matrix) SV}

(5) Hypothesis 2 (H2)

\textit{Sentential objects are not obligatorily separated by prosodic means from the preceding matrix clause}

In order to be able to compare the influence of sentential objects with the influence of DP objects on the phrasing of S and V, as described in H1, we first need the percentage value of the (SV)/(O) groupings of SVO structures with prosodically heavy DP objects in D’Imperio et al. (2005). It is 33.1\%.\footnote{This average is calculated from Tables 3–5 in D’Imperio et al. (2005:68f.).} H1 is validated if sentential objects increase the value of (SV) phrasing, i.e. the value has to be significantly higher than 33.1\%. The present study confirms H1. Matrix (SV) phrasing increases to 48.1\% with CPs consisting of one \( v \) and to 56\% with CPs consisting of three \( v \).

H2 is based on the question of whether embedded object clauses are prosodically separated from the matrix clause. Previous research argues that object clauses are not obligatorily separated from the preceding matrix clause. Downing (1970), for example, argues that only the edges of root clauses are obligatorily aligned with larger prosodic boundaries – meaning that the left edge of an embedded clause does not have an obligatory boundary. This position is adopted in Nespor and Vogel (1986/2007). Even though long intonational phrases may be restructured such that a boundary is inserted between a main verb and its clausal object argument, this boundary is not obligatory (Nespor and Vogel, 1986/2007:189, 198f.). This is also confirmed by Truckenbrodt (2005), who does not find any evidence of an intonational-phase boundary preceding an embedded object clause. H2 is validated if there is no obligatory prosodic boundary between the matrix verb and the embedded clause. If there is a boundary in 100% of the instances, the boundary is presumably obligatory and H2 is falsified. My results show that Catalan has a significant tendency to place a boundary before the embedded object clause; nearly 80\% of all embedded clauses are prosodically separated from their matrix clause.

3. Foundations for the prosodic analysis

This section provides the reader with the five boundary cues used in the study: continuation rise, sustained pitch, preboundary lengthening, pauses, and complex boundary tones. Before presenting the boundary cues in greater detail, the theoretical framework will be introduced.

3.1. The Cat_ToBI transcription system of intonation and phrasing

In the present study, the Tone and Break Index system for Catalan (Cat_ToBI; Prieto et al., 2009; Prieto, in press; Aguilar et al., 2009) is generally taken as the model for transcription. This system is itself based on the Autosegmental-Metrical (AM) model (Pierrehumbert, 1980; Ladd, 2008). I follow Aguilar et al. (2009), Prieto et al. (2009), and Prieto (in press) in assuming two levels between the prosodic word (\( \omega \)) and the Utterance (U): Intermediate Phrase (ip) and Intonational Phrase (IntP). However, I introduce the term \textit{prosodic phrase} as a hypernym for ip and IntP, since there is often no need to distinguish between the two levels.\footnote{As usual in ToBI systems, I assume three types of tonal targets: pitch accents (which associate with...}
metrical strong syllables), ip-boundary tones (which associate with the edge of an ip; labeled as T-), and IntP-boundary tones (which associate with the edge of an IntP; labeled as T%). I use the term edge tone as a hypernym for ip-boundary tone and IntP-boundary tone. With Aguilar et al. (2009) I assume six pitch accents for Catalan: the two monotonal accents L* (‘low’) and H* (‘high’) and four bitonal accents: L+H*, L*>H*, L*+H, and H*>L*.7 The predominant choice for prenuclear accents in Catalan broad-focus statements is L*>H*; the typical nuclear accent in those statements is L*. As for edge tones, I assume the tones given in Cat_ToBI for ip and IntP, namely L-, M-, and H-, as well as L%, M%, and H% (the mid tone M, however, does not show up in my data). In line with standard assumptions but in contrast to the recent Cat_ToBI versions, I assume that the combination of these tones (such as H–H%, L–L%, L–H% etc.) automatically signals a boundary on the IntP level.8

3.2. Boundary cues in Catalan

Table 1 presents a general overview of the assumed boundary cues and associates each of them with one of the two prosodic levels. While a continuation pitch, a sustained pitch, and preboundary lengthening are taken as cues for the ip level, pauses and complex boundary tones are taken as cues for the IntP level.9 The boundary cues are successively described in what follows.

A continuation rise is a preboundary stretch realized as a rise from the last stressed syllable into the syllable located at the boundary. Fig. 1 is a schematic representation of such a rise occurring on a proparoxytonic word (meaning the stressed syllable σ₁ is followed by two posttonic syllables σ₁ and σ₂). The diagonal line represents the F0 contour of this word. The L+H* accent is realized on the stressed syllable, i.e. there is a rising pitch movement during that syllable (which is marked by a continuation rise on a proparoxytonic word."

---

\[ \sigma_1 \]  

\[ \sigma_2 \]  

**Fig. 1.** Schematic diagram of a continuation rise on a proparoxytonic word.

---

Table 1

<table>
<thead>
<tr>
<th>Prosodic constituent</th>
<th>Intermediate phrase (ip)</th>
<th>Intonational phrase (IntP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boundary cues</td>
<td>• Continuation rise</td>
<td>• Pauses</td>
</tr>
<tr>
<td></td>
<td>• Sustained pitch</td>
<td>• Complex boundary tones</td>
</tr>
<tr>
<td></td>
<td>• Preboundary lengthening</td>
<td></td>
</tr>
</tbody>
</table>

---

7 The diacritic ‘-’ has been used in ToBI systems for languages such as English (Beckman et al., 2005:23), Serbo-Croatian (Godjevac, 2005:159), Korean (Jun, 2005:225), and Spanish (Prieto and Roseano, 2010). In using L*>H* to mark a delayed F0 peak, I follow Aguilar et al. (2009) and Prieto (in press). I refer to Prieto et al. (2005) and Face and Prieto (2007) for arguments for a three-way distinction among rising accents (such as L*>H, L>H, and L>H*).

8 Cf. Feldhausen (2010:ch. 2.2) for a detailed overview of the deviations between the present tonal analysis and the one proposed in the recent Cat_ToBI versions such as Aguilar et al. (2000), Prieto et al. (2009), and Prieto (in press).

9 A reviewer points out that the difference between ips and IntPs might be a gradual one and not as clear-cut as suggested in Table 1. S/he gives two arguments. The first argument concerns parentheticals. According to her/him, parentheticals are not necessarily a pause or a complex boundary tone. I would like to object here that the pattern of parentheticals is not as easy as the reviewer describes. First, Dehé (2009) shows that there is a certain variability in the exact position of the boundaries. Shorter parentheticals, such as Comment Clauses (i.e. reporting verbs, vocatives and question tags) “do not show a strong tendency to be phrased separately” (Dehé, 2009:610). Second, Fodor (2002) reports an ambiguity in the “attachment of a relative clause (RC) to a complex NP with two competing noun hosts.” She argues that “the presence or absence of a prosodic break before the RC does affect RC-attachment tendencies” (cf. Fodor, 2002 and references cited there).

As a second argument the reviewer says that the difference between ip and IntP may be detected by a difference in preboundary lengthening (longer before an IntP boundary than before an ip boundary). However, as mentioned in footnote 6, recent studies find no significant difference between preboundary syllables at ip and IntP edges in Catalan and Spanish (cf. also Féry et al., 2011:35 for similar findings in French).

Regardless, I surely acknowledge that the proposed assignment of boundary cues might be too strict. Nevertheless, it is not atypical to assign a certain boundary tone to a specific level. Aguilar et al. (2009), for example, rely on the one hand on perceptual arguments for distinguishing between ip and IntP (“In Catalan prosodic transcriptions, transcribers clearly distinguish between two levels of degree of perceived disjuncture”), and they rely on the other hand on tonal marking cues. As for tonal marking cues, they state, “Like the intonational phrase, the intermediate phrase is tonally marked after its final pitch accent (yet not as strongly as the intonational phrase), but the inventory of boundary tones that appear in this position is of a different (but partially overlapping) class. Typically, H- boundary tones, also called ‘continuation rises’, mark the end of an intermediate phrase.” In the present study, I strongly rely on tonal markings and in order to make my judgments clear and comprehensible, I propose the assignment given in Table 1.
the thick part of the diagonal line).\textsuperscript{10} The rise does not end on $\sigma_s$, though, but continues until the word boundary at the end of $\sigma_2$ (the thick rightmost vertical line marked by $\omega$). The rise on the posttonic syllables is caused by the High boundary tone. The term \textit{continuation rise} is descriptive and its equivalent in the AM framework is the high boundary tone H-.\textsuperscript{11}

A \textit{sustained pitch} is similar to a continuation rise, but the preboundary stretch is realized differently. The rise on the stressed syllable is followed by a high plateau up to the boundary: see Fig. 2. I label the sustained pitch with !H-, a downstepped high edge tone, in order to signal that F0 remains high but does not continue rising (cf. Gabriel et al., 2011:163, and Pešková et al., to appear, for the same analysis; while Prieto, in press, marks the sustained pitch with M-).

\textit{Preboundary lengthening} refers to the longer duration of syllables that precede a boundary. Previous research shows that the duration of preboundary syllables is significantly longer than the duration of non-preboundary syllables and that this longer duration signals the boundary (see Estebas-Vilaplana, 2000:120 and Astruc, 2005:153 for Catalan; Medina Murillo, 2005 for Spanish; Prieto et al., 2010a,b for both languages). Figs. 3 and 4 illustrate the difference in length of (non-) preboundary syllables. In Fig. 3 the noun \textit{l'alfa`brega} ‘the basil’ is not in a position preceding a boundary, while in Fig. 4 it is right before an IntP break (marked with H% and break index 4). Whereas the last syllable \textit{ga} has a duration of only 92 milliseconds (ms) in Fig. 3, it has a duration of 168 ms in Fig. 4 (both positions marked by an arrow).

A further boundary cue is \textit{pauses}. They are defined as a stretch of silence (Frota et al., 2007) or a major F0 break (Estebas-Vilaplana, 2000:118). Here, I distinguish between two kinds of pauses. The first kind comprises \textit{stretches of silence} that are

\textsuperscript{10} A reviewer remarks that it is always difficult to tell whether there is a L+H* or a L+H* pitch accent in the configurations presented in Fig. 1, since there is no clear turning point. I completely agree. To avoid this problem, I have chosen, as far as continuation rises are concerned, to follow Aguilar et al. (2009) in assuming that L+H* cannot be found in nuclear position.

\textsuperscript{11} In Selkirk (1984:288) the continuation rise is taken as a phonetic cue for an IntP. I take this rise as simply signaling an ip break. Only if the continuation rise is paired with a cue for IntPs (a pause in the present study) is it located at an IntP boundary; in that case, it is marked H-H%.
visible in F0 due to an interruption of the pitch track. The stretch of silence has to be longer than 100 ms, a period of time which Astruc (2005:153) considers long enough to be perceived as a pause, i.e. as a major prosodic break in this type of ‘read, pre-planned, non-spontaneous data’. The second kind of pause is what I call an audible pause. In general, this type of pause comprises breaks that can be perceived audibly as a pause, but that are not visible in the pitch track nor in the speech wave (see Feldhausen, 2010:54 for an illustration). In addition, stretches of silence that are shorter than 100 ms are also classified as audible pauses.

**Complex boundary tones** are the last boundary cue in the present study. Here, I only call L-H% a complex boundary tone – even though any combination of two tones results in a complex tone T-T% (Beckman and Pierrehumbert, 1986:288; Ladd, 1996:98; Gussenhoven, 2004:132).12 A schematic representation of the complex boundary tone L-H% is given in Fig. 5: corresponding natural pitch contours can be seen in Figs. 9 and 10 in section 5.1.2. In Fig. 5 the peak is reached at the onset of the first posttonic syllable $\sigma_1$. Then the pitch falls and reaches its dip at the onset of the second posttonic syllable ($\sigma_2$). After that the contour rises again until the end of the word (marked by $\omega$). This fall-rise is expressed by L-H%.

My assumption that continuation rises, sustained pitches, and preboundary lengthening signal IP boundaries rather than IntP boundaries is based on the following reasoning: (a) according to the Strict Layer Hypothesis (Nespor and Vogel, 1986/2007; Selkirk, 1986) an IntP boundary entails a boundary of the lower-level prosodic domain ip, and (b) continuation rises, sustained pitches, and preboundary lengthening can occur at the level of ip without a higher-level prosodic phrase boundary (Prieto, 2005; Aguilar et al., 2009); thus, when these cues do occur at an IntP boundary, they actually appear at the right edge of an ip.13 The two remaining boundary cues, pauses and complex boundary tones, only occur at IntP boundaries. They are consequently taken as signaling IntP boundaries.

---

12 Based on Beckman and Pierrehumbert’s (1986:288) reanalysis of the phrase accent T- as an ip-boundary tone, the ToBI transcription system (Silverman et al., 1992) represents T- in IntP-final position together with the IntP-boundary tone T% as a complex symbol T-T% (cf. also Ladd, 1996:98; Gussenhoven, 2004:132). As such, it is the standard assumption for English that a complex boundary tone signals a boundary of an intonational phrase.

13 For details on tonal cues see last paragraph of fn. 9; for details on preboundary lengthening see fn. 6.
4. Methodology

This section outlines the production experiments. In order to test H1 and H2 two experiments were conducted. Experiment 1 only includes sentences of the type presented in (3), i.e. their sentential object consists of three $\omega$. Its methodology is presented in section 4.1. Experiment 2 only includes structures of the type presented in (2), with a sentential object consisting of one $\omega$, and it is described in section 4.2.

4.1. Experiment 1 (3 $\omega$)

4.1.1. Location and subjects

This experiment was conducted in Berlin and Hamburg. Using a Marantz hard-disk recorder (CDR310) and a Sennheiser microphone (ME64), ten subjects (seven female and three male speakers) ranging in age from 20 to 39 years were recorded once. The sentences were recorded directly as .wav files (sample rate 22.050 Hz) and F0 tracks were analyzed using PRAAT (Boersma and Weenink, 1992–2011). Subsequently, the target sentences were analyzed using pitch tracks and spectrograms to guide the segmentation and the text-to-tune alignment. All speakers were native speakers of Central Catalan. They were exchange students or other people who had not been living in Germany for a long time. All subjects were entirely unaware of the purpose of the experiment.

4.1.2. Material

This experiment is devoted to sentential objects that consist of at least three prosodic words. The syntactic structure of the embedded clause is characterized by a transitive verb whose external argument is overtly realized. The embedded subject and object can, however, vary in their prosodic weight. They can be short and thus consist of only one $\omega$ or they can be long and consist of at least two $\omega$. In consequence, there are four conditions: (a) short S / short O; (b) short S / long O; (c) long S / short O; and (d) long S / long O (see (6)). The target sentences thus reflect the finding by D’Imperio et al. (2005:69ff.) that prosodic branchingness of S and O influences phrasing decisions.\footnote{The four target sentences in (6) are not only prosodically but also syntactically branching; cf. D’Imperio et al. (2005) for experiments setting these two factors apart. For the current purpose, the ambiguity in the branching condition is irrelevant, since syntactically branching constituents are always prosodically branching.} The prosodic weight of the matrix subject does not vary and consists of one prosodic word throughout the conditions. Since a (prosodically) branching subject (S $\geq 1 \omega$) typically calls for a boundary at its right edge (i.e. (S)(VO); D’Imperio et al., 2005:83), it would be impossible to see an influence on the matrix SV grouping by the CP object if the matrix S were branching. There are three sets (A, B, and C) of sentences showing the four conditions. Altogether, there are 120 sentences (3 sets of sentences $\times$ 4 conditions $\times$ 10 speakers). All sentences are neutral declarative SVO utterances.\footnote{‘Neutral’ means that the SVO sentences are taken to be all-new sentences. A sentence counts as all-new/all-focus if it is an appropriate answer to the question “What happened?” (cf. Krifka, 2007:23). By speaking of ‘neutral declarative SVO utterances’ I implicitly assume that preverbal subjects are not automatically taken as topics; otherwise they could not be ‘neutral’ (see Sheehan, 2006:75ff. and López, 2009:132 for a discussion in favor of this assumption).} In (6), one complete set (set A) is presented (for sets B and C, see Appendix A).

(6) Set A

Question:

No et trobes bé? Em sembla que estàs de mal humor.

‘Do you feel bad? You seem to be in a bad mood. What happened?’

Target sentences:

a. Condition: short S / short O (in the embedded clause)

La Bàrbara suposa que l’àguila robà el ratolí.

‘Barbara assumes that the eagle stole the mouse.’

b. Condition: short S / long O (in the embedded clause)

La Bàrbara suposa que l’àguila robà el ratolí del meu germà.

‘Barbara assumes that the eagle stole my brother’s mouse.’

\footnote{\footnotetext}
c. Condition: long S / short O (in the embedded clause)

La Bàrbara suposa que la meva àguila robà el ratolí.
Barbara assumes that my eagle stole the mouse.

The material was further controlled with respect to word stress (mainly proparoxytonic words in the subject positions; cf. Bàrbara and àguila 'eagle' in (6)), with respect to segmental aspects (mainly sonorants; cf. àguila ['ayla]), and with respect to syllable structure (mainly CV; cf. àgi.lu). Having controlled for these factors, the materials were ideal to facilitate pitch tracking and to allow for enough space between pitch accent and edge tone to avoid tonal crowding. Along with the target sentences, further sentences with a similar set-up (i.e. a context question with specific information structural demands for the answer) were used as filler clauses.16

4.1.3. Procedure

The subjects were placed in a quiet room in front of a computer and the stimuli (target sentences and fillers) were presented in a PowerPoint file in a pseudo-randomized order. The context question and the stimuli were presented together on a single slide. The subjects were told to read out the stimuli at a normal speech rate only after they understood the question. The question was presented in two different ways. First, it was visually presented on the computer screen and the subjects were told to read the question to themselves in order to understand it. Second, the question was presented auditorily (triggered by pressing a button) and the subjects had to listen to the question. The recorded question was uttered at a normal speech rate by a native speaker of Central Catalan. This procedure has several advantages: the subjects can read and hear the question; by reading the question, the subjects are able to understand the context correctly; by hearing the question, the subjects are put in a more conversation-like frame of mind, despite the unnaturalness of the laboratory situation. In addition, speakers who do not read the question completely at least have to hear it completely; and finally, by recording the context question in advance, it is possible to control the way the question is uttered and to make sure that the question matches the intended context.

4.2. Experiment 2 (1 ω)

4.2.1. Location and subjects

This experiment was conducted in Hamburg. Three female subjects, ranging in age from 22 to 38 years, were recorded twice using a hard-disk recorder Marantz PMD671 and the Sennheiser microphone ME64 (the sample rate and the analysis with PRAAT were the same as in Experiment 1). All speakers were native speakers of Central Catalan and they were entirely naïve as to the purpose of the experiment.

4.2.2. Material

Experiment 2 is devoted to sentential objects which consist of only one prosodic word. The syntactic structure of the embedded clause is characterized by an intransitive verb whose external argument is tacit (i.e. realized by pro in generative terms). Due to the fact that Experiment 2 needs to emulate the prosodic pattern of (1a), the prosodic weight of the CP object does not vary. Consequently there is a single condition (O = 1 ω) so that the whole SVO structure consists of exactly three ω; see (7). The matrix sentences consist of two ω and are parallel in structure with and lexically identical or similar to the three sets of Experiment 1 (La Bàrbara suposa... ‘Barbara assumes’, La Sílvia no ha mencionat... ‘Silvia did not say’, and El pare ha dit... ‘The father has said...’). Since there was just one condition for the object clause, each matrix clause was combined with three different CPs; cf. (7) and Appendix B. The context question Qué passa? ‘What happens?’ triggers an all-focus reading of the target sentences.

16 The filler clauses consisted of sentences used for other studies, which are published in Feldhausen (2010). One set of sentences consisted of simple SVO structures set up along the lines proposed in D’Imperio et al. (2005), cf. Feldhausen (2010:67ff.). This set thus counts as a control condition for the present study. The results of the simple SVO experiment are similar to the findings of D’Imperio et al. (2005). Due to these results speaker variation is to be excluded as an explanation for the effect.
(7)  a. Question:
Qué passa?
write happen.3SG
‘What happens?’

b. Target Sentences:
La Bàrbara suposa que dormen.
the B. assume.3SG that sleep.3PL
‘Barbara assumes that they sleep.’

La Bàrbara suposa que van trucar.
the B. assume.3SG that PAST call.INF
‘Barbara assumes that they called.’

La Bàrbara suposa que menteixen.
the B. assume.3SG that lie.3PL
‘Barbara assumes that they lie.’

The subjects were thus confronted with 9 target sentences (3 ‘sets’ × 3 sentences with lexically different object clauses). I furthermore added 9 filler sentences, which also represented a possible answer to the question Qué passa? ‘What happens?’. Each subject read the sentences twice, yielding 54 target sentences (3 speakers × 9 target sentences × 2 repetitions).

4.2.3. Procedure
The subjects were placed in a quiet room in front of a computer and the stimuli were presented in a PowerPoint file in a pseudo-randomized order. The context question and the target sentence were presented together on a single slide. The subjects were told to first silently read the sentences and then to read the target sentences out loud at a normal speech rate. As in Experiment 1, the subjects were told to read the sentences aloud in a conversational style without being given any specific instructions regarding the phrasing. A short practice session at the beginning of the experiment was included.

5. Results
This section presents the results of the two production experiments. It is shown that the rate of matrix (SV) phrasing in both experiments is substantially higher than with long DP objects. This suggests that the presence of the CP in syntax affects the prosodic structure – irrespective of its actual prosodic weight. It is furthermore shown that the intonational grouping of complex sentences allows for variation and that this variation is frequency-dependent. An Optimality Theory analysis to account for the frequency distribution of the various phrasing patterns is provided in section 6.

Section 5 has the following structure. First, the results of Experiment 1 are given in detail (section 5.1). Second, the results of Experiment 2 are given (section 5.2). A final discussion of the results is presented in section 5.3.

5.1. Results – Experiment 1 (3 ω)

Experiment 1 yields several results: (a) the number of matrix (SV) is higher with sentential objects than with DP objects in previous research; (b) CP objects are typically separated from the matrix clause by a prosodic boundary; (c) ip and IntP boundaries both appear sentence-internally; (d) the complex boundary tone is the most frequent cue at the IntP level, while the continuation rise is the most frequent cue at the ip level; and (e) preboundary syllables are significantly longer than non-preboundary ones.

Section 5.1.1 is devoted to tonal and durational boundaries. The phrasing patterns are presented in section 5.1.2. The results of Experiment 1 are discussed in section 5.1.3.

5.1.1. Boundary cues
Whereas tonal cues are distributed complementarily (e.g. an ip boundary may be realized as a sustained pitch or as a continuation rise, but not both), lengthening of the preboundary syllable might co-occur with a tonal boundary marker. For this reason, the tonal and durational boundary cues are considered separately.

17 Due to the shortness of the context question I refrained from letting the subjects hear it.
5.1.1.1. Tonal boundary cues. Sentence-internal breaks are realized by both ip- and IntP-boundary tones. This result contrasts with the assumption that IntP boundaries cannot appear sentence-internally (cf. i.a. Aguilar et al., 2009). Furthermore, the prosodic breaks are primarily marked by high edge tones (thereby confirming results of Frota et al., 2007).

In the data, there are altogether 255 sentence-internal boundaries (Table 2). Of these, 55.3% are realized by ip-boundary tones and the remaining 44.7% are realized by IntP-boundary tones. The location of the boundaries, and the resulting prosodic groupings, are presented in section 5.1.2.

The total of IntP boundaries and the total of ip boundaries in Table 2 are split up in Table 3 to illustrate their realization in detail. Row 1 (labeled ‘type’) shows the different possible realizations of boundaries at the two prosodic levels. Row 2 gives the absolute occurrences for each type and the corresponding percentage value. The percentages are calculated for each prosodic level (ip and IntP), since this gives an idea of how frequently a cue is used to signal an ip or an IntP boundary.

At the IntP level, the most common phonetic realization of the boundary is a complex boundary tone: 70 (i.e. 13 + 57) boundaries are realized by a complex tone. This group constitutes 61.4% of the 114 boundaries. High boundary tones (H%), followed by either a visible or audible pause, constitute the second most frequent IntP realization (40 realizations [15 + 25]; or 35.1%). Low boundary tones (L%) are virtually never realized sentence-internally at the IntP level (only four times). As for the ip level, high as well as low boundary tones exist: the low tone L- constitutes 14.2% and the sustained pitch (!H-) constitutes 12.8% of boundaries at this level. The continuation rise H-, however, is the most common realization (73% of all 141 ip boundaries). If we include the high IntP-boundary tones (H%), the total number of continuation rises increases (103 = 15 + 25 = 143 out of 255 boundaries in total). The total further increases when one even considers the high tone of the complex boundary tone L-H% to represent a continuation rise (143 + 70 = 213 out of 255 boundaries). This corresponds to the findings of Frota et al. (2007:135), who state that ‘prosodic breaks in Romance are predominantly marked by a High boundary tone. The preboundary stretch tends to be realized as a continuation rise’.

5.1.1.2. Preboundary lengthening. As for preboundary lengthening, the results show that the length of a given syllable is significantly longer in the preboundary position than in the non-preboundary position. In what follows, I first briefly describe the scrutinized syllables. After that, I present the average values for syllable duration for all speakers (in milliseconds and percentages), followed by a statistical evaluation of the durational difference (paired t-test).

Lengthening is relational and for this reason, to demonstrate preboundary lengthening, the length of preboundary syllables has to be compared to the length of syllables in non-preboundary positions. Unlike Feldhausen (2010:78ff.), who examines syllable length across different types of CV syllables (and thereby confirms previous results that find that preboundary syllables are longer than non-preboundary ones), the present study follows the methodology of Gabriel et al. (2011:173ff.) and concentrates on only one single syllable, namely <la> [la]. The syllable occurs in set A (see section 4.1.2) and set B (see Appendix A), both in a typical preboundary position, at the end of the embedded subject (<aigüa ‘eagle’ and Ángela), and in a position where it hardly ever experiences lengthening, namely at the very beginning of the target sentences (in the form of the definite feminine article: La Bárbara, set A; La Silvia, . . . , set B). The results confirm the previously observed pattern: Preboundary la is significantly longer than its non-preboundary counterpart.

The bar diagram in Fig. 6 illustrates the clear difference between preboundary and non-preboundary syllables by giving their average length expressed in milliseconds. The average for all ten speakers is given in the last column, indicated by . The preboundary syllables have an average length of 178.4 ms, while the non-preboundary syllables, i.e. the syllables in

Table 2
Number of occurrences and percentages of sentence-internal IntP and ip boundaries in the complex SVO structures.

<table>
<thead>
<tr>
<th>Type</th>
<th>IntP</th>
<th>ip</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBT</td>
<td>114 (44.7%)</td>
<td>141 (55.3%)</td>
</tr>
</tbody>
</table>

Table 3
Number of occurrences (occ.) and percentage (%) of each type of sentence-internal boundary (where CBT = complex boundary tone, CR = continuation rise, SP = sustained pitch, (P) = pause; also H = high tone, L = low tone).

<table>
<thead>
<tr>
<th>Type</th>
<th>IntP</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visible</td>
<td>Audible</td>
<td></td>
</tr>
<tr>
<td>H%</td>
<td>L%</td>
<td>H%</td>
</tr>
<tr>
<td>occ.</td>
<td>15</td>
<td>1</td>
</tr>
<tr>
<td>%</td>
<td>13.2%</td>
<td>0.8%</td>
</tr>
<tr>
<td>Total</td>
<td>114 IntP boundaries (~100%)</td>
<td>141 ip boundaries (~100%)</td>
</tr>
</tbody>
</table>

5.1.1.2. Preboundary lengthening. As for preboundary lengthening, the results show that the length of a given syllable is significantly longer in the preboundary position than in the non-preboundary position. In what follows, I first briefly describe the scrutinized syllables. After that, I present the average values for syllable duration for all speakers (in milliseconds and percentages), followed by a statistical evaluation of the durational difference (paired t-test).

Lengthening is relational and for this reason, to demonstrate preboundary lengthening, the length of preboundary syllables has to be compared to the length of syllables in non-preboundary positions. Unlike Feldhausen (2010:78ff.), who examines syllable length across different types of CV syllables (and thereby confirms previous results that find that preboundary syllables are longer than non-preboundary ones), the present study follows the methodology of Gabriel et al. (2011:173ff.) and concentrates on only one single syllable, namely <la> [la]. The syllable occurs in set A (see section 4.1.2) and set B (see Appendix A), both in a typical preboundary position, at the end of the embedded subject (<aigüa ‘eagle’ and Ángela), and in a position where it hardly ever experiences lengthening, namely at the very beginning of the target sentences (in the form of the definite feminine article: La Bárbara, set A; La Silvia, . . . , set B). The results confirm the previously observed pattern: Preboundary la is significantly longer than its non-preboundary counterpart.

The bar diagram in Fig. 6 illustrates the clear difference between preboundary and non-preboundary syllables by giving their average length expressed in milliseconds. The average for all ten speakers is given in the last column, indicated by . The preboundary syllables have an average length of 178.4 ms, while the non-preboundary syllables, i.e. the syllables in
IntP-initial position, have a length of 118.5 ms. Turning to single speakers, it can be seen that, in general, there is no enormous interspeaker variation: The non-preboundary syllable normally has an average length between 100 and 130 ms, while the preboundary syllable is between 40 to 80 ms longer. Only speaker MO has a difference of more than 120 ms between the two syllables, while speaker RS shows only a very small difference. A paired \( t \)-test on scaled data shows that the durational differences are significant. For this test, I normalized the absolute values to avoid confounding influences induced by extreme values or by differences in speech rate (the normalized values are given in Appendix C). The syllable duration was normalized to sentence duration for each speaker using the formula \( d' = d/s \) (i.e. the normalized duration \( d' \) of a CV syllable is the proportion of its absolute duration \( d \) to the sentence duration \( s \)). \( d' \) was then multiplied by 1000 to avoid decimal places. Statistical comparison of the normalized syllable durations using paired \( t \)-tests yields the result that syllables in preboundary position are significantly longer than non-preboundary syllables (\( t = -9.726, \ p < 0.01 \)). Shapiro–Wilk tests showed that all data sets had a normal distribution (non-preboundary syllables: \( p = .803 \), preboundary syllables: \( p = .099 \)). The boxplots of the normalized syllable durations in Fig. 7 illustrate these significant differences. While the non-preboundary syllables have an average normalized duration of 35, the average for the preboundary syllables is 52.7. Furthermore, whereas non-preboundary syllables have a small variance (from 32 to 39) and a small standard deviation (SD = 2), preboundary syllables have a high variance (from 48 to 63) and a high standard deviation (SD = 4.9).

5.1.2. Phrasing patterns

In this section, the intonational groupings of the SVO structures of Experiment 1 are presented. The results show that (a) prosodically heavy CP objects increase matrix (SV) phrasing, (b) there is variation in the intonational grouping of the complex SVO structures, (c) 80% of the CP objects are prosodically separated from the matrix clause, and (d) in embedded clauses, S and V never phrase together.

The phrasing patterns are presented in Fig. 8. There are four groupings. The black bars give the percentage values of these groupings. The word order of the complex SVO structures is formulated in detail using the notation SVqSVO, which

---

19 See Appendix C for a table showing the millisecond values of each of the ten speakers for each sentence.

20 In fact, there were eight further groupings; due to low numbers, I ignore them in the analysis and in what follows. The four groupings I discuss here represent more than 83% of the data (see Feldhausen, 2010:85ff., for an overview of the eight seldom realized groupings). For the sake of convenience, I assume for further calculation, beginning in Fig. 8 (and with the exception of Fig. 13) that the four main groupings constitute 100%.
ignores prosodic weight. The symbol 'q' stands for the complementizer que 'that'. SV preceding q represents matrix subject and verb; SVO following q represents the embedded subject, verb, and object; the syntactic structure is [SV [qSVO]CP2]CP1. As before, parentheses ‘()’ represent prosodic grouping.

The percentage values in Fig. 8 reveal that the most common phrasing pattern is (SV)(qS)(VO), in which the matrix subject and verb are phrased together, followed by the complementizer and embedded subject, followed by the embedded verb and object (56%; cf. Fig. 9). The second favored phrasing is (S)(V)(qS)(VO), which comes to 24%. This phrasing pattern differs from the former only in that the matrix subject and verb are phrased separately (cf. Fig. 10). Thus, the two most common groupings place a boundary between the matrix verb and the object clause. In total, the object clause is separated 80% of the time (56% + 24%). As shown in Table 4, 60% of the prosodic breaks preceding object clauses are ip boundaries and 40% are IntP boundaries. The two remaining groupings presented in Fig. 8 are (SVqS)(VO) and (S)(VqS)(VO). They are both at 10% (cf. Figs. 11 and 12).21 Despite the differences between the four groupings, there are also some parallels. Interestingly, the parallels are specific to the embedded clause. First, the embedded subject is always followed by a boundary. This holds irrespective of the weight of the embedded constituents (a detailed explanation of the effect of the different weight conditions is presented in Fig. 13). Second, the embedded verb and object always phrase together.22

Fig. 9 offers one representative pitch track of the most common phrasing pattern (SV)(qS)(VO). There is no boundary after the matrix subject Silvia. The pitch contour on Silvia starts falling immediately after the H tone of the pitch accent L+>H* and is not followed by a boundary tone. Throughout the figures, break index 1, which signals a word boundary, is used to indicate

---

21 Most of the (SVqS)(VO) groupings result from the fact that some boundaries have been classified as unclear.
22 In the overall data, though, there are some cases in which embedded S phrases with following material or in which embedded V and O are separated by a boundary. These phrasings only occur in the eight seldom realized groupings (see fn. 20 for further details).
that there is no phrasal prosodic break at the ip or IntP level (break index 3 and 4 respectively). A phrasal prosodic break can be observed, however, after the verb. This break is signaled by a pause of 100 ms and a continuation rise (here H-H%). The next break is located after the embedded subject l’Àngela ‘the Àngela’ and it is realized by the complex boundary tone L-H% combined with a pause of 141 ms.

Table 4
Percentages of boundary types immediately preceding the embedded clause in the (SV)(qS)(VO) and (S)(V)(qS)(VO) groupings (all ten speakers, all three sets).

<table>
<thead>
<tr>
<th></th>
<th>(SV)(qS)(VO)</th>
<th>(S)(V)(qS)(VO)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>ip</td>
<td>31 (55%)</td>
<td>17 (70%)</td>
<td>48 (60%)</td>
</tr>
<tr>
<td>IntP</td>
<td>25 (45%)</td>
<td>7 (30%)</td>
<td>32 (40%)</td>
</tr>
</tbody>
</table>

Fig. 10. (SV)(qS)(VO) phrasing – Waveform, spectrogram, and F0 trace for the sentence La Bàrbara suposa que la meva àguila robà el ratoli ‘Barbara assumes that my eagle stole the mouse’ of speaker MO (sentence 35_Emb_MO).

Fig. 11. (SVqS)(VO) phrasing – Waveform, spectrogram, and F0 trace for the sentence El pare va dir que l’Amèlia se n’ha anat a Màlaga ‘The father said that Amèlia went to Malaga’ of speaker IS (sentence 41_Emb_IS).
Fig. 10 shows the pitch track of the second most common grouping, (S)(V)(qS)(VO). The break between the matrix subject and verb is realized by L-H%, as are the two further breaks.23

A corresponding pitch track of the grouping (SVqS)(VO) is given in Fig. 11. The only intonational break is located after the embedded subject el Ame`lia ‘the Ame`lia’ and is realized by H- (i.e. continuation rise). No boundary can be found after the matrix subject el pare ‘the father’, since the delayed peak of L+H* is located on the following word va (PST.3SG). There is no break between the matrix verb dir ‘say’ and the complementizer que ‘that’ either. The F0 contour in this position shows a transition between the verb’s pitch accent and the low tone of the embedded subject.

An example of (S)(VqS)(VO) appears in Fig. 12. Both sentence-internal breaks are marked by a continuation rise (H- and H-H% plus pause).

Fig. 13 presents the frequencies of these four phrasing patterns in the four different weight conditions,24 which involve varying the weight of the embedded S and O. The predominant grouping (SV)(qS)(VO) is the most common grouping in the

---

23 One interesting phonetic characteristic should be noted here. If the end of the word is marked by a complex boundary tone, the peak of the pitch accent is not delayed but aligns with the right edge of the strong syllable. This holds e.g. for the syllable l’A`n of l’A`ngela in Fig. 9 and for the syllable Bárbara in Fig. 10. On the other hand, if the pitch accent is not followed by a complex boundary tone, a delayed peak is possible. This is not a generalization: it is a description of the data at hand. Further research has to be done in order to generalize this statement.

24 If the values for each condition are added, the total is not 100%. This is due to the fact that the difference represents the percentage of the eight seldom realized groupings (cf. fn. 20), which are not listed in the figure.
three other conditions short S / short O (~47%), long S / short O (60%), and long S / long O (60%). However, it is not the most common in the short S / long O condition (only 20%); rather, (S)(V)(qS)(VO) is what is realized most often (30%). In addition, the two least common groupings (SVqS)(VO) and (S)(VqS)(VO) have the same relatively high frequency in this condition, while in the other three conditions, they have relatively lower and dissimilar frequencies. However, the condition short S / long O, known for its special effect on grouping in simple SVO structures (see D’Imperio et al., 2005), does not change the grouping of the embedded constituents: Throughout all conditions, the embedded subject is phrased separately from following V and O, which in turn phrase together.

5.1.3. Discussion

In section 2 two hypotheses were formulated, H1 and H2. H1 says that sentential objects increase the frequency of single-group phrasing of matrix SV. H2 says that sentential objects are not obligatorily separated prosodically from the preceding clause. Both hypotheses are borne out in Experiment 1. As for H1, the matrix (SV) phrasing in complex structures totals 56%, whereas it reaches only 33.1% in D’Imperio et al. (2005). As for H2, the results show that the boundary preceding the embedded clause is not obligatory, even though Catalan speakers seem to prefer a boundary in this position. Previous studies such as Downing (1970), Nespor and Vogel (1986/2007), and Truckenbrodt (2005) agree on the non-existence of obligatory IntP boundaries before object clauses. My results conform with their findings in so far as only 80% of the object clauses are separated (Fig. 8) and among them only 40% are separated by IntP boundaries (Table 4). However, even though the high percentage of 80% does not reflect obligatoriness, Catalan clearly tends to separate object clauses from preceding material.

The results further show that there is great variability as to whether prosodic constituents have an ip- or IntP-boundary tone. This variation of edge tones seems to be normal in Catalan. Astruc (2005), for example, shows that sentential adverbs are separated by (low) ip-boundary tones (32.7%) nearly as often as by (low) IntP-boundary tones (33.2%; see Astruc, 2005:154: Table 4.3). Feldhausen (2010) shows the same variation for Catalan left-dislocations, although there is a slight preference for IntP boundaries (Feldhausen, 2010:166). This variation in edge tones might be the reason why (a) the Catalan boundary cues described in Frota et al. (2007) are not specific as to the exact level of the prosodic hierarchy, and (b) only one type of edge tone (namely T%) is assumed in the Cat_TOBI version of Prieto et al. (2009).

The phrasing results across the four different conditions (Fig. 13) show that the prosodic weight of the embedded constituents does not play an important role in the phrasing of complex SVO structures. Only the condition short S / long O slightly varies the order of the four groupings, but this effect does not change the general picture. In the remainder of the paper I therefore ignore the different conditions. Even within embedded clauses, prosodic weight does not play a role. Unlike the findings for main clauses in D’Imperio et al. (2005), a long DP object never leads to a (SV) grouping in embedded SVO clauses. The present data thus shows that the (SV) grouping is a root phenomenon; it never shows up in embedded clauses.

Whether it is the prosodic weight of the sentential object (consisting of three or more ω) or its syntactic status (CP) that is responsible for the (SV)(O) phrasing has yet to be established. Each object clause in Experiment 1 is composed of a transitive verb with an overtly realized subject and is thus prosodically branching. Therefore, no distinction can be made between the two factors prosodic weight and syntactic status. Separating the two factors, the test sentences will have to include a sentential object consisting of only one prosodic word, as in (2), repeated here as (8). If prosodic weight is the key factor, the phrasing of (8) should be (S)(VO) (cf. (1a), repeated here as (9a)), since the object is light. If syntactic status is crucial, the phrasing should be (SV)(O) (cf. (1b), repeated here as (9b)), because of the CP object. Experiment 2 has been designed to address this topic.

(8) Sentential object consisting of one prosodic word

(S) (V O) <= pattern (9a)
(S V) (O) <= pattern (9b)

[La Maria i suposa [que proj dorm.]CP2]CP1
the M. assume.3SG that sleep.3SG
‘Mary assumes that (Peter) sleeps.’

(9) Simple SVO structures (main results from D’Imperio et al., 2005)

( S ) ( V O )
a. L’àguila robà el ratolí.
the.eagle steal.3SG.PST the mouse
‘The eagle stole the mouse’
b. L’àguila robà el ratolí del meu germà.
the.eagle steal.3SG.PST the mouse of.the my brother
‘The eagle stole my brother’s mouse.’
5.2. Results – Experiment 2 (1 ω)

Experiment 2 yields two main results: First, (SV)(O) is the most common phrasing pattern with prosodically light sentential O; and second, variation shows up with the same distribution as in Experiment 1.25

As Fig. 14 shows, there are altogether three different prosodic groupings attested (they are shown in the order of frequency). The most common phrasing pattern is (SV)(O), 48.1%, followed by (S)(V)(O), 31.5%, followed by the third and last grouping (S)(VO), 20.4%. The fourth logically possible grouping, (SVO), is not attested in the data.

Figs. 15 and 16 illustrate the most common phrasing pattern of Experiment 2: (SV)(O). In Fig. 15 a sustained pitch (!H-) is realized right after dit ‘said’ and in Fig. 16 a continuation rise (H-) separates the matrix verb suposa ‘suppose’ from the embedded clause que van trucar ‘that they called’.

The pitch track in Fig. 17 represents the second most common grouping (S)(V)(O). The matrix subject La Silvia ‘Silvia’ is immediately followed by a continuation rise and a pause of 197 ms (including the drwaled end of the last vocal), i.e. a break at the IntP level (break index 4). The next boundary is located between the verb mencionat ‘mentioned’ and the embedded clause que mentien ‘that they lied’ and it is realized as a continuation rise H-.

Fig. 18 illustrates the least common grouping (S)(VO). After the first rising accent L+H* on the antepenultimate syllable of Bárbara the contour keeps rising until the boundary tone H-. In the remainder of the sentence, there is no further boundary and thus the sentential object que menteixen ‘that they lie’ prosodically groups together with the preceding matrix verb suposa ‘suppose’.

---

25 In the presentation of the results of Experiment 2, I concentrate solely on the phrasing patterns. The distribution of tonal and non-tonal boundary cues was described in Experiment 1, and there are no important differences in Experiment 2.
Finally, I compare the overall phrasing patterns attested in Experiments 1 and 2. As shown in Table 5, the matrix phrasing patterns are identical in the two experiments, and so is their order of frequency. Matrix (SV) phrasing represents the most common pattern in both experiments, followed by (S)(V), followed by (S)(V...]. The fourth grouping, (SV...], only occurs in Experiment 1. It does not occur in Experiment 2. Since no internal phrase boundaries are possible in an embedded clause consisting of only one prosodic word, having no boundary between matrix S and V or between matrix V and the embedded clause would mean that the entire complex SVO structure would group as one single prosodic phrase. This is not attested in the data. When the sentential object consists of only one prosodic word, there always appears a sentence-internal boundary in the matrix clause.

5.3. Discussion and intermediate conclusion

In Experiment 2, the two hypotheses H1 and H2 are again borne out. It is furthermore shown that the syntactic status of the sentential object is the key factor for matrix (SV) grouping.
As for H1, matrix (SV) phrasing in complex structures totals 48.1%, represented by the grouping (SV)(O). It is thus a bit lower than in Experiment 1, but still substantially higher than in D’Imperio et al. (2005). As for H2, 79.6% of the sentential objects are separated from the preceding matrix clause by a prosodic boundary. As in Experiment 1, the boundary is not obligatory (confirming H2), but the tendency toward separation is clear. Experiment 2 shows that it is the syntactic status of the object clause rather than prosodic weight, which is responsible for matrix (SV) phrasing. If it were prosodic weight, (S)(VO) rather than (SV)(O) should be the most common phrasing pattern. Even though the sentential object consists of only one prosodic word – and hence has the same prosodic weight as the DP object in (1a) – the phrasing strongly tends to conform to the pattern induced by a prosodically heavy object, namely (SV)(O). This suggests that the syntactic category of the object strongly influences the prosodic structure, entirely independently of its actual prosodic weight. Prieto (2005) argues that Catalan is sensitive to eurhythmic constraints. Along with the effect of a prosodically heavy DP object on phrasing, she shows that there is a binary length requirement, which has a prominent role in Catalan (Prieto’s 2005:205 Max-Bin-End constraint; cf. also (12) in section 6.1). The present study, however, shows that the syntactic factor ‘CP object’ can override eurhythmic effects (which themselves override other syntactic factors). Consequently, Catalan is clearly a language in which prosodic/eurhythmic constraints interact with syntactic constraints in a balanced way.

In sum, experiments 1 and 2 thus show that there is variation in prosodic grouping, but that (SV)(O)CP2) is the most common phrasing pattern irrespective of prosodic weight. The main phrasing characteristics demonstrated in the experiments for complex SVO structures are listed here:

**Characteristics based on Experiments 1 and 2:**

i. the embedded object clause is normally separated from the matrix clause
ii. the matrix subject has the tendency to phrase with the matrix verb

**Characteristics based on Experiment 1 only:**

iii. the embedded object is phrased with the embedded verb
iv. in most cases the embedded subject is phrased alone

---

Table 5
Comparison of the matrix phrasing pattern of both experiments (ignoring the details of the phrasing pattern of the embedded object of Experiment 1). Parentheses represent prosodic phrasing, and square brackets indicate relevant syntactic edges.

<table>
<thead>
<tr>
<th></th>
<th>Exp. 1</th>
<th>Exp. 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>(SV)(O)CP2</td>
<td>(SV)(O)CP2</td>
</tr>
<tr>
<td>2.</td>
<td>(S)(V)CP2</td>
<td>(S)(V)CP2</td>
</tr>
<tr>
<td>3.</td>
<td>(SV)(O)CP2</td>
<td>(SV)(O)CP2</td>
</tr>
<tr>
<td>4.</td>
<td>(SV)(O)CP2</td>
<td>(SV)(O)CP2</td>
</tr>
</tbody>
</table>

As for H1, matrix (SV) phrasing in complex structures totals 48.1%, represented by the grouping (SV)(O). It is thus a bit lower than in Experiment 1, but still substantially higher than in D’Imperio et al. (2005). As for H2, 79.6% of the sentential objects are separated from the preceding matrix clause by a prosodic boundary. As in Experiment 1, the boundary is not obligatory (confirming H2), but the tendency toward separation is clear.

Experiment 2 shows that it is the syntactic status of the object clause rather than prosodic weight, which is responsible for matrix (SV) phrasing. If it were prosodic weight, (S)(VO) rather than (SV)(O) should be the most common phrasing pattern. Even though the sentential object consists of only one prosodic word – and hence has the same prosodic weight as the DP object in (1a) – the phrasing strongly tends to conform to the pattern induced by a prosodically heavy object, namely (SV)(O). This suggests that the syntactic category of the object strongly influences the prosodic structure, entirely independently of its actual prosodic weight. Prieto (2005) argues that Catalan is sensitive to eurhythmic constraints. Along with the effect of a prosodically heavy DP object on phrasing, she shows that there is a binary length requirement, which has a prominent role in Catalan (Prieto’s 2005:205 Max-Bin-End constraint; cf. also (12) in section 6.1). The present study, however, shows that the syntactic factor ‘CP object’ can override eurhythmic effects (which themselves override other syntactic factors). Consequently, Catalan is clearly a language in which prosodic/eurhythmic constraints interact with syntactic constraints in a balanced way.

In sum, experiments 1 and 2 thus show that there is variation in prosodic grouping, but that (SV)(O)CP2) is the most common phrasing pattern irrespective of prosodic weight. The main phrasing characteristics demonstrated in the experiments for complex SVO structures are listed here:

**Characteristics based on Experiments 1 and 2:**

i. the embedded object clause is normally separated from the matrix clause
ii. the matrix subject has the tendency to phrase with the matrix verb

**Characteristics based on Experiment 1 only:**

iii. the embedded object is phrased with the embedded verb
iv. in most cases the embedded subject is phrased alone
Any analysis of the phrasing of complex SVO structures in Catalan has to account for these characteristics. In section 6, such an analysis is given.

6. Theoretical approach

This section offers an analysis of the prosodic phrasing of Catalan complex SVO structures. The analysis is grounded in a variant of Optimality Theory (Prince and Smolensky, 1993/2004) called stochastic OT (Boersma and Hayes, 2001). It is further based on Prieto’s (2005) account of phrasing in simple SVO sentences and includes her three constraints MAX-BIN-END >> MIN-N-PHRASES >> ALIGN-XP,R. A further constraint – ALIGN-CP,L – is proposed, which is ranked below MAX-BIN-END but higher than MIN-N-PHRASES and ALIGN-XP,R, as shown in (10). The constraint ALIGN-CP,L aligns the left edge of a CP with the left edge of a prosodic phrase and thus accounts for the tendency for the embedded clause to be prosodically separated from the matrix clause.

(10) MAX-BIN-END >> ALIGN-CP,L >> MIN-N-PHRASES >> ALIGN-XP,R

By using the stochastic-OT framework (Boersma and Hayes, 2001), in which the constraints are ranked on a continuous ranking scale, the model can account for the variation in the data of both experiments. I argue that ALIGN-CP,L, MIN-N-PHRASES, and ALIGN-XP,R overlap, and the actual ranking of the constraints will sometimes be the reverse of their ‘normal’ ranking. The different rankings derived from the underlying form in (10) account for the data.

In section 6.1, the four OT constraints are introduced. Section 6.2 proposes that four different rankings can account for the different groupings of Experiments 1 and 2 and that the one that accounts for the most common groupings constitutes the underlying ranking. In section 6.3 the different rankings are modeled in the framework of stochastic OT (Boersma, 1998; Boersma and Hayes, 2001).

6.1. The constraints

In order to account for the prosodic phrasing of SVO structures with a sentential object, four constraints are needed. Previous research (Prieto, 2005) has shown the first three constraints to be relevant for the prosodic grouping of simple SVO structures. In this section I depart from Prieto (2005) in two respects: first, by assuming ALIGN-CP,L, and second, by claiming that the three constraints ALIGN-CP,L, MIN-N-PHRASES, and ALIGN-XP,R are overlapping constraints (in stochastic-OT terms) and consequently show free variation.

1) ALIGN-XP,R: Selkirk’s (1986, 1996, 2000) end-based/Alignment theory predicts ‘anchor points’ where both syntactic and prosodic structures coincide.26 The interface constraints require that the edge of a maximal projection in the surface syntactic structure (i.e. XP) coincide with (i.e. align with) the edge of a prosodic constituent. It can be the right edges that are aligned with each other or the left edges. Languages are characterized as having either right-alignment (e.g. Kisseberth and Abasheikh, 1974, and Selkirk, 1986 for Chi Miwi:ni) or left-alignment (Selkirk and Shen, 1990 for Chinese; Selkirk and Tateishi, 1991 for Japanese). The OT version of the end-based theory is given in (11).

(11) Alignment Constraints

a. ALIGN-XP,R: \( \text{ALIGN (XP, R; } \phi, R) \)

‘Align the right edge of a syntactic XP to the right edge of \( \phi \).’

b. ALIGN-XP,L: \( \text{ALIGN (XP, L; } \phi, L) \)

‘Align the left edge of a syntactic XP to the left edge of \( \phi \).’

26 In Selkirk (2009a,b) a new theory of the syntax-prosody interface is proposed, which is labeled Match Theory. In the Alignment theory of Selkirk (1996), the framework adopted in this study, only the right or the left edge of a syntactic constituent is aligned to the corresponding edge of a prosodic constituent. In contrast, the Match Theory of syntactic-prosodic structure faithfulness requires full correspondence between syntactic and prosodic constituents, causing ‘a matching up of the constituents themselves’ (Selkirk, 2009a:section 2). This is achieved, on the one hand, by S–P faithfulness constraints, which require ‘that syntactic constituency be faithfully reflected in prosodic constituency’, and, on the other hand, by P–S faithfulness constraints, which require that ‘prosodic constituency be a faithful reflection of syntactic constituency’ (Selkirk, 2009a:section 2).

In addition to Selkirk’s accounts, there are other important approaches: Truckenbrodt’s (1995, 1999, 2005) WRAP constraints, and the relation of phases (CP, VP) to prosodic structure (Ishihara, 2004, 2007; Kratzer and Selkirk, 2007). However, these approaches are not of importance for accounting for the data of the present study. As for the WRAP constraints, Feldhausen (2010:13) argues that WRAP–XP has no effect on the phrasing of complex clauses. As for the phase-based account, it is difficult to see how having phases trigger boundary placement can account, first, for the variation in the data and, second, for the location of the boundary. In the present data, there are instances in which no boundary appears at the left edge of a CP. This optionality must be accounted for. Furthermore, at the end of a phase, the complement of the phase head is sent away to Spell-Out. The complement of the phrase head C is TP (or IP) and its right and left edge would thus be available for prosodic boundary placement. However, in the present data, a prosodic boundary is needed at the left edge of the CP. I leave the other approaches aside and concentrate on a stochastic-OT account.
The constraints in (11) are taken as ranked and violable constraints (Selkirk, 1996). In addition, they are universal in this theory. This becomes important for the new constraint ALIGN-CP,L (see below). The constraint ALIGN-XP,R is taken to be active in Catalan. As far as my data are concerned, it has the effect of placing a prosodic phrase boundary after the subject constituent.

2) **MAX-BIN-END**: Previous work has shown that factors such as constituent weight, i.e. the number of syllables or of prosodic words in a given constituent, and speech rate, also play a major role in sentence phonology. A heavy prosodic constituent has a greater tendency to be phrased independently (Ghini, 1993; Nespor and Vogel, 1986/2007; Zec and Inkelas, 1990, among others). The constraint MAX-BIN-END (Prieto, 2005) is a more restricted version of MAX-BIN (‘P-Phrases consist of maximally two prosodic words’: Sandalo and Truckenbrodt, 2002:295; cf. also Selkirk, 2000:244). MAX-BIN-END, (12), limits this prosodic binarity to prosodic phrases that contain the main stress of the utterance.

(12) Restricted prosodic binarity in Catalan (Prieto, 2005:205)

**MAX-BIN-END**

‘P-Phrases containing the main stress of the utterance consist of maximally two prosodic words.’

Since the main stress of all-new utterances in Catalan is on the last stressed syllable, MAX-BIN-END applies only to the sentence-final prosodic phrase in such utterances. Prosodic binarity needs to be restricted to the end of sentences because in Catalan there are longer p- phrases in non-final positions than in final positions, (13). The constraint sets only a maximal number of prosodic words. If the sentence-final phrase consists of only one prosodic word, as in (13), MAX-BIN-END is not violated.

(13) Phrasing of subjects with more than two prosodic words (Prieto, 2005:217)

<table>
<thead>
<tr>
<th>Subject</th>
<th>noun phrase</th>
<th>verb</th>
<th>object</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. (Els veïns catalans de l’Ebre) (s’enfaden)</td>
<td>The neighbors Catalan of the Ebre (River) get angry</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For my data, the constraint MAX-BIN-END is responsible for guaranteeing that, in the embedded clause, the subject, the verb, and the object do not get phrased together.27

3) **MIN-N-PHRASES**: The next non-syntactic factor influencing phrasing decisions concerns the rate of speech (Frascarelli, 2000; Jun, 2003; Nespor and Vogel, 1986/2007, among others). The faster a sentence is uttered, the longer the IntPs of this utterance tend to be, i.e. the utterance tends to be broken down into fewer IntPs (Nespor and Vogel, 1986/2007:195). Prince and Smolensky (1993:25, fn.13) set up a family of constraints called *STRUC that ensures that structure is constructed minimally. *P-PHRASE, (14a), which is part of the *STRUC family, seeks to avoid phonological phrases altogether (cf. Truckenbrodt, 1999:228, 2002:274). Elordieta et al. (2005:133) and Féry (2007) have corresponding constraints that place a ban on intonational phrases: *INTBREAK and *I-PHRASE respectively, (14b).

(14) Constraints punishing p- phrases and IntPs (taken from Féry, 2007)

<table>
<thead>
<tr>
<th>Constraint</th>
<th>P-Phrase</th>
<th>I-Phrase</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. *P-PHRASE:</td>
<td>‘No phonological phrase’</td>
<td></td>
</tr>
<tr>
<td>b. *I-PHRASE:</td>
<td>‘No intonational phrase’</td>
<td></td>
</tr>
</tbody>
</table>

These constraints have the effect of punishing additional structure. The corresponding constraint in Prieto (2005) is called MIN-N-PHRASES and is defined in (15).

(15) Prieto’s (2005:216) constraint for avoiding p- phrases

**MIN-N-PHRASES** (rapid speech)

‘Minimize the number of phrases (rapid speech).’

As far as my data are concerned, this constraint has the effect just described: it minimizes the number of prosodic phrases in an utterance. However, in my analysis the constraint is not limited to a rapid speech rate.28

---

27 The careful reader might object that MAX-BIN-END is violated when the object of the embedded clause is (prosodically) branching. A possible solution to this shortcoming would be the high-ranked constraint IDENT-vP established in Feldhausen (2010:118). This constraint requires the right and the left edge of vP to coincide with the right and the left edge of a prosodic phrase. Thus it is independent of the actual length of the object. I refer to Feldhausen (2010) for a possible analysis which integrates this constraint.

28 I refrain from limiting the constraint to only rapid speech for conceptual reasons. The *STRUC family is not constrained to rapid or low speech but seeks to avoid (prosodic) structure altogether. As Elordieta et al. (2005:133) write, *STRUC is a ‘general constraint that penalises the insertion of intonational breaks’. Féry (2007) shows that the effect of the constraints can be controlled by either a high or a low ranking: *P-PHRASE is ranked higher than *I-PHRASE.
4) \textit{ALIGN-CP,L}: Experiments 1 and 2 show that embedded object clauses are typically separated from the matrix clause. Based on this empirical result, I propose the constraint \textit{ALIGN-CP,L} (16), which requires a prosodic boundary preceding an (embedded) clause.

\begin{enumerate}
\item Further constraint of the alignment family
\item \textit{ALIGN-CP,L}: ALIGN(CP,Left; PrP,Left)
\item ‘Align the left edge of a CP to the left edge of a prosodic phrase (PrP).’
\end{enumerate}

Having a constraint that reflects the influence of morphosyntactic structure on prosodic structure at sentence level is not new. Based on the proposal by Downing (1970), who argued that intonational phrases obligatorily align with root sentences (including object clauses, but excluding adverbial constituents and other preposed, postposed, and interposed constituents), Gussenhoven (2004:167) introduced the OT constraint \textit{ALIGN(S, i)}, which aligns the right edge of every S(entence) with the right edge of IntP. My version, however, aligns the left edge of the CP with the left edge of a prosodic phrase.\textsuperscript{29} The possibility of having both alignment of the right edge and alignment of the left edge of a prosodic constituent in one and the same language is extensively argued for by de Lacy (2003) for the Polynesian language Māori (and cf. Gussenhoven, 2004:285 for a similar assumption with evidence from English). de Lacy (2003) invokes the notion of universality by concluding that in all grammars all constraints are present. He states that a grammar cannot choose between, for example, \textit{ALIGN(XP,Left; φ,Left)} and \textit{ALIGN(XP,Right; φ,Right)}, but that both constraints are present in every grammar (de Lacy, 2003:60). However, it is possible that a constraint in a given language is rendered inactive (de Lacy, 2003:70ff.) by higher-ranked constraints, so that the language behaves as if there were only left-alignment or right-alignment. The Catalan data presented in this study is in line with de Lacy (2003) and Gussenhoven (2004). We can thus conclude that the constraint \textit{ALIGN-CP,L} is active in Catalan. As Truckenbrodt (1999:228) notes, alignment constraints that are ranked below \textit{*P-PHRASE} are rendered inactive because a boundary favored by \textit{ALIGN} is punished by \textit{*P-PHRASE}. Thus, the universality argued for in de Lacy (2003) does not exclude languages in which, for example, \textit{ALIGN-XP,L} never shows up. In these languages \textit{ALIGN-XP,L} is simply ranked below \textit{*P-PHRASE} (or \textit{*I-PHRASE}).

6.2. The constraint hierarchy

In this section four rankings are introduced. I first show how different rankings of the constraints can account for the different groupings. I then propose one underlying constraint hierarchy from which all the other orderings are derived.

With different orders of the four OT constraints, different candidates win. The four rankings in (17) are able to account for the four different groupings attested in Experiment 1 as well as the three groupings in Experiment 2 (Figs. 8 and 14 respectively). The lowercase letters below the constraints help the reader identify the change in order. In addition, they signal the ‘selection point’ in the stochastic-OT approach developed in section 6.3 below. The percentage values are listed, because they are important for the stochastic-OT approach.

\begin{enumerate}
\item Ranking for (SV)(qS)(VO) [56%] and for (SV)(O) [48.1%]
\item \begin{tabular}{l}
\textit{MAX-BIN-END} >> \textit{ALIGN-CP,L} >> \textit{MIN-N-PHRASES} >> \textit{ALIGN-XP,R}
\end{tabular}
\item Ranking for (S)(V)(qS)(VO) [24%] and for (S)(V)(O) [31.5%]
\item \begin{tabular}{l}
\textit{MAX-BIN-END} >> \textit{ALIGN-CP,L} >> \textit{ALIGN-XP,R} >> \textit{MIN-N-PHRASES}
\end{tabular}
\item Ranking for (SVqS)(VO) [10%] and for (S)(VO) [20.4%]
\item \begin{tabular}{l}
\textit{MAX-BIN-END} >> \textit{MIN-N-PHRASES} >> \textit{ALIGN-XP,R} >> \textit{ALIGN-CP,L}
\end{tabular}
\item Ranking for (SVqS)(VO) [10%]
\item \begin{tabular}{l}
\textit{MAX-BIN-END} >> \textit{ALIGN-XP,R} >> \textit{MIN-N-PHRASES} >> \textit{ALIGN-CP,L}
\end{tabular}
\end{enumerate}

Ranking (17a), illustrated in Tables 6 and 7 for Experiments 1 and 2 respectively, accounts for the most common phrasing patterns. Table 6 shows the competition between the four groupings of Experiment 1 (candidates a–d) and two hypothetical phrasings (e and f). The highest-ranked constraint, \textit{MAX-BIN-END}, is violated only by candidate e, thus the evaluation is passed onto the next constraint \textit{ALIGN-CP,L}. This constraint is violated by candidates c, d, e, and f because these candidates do not have a boundary between matrix V and the complementizer q at the left edge of CP. But there are still two possible winners: a and b. The next constraint, \textit{MIN-N-PHRASES}, is able to decide between them. Candidate a wins, because it violates the

\textsuperscript{29} \textit{ALIGN-CP,L} might be reminiscent of the constraint \textit{ALIGN-ExtVP,L} of Elordieta et al. (2005), which aligns the left edge of an extended projection of VP to the left edge of a Major Phrase (i.e. prosodic phrase). Even though CP is an extended projection of VP, the two constraints differ in one crucial aspect. While \textit{ALIGN-ExtVP,L} also inserts a boundary between CP and TP (cf. Elordieta et al., 2005:129), \textit{ALIGN-CP,L} only inserts a boundary preceding the CP – exactly what is needed here.
constraint less severely than candidate b. The lowest-ranked constraint, ALIGN-XP,R, is not of importance here as the decision has already been made.

The same ranking accounts for the most common grouping of Experiment 2, (SV)(O), see Table 7. While candidates c and d immediately violate MAX-BIN-END and ALIGN-CP,L, the decision between candidates a and b is taken as before by MINIMIZE-N-PHRASES: since candidate a has fewer violations of MINIMIZE-N-PHRASES, (SV)(O) wins.

The ranking in (17b) accounts for the second most frequent phrasing patterns of Experiment 1 and 2. As shown in Table 8 the winning candidate of Experiment 1 with this ranking is (S)(V)(qS)(VO). Table 9 shows that the winning candidate of Experiment 2 is (S)(V)(O). The difference between Table 6/Table 7 and Table 8/Table 9 is the order of MINIMIZE-N-PHRASES and ALIGN-XP,R. Since the latter is ranked higher in Table 8/Table 9, it is this constraint that decides between candidates a and b. Candidate b wins because it does not violate ALIGN-XP,R, whereas candidate a violates this constraint because the matrix subject does not right-align with a prosodic boundary.

Table 6
Constraint ranking for the most common phrasing pattern (SV)(qS)(VO) of Experiment 1 (‘Sm’: matrix subject; ‘Se’: embedded subject).

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>56% (SV)(qS)(VO)</td>
<td>3</td>
<td>$S_m$</td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>24% (S)(V)(qS)(VO)</td>
<td>4</td>
<td>$S_m$</td>
<td></td>
</tr>
<tr>
<td>c</td>
<td>10% (SVqS)(qS)(VO)</td>
<td>*</td>
<td>2</td>
<td>$S_m$</td>
</tr>
<tr>
<td>d</td>
<td>10% (S)(VqS)(VO)</td>
<td>*</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>e</td>
<td>(SVqSVO)</td>
<td>*</td>
<td>1</td>
<td>$S_mS_e$</td>
</tr>
<tr>
<td>f</td>
<td>(SVqSV)(O)</td>
<td>*</td>
<td>2</td>
<td>$S_mS_e$</td>
</tr>
</tbody>
</table>

Table 7
Constraint ranking for the most common phrasing pattern (SV)(O) of Experiment 2.

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>48.1% (SV)(O)</td>
<td>2</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>31.5% (S)(V)(O)</td>
<td>31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c</td>
<td>20.4% (S)(V)(O)</td>
<td>*</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>d</td>
<td>0% (SV)(O)</td>
<td>*</td>
<td>1</td>
<td>*</td>
</tr>
</tbody>
</table>

Table 8
Constraint ranking for the phrasing pattern (S)(V)(qS)(VO) of Experiment 1.

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>SV[O]</td>
<td>MAX-BIN-END</td>
<td>ALIGN-CP,L</td>
<td>ALIGN-XP,R</td>
<td>MINIMIZE-N-PHRASES</td>
</tr>
<tr>
<td>a</td>
<td>56% (SV)(qS)(VO)</td>
<td>$S_m$</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>24% (S)(V)(qS)(VO)</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c</td>
<td>10% (SVqS)(qS)(VO)</td>
<td>*</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>d</td>
<td>10% (S)(VqS)(VO)</td>
<td>*</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>e</td>
<td>(SVqSVO)</td>
<td>*</td>
<td>1</td>
<td>$S_mS_e$</td>
</tr>
<tr>
<td>f</td>
<td>(SVqSV)(O)</td>
<td>*</td>
<td>2</td>
<td>$S_mS_e$</td>
</tr>
</tbody>
</table>

Table 9
Constraint ranking for the phrasing pattern (S)(V)(O) of Experiment 2.

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>SV[O]</td>
<td>MAX-BIN-END</td>
<td>ALIGN-CP,L</td>
<td>ALIGN-XP,R</td>
<td>MINIMIZE-N-PHRASES</td>
</tr>
<tr>
<td>a</td>
<td>48.1% (SV)(O)</td>
<td>*</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>31.5% (S)(V)(O)</td>
<td>*</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>c</td>
<td>20.4% (S)(V)(O)</td>
<td>*</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>d</td>
<td>0% (SVO)</td>
<td>*</td>
<td>*</td>
<td>1</td>
</tr>
</tbody>
</table>
The ranking in (17c) leads to the winning candidate c, representing the third most common prosodic grouping in both experiments: see Tables 10 and 11. The difference between Tables 6 and 10 (Experiment 1) as well as between Tables 7 and 11 (Experiment 2) is in the position of ALIGN-CP,L. It is ranked last in the present ranking, below MIN-N-PHRASES and ALIGN-XP,R.

As for Table 10, after the evaluation process has passed onto MIN-N-PHRASES, two candidates remain: c and f. Candidate c wins since candidate f violates ALIGN-XP,R one time more than candidate c does: not only does the matrix subject in candidate f fail to right-align with a prosodic phrase, but so does the embedded subject. Table 10 shows the importance of MAX-BIN-END being ranked higher than MIN-N-PHRASES. If it were lower than MIN-N-PHRASES, candidate e would wrongly win. However, in complex sentences the single phrasing of the whole clause is never attested in my data. Thus such a phrasing is very unlikely, if not impossible. In order to account for this, MAX-BIN-END is ranked higher than MIN-N-PHRASES.

In Experiment 2, the third most common phrasing pattern is (S)(VO), the winning candidate in Table 11. Due to the fact that MIN-N-PHRASES is ranked relatively high, candidate b, with three prosodic phrases, cannot win. Since candidate a and c both violate MIN-N-PHRASES twice, the decision is passed down to ALIGN-XP,R. Candidate c wins, because it respects the alignment constraint.

Tables 7, 9 and 11 demonstrate that by the first three constraint rankings, namely (17a–c), all the groupings attest in Experiment 2 are accounted for. This set of constraint rankings also accounts for the non-occurrence of the grouping (SVO): (SVO) always violates MAX-BIN-END, and in all four rankings, this constraint is strictly ranked above the other three constraints ALIGN-CP,L, MIN-N-PHRASES, and ALIGN-XP,R. Thus there is no chance for (SVO) to be the optimal candidate.

In contrast to Experiment 2, there is a fourth phrasing pattern in Experiment 1. The ranking in (17d), illustrated in Table 12, accounts for this grouping. Due to the fact that ALIGN-CP,L is the lowest-ranked constraint and MIN-N-PHRASES is ranked lower than ALIGN-XP,R, two candidates, b and d, survive the evaluation until MIN-N-PHRASES. Since candidate b violates the constraint more often than candidate d, the latter wins.
The preceding tables show that changing the order of the three constraints ALIGN-CP,L, MIN-N-PHRASES, and ALIGN-XP,R results in different winning candidates. However, it would be inappropriate to assume that there are four different grammars in Central Catalan for the different groupings; rather, one phonological input has several outputs. Such a situation is normally described as ‘variation’ or ‘optionality’. I take (17a), here repeated as (18), as the underlying form for the different groupings.

(18) Underlying constraint hierarchy for embedded SVO phrasing

\[
\text{MAX-BIN-END} >> \text{ALIGN-CP,L} >> \text{MIN-N-PHRASES} >> \text{ALIGN-XP,R}
\]

Two aspects speak in favor of (18). First, this ranking agrees with the order of constraints proposed for simple SVO structures in Prieto (2005) and modified in Feldhausen (2010:110), namely MAX-BIN-END >> MIN-N-PHRASES >> ALIGN-XP,R. Only one addition is proposed, placing ALIGN-CP,L after MAX-BIN-END. Thus the analysis is an extension of the approach by Prieto (2005); in addition, it reflects the close relationship between the simple and complex SVO structures. Second, the ranking accounts for the most common phrasing patterns of both experiments: it is common practice to take the form with the widest distribution as underlying.30

6.3. Optionality in Optimality Theory – stochastic OT

Classical generative models are non-probabilistic and for this reason any given sequence is either grammatical or completely impossible (Pierrehumbert, 2001:195). However, the results in section 5 show that there is frequency-dependent variation in the present data. These results can be accounted for in a model such as stochastic OT (Boersma and Hayes, 2001), a probabilistic model for capturing gradient optionality. First, the model is briefly introduced. Next, it is applied to the present data.

6.3.1. The Stochastic Model by Boersma and Hayes (2001)

Two notions are important for stochastic OT (Boersma, 1998; Boersma and Hayes, 2001): the continuous ranking scale and the stochastic candidate evaluation. Instead of assuming a strict ranking scale, stochastic OT assumes a continuous one. Constraints have a certain ranking value. Higher values correspond to higher-ranked constraints, lower values correspond to lower-ranked constraints; see (19). The concrete ranking values 120, 100, 85, 80 are for demonstration and simply represent the distance relationship between the constraints. A shorter distance between two constraints implies that the relative ranking of the constraints is less fixed (e.g. 85–80 in contrast to 120–100). Less fixed orders have a bigger role to play in accounting for variation. Boersma (1998) and Boersma and Hayes (2001:47) suggest that constraints are not single points, but act as if they are associated with ranges of values (21). This happens due to a temporary perturbation of the position of each constraint by a random positive or negative value at evaluation time (i.e. the time when the candidates in an OT table have to be evaluated in order to determine a winner). This is illustrated in (20). The concrete value that is used for a single constraint is called the selection point. This point is indicated by the straight vertical line in (21); the line is marked ‘b’ at the end, meaning that this is the selection point for constraint b. The selection point can differ from the ranking value (dotted line). The latter is the center of the range, i.e. the value more permanently associated with the constraint (e.g. example value ‘100’ in (20) and (21)).

30 Cf. for example the motivation for choosing /c/ as the phoneme of the two German allophones [ç] and [x] (Hall, 2000:64).
A selection point near the center (i.e. the ranking value) is more probable than a selection point far away from the center because constraint ranges are interpreted as probability distributions. They can then account for noisy events that are described with a normal (= Gaussian) distribution. They can then account for noisy events that are described with a normal (= Gaussian) distribution (for details see Boersma and Hayes, 2001:48). The grey box around the center in (21) represents the standard deviation, in which most of the values drawn from a normal distribution are located. In stochastic OT, every constraint has the same standard deviation.

The fact that the order of constraints is not totally fixed becomes important when their distance is relatively short. If the distance is short enough, two (or more) constraints overlap, i.e. their standard deviations overlap. Due to the fact that at evaluation time it is possible to choose the selection points from anywhere within the two given constraints, the ranking of the constraints usually results in the ‘normal’ ranking order, but sometimes the reverse order is the result. Free variation arises due to overlapping constraints because they can generate multiple output forms from a single underlying form. The more the constraints overlap, the more probable a reverse ranking is. This means that in a certain percentage of the evaluations (depending on the amount of overlap) a lower-ranked constraint is given precedence over a higher-ranked constraint and a suboptimal candidate wins. Constraint ranges are hence interpreted as probability distributions.

In order to know what ranking value the necessary constraints have, Boersma and Hayes (2001) developed the Gradual Learning Algorithm (GLA), an algorithm for learning OT constraint ranking. The GLA requires two inputs: OT constraints and the frequencies of distribution. The second input enables the model to include the results stemming from empirical data. Now the process of learning an appropriate constraint ranking consists merely of finding a workable set of ranking values on a continuous scale. The GLA calculates the location of the constraints relative to each other.31

6.3.2. Application of the Gradual Learning Algorithm

In this section, the GLA is applied to complex SVO structures. First, overlapping constraints are identified. Second, application of the GLA provides the ranking values that derive the frequency effects.

The underlying constraint hierarchy was given in (18), the four different orders of the constraints in (17). As can be seen from (17), MAX-BIN-END is always the highest-ranked constraint, whereas the remaining three constraints vary their positions. I claim that while the ranking of the lower three constraints is as in (18), the distance between them is so short that their ranges overlap. The corresponding Catalan grammar is pictured in (22).

(22) Constraint hierarchy for Catalan clauses with sentential objects

The differences in the height and length of the constraint ranges in (22) have no meaning, but are for the sake of clarity. The constraints still have the same standard deviation. The area of overlap of the three constraints allows for any order between ALIGN-CP,L, MIN-N-PHRASES, and ALIGN-XP,R to be generated. This is exactly what is needed for deriving the four different rankings of (17).32 As an example, the order of the selection for the fourth grouping of Experiment 1 (S) (VqS)(VO) is illustrated in (23). The corresponding constraint ranking (17d) is repeated below for the sake of convenience.

(23) Order of selection points for the phrasing (S)(VqS)(VO)

(17d) MAX-BIN-END >> ALIGN-XP >> MIN-N-PHRASES >> ALIGN-CP,L

31 In Boersma and Hayes (2001), the empirical application of the GLA is illustrated with examples of free variation of glottal stop and glides in Ilokano (an Austronesian language of the northern Philippines), of output frequency in the Finnish genitive plural, and of gradient well-formedness judgments of English light and dark /l/.

32 There are six possible permutations, although only four are needed to account for the variation in Experiment 1 and only three are needed for Experiment 2. As the reader might easily determine, the remaining permutations generate existing groupings (a >> c >> b >> d and a >> d >> b >> c for groupings attested in both experiments, and a >> d >> c >> b for Experiment 2) and they therefore do not pose a problem for my approach.
Application of the GLA endows the hierarchy with concrete ranking values for the constraints. Applying the GLA to the data of Experiment 1 yields the possible values presented in (24). The associated frequency prediction in (25) is given in percentages to facilitate comparison with the empirical results. The values in parentheses reflect the absolute numbers of each winning candidate out of the 100,000 pieces of learning data (i.e. 56,254 + 23,975 + 9,917 + 9,854 = 100,000).\(^{33}\)

(24) Ranking values proposed by GLA
a. MAX-Bin-End 113.494
b. ALIGN-CP,L 108.814
c. MIN-N-Phrases 106.484
d. ALIGN-XP,R 105.295

(25) GLA frequency prediction

<table>
<thead>
<tr>
<th>Frequency prediction</th>
<th>Empirical results</th>
</tr>
</thead>
<tbody>
<tr>
<td>(SV)(qS)(VO)</td>
<td>56.25% (56254)</td>
</tr>
<tr>
<td>(S)(V)(qS)(VO)</td>
<td>23.97% (23975)</td>
</tr>
<tr>
<td>(S)(VqS)(VO)</td>
<td>09.92% (09917)</td>
</tr>
<tr>
<td>(S)(VqS)(VO)</td>
<td>09.85% (09854)</td>
</tr>
</tbody>
</table>

The application of the learning algorithm shows that the proposed grammar accounts for the data. This paper thus demonstrates that stochastic grammars can be used in (intonational) phonology. As shown here, variation in phrasing is rather the norm than an exception, and a grammatical model should be able to account for the empirical facts. The GLA proposed by Boersma and Hayes (2001) is such a model and the empirical frequency values of the present data are captured without difficulty.\(^{34}\)

7. Summary and conclusion

The present study shows that the syntactic projection CP strongly influences prosodic structure – irrespective of its actual prosodic weight. It further shows that syntactic and prosodic/eurhythmic constraints interact in Catalan and jointly regulate prosodic phrasing. The most common phrasing pattern in simple SVO structures is (S)(VO). However, (SV)(O) is possible when the DP object consists of two or more \(\omega\), i.e. if it is prosodically heavy. And as shown here, in the case of a sentential object, (SV)(O) is again the preferred grouping. This is the case for both prosodically heavy (3 or more \(\omega\)) and prosodically light (1 \(\omega\)) CP objects. The separate grouping of heavy CP objects might be taken as no great surprise: Catalan has a ‘strong tendency to balance the length of the prosodic constituents’ (D’Imperio et al., 2005:71) and a long CP complement is prosodically as heavy as a long DP complement. However, this parallel no longer holds in structures with a light CP object. Considerations of prosodic weight alone would predict a phrasing pattern parallel to that of simple SVO clauses with a prosodically light DP object. However, (SV)(O) remains the most common grouping. Thus, sentential objects with one prosodic word show the same pattern as prosodically heavy complements. This strongly suggests that syntactic structure has a direct effect, independent of prosodic heaviness: A CP object, whether light or heavy, is typically preceded by a prosodic boundary. To account for the phrasing pattern with CP objects, the constraint ALIGN-CP,L is introduced, which induces a boundary immediately preceding the complementizer que ‘that’.

\(^{33}\) For the present study I used the Gradual Learning Algorithm as implemented in PRAAT (Boersma and Weenink, 1992–2011). As described, the goal of the GLA is to ‘locate an empirically appropriate ranking value for every constraint’ (Boersma and Hayes, 2001:51). This is done in six stages, including an initial state, a four-step procedure, and a final state. In the initial state, the constraints have arbitrarily chosen ranking values. Here, I chose the value 100 for each constraint. In the first step, the algorithm is presented with a fully specified input–output pair \(\Phi\alpha\), i.e. it is presented with a learning datum. I presented the underlying form S(VqSVO) and the percentage values of the four groupings (as given in Fig. 8). The second step consists of generating an output \(\alpha'\) for the input \(\alpha\). In the third step, comparison of \(\alpha'\) and \(\alpha\) determines whether there is a mismatch between them. If so, the algorithm has to be adjusted (which is the next step). Otherwise, nothing happens, because the algorithm is error-driven. The fourth and final step in the procedure consists of a moderate adjustment of the ranking values. All constraints favoring \(\Phi\alpha\) over \(\Phi\alpha'\) will be increased by a small predefined numerical amount (called “plasticity”), while the constraints favoring \(\Phi\alpha'\) over \(\Phi\alpha\) will be decreased. The final state includes a repetition of the four-step procedure until the constraint values stabilize. For the present study, I follow the description in Boersma (1999:32). The first exposure consists of 1000 pieces of learning data with a plasticity of 0.1. After that, the GLA is given 10,000 pieces with the same plasticity as before. Finally, I supplied 100,000 new data with a plasticity of 0.001. For a detailed description of the GLA, cf. Boersma (1997, 1998:ch.15, 1999:32f.), Boersma and Hayes (2001:51ff.), Jäger (2003:ch.6) or the GLA homepage <http://www.fon.hum.uva.nl/paul/gla/>.

\(^{34}\) Since the algorithm learns from the beginning with each application, the resulting ranking-value proposals differ with each application (cf. the values given in Feldhausen, 2010:125, which differ from the ones presented here, even though the starting values for each constraint were again set to 100). However, every application of the GLA to this data should lead to a shorter distance between the three lowest constraints, regardless of the exact ranking values proposed.
The results of the complex-SVO experiments have further shown that there is variation in the prosodic grouping of structures with embedded complement clauses. The variation is modeled in the stochastic-OT framework (Boersma and Hayes, 2001) and ALIGN-CP,L, MIN-N-PHRASES, and ALIGN-XP,R are taken to overlap on the continuous ranking scale. The general constraint hierarchy I propose is as follows:

\[
\text{MAX-BIN-END} \gg \text{ALIGN-CP,L} \gg \text{MIN-N-PHRASES} \gg \text{ALIGN-XP,R}
\]

The (SV) phrasing of the matrix clause is not derived from a specific constraint reflecting the length of the object. It is derived from the interaction of three constraints: low-ranked ALIGN-XP,R and high-ranked ALIGN-CP,L and MIN-N-PHRASES.

The present work offers several directions for further research. The experiments here were limited to object clauses in which no word order variation appeared (thus being either qSVO or simply qV). These clauses are typically separated from the preceding matrix clause by a prosodic boundary. However, Feldhausen (2010:ch.5) shows that object clauses with clitic left-dislocations (CLLDs) phrase differently. Embedded dislocations phrase with the preceding matrix clause and thus no boundary occurs between the matrix clause and the complementizer. Below, (26a) depicts the basic result of the present study; (26b) compares embedded CLLDs: a prosodic boundary is located only after the CLLD constituent. Thus one can conclude that the presence of a dislocated element interferes with the tendency to phrase an object clause separately. For future work, it would be interesting to see how the presented OT analysis can integrate these data.

(26) Different prosodic groupings of object clauses with and without CLLD

\[
\begin{align*}
\text{a. ( Matrix ) } & \text{ embedded clause…} \quad \ll \text{ Prosodic Structure} \\
& \quad \ll \text{ Syntactic Structure} \\
\text{b. ( Matrix } & \text{ emb. CLLD } \text{ ( emb. clause…} \quad \ll \text{ Prosodic Structure} \\
& \quad \ll \text{ Syntactic Structure} \\
\end{align*}
\]

Further important direction for future work concerns levels of the prosodic hierarchy and their corresponding edge tones. In the present study the term prosodic phrase is used as a hypernym for intermediate phrase and intonational phrase. The reason for this is the fact that no clear criteria can be established for predicting when an ip or an IntP boundary appears – even though speakers clearly perceive differences in boundary strength. The version of Cat_ToBI in Prieto et al. (2009) deals with this difficulty in another way. Instead of using the term prosodic phrase, only one type of edge tone is assumed, namely T%. This tone is assigned to any boundary, whether it is break-index level 3 or 4. Thus both accounts (Prieto et al., 2009, and the present one) combine the two prosodic levels. Future studies will hopefully shed more light on whether two different prosodic levels do in fact exist and if so, whether they are marked by different sets of phonetic cues.

Acknowledgements

I would like to thank the audiences of the PSI2 workshop (ZAS Berlin, June 2008), the TIE3 conference (Lisbon, September 2008), and the XXIV Encuentro Internacional de la Asociación de Jóvenes Lingüistas (Barcelona, May 2009) for their comments and questions. In particular, I would like to thank Nicole Dehé, Élisabeth Delais-Roussarie, Gorka Elordieta, Caroline Fény, Izarbe García Sánchez, Shinichiro Ishihara, Brechtje Post, Pilar Prieto, Johan Rooryck, Vieri Samek-Lodovici, and Francesc Josep Torres, as well as Ariadna Benet, Susana Cortés, Elena Kireva, and Nina Nanula. My gratitude also goes to Tomke Brüggemann for her statistical support and to David Hill for his assistance with editing. I am especially indebted to two anonymous reviewers for their detailed, helpful and constructive comments and suggestions. It goes without saying that all errors remain my own.

This article is a revised and expanded version of chapter 3 of Feldhausen (2010). It significantly differs from the book chapter by virtue of an additional, new experiment (i.e. the one on CP objects consisting of only one prosodic word), the results of which allow for stronger major conclusions. Next to this, further important changes have been made: e.g., the section on preboundary lengthening has been completely revised and has undergone a statistical evaluation, the application of the Gradual Learning Algorithm is described in greater detail, and certain tables have been modified to allow for a better presentation of the data.

Appendix A

One set of target sentences used in Experiment 1 was given in section 4.1.2 above. The remaining two sets are given here. Parentheses around constituents indicate that these constituents are only inserted in the long S or long O condition.
Appendix B

This appendix provides the remaining sets of target sentences used in Experiment 2.

1. Target sentences of the second set:

La Sílvia no ha mencionat que dormien.
the S. not PST mention that sleep.PAST.3PL

'Silvia did not mention that they were sleeping.'

La Sílvia no ha mencionat que van trucar.
the S. not PST mention that PAST call. INF

'Silvia did not mention that they called.'

La Sílvia no ha mencionat que mentien.
the S. not PST mention that lie. PAST.3PL

'Silvia did not mention that they lied.'

2. Target sentences of the third set:

El pare ha dit que dormien.
the father has say. PTCP that sleep. PAST.3PL

'The father said that they were sleeping.'

El pare ha dit que van trucar.
the father has say. PTCP that PAST call. INF

'The father said that they called.'

El pare ha dit que mentien.
the father has say. PTCP that lie. PAST.3PL

'The father said that they lied.'
Appendix C

Here the durations of the individual la syllable are given. Table A presents the actual values in milliseconds and Table B presents the normalized values. Recall that syllable duration was normalized to sentence duration for each speaker using the formula $d' = d/s$ (i.e. the normalized duration $d'$ of a CV syllable is the proportion of its absolute duration $d$ to the sentence duration $s$). $d'$ was then multiplied by 1000 to avoid decimal places. For example, for speaker CP, sentence 26, $D$ syllable, $d' = 135/4569 = 0.02954$; this value is then multiplied by 1000 ($0.02954 \times 1000 = 29.54$) and rounded up to 30. For the corresponding $\sigma$ syllable the equation is: $156/4569 = 34$.

Table A

<table>
<thead>
<tr>
<th>sent.</th>
<th>AT</th>
<th>CB</th>
<th>CP</th>
<th>DS</th>
<th>GM</th>
<th>GV</th>
<th>IS</th>
<th>MM</th>
<th>MO</th>
<th>RS</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D'$</td>
<td>25</td>
<td>137</td>
<td>110</td>
<td>138</td>
<td>128</td>
<td>115</td>
<td>129</td>
<td>89</td>
<td>118</td>
<td>134</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>173</td>
<td>/</td>
<td>172</td>
<td>198</td>
<td>/</td>
<td>/</td>
<td>178</td>
<td>125</td>
<td>305</td>
<td>/</td>
</tr>
<tr>
<td>diff.</td>
<td>36</td>
<td>34</td>
<td>34</td>
<td>70</td>
<td>/</td>
<td>89</td>
<td>7</td>
<td>171</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>s</td>
<td>3382</td>
<td>3431</td>
<td>3877</td>
<td>3918</td>
<td>3613</td>
<td>3161</td>
<td>3015</td>
<td>3469</td>
<td>3559</td>
<td>3642</td>
</tr>
<tr>
<td>$D'$</td>
<td>26</td>
<td>189</td>
<td>137</td>
<td>135</td>
<td>157</td>
<td>99</td>
<td>116</td>
<td>86</td>
<td>86</td>
<td>97</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>192</td>
<td>/</td>
<td>156</td>
<td>212</td>
<td>141</td>
<td>184</td>
<td>162</td>
<td>115</td>
<td>270</td>
<td>169</td>
</tr>
<tr>
<td>diff.</td>
<td>3</td>
<td>21</td>
<td>55</td>
<td>42</td>
<td>68</td>
<td>76</td>
<td>29</td>
<td>173</td>
<td>19</td>
<td>/</td>
</tr>
<tr>
<td>s</td>
<td>4556</td>
<td>4331</td>
<td>4569</td>
<td>4193</td>
<td>4479</td>
<td>4170</td>
<td>3764</td>
<td>3551</td>
<td>4368</td>
<td>4306</td>
</tr>
<tr>
<td>$D'$</td>
<td>27</td>
<td>125</td>
<td>124</td>
<td>122</td>
<td>106</td>
<td>107</td>
<td>74</td>
<td>93</td>
<td>124</td>
<td>132</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>150</td>
<td>/</td>
<td>201</td>
<td>154</td>
<td>209</td>
<td>170</td>
<td>/</td>
<td>/</td>
<td>278</td>
<td>149</td>
</tr>
<tr>
<td>diff.</td>
<td>25</td>
<td>79</td>
<td>28</td>
<td>103</td>
<td>63</td>
<td>154</td>
<td>17</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>s</td>
<td>3535</td>
<td>3375</td>
<td>4281</td>
<td>3965</td>
<td>4176</td>
<td>3903</td>
<td>3304</td>
<td>3571</td>
<td>3638</td>
<td>3982</td>
</tr>
<tr>
<td>$D'$</td>
<td>28</td>
<td>137</td>
<td>87</td>
<td>132</td>
<td>145</td>
<td>124</td>
<td>130</td>
<td>116</td>
<td>64</td>
<td>123</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>165</td>
<td>145</td>
<td>163</td>
<td>170</td>
<td>174</td>
<td>163</td>
<td>120</td>
<td>102</td>
<td>238</td>
<td>146</td>
</tr>
<tr>
<td>diff.</td>
<td>28</td>
<td>58</td>
<td>31</td>
<td>25</td>
<td>50</td>
<td>33</td>
<td>4</td>
<td>38</td>
<td>115</td>
<td>16</td>
</tr>
<tr>
<td>s</td>
<td>4102</td>
<td>3881</td>
<td>4836</td>
<td>4469</td>
<td>4611</td>
<td>4174</td>
<td>3858</td>
<td>3814</td>
<td>4153</td>
<td>4396</td>
</tr>
<tr>
<td>$D'$</td>
<td>33</td>
<td>102</td>
<td>90</td>
<td>101</td>
<td>129</td>
<td>111</td>
<td>123</td>
<td>109</td>
<td>121</td>
<td>134</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>160</td>
<td>/</td>
<td>126</td>
<td>/</td>
<td>209</td>
<td>144</td>
<td>126</td>
<td>112</td>
<td>134</td>
<td>165</td>
</tr>
<tr>
<td>diff.</td>
<td>58</td>
<td>25</td>
<td>80</td>
<td>33</td>
<td>3</td>
<td>13</td>
<td>31</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>s</td>
<td>2381</td>
<td>2142</td>
<td>2475</td>
<td>2725</td>
<td>2945</td>
<td>2687</td>
<td>2196</td>
<td>2227</td>
<td>2512</td>
<td>2508</td>
</tr>
<tr>
<td>$D'$</td>
<td>34</td>
<td>116</td>
<td>124</td>
<td>141</td>
<td>132</td>
<td>128</td>
<td>89</td>
<td>109</td>
<td>98</td>
<td>112</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>221</td>
<td>131</td>
<td>229</td>
<td>263</td>
<td>/</td>
<td>159</td>
<td>166</td>
<td>/</td>
<td>221</td>
<td>174</td>
</tr>
<tr>
<td>diff.</td>
<td>105</td>
<td>7</td>
<td>88</td>
<td>131</td>
<td>70</td>
<td>57</td>
<td>109</td>
<td>42</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>s</td>
<td>3300</td>
<td>3007</td>
<td>3449</td>
<td>3458</td>
<td>3373</td>
<td>3170</td>
<td>2975</td>
<td>2770</td>
<td>3363</td>
<td>3305</td>
</tr>
<tr>
<td>$D'$</td>
<td>35</td>
<td>115</td>
<td>100</td>
<td>98</td>
<td>107</td>
<td>119</td>
<td>101</td>
<td>103</td>
<td>134</td>
<td>137</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>202</td>
<td>155</td>
<td>222</td>
<td>214</td>
<td>164</td>
<td>314</td>
<td>109</td>
<td>128</td>
<td>305</td>
<td>174</td>
</tr>
<tr>
<td>diff.</td>
<td>87</td>
<td>55</td>
<td>124</td>
<td>119</td>
<td>57</td>
<td>195</td>
<td>8</td>
<td>25</td>
<td>171</td>
<td>37</td>
</tr>
<tr>
<td>s</td>
<td>2733</td>
<td>2685</td>
<td>3067</td>
<td>2755</td>
<td>2878</td>
<td>2832</td>
<td>2420</td>
<td>2592</td>
<td>3236</td>
<td>2862</td>
</tr>
<tr>
<td>$D'$</td>
<td>36</td>
<td>79</td>
<td>130</td>
<td>152</td>
<td>154</td>
<td>157</td>
<td>125</td>
<td>119</td>
<td>97</td>
<td>108</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>/</td>
<td>163</td>
<td>192</td>
<td>274</td>
<td>204</td>
<td>149</td>
<td>170</td>
<td>261</td>
<td>182</td>
<td>/</td>
</tr>
<tr>
<td>diff.</td>
<td>3410</td>
<td>3083</td>
<td>3647</td>
<td>3898</td>
<td>4105</td>
<td>3443</td>
<td>3078</td>
<td>3199</td>
<td>3833</td>
<td>3905</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>all</td>
<td>125</td>
<td>113</td>
<td>127</td>
<td>129</td>
<td>121</td>
<td>116</td>
<td>102</td>
<td>96</td>
<td>119</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>all</td>
<td>180</td>
<td>149</td>
<td>183</td>
<td>212</td>
<td>184</td>
<td>183</td>
<td>147</td>
<td>141</td>
<td>242</td>
</tr>
</tbody>
</table>

Table B

Normalized duration of la syllables.

<table>
<thead>
<tr>
<th>sent.</th>
<th>AT</th>
<th>CB</th>
<th>CP</th>
<th>DS</th>
<th>GM</th>
<th>GV</th>
<th>IS</th>
<th>MM</th>
<th>MO</th>
<th>RS</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D'$</td>
<td>25</td>
<td>41</td>
<td>32</td>
<td>36</td>
<td>33</td>
<td>32</td>
<td>41</td>
<td>30</td>
<td>34</td>
<td>38</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>51</td>
<td>/</td>
<td>44</td>
<td>51</td>
<td>/</td>
<td>/</td>
<td>59</td>
<td>36</td>
<td>86</td>
<td>/</td>
</tr>
<tr>
<td>diff.</td>
<td>10</td>
<td>8</td>
<td>18</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>29</td>
<td>2</td>
<td>4</td>
<td>/</td>
</tr>
<tr>
<td>$D'$</td>
<td>26</td>
<td>41</td>
<td>32</td>
<td>30</td>
<td>37</td>
<td>22</td>
<td>28</td>
<td>23</td>
<td>24</td>
<td>35</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>42</td>
<td>/</td>
<td>34</td>
<td>51</td>
<td>31</td>
<td>44</td>
<td>43</td>
<td>62</td>
<td>26</td>
<td>39</td>
</tr>
<tr>
<td>diff.</td>
<td>1</td>
<td>4</td>
<td>14</td>
<td>9</td>
<td>16</td>
<td>20</td>
<td>38</td>
<td>4</td>
<td>4</td>
<td>/</td>
</tr>
<tr>
<td>$D'$</td>
<td>27</td>
<td>35</td>
<td>37</td>
<td>28</td>
<td>32</td>
<td>25</td>
<td>27</td>
<td>22</td>
<td>26</td>
<td>34</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>42</td>
<td>/</td>
<td>47</td>
<td>39</td>
<td>50</td>
<td>44</td>
<td>/</td>
<td>/</td>
<td>76</td>
<td>37</td>
</tr>
<tr>
<td>diff.</td>
<td>7</td>
<td>19</td>
<td>7</td>
<td>25</td>
<td>17</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>42</td>
<td>4</td>
</tr>
</tbody>
</table>