

VARIATION IN THE INTONATION OF EXTRA-SENTENTIAL ELEMENTS*

LLUÏSA ASTRUC-AGUILERA & FRANCIS NOLAN
University of Cambridge

Abstract

Extra-sentential elements usually form independent intonational domains. Due to this property, they have been used in the phonological literature to define the intonational phrase. This study, using English and Catalan empirical data collected in three experiments, shows significant variation in the phrasing and accentuation of these constructions. We argue that their role is primarily semantic: they are supplementary, semantically non-restrictive, and anaphorically linked to their referent. Prosody signals their grammatical function by means of independent phrasing, by reductions in pitch span leading to total deaccentuation, and/or by tonal reduplication. We argue that both tonal and junctural cues are used in combination to mark extra-sentential elements as external to the phrase.

1. Introduction

The group of so-called extra-sentential elements (ESEs) includes phrases, such as dislocated phrases (*They are crazy, those Romans*), and words, such as vocatives (*Thanks, sir*), and sentential adverbs (*obviously*). Extra-sentential elements have been used in the phonological literature to define the intonational phrase, the level of the prosodic hierarchy that is most relevant to intonation. Nespor and Vogel claim that there are certain types of constructions such as “parenthetical expressions, non-restrictive relative clauses, tag questions, vocatives, expletives, and certain moved elements” that usually form independent intonation domains (Nespor & Vogel 1986:188). The traditional assumption is that ESEs are syntactically independent, just as they are also prosodically independent (e.g. Pierrehumbert 1980; Nespor & Vogel 1986; Nespor 1993), and thus they have been used as evidence of the effect of syntactic constraints upon phrasing (Cooper & Paccia-Cooper 1980; Nespor & Vogel 1986; among others). Recent work has refined this account by proposing that the boundaries of extra-sentential elements are compulsorily aligned with

* Many thanks to Eva Estebas-Vilaplana and to an anonymous reviewer for their very useful comments on an earlier version of this article. All the errors that remain are ours. We wish to thank our informants for kindly donating their time. The research reported here was partly supported by a scholarship *Batista i Roca* (*Generalitat de Catalunya*).

those of the intonational phrase and by developing specific mechanisms to generate this alignment (Selkirk 1984; Truckendbrodt 1995, 1999). One possible problem is that extra-sentential elements show syntactic, phonetic, and perhaps also phonological variation that has not yet been studied in depth.

Syntactically, ESEs do not fall into a single class. Appositions are syntactically governed because they are attached to the NP they modify while parentheses and vocatives are commonly analysed as being external to the syntactic structure. The status of non-restrictive relatives is subject to debate: for Emonds (1979), they are attached to S, the root sentence; for Safir (1986), they are attached at the level of Logical Form; for Fabb (1990), they belong to the level of discourse structure; and for Kempson, Meyer-Viol & Dov Gabbay (2000), they are syntactic constituents. The status of other categories such as epithets (*He wouldn't lend me his car, the bastard*) has not been addressed in the literature. Consequently, the first difficulty in the study of ESEs arises when deciding which elements belong to this class. Drawing up a preliminary list of ESEs is a necessary first step in the investigation, even if this might not be a comprehensive list (which has never been attempted in previous studies, as far as we know). The first research question would be: which elements belong to this class and what should be the criteria for membership?

Another problem is that ESEs show puzzling differences between elements which are initial in the sentence and elements which are not. Initial elements receive a normal intonation, while non-initial elements receive an intonation which is tonally subordinated to that of the main phrase. This tonal subordination is manifested either by reductions in prominence, leading even to total deaccentuation (Bonet 1984:31-32, 90; Ladd 1980:163-165; Liberman 1975:182-184; Gussenhoven 1985, 1993, 2004:290-295; Beckman & Pierrehumbert 1986:293-298; Nespor 1993:265; Prieto 2002a:409ff) and/or by tonal reduplication (Gussenhoven 1985:107, 2004:315-316; Bonet 1984:30, 34; Recasens 1993:214; Prieto 2002a:428-430). It is not clear, however, why the same element in the same phrase should be pronounced in a different way depending on whether it appears in first position or not, as it is not clear whether the different mechanisms used for signalling tonal subordination also have different functions. Tonal subordination of this kind, often right across a prosodic boundary, constitutes a challenge for intonational theory (Ladd 1996:246).

A further problem, as has been observed in the literature and also in the results of a previous study (Astruc 2003), is the evidence of prosodic variation among ESEs. However, divergences in the literature can also be attributed to underlying theoretical discrepancies. Therefore, the main goal of this study is to carry out a detailed quantitative study of the phrasing and accentuation of the main categories of ESEs, thus answering the research questions: are ESEs accented? Do they always form independent intonational phrases? Is there real variation in the prosodic form of ESEs, as suggested by descriptions in the literature?

Variation in the prosodic form of these categories, within or across the languages in the study, would point to the inadequacy of the usually held view that the prosodic form of ESEs is determined by their syntactic form. The apparent lack of prosodic homogeneity casts doubt on the commonly assumed view that ESEs should compulsorily form independent phrases, since this prosodic property is taken to follow from their assumed syntactic independence. The reported asymmetry between initial and non-initial elements casts further doubts on this.

The structure of this article is as follows: in Section 1.1 we examine the core categories of ESEs, discuss the criteria for membership, and review previous studies. In Section 2 we present the overall objectives of the study, and we describe the main phonological characteristics of ESEs in English and in Catalan. In Section 3 we describe two experiments involving right-dislocated phrases in Catalan. In Section 4 we summarize the results and draw some conclusions.

1.1 *Members of the class of ESEs*

The decision adopted in this study is to take as the core members of the class of ESEs those categories identified on the basis of grammatical (Huddleston & Pullum 2002) and prosodic (see Nespor & Vogel 1986:188) criteria. The intersection of the two sets of categories yields the categories in (1)¹:

- (1) Dislocated phrases, sentential adverbs, non-restrictive relative clauses, appositions, parentheses, epithets, quotations, vocatives.

The starting assumption is that ESEs do not form a homogenous class from a grammatical point of view. Dislocated phrases and sentential adverbs are clearly syntactic constituents. Non-restrictive relatives, parentheses, sentential adverbs, appositions, and epithets are semantically governed, since they are semantically linked to the element they modify, while direct speech markers and vocatives are pragmatically governed. The lack of grammatical homogeneity of ESEs compromises the commonly assumed view that they form a single category. Two alternative explanations can be considered: one, that there is no motivation to treat ESEs as a single category; and two, that the core property that defines the set resides somewhere else, namely, in their structural role. In this paper we will argue for the second explanation and we will cite as evidence Catalan and English data.

¹ Question tags and interjections appear in both lists, but they were not included in the study, because the former are only found in English and the latter are susceptible to much paralinguistic variation.

1.2 *ESEs as supplements*

In Huddleston and Pullum (2002) ESEs have the role of ‘supplements’, since they add information that supplements but does not restrict or modify the propositional content of the main clause. They are external to the syntactic structure, as illustrated by this example:

- (2) Pat, the life and the soul of the party, had invited all the neighbours.

If *the life and the soul of the party* is removed, neither the grammaticality nor the truth conditions of the sentence are affected. Therefore, the parenthetical clause is optional: it has the function of a supplement which is semantically related to the clause with which it co-occurs, the ‘anchor’. Owing to their lack of integration into the syntactic structure, supplements are semantically non-restrictive. As seen in (2), the supplement *the life and soul of the party* does not serve to distinguish one referent from another: it does not restrict the denotation of *Pat*, the head nominal. With the possible exception of sentential adverbs², the syntactic peripherality of ESEs neatly corresponds to their function as semantic supplements, in that ESEs show a clear correspondence between grammatical form and grammatical function. It remains to be seen whether there is also a correspondence between grammatical form and prosodic form. In the next section, we are going to examine what has been said about the phrasing and intonation of ESEs in previous studies.

1.3 *Previous intonational studies*

ESEs have attracted ample attention in English (Lieberman 1975; Pierrehumbert 1980; Beckman & Pierrehumbert 1986; Pierrehumbert & Hirschberg 1994; Ladd 1980, 1996; Bing 1985; Gussenhoven 1985), in Catalan (Bonet 1984; Prieto 2002a, 2002b; Emonds 1979; Recasens 1993; Payà 2002, 2003), and also in other languages such as French (Wunderli 1987; Martin 1987; Delattre 1972; Fagyal 2002a, 2002b), Italian (Nespor 1993; Nespor & Vogel 1986), Spanish (Zubizarreta 1998).

There is widespread agreement as regards the phrasing of ESEs into independent units. Most studies describe ESEs as separated by audible prosodic breaks, be these tonal boundaries or pauses. There are different views about which prosodic unit ESEs belong to. Proposals range from analysing ESEs as ‘enclitic’ phrases, which are both separated from, and tonally attached to, the main phrase (Trim 1959; Lieberman 1975; Pierrehumbert 1980; Gussenhoven 2004), to proposals treating them as intermediate phrases (Beckman & Pierrehumbert 1986), and to proposals describing them as full-

² Certain types of sentential adverb (modal adverbs such as *possibly*) do restrict and/or modify the truth conditions of the proposition (Bing 1985; Allerton & Cruttenden 1974). Sentential adverbs, therefore, appear to have functional characteristics that set them apart from the other ESEs (see Astruc 2005; Astruc & Nolan in press for a more detailed description).

fledged intonational phrases (Ladd 1996; Zubizarreta 1998). However, there is much less agreement in the literature regarding their accentuation. Certain categories such as epithets and direct speech markers (Ladd 1980:164-165³) tend to appear as deaccented. Categories generally described as accented are appositions and non-restrictive relatives (with the main exceptions of Pierrehumbert & Hirschberg 1990; Pierrehumbert 1980; Cruttenden 1997). Vocatives are described as deaccented in English (Lieberman 1975; Ladd 1980, 1996; Beckman & Pierrehumbert 1986; Bing 1985; Pierrehumbert & Hirschberg 1990) but as accented in Catalan (Bonet 1984:28-29, 60-61, 88; Recasens 1993:211, 214; Prieto 2002a:428-429). Cross-linguistic variation in the frequency of accentuation of certain categories should not come as a surprise since it is well known that English uses accentuation to a greater extent than Catalan does. Intra-linguistic variation, however, can prove difficult to accommodate in a model that treats ESEs as a single class, such as the model aimed at in this study, unless it can be proved that divergences in the literature have a theoretical basis. This is not unlikely, since, by their very nature, ESEs provide abundant examples of mismatches between phrasing and intonation that are not easily accommodated within the framework assumed in this study (the Autosegmental Metrical approach, henceforth AM), which does not admit the existence of intonational units that are both independent and deaccented.

2. *Experiment 1: comparative production study*

2.1 *Goals and methodology*

The main goal of this study is to analyse the phrasing and intonation of the different elements in order to elucidate whether the categories traditionally considered as ESEs, and which are hypothesised to form a single pragmatic category, also show a consistent prosodic behaviour. Narrowing down the main research question, we can ask: are ESEs accented? Do they always form independent intonational phrases? Is there real variation in the prosodic form of ESEs, as seems to emerge from descriptions in the literature? Is it true that there is an asymmetry between the elements of the left periphery and those of the right periphery? In order to address these research questions, the experimental design has covered three studies: a general comparative study of the phrasing and intonation of the core types of ESEs in English and Catalan, two quantitative studies to find evidence that some ESEs are completely deaccented, and a further quantitative study (which is not reported here, for reasons of space) to test the hypothesis that English and Catalan use different strategies for marking ESEs prosodically.

³ But see Ladd 1996:219-221, where direct speech markers are described as underspecified regarding accentuation: they can receive either pitch accents or phrase tones.

A corpus of 605 phrases was gathered⁴ containing the target ESEs listed in Example (1), and which included the main categories of ESEs usually discussed in syntactic and phonological studies. The English part consisted of 462 items: twenty-nine sentences each read twice by eight Southern British English speakers. The Catalan part consisted of 143 items: fifty sentences, each read once by three Central Catalan speakers. Recordings were conducted in a sound-proof studio, using a DAT recorder. In both corpora, some tokens had to be discarded because of reading errors. Examples of the target ESEs are shown in (3):

- (3) a. Dislocated phrases: *Ella només feia que pensar en sa mare, la bona de la Norma* (“She was thinking all the time about her mother, good Norma”).
- b. Sentential adverbs: *No saben comportar-se, honradament* (“They can’t behave, honestly”).
- c. Non-restrictive relative clauses: *Una de les noies, que es diu Norma, es va posar malalta* (“One of the girls, whose name is Norma, got ill”).
- d. Appositions: *Molt millor que es quedin amb son pare, el Norman*. (“They’d better stay with their dad, Norman”).
- e. Parentheses: *I es va gastar els diners, amb gran alegria, fent un viatge a una illa tropical* (“And she spent the money, with great joy, by taking a trip to a tropical island”).
- f. Epithets: *Acabo de veure el meu ex, el cabró* (“I’ve just seen my ex, the bastard”).
- g. Quotation markers: *Com va anar el viatge?—l’Alma demana a la Mariana* (“‘How was the trip?’ Alma asks Mariana”).
- h. Vocatives: *L’Anna va guanyar-la, Manu* (“Anna won it, Manu”).

2.2 Results

The data was digitised at 16 kHz and subsequently analysed by the first author by listening to the recording and looking at pitch traces obtained with Praat 4.1.21. The analysis comprised identifying and labelling the type and location of pitch accents and prosodic breaks. In this study, phrasing and accentuation are considered to be independent systems, as proposed by Trim (1959) and Ladd (1980:164), among others. Phrasing is taken to be based on junctural cues without tonal movement being obligatory. The criteria for deciding that a prosodic boundary was present were the presence of junctural cues, such as pauses and/or pre-boundary lengthening. The criteria for deciding that a stretch of speech was deaccented were the presence of flat F0 and reduced amplitude. A pause was defined here as a period of silence not caused

⁴ The Catalan data was collected in three stages: a pilot experiment, and two further studies. The English data was recorded later, together with a study on sentential adverbs in which ESEs were used as distractors (see Astruc 2005; Astruc & Nolan in press). The imbalance in the number of examples in the two corpora does not affect the results of the phonological analysis.

by the presence of obstruents of 100 ms or longer. Anything shorter than 100 ms was considered pre-boundary lengthening. We used a stylised representation for pitch accents, combined with a numbered coding system (for both accents and prosodic breaks) that was later used for entering the data in SPSS. The analyses were repeated two further times, at intervals, without significant inconsistencies emerging. A subset of the data was checked by the second author and the analyses mostly confirmed.

The following generalizations can be drawn about the main prosodic properties of ESEs. First, ESEs tend to be prosodically independent of the main phrase, except for appositions and vocatives. In these two cases, prosodic variation seems to correlate with a dual communicative function (see Section 2.2.3). Second, there is a prosodic difference between initial and non-initial ESEs. In initial position, ESEs are both rhythmically and intonationally independent (again, with the exception of vocatives and appositions), while in medial and final position they are tonally subordinated to the main phrase. Third, this tonal subordination is carried out by means of two principal strategies:

- (4) a. by reductions in pitch span leading to total deaccentuation.
- b. by reduplicating the contour of the main phrase, which can be accompanied by the use of an overall lower pitch level and a much lower voice volume.

Dislocated phrases, quotations markers, and epithets follow strategy (4a); that is, they show deaccentuation. Parentheses, non-restrictive relatives, and appositions follow strategy (4b): they show tonal reduplication at a lower level and with a compressed pitch range. Sentential adverbs and vocatives do not fit clearly into either (4a) or (4b).

2.2.1 Dislocated phrases, quotations markers, and epithets. In English, dislocated phrases form independent phrases 88% of the time, both in initial and final position (see Figure 1). In final position they are always deaccented (100% of the time) and are usually followed by a final rise (70% of the time). In the Catalan corpus, they always form independent intonational phrases (see Catalan examples in Section 3), and they are nearly always deaccented (95%). Unlike in English, they are not followed by a final rise.

Direct speech markers, in both English and Catalan, are nearly always deaccented (86%) and they always form independent phrases (100%), which are usually set off by pauses (65% of the time) (see Figure 2).

In English, epithets always form independent phrases, are nearly always deaccented (86%) and are frequently followed by a rise (68%). This final pitch movement is assumed to be not an accent but a boundary tone. Figure 3 shows the intonation that typically corresponds to English epithets.

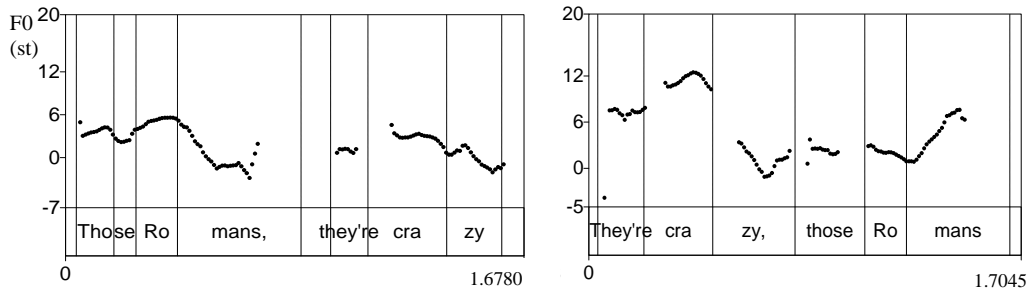


Figure 1: In the left panel, an example of a left-dislocated subject in English. In the right panel, an example of a right-dislocated subject (*st* = semitones).

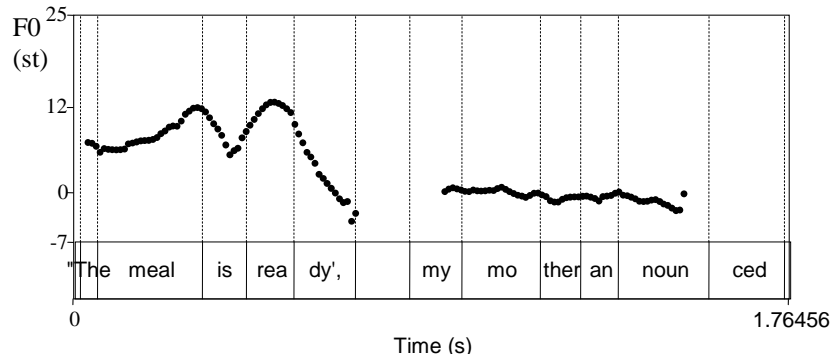


Figure 2: Example of a direct speech marker in English.

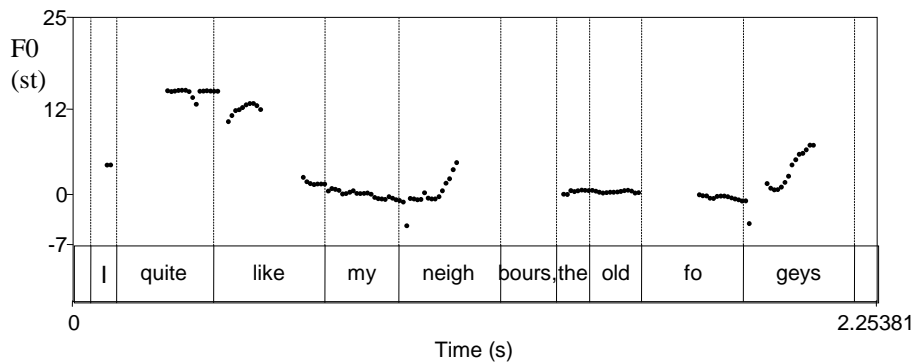


Figure 3: Example of epithet in English.

Unlike in English, epithets in Catalan⁵ may be accented (50% of the time) and they never carry a final rise. The epithets in the corpus are at times produced in such a way that they form part of the same intonation unit as the main phrase (50%), and at other times produced so that they form a separate intonation unit and are accented (although with a very reduced pitch range). Even in those cases where the speaker added extra emphasis, the emphatic

⁵ Epithets, however, can carry an emotional load, in which case they are accented, irrespective of language. In the data, the few occurrences of accenting correspond to the same phrase, *I've just caught a glimpse of my ex, the bastard* and its interrogative counterpart *Have you seen my ex, the bastard?*

accent had a pitch range subordinated to that of the main phrase, as can be seen in the pitch trace in Figure 4.

Parallel lines have been placed by hand to facilitate the comparison of pitch range and level in the main clause and epithet. From the pitch trace it is easy to appreciate that the epithet is pronounced at a lower level and with a narrower pitch range than the main phrase, and the speech wave indicates that it also has less intensity.

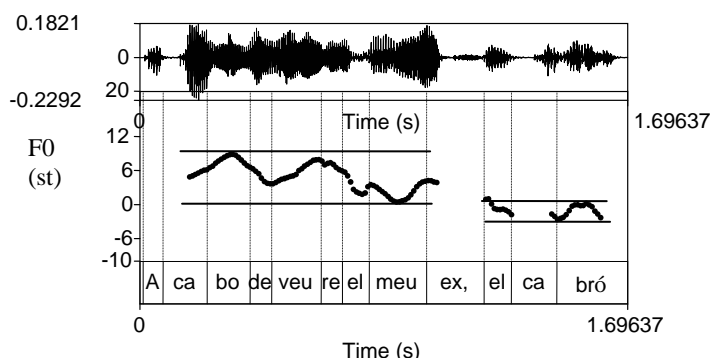


Figure 4: Example of epithet in Catalan: *Acabo de veure el meu ex, el cabró* (“I’ve just seen my ex, the bastard”).

2.2.2 *Parentheses, non-restrictive relatives, appositions, and vocatives.* Parentheses, non-restrictive relatives, and appositions follow strategy (5b): they show tonal reduplication at a lower level and with a compressed pitch range. Parentheses in English usually form independent phrases (97%, with pauses occurring 60% of the time), and are mostly accented (83%), often with a reduplicating contour (53% of the time). In Catalan, parentheses always form independent units which are mostly demarcated by pauses (58%) (see Figure 5). This is in line with descriptions in previous work (Payà 2002, 2003; Prieto 2002a, 2002b), just as also is the fact that they are mostly accented, with a low register and compressed pitch range.

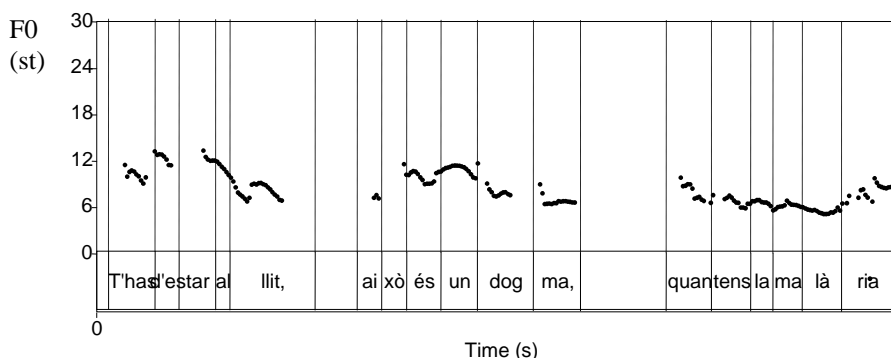


Figure 5: Example of parenthesis in Catalan: *T’has d’estar al llit, això és un dogma, quan tens la malària* (“You have to stay in bed, that’s a rule, when you have malaria”).

One of the first things that is observed about non-restrictive relative clauses is that they show more cohesion with the phrase to their left than with that to their right. Both in English and Catalan, the non-restrictive relative clause reduplicates the tonal pattern of the left-hand clause (60% of the time in Catalan, 53% in English) (see Figure 6). They are mostly prosodically detached (88% in English, 100% in Catalan) by phrase tones or pauses, and in those cases when only one pause is present, the pause occurs between the non-restrictive relative clause and the right-hand clause. This behaviour is not surprising since non-restrictive relative clauses are closely grammatically linked to their immediately preceding clause, their ‘anchor’, which contains their referent and with which they agree in number, and, in Catalan, also gender.

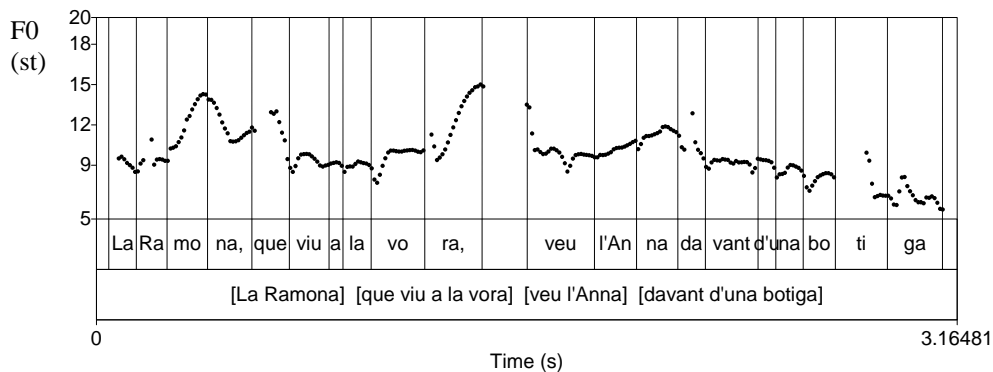


Figure 6: Example of a non-restrictive relative clause in Catalan: La Ramona, que viu a la vora, veu l'Anna davant d'una botiga (“*Ramona, who lives nearby, sees Anna in front of a shop*”).

Similar to the case of non-restrictive relative clauses, appositions, both in English (88% of the time) and in Catalan (83% of the time) tend to form independent intonational phrases that reduplicate the contour of their anchor (Figure 7)⁶.

Vocatives show a different behaviour in English and Catalan. In English they are mostly deaccented, while in Catalan they are accented 50% of the time (see Figure 8). As noted by Gussenhoven (1985) and Cruttenden (1997), vocatives have either an attention-catching function (in which case they are accented and receive an intonational contour of their own), or an expressive function (in which case they are deaccented and belong to the same intonational unit as the main phrase).

⁶ Appositions that provide identifying information about the referent are interpreted as ‘appositive modifiers’ and do not receive independent contours. For instance, *Norman* in *This is my husband Norman* would be interpreted as an ‘appositive modifier’, that is, a way of identifying this person, whereas in *This is my husband, Norman* would be interpreted as a ‘supplement’, that is, as additional information about the referent. (Huddleston and Pullum 2002:1064, 447). This distinction would correspond to that proposed by Gussenhoven (2004:290-292) between ‘incorporating’ and ‘enclitized’ ESEs.

Sentential adverbs, and adverbs in general, are more heterogeneous prosodically than any of the other categories, just as they are also more heterogeneous semantically than any of the other parts of speech. The examples in the corpus presented variation in their accentuation, but they were very consistent as regards to their phrasing: they nearly always formed independent sentences. They behave, thus, according to what is predicted in the literature about ESEs in general. Their detailed description, however, exceeds the scope of this article (see Astruc 2005; Astruc & Nolan in press).

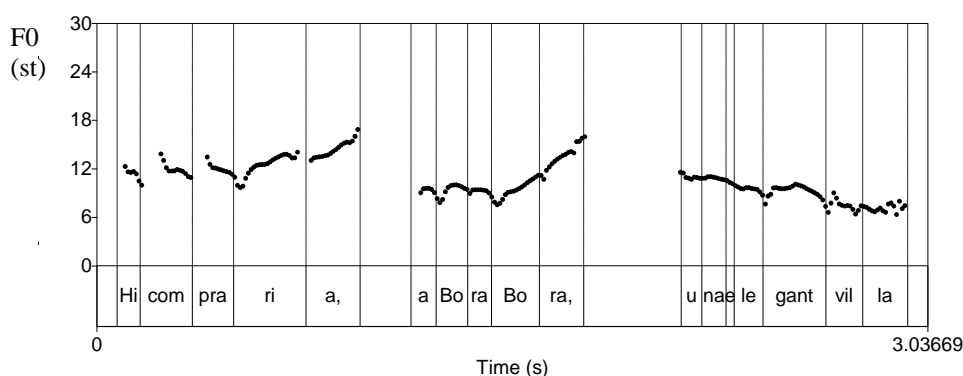


Figure 7: Example of an apposition in Catalan: *Hi compraria, a Bora Bora, una elegant vil·la* (“I would buy an elegant villa there, in Bora Bora”).

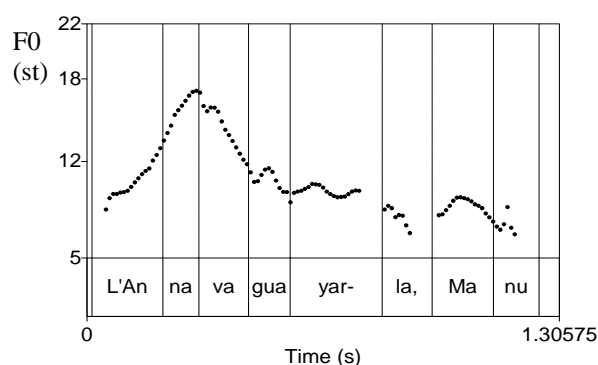


Figure 8: Example of a Catalan vocative in final position: *L'Anna va guanyar-la, Manu* (“Anna won it, Manu”).

3. Experiments 2 and 3: Right-dislocated phrases in Catalan

3.1 Introduction

Right-dislocated phrases are clauses such as *those girls* in *They are nice, those girls*. Their main characteristic is the presence of a co-referential element within the main phrase (in this case, the pronoun *they*) which is linked to the dislocated NP (*those girls*). The main function of the right-dislocated phrase is that of introducing background information in a position where a high informative content would normally be expected (Huddleston & Pullum 2002; Geluykens 1992, 1994; Lambrecht 1981, 1994). In languages such as Catalan, which mainly use syntactic changes to signal focus, right- and left-dislocations

serve the main function of removing background information out of the main clause, so that the focal accent coincides with the last pitch accent of the main phrase (Vallduví 1990, 1994, and elsewhere).

Previous work on Catalan intonation has made no strong claims about the accentual status of right-dislocated phrases. It is implied that they have an accentual pattern of their own, though compressed and subordinated to that of the main phrase (Prieto 2002a; Recasens 1993; Payà 2003). However, they have also been described as lacking prominence (Bonet 1984:31-32, 90). Therefore, empirical testing is needed in order to support the claim that right-dislocations are really deaccented. It was noted in the data from Experiment 1 that there were several instances of ‘miniature accents’ following the focal accent; little ‘bumps’ with an excursion size of about 1 semitone. Post-focal accents have been identified so far in narrow focus sentences in Catalan (Estebas-Vilaplana 2000), in Spanish (Zubizarreta 1998), and in Italian (Grice 1995; D’Imperio 2002). It is doubtful, though, whether these ‘bumps’ in the present data correspond to a reduced and subordinated pitch accent, or whether they are a mere side-effect of the higher subglottal pressure associated with a stressed syllable.

The main goal of the two experiments reported here is, therefore, to provide quantitative evidence that right-dislocations are indeed deaccented in Catalan, so that this analysis can be extended to the English data and to the other structures similarly described as deaccented in final position, that is, epithets, reported speech markers, parentheses, and some sentential adverbs. Experiments 2 and 3 were designed as separate experiments, although they were recorded in a single session, and their aim was to quantify the scaling of the target syllable in the dislocated element, in order to assess whether it receives a true pitch accent or not.

3.2 *Experiment 2*

3.2.1 *Experimental material.* The corpus consisted of fourteen sentences each read twice by four speakers under four experimental noise conditions, thus yielding 448 sentences in total, of which 224 contained appositions and 224 right-dislocations. The fourteen sentences thus consisted of seven minimal pairs (see Appendix), each containing an apposition and a dislocation, with both elements of each pair introduced by a short background text:

- (5) a. Apposition: —*La mama va veure els nuvis, abans de la boda?* (“Did Mum see the bride and the groom before the wedding?”)
 —*Va veure la Núria, la núvia* (“She saw Núria, the bride”)
 b. Right-dislocated subject: —*Qui és l’amiga que va anar a veure la núvia?* (“Who’s the friend that the bride went to see?”)
 —*Va veure la Núria, la núvia* (“She saw Núria, the bride”)

Two sets of comparisons were planned: (1) comparison of the stressed syllables in the apposition and dislocation, that is the initial syllable *nú* in *Nú-*

ria in (5a) and (5b); (2) comparison of the most prominent syllables (the penultimate ones) in the main phrase and dislocation, that is, the comparison of *Nú-ria* with *nú-via*⁷. As can be seen, the pragmatic context is more rigorously controlled than in the comparative study described in Section 2, though both aimed at being as near to natural sounding speech as possible. The seven minimal pairs were repeated twice and mixed in random order with fillers and distractors. Among the fillers were the phrases used in Experiment 3, which will be described in Section 3.3.

3.2.2 *Methodology*. Masking noise was used to elicit an increase in voice volume which in turn induced an increase in pitch (the so-called Lombard effect: see Lane & Tranel 1971; Junqua 1996. See detailed description of the procedure in Astruc 2005, Chapter 3). It was expected that there would be an increase in the scaling of potential pitch accents. However, preliminary results showed a general raising in pitch level instead of the expected local raise in pitch. This prompted a change in methodology in Experiment 2.

The first author carried out a phonological analysis, quantifying the cases of prosodic separation and the types of prosodic breaks used, as well as the instances of accentuation and deaccentuation. This was done by carefully listening to the recordings and by looking at the pitch traces obtained with Praat (4.1.21) following the criteria described in Section 2.2. This analysis was repeated twice, with an interval of a few weeks separating each analysis, and without having the annotations of the previous analysis at hand, and a portion of the data was checked by the second author, without finding substantial divergences.

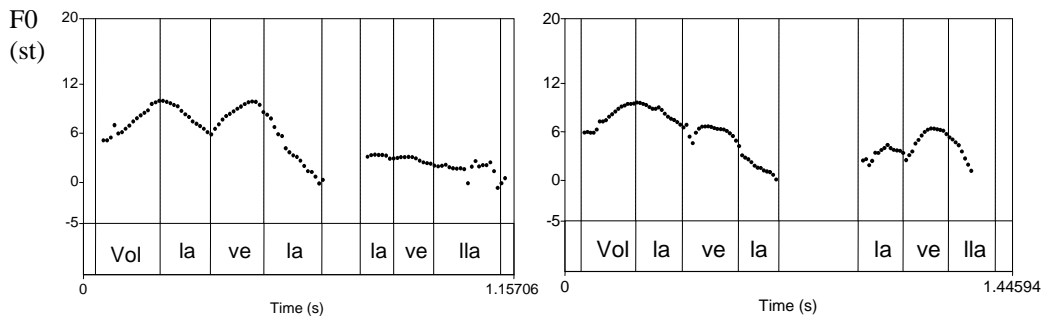


Figure 9: A minimal pair in Catalan. Left-hand panel, an example of right-dislocation. Right-hand panel, the apposition counterpart.

Figure 9 presents in the left-hand panel the phrase *Vol la vela, la vella* (“She wants the sail, the old lady [does]”), with a right-dislocated NP subject. This interpretation was elicited by a sentence about Mary, an old lady who has

⁷ The planned comparisons were not carried out in the end because a better method for testing the accentuation of the right-dislocations was devised and applied in the following experiment, as reported in Section 3.3.

a ship in a bottle that is missing a sail, followed by the question *What does the old lady want?* In the right-hand panel, we see the appositive interpretation of the same segmental string, *Vol la vela, la vella*. Subjects were presented with the context of a fisherman fixing his boat and working on the sail, which was old. To the question *What does the fisherman want?*, the answer was: *[He] wants the sail, the old one*. Both appositions and dislocations are set off by prosodic boundaries, that is, by lengthening, tonal movements, and/or pauses. The main difference between them is that appositions receive a contour that reduplicates the contour of the main phrase at a lower voice level. Such reduplication is observed in 60% of the data.

3.2.3 Right-dislocations: Phrasing and intonation. In this section we will present an overview of the phrasing and intonation of the 224 right-dislocated phrases, which constitute half of the corpus. The analysis was performed by carefully listening to the recordings and examining the pitch traces. Following the criteria described in Sections 2.2 and 3.2.2, it was decided that a right-dislocation was deaccented if it sounded less loud than the main phrase and without any perceivable pitch movement. (A more detailed quantification of the scaling of the stressed syllables of the dislocated phrases will be undertaken in the experiment reported in Section 3.3, which uses a different methodological approach). The criterion for deciding whether the main phrase and the right-dislocated NP formed independent units or not was the presence of lengthening, tonal movements, and/or pauses. When the constituent started with a vowel, there was also creakiness at the end of prosodic constituents and glottalization at the beginning. If any of these indications of a prosodic break was found, it was decided that they formed independent phrases. That is, ‘independent units’ includes both intermediate and intonational phrases. According to this criterion, right-dislocations formed independent units 70% of the time. This percentage includes a sizeable degree of inter-speaker variation, as can be seen in the bar graph presented in Figure 10.

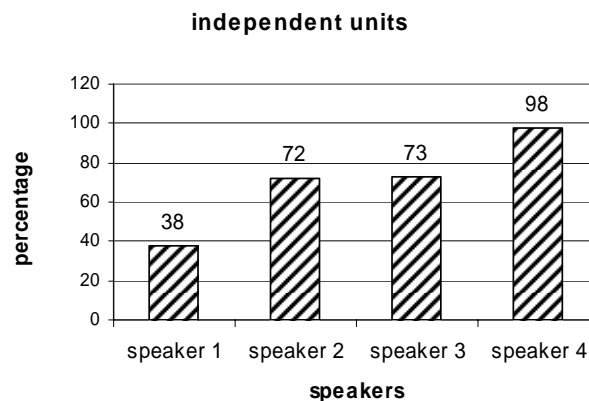


Figure 10: Percentages of right-dislocated phrases forming separate units (y-axis) for each speaker (x-axis).

In Figure 10, the first speaker shows a marked tendency to produce the right-dislocated phrase in the same unit as the main phrase, (forming separate units only 38% of the time). On the other hand, nearly all (98%) of the productions of the last speaker fall into separate units. The other two speakers show a very similar behaviour, with 72% of the productions of speaker 2 and 73% of speaker 3 forming prosodically independent units.

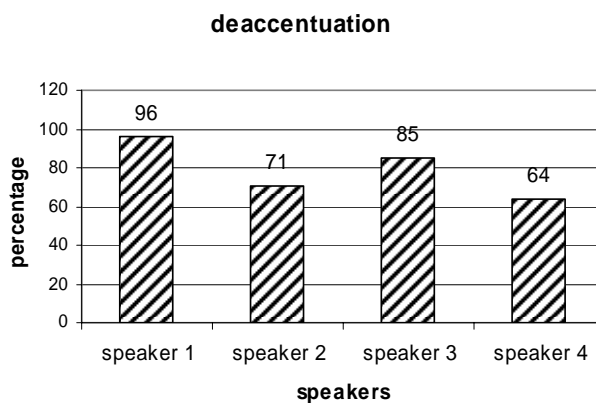


Figure 11: Percentages of deaccented right-dislocated phrases (y-axis) for each speaker (x-axis).

Figure 11 shows that, on average, speakers deaccented 79% of right-dislocations, with the highest percentage belonging to speaker 1 (96% of deaccented cases), and the lowest to speaker 4 (64%), with the other two speakers 3 and 4 showing 71% and 85% deaccented tokens respectively.

3.2.4 Discussion. Accentual cues seem to be slightly stronger than phrasing cues, since 79% of the cases appear as deaccented and the degree of inter-speaker variation is lower than with phrasing cues (70% of phrases form independent units and inter-speaker variation is much higher). One possible explanation for the remarkable inter-speaker differences in phrasing may be the speakers' different reading styles. The reading pace of speaker 1 is quite fast, while that of speaker 4 is rather slow, and that of the other two speakers can be considered normal. As has been suggested in work by Cooper and Paccia-Cooper (1980:189), the slower the reading, the more likely the speaker is to break utterances into separate prosodic phrases. Speakers who read fast make fewer prosodic breaks and deaccent more frequently, as is the case with speaker 1, the fastest reader. By contrast, speaker 4, the slowest reader, tends to produce his right-dislocated phrases in a separate unit, and he also shows a greater tendency to accent them. This seems to point to a trade-off between phrasing and accentuation in the prosodic form of right-dislocated phrases (at least in read, pre-planned speech). However, further study is needed to examine how differences in phrasing correlate with differences in reading style and with the increasing degrees of vocal effort.

3.3 *Experiment 3*

3.3.1 *Methodology*. In Experiment 3, instead of using masking noise to elicit variations in pitch range, the levels of prosodic prominence were manipulated. The experimental material (see Appendix) was highly controlled. The target syllables were the initial syllable in disyllabic words with either stress on the second syllable (as in *Vi^la*, a surname), or stress on the first syllable (as in ^l*Vila*, the nickname of a football club), as well as the initial syllable in tetrasyllabic words with secondary stress on the first syllable and primary stress on the third one (as in ₁*Vila^lbe^lla*, a place name and also the name of a football club). In this way, the target syllables have identical segmental composition but different degrees of stress—unstressed, primary stress, and secondary stress⁸. Therefore, the three stress conditions are:

(6) *Stress conditions*

- a. Stress 0: *Vi^la* (“surname”)
- b. Stress 1: ^l*Vila* (“nickname of a football club”)
- c. Stress 2: ₁*Vila^lbe^lla* (“a football club”)

There were twelve such words in total embedded in three right-dislocated subject phrases and in three right-dislocated object phrases (see Appendix). The material was balanced to level out vowel-specific pitch differences (Lehiste 1970), so that half of the target syllables contained high vowels and half contained low and central vowels. Both the information structure and the semantic/pragmatic context were kept constant. To this effect, the sentences were introduced by a question calculated to elicit an out-of-focus interpretation, as in:

- (7) a. *Va guanyar la lliga, el Vilabella?* (“Did they win the league, Vilabella?”)
 b. *Va guanyar-la, el Vila* (“They won it, Vila”)

The target structures were mixed in random order with other phrases intended to act as ‘distractors’ to prevent subjects from falling into a repetitive reading style. Thus prepared, the text was read by six Central Catalan speakers, three males and three females (see Astruc 2003). It was expected that syllables bearing stress, whether primary or secondary, would be scaled higher than their unstressed counterparts. But if syllables with primary stress were significantly higher than those with just secondary stress, this would indicate that they receive real pitch accents and not mere stress effects.

⁸ Three or more unstressed syllables, as in *el-Vi-la-^lbe-lla*, are not allowed in Catalan and a support stress has to be re-introduced (Oliva & Serra 2002; but see also Mascaró 2002).

3.3.2 *Results*. With regards to phrasing, it was found that dislocations, as in Experiment 1, tended to form separate tonal units. This can be observed in the two pitch traces in Figure 12.

In both cases there is a pause between the main phrase and the dislocated phrase, which has a rather flat pitch range compared with that of the main phrase. The dislocated sentences appeared to be deaccented, both acoustically and instrumentally, as shown in Figure 13.

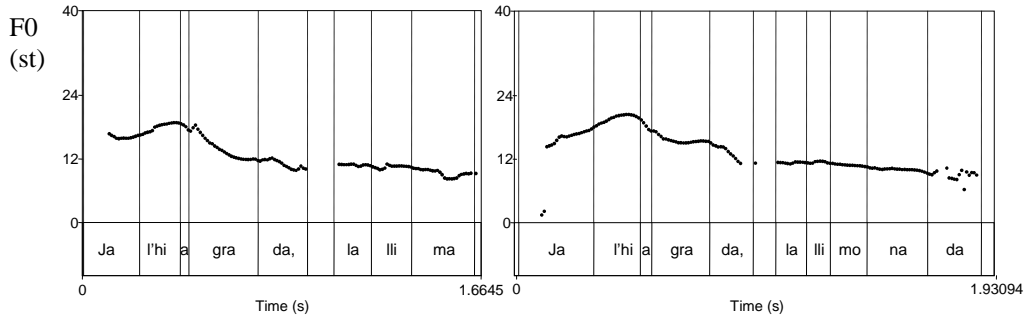


Figure 12: *Phrasing in two right-dislocations in Catalan: Ja li agrada, la llima (“S/he likes it, lime”); ja li agrada, la llimonada (“S/he likes it, lemonade”).*

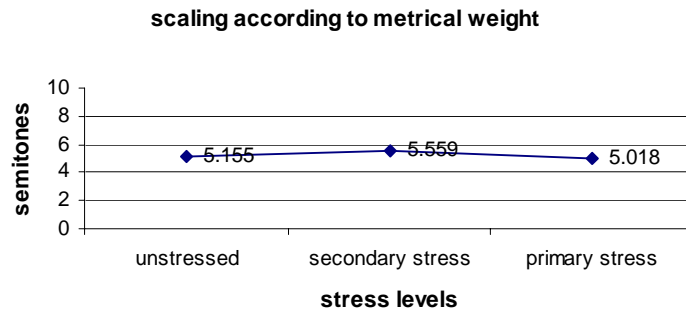


Figure 13: *Scaling of the target syllables. Fundamental frequency in semitones (on the y-axis) and three levels of stress (on the x-axis).*

As Figure 13 shows, the F0 level measured over the unstressed syllables is virtually identical to that of the target primary stressed syllable. This can be taken as evidence against the existence of pitch accents, which is further confirmed statistically by a one-factor repeated-measures mixed ANOVA run on the data of all six speakers for the three stress conditions. The ANOVA provides no significant evidence of effects of stress level upon scaling ($F(2,10) = 0.547, p > 0.05$). The 0.50 semitones difference between the syllables with secondary stress and those with primary stress, apart from not being significant, is contrary to the initial hypothesis that syllables with primary stress would have a higher scaling. The lower scaling of all primary stressed syllables is interpreted as an artifact of the experimental procedure used, because unstressed syllables (*lli¹mona* “lemon”) and syllables with primary stress (*lli¹ma* “lime”) are shorter than syllables with secondary stress (*llimo¹nada* “lemonade”) and so the measurement point is earlier in the overall pitch downtrend.

There is no support for the hypothesis that right-dislocations are accented since the differences between the three stress conditions are fairly modest and, furthermore, in a direction opposite to the initial hypothesis, which was that primary stressed syllables should be higher than secondary stressed ones. In this case, secondary stressed syllables, which appear in initial position in tetrasyllabic words, are lower than primary stressed ones that belong to disyllabic words. This is not interpreted as an indication of the existence of low pitch accents. Such a possibility is discarded, first, on auditory grounds, and, second, following the outcome of the statistical analysis which shows no significant differences in scaling under the three different stress conditions.

4. *Summary and conclusion*

Phonological studies have traditionally considered ESEs as forming syntactically and prosodically independent units. However, as argued in Section 1.2, they do not form a homogeneous grammatical category, even though they have certain common characteristics which are semantic in nature: their semantic scope encompasses the whole sentence, and most ESEs (except sentential adverbs) also share the semantic function of adding supplementary information. The question is, how does this functional role relate to their prosodic form?

First of all, our view is that the prosodic behaviour of dislocations and extra-sentential elements in general is governed by general principles of information structure and textual organization, as suggested by Liberman (1975:185), and as is also implicitly contained in Ladd's notion of 'structural pitch range effects' (that is, downstep and pitch range shifts) to signal syntactic and textual structure (Ladd 1996:279). Our view is that the role of ESEs is primarily semantic (either that of signalling sentence-wide semantic scope, as sentential adverbs do, or that of signalling an anaphoric connection to their referent, as most ESEs do), and that such a semantic role is signalled prosodically by means of tonal and/or junctural cues.

The purpose of Experiment 1 was to establish whether ESEs always form independent tonal units and are always deaccented, or rather show variation in their phrasing and intonation. The answer is *yes*: they show both types of variation, which can be taken as an indication of an on-going trade-off between prosodic independence and tonal subordination to cue the peripheral status of ESEs. Experiment 2 used masking noise to study the accentuation of Catalan appositions and right-dislocations, which were showed to differ in the way they signalled tonal subordination (appositions, with reduplication; right-dislocations, with deaccenting). There was also observed in both Experiment 1 and Experiment 2 a potential correlation between phrasing and speaking rate, so that the faster the rate, the lower the occurrence of prosodic breaks. A tightly controlled methodology is needed to find confirmation for this trend, which was not confirmed statistically, perhaps owing to inter-speaker variation. Unfortunately, the masking noise technique used in Experiment 2 did not yield quantitative confirmation of the analysis. Experiment 3 followed a

much stricter methodology, which involved measuring the F0 excursion of syllables with different degrees of prosodic prominence. With this method, it was possible to conclude that right-dislocated phrases, when tightly controlled for contextual factors, are totally deaccented.

Prosodically, ESEs signal their usual peripheral role in the sentence either by being totally deaccented or, if accented, by means of a dramatic compression in pitch range coupled with prosodic separation, and often, by a combination of both strategies. Prosodic separation, in fact, is not strictly compulsory. About 70% of the tokens in the corpus were split into two units. Most of them, but not all, were also deaccented. Deaccentuation only seemed to be compulsory in those cases in which the ESEs and the main phrase belonged to the same prosodic unit, and there was scope for ambiguity. This behaviour hints at the existence of a trading relationship between rhythm and melody, a notion that can be traced back to Trim (1959) and Ladd (1980:164). If further evidence were found, standard AM will have to account for it. A possible solution (in line with Beckman & Pierrehumbert 1986 and with Truckenbrodt 1995, 1999) would be to analyse ESEs as intermediate phrases, with obligatory phrase accents and boundary tones but with optional pitch accents. In the absence of pitch accents, phrase accents would spread from right to left. When pitch accents are present, tonal subordination mechanisms such as reductions in pitch range or tonal reduplication would operate within the domain of the intermediate phrase to signal the subordinated grammatical role of ESEs.

References

- Allerton, David J. & Alan Cruttenden. 1974. "English sentence adverbials: Their syntax and their intonation in British English". *Lingua* 34. 1-30.
- Astruc, Lluïsa. 2003. "Right-dislocations in Catalan". *Proceedings of the 15th International Congress of the Phonetic Sciences* ed. by Maria-Josep Solé, Daniel Recasens & Joaquín Romero. 1265-1268. Barcelona: Causal Productions.
- , 2005. The intonation of extra-sentential elements in Catalan & English. PhD diss., University of Cambridge (available from www.astruc.info).
- & Francis Nolan. In press. "Variation in the intonation of sentential adverbs". *Tones and Tunes* ed. by Tomas Riad & Carlos Gussenhoven, vol. I. Berlin & New York: Mouton de Gruyter.
- Beckman, Mary & Janet Pierrehumbert. 1986. "Intonational structure in Japanese and English". *Phonology Yearbook* 3. 225-310.
- Bing, Janet M. 1985. *Aspects of English Prosody*. New York & London: Garland Publishing.
- Bonet, Eulàlia. 1984. Aproximació a l'entonació del català central. [A preliminary investigation of the intonation of Central Catalan]. Master's thesis, Universitat Autònoma de Barcelona.

- Cooper, William E. & Jeanne Paccia-Cooper. 1980. *Syntax and Speech*. Cambridge, Mass. & London: Harvard University Press.
- Cruttenden, Alan. 1997. *Intonation*. (2nd edition). Cambridge: Cambridge University Press.
- Delattre, Pierre. 1972. "The distinctive function of intonation". *Intonation: Selected Readings* ed. by Dwight Bolinger. 159-174. Harmondsworth: Penguin.
- D'Imperio, Mariapaola. 2002. "Italian intonation: An overview and some questions". *Probus* 14. 37-69.
- Emonds, Joseph E. 1979. "Appositive relatives have no properties". *Linguistic Inquiry* 10. 211-243.
- Estebas-Vilaplana, Eva. 2000. The use and realisation of accentual focus in Central Catalan with a comparison to English. PhD diss., University College London.
- Fabb, Nigel. 1990. "The difference between English restrictive and non-restrictive relative clauses". *Journal of Linguistics* 26. 57-77.
- Fagyal, Zsuzsanna. 2002a. "Prosodic boundaries in the vicinity of utterance-medial parentheticals in French". *Probus* 14. 93-111.
- , 2002b. "Tonal template for background information: The scaling of pitch in utterance-medial parentheticals in French". *Proceedings of the 1st International Conference on Speech Prosody* ed. by Bernard Bel & Isabelle Marlien. 279-282. Aubenas d'Ardèche: Lienhart.
- Geluykens, Ronald. 1992. *From Discourse Process to Grammatical Construction: On Left-dislocation in English*. Amsterdam: John Benjamins.
- , 1994. *The Pragmatics of Discourse Anaphora in English: Evidence from Conversational Repair*. Berlin: Mouton de Gruyter.
- Grice, Martine. 1995. *The Intonation of Interrogation in Palermo Italian: Implications for Intonation Theory*. Tübingen: Niemeyer.
- Gussenhoven, Carlos. 1985. "The intonation of George and Mildred: Post-nuclear generalizations". *Intonation in Discourse* ed. by Catherine Johns-Lewis. 77-123. San Diego: College Hill.
- , 1993. *On the Grammar and Semantics of Sentence Adverbs*. Dordrecht: Foris.
- , 2004. *The Phonology of Tone and Intonation*. Cambridge: Cambridge University Press.
- Huddleston, Rodney & Geoffrey K. Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- Junqua, Jean-Claude. 1996. "The influence of acoustics on speech production: A noise-induced stress phenomenon known as the Lombard reflex". *Speech Communication* 20. 13-22.
- Kempson, Ruth, Wilfried Meyer-Viol & Dov Gabbay. 2000. *Dynamic Syntax: The Flow of Language Understanding*. Oxford: Blackwell.
- Ladd, D. Robert. 1980. *The Structure of Intonational Meaning: Evidence from English*. Bloomington & London: Indiana University Press.

- . 1996. *Intonational Phonology*. Cambridge: Cambridge University Press.
- Lambrecht, Knud. 1981. *Topic, Antitopic, and Verb Agreement in Non-standard French*. Amsterdam: John Benjamins.
- . 1994. *Information Structure and Sentence Form: Topic, Focus, and the Mental Representation of Discourse Referents*. Cambridge: Cambridge University Press.
- Lane, Harlan & Bernard Tranel. 1971. "The Lombard Sign and the role of hearing in speech". *Journal of Speech & Hearing* 14. 677-709.
- Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge, Mass. & London: The MIT Press.
- Lieberman, Mark. 1975. *The Intonational System of English*. Bloomington, Ind.: Indiana University Linguistics Club.
- Martin, Pierre. 1987. "Prosodic and rhythmic structures in French". *Linguistic Inquiry* 8. 249-336.
- Mascaró, Joan. 2002. "Redució vocàlica". *Gramàtica del català contemporani* ed. by Joan Solà, Maria-Rosa Lloret, Joan Mascaró & Manuel Pérez Saldanya. 89-123. Barcelona: Editorial Empúries.
- Nespor, Marina. 1993. *Fonologia: Le strutture del linguaggio*. Bologna: Il Mulino.
- & Irene Vogel. 1986. *Prosodic Phonology*. Dordrecht: Foris Publications.
- Oliva, Salvador & Pep Serra. 2002. "Accent". *Gramàtica del català contemporani* ed. by Joan Solà, Maria-Rosa Lloret, Joan Mascaró & Manuel Pérez Saldanya. 345-391. Barcelona: Editorial Empúries.
- Payà, Marta. 2002. "Incidental clauses in spoken Catalan: prosodic characterization and pragmatic function". *Proceedings of the 1st International Conference on Speech Prosody* ed. by Bernard Bel & Isabelle Marlien. 559-562. Aubenas d'Ardèche: Lienhart.
- 2003. "Prosody and pragmatics in parenthetical insertions in Catalan". *Catalan Journal of Linguistics* 2. 207-227.
- Pierrehumbert, Janet. 1980. The phonetics and phonology of English intonation. PhD diss., MIT.
- & Julia Hirschberg. 1990. "The meaning of intonational contours in the interpretation of discourse". *Intentions in Communication* ed. by Philip R. Cohen, Jenny L. Morgan & Martha E. Pollack. 271-311. Cambridge, Mass. & London: The MIT Press.
- Prieto, Pilar. 2002a. "Entonació". *Gramàtica del català contemporani* ed. by Joan Solà, Maria-Rosa Lloret, Joan Mascaró & Manuel Pérez Saldanya. 393-462. Barcelona: Editorial Empúries.
- , 2002b. *Entonació: Models, teoria, mètodes*. Barcelona: Ariel.
- Recasens, Daniel. 1993. *Fonètica i fonologia*. Barcelona: Enciclopèdia Catalana.
- Safir, Ken. 1986. "Relative clauses in a theory of binding and levels". *Linguistic Inquiry* 17. 663-683.

- Selkirk, Elisabeth. 1984. *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, Mass.: The MIT Press.
- Trim, John L. 1959. "Major and minor tone groups in English." *Le Maître Phonétique* 112. 26-29.
- Truckenbrodt, Hubert. 1995. Phonological phrases: Their relation to syntax, focus, and prominence. PhD diss., MIT.
- , 1999. "On the relation between syntactic phrases and phonological phrases". *Linguistic Inquiry* 30:2. 219-255.
- Vallduví, Enric. 1990. "The role of plasticity in the association of focus and prominence". *Proceeding of the Eastern States Conference on Linguistics* 7 ed. by Yongkyoon No & Mark Libucha. 295-306. Columbus, Ohio: Ohio State University.
- , 1994. "Detachment in Catalan and information packaging". *Journal of Pragmatics* 22. 573-601.
- Wunderli, Peter. 1987. *L'intonation des sequences extraposées en français*. Tübingen: Gunter Narr Verlag.
- Zubizarreta, María Luisa. 1998. *Prosody, Focus, and Word Order*. Cambridge, Mass. & London, England: The MIT Press.

Appendix

Experiment 2

Quasi-minimal pairs of apposition and right-dislocated phrases in Catalan (with translations in English).

Appositions

- És la Mona, la dona. Sí, es diu així: Mona.*
("She is Mona, the wife. Yes, that's her name, Mona.")
- Va veure la Núria, la núvia.* ("She saw Núria, the bride.")
- És Móra, la Nova. No l'altra Móra, la Móra d'Ebre.* ("This is Móra, la Nova. Not the other Móra, Móra d'Ebre.")
- Fan el drama, 'La Mama'.* ("They are showing the drama, 'The Mum'.")
- Vol la nena, la Lena.* ("She wants the little girl, Lena.")
- Vol la vela, la vella.* ("He wants the sail, the old one.")
- Han posat Dallas, el drama.* ("They have shown Dallas, the drama.")

Dislocations

- És mona, la dona. Però no sembla pas gaire agradable.* ("She's cute, the wife [is]. But she doesn't seem very nice.")
- Va veure la Núria, la núvia.* ("She saw Núria, the bride [did].")
- És mora, la nova.* ("She's an Arab, the new one [is].")
- Fa drama, la mama.* ("She's making a drama out of it, mother [is].")
- Vol una nena, la Lena.* ("She wants a girl, Lena [does].")
- Vol la vela, la vella.* ("She wants the sail, the old lady [does].")
- Va passar a Dallas, el drama.* ("It happened in Dallas, the drama [did].")

Experiment 3

Right-dislocated phrases in Catalan with translations in English.

- Ja li agrada, la mel* ("S/he likes it, honey").
Ja li agrada, el meló ("S/he likes it, watermelon").

Ja li agrada, la melonada (“S/he likes it, watermelon juice”).

Ja en vol, de mel (“S/he wants some, honey”).

Ja en vol, de meló (“S/he wants some, melon”).

Ja en vol, de melonada (“S/he wants some, melon juice”).

Va guanyar-la, la mare (“She won it, the mother [did]”).

Va guanyar-la, la mamà (“S/he won it, the mama [did]”).

Va guanyar-la, la Mamabona (“They won it, the Mamabona [did]”).

Ja li agrada, la llima (“S/he likes it, lime”).

Ja li agrada, la llimona (“S/he likes it, lemon”).

Ja li agrada, la llimonada (“S/he likes it, lemonade”).

Ja en menja, de llima (“S/he eats it, lime”).

Ja en menja, de llimona (“S/he eats it, lemon”).

Ja en beu, de llimonada (“S/he drinks it, lemonade”).

Va guanyar-la, el Vila (“They won it, the Vila [did]”).

Va guanyar-la, el Vilà (“He won it, Vilà [did]”).

Va guanyar-la, el Vilabella (“They won it, the Vilabella [did]”).

