

# Variation in the intonation of extra-sentential elements

*Lluïsa Astruc-Aguilera*  
*Francis Nolan*

## 1. Introduction

The group of so-called extra-sentential elements<sup>1</sup> includes phrases, such as dislocated phrases ('They are crazy, those Romans'), and words, such as vocatives ('Thanks, sir'), and sentential adverbs ('obviously'). As described in recent work (Huddleston and Pullum 2002) their main common characteristics are that they belong to the periphery of the sentence, that they have semantic scope over the whole sentence, and that they form independent intonational phrases.

Extra-sentential elements have traditionally played a central role in intonation studies since they have been widely used in the phonological literature to define the intonational phrase, which is the level of the prosodic hierarchy that is most relevant to intonation. As pointed out by Nespor and Vogel, extra-sentential elements (henceforth ESEs) usually form independent intonational domains:

One of the first things that is usually observed in relation to intonation contours is that there are certain types of constructions that seem to form intonation domains on their own. These constructions include parenthetical expressions, nonrestrictive relative clauses, tag questions, vocatives, expletives, and certain moved elements [...]. (Nespor and Vogel 1986: 188)

The traditional assumption is that ESEs are syntactically independent, just as they are also prosodically independent (e.g. Pierrehumbert 1980, Nespor 1993, Nespor and Vogel 1986), and thus they have been used as evidence of the effect of syntactic constraints upon phrasing (Cooper and Paccia-Cooper 1980, Nespor and Vogel 1996, among others). Recent work on phonological phrasing has refined this account by proposing that the boundaries of extra-sentential elements are compulsorily aligned with those of the intonational phrase and by developing specific mechanisms to generate this alignment (Selkirk 1984; Truckendbrodt 1995, 1999). One possible problem with the account of ESEs as being independent, both syntactically and prosodically, with the latter assumed to be a consequence of the former, 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 (1976, 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 et al. (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?

---

<sup>1</sup> Extra-sentential elements (Gussenhoven 2004) have also been called, among other terms, 'sentence-external elements' (Ladd 1996), 'tags' (Lieberman 1975, Ladd 1980), and 'parentheses' (Fagyal 2002a, 2002b).

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, 164-165; Liberman 1975: 182-184; Gussenhoven 1985, 1993, 2004: 290-295; Beckman and Pierrehumbert 1980: 293-298; Nespov 1993: 265; Prieto 2002b: 409ff) and/or by tonal reduplication (Gussenhoven 1985: 107; 2004: 315-316; Bonet 1984: 30, 34; Recasens 1993: 214; Prieto 2002b: 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<sup>2</sup>.

A further problem, as has been observed in the literature (see section 1.4) and also in the results of a previous study (Astruc 2003b), 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 seems to emerge from 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.

The structure of this article is as follows: In section 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 a cross-linguistic production experiment. In section 3, we describe two experiments involving right-dislocated phrases in Catalan. In section section 4 we summarize the results and draw some conclusions.

## 1.2. 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 criteria in the Cambridge Grammar of the English Language (Huddleston and Pullum 2002) and on the basis of prosodic criteria in phonological studies (see Nespov and Vogel's list, above, which includes non-restrictive relative clauses, tag questions, vocatives, expletives, and dislocated phrases). The intersection of both sets of categories yielded the categories in (1)<sup>3</sup>:

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

---

<sup>2</sup> 'One of the most widely ignored in problems in intonational phonology is the problem of *dependencies* between intonational phrases, such as the intonational dependency between main sentences and tags [...] there is no convincing account of how such dependencies work anywhere in the literature' (Ladd 1996: 246).

<sup>3</sup> 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.

The starting assumption is that ESEs do not form a homogenous class from a grammatical point of view. Rather, ESEs are formed by categories that belong to different components of the grammar. Dislocated phrases and sentential adverbs are clearly syntactic constituents. Non-restrictive relatives, parenthesis, 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.

### 1.3. ESEs as supplements

In Huddleston and Pullum (2002) ESEs have the role of ‘supplements’, that is, they add information that supplements, but does not restrict or modify, the propositional content of the main clause. They are, therefore, 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, the sentence reads ‘Pat had invited all the neighbours’. Since neither the grammaticality nor the truth conditions of the sentence are affected, we can conclude that the parenthetical clause is truly optional and supplementary.

Another important characteristic is that supplements are semantically related to the clause with which they co-occur, their ‘anchor’. Owing to their lack of integration into the syntactic structure, supplements are semantically non-restrictive. As seen in (3), 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. This contrast is perhaps easier to observe in the distinction between the two types of relative constructions in the examples below (Huddleston and Pullum 2002):

- (3) a. The necklace *which her mother gave her* was in the safe  
b. The necklace, *which her mother gave her*, was in the safe

In (3a), the relative clause modifies the head noun *necklace* and serves to identify which necklace is being referred to (only the one that her mother gave her). On the other hand, the referent of (3b) is assumed to be easily identifiable, independent of the information provided by the relative clause, which only adds supplementary detail. If this were to be removed, the sentence would still make perfect sense. However, sentential adverbs, one of the core categories of ESE, are not served well by this account. As already noted by Bing (1986) and by Allerton and Cruttenden (1974, 1976, 1978) certain types of sentential adverb (modal adverbs such as ‘possibly’) do restrict and/or modify the truth conditions of the proposition expressed in the main sentence. Sentential adverbs, therefore, appear to have functional characteristics that sets them apart from the other ESEs (see Astruc 2005a, and Astruc and Nolan in press for more detailed description of sentential adverbs).

With the possible exception of sentential adverbs, 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.

### 1.4 Previous intonational studies

The prosodic form of ESEs has been studied in a variety of languages, such as French (Wunderli 1987, Martin 1987, Delattre 1972, Fagyal 2002a, 2002b) and Italian (Nespor 1993, Nespor and Vogel 1986), Spanish (Zubizarreta 1998), and Danish (Hansen 2002). ESEs have attracted ample attention in English (Lieberman 1975; Pierrehumbert 1980; Beckman and Pierrehumbert 1986; Pierrehumbert and Hirschberg 1994; Ladd 1980, 1996; Bing 1984, 1985; Gussenhoven 1985), and also in Catalan (Bonet 1984; Prieto 2002b, 2002c; Recasens 1993; Payà 2002, 2003).

Pre-theoretically, 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 and Pierrehumbert 1986), and to proposals describing them as full-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: 164-165<sup>4</sup>) tend to appear as deaccented. Categories generally described as accented are appositions and non-restrictive relatives (with the main exceptions of Pierrehumbert and Hirschberg 1990, Pierrehumbert 1980, and Cruttenden 1997). Vocatives are described as deaccented in English (Lieberman 1975; Ladd 1980, 1996; Firbas 1980; Beckman and Pierrehumbert 1986; Bing 1984, 1985; Pierrehumbert and Hirschberg 1992) but as accented in Catalan (Bonet 1984: 28-29, 60-61, 88; Recasens 1993: 211, 214; Prieto 2002b: 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. General comparative 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 (which is not reported here, for reasons of space).

---

<sup>4</sup> 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.

## 2.2 General comparative study of the prosodic characteristics of extra-sentential elements

For this study, a corpus of 605 phrases was gathered (462 phrases in English and 143 in Catalan) containing the target ESEs listed in Example (1), and which included the main categories of ESEs usually discussed in syntactic and phonological studies.

The target ESEs appeared in different positions in the sentence, in sentences with identical meaning in English and Catalan:

- (4) 'L'Anna va guanyar-la, *Manu*'  
'Anna won it, *Manu*'

The English part was formed by 462 phrases: 29 examples which were read twice by eight Southern British English speakers. Of these, two tokens had to be discarded because of reading errors. The target sentences were mixed in pseudo-random order with sentences containing adverbs which were intended for another experiment. Likewise, following the standard practice in experimental design, other phrases containing completely unrelated grammatical constructions, so-called 'distractors', were also inserted, to prevent readers from falling into a monotonous reading style.

The Catalan part was formed by 143 phrases (50 sentences, read in roughly equivalent proportions by 3 speakers), which were gathered at three different stages. The first stage was a pilot experiment intended to investigate the effects of changes in pitch range upon ESEs. Later on, more data on dislocations, appositions, and sentential adverbs were gathered, as part of two further studies (Astruc 2004, Astruc and Nolan to appear).

### 2.2.1 Results

The data was digitised at 16Kz and subsequently analysed by the first author by listening and by looking at pitch traces obtained with Praat 4.1.21. 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.

In this study, phrasing and accentuation are considered to be independent systems, as proposed by Trim (1959) and by 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<sup>5</sup>). The criteria for deciding that a stretch of speech was deaccented were the following acoustic correlates of lack of prosodic prominence on the stressed syllable: flat F0, reduced amplitude, and shallower spectral tilt.

The following generalizations can be drawn about the main prosodic properties of ESEs. First, all categories of ESEs tend to be prosodically independent of the main phrase. The exceptions are appositions and vocatives, which at times may appear integrated in the main phrase. In both 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. ESEs in initial position 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:

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

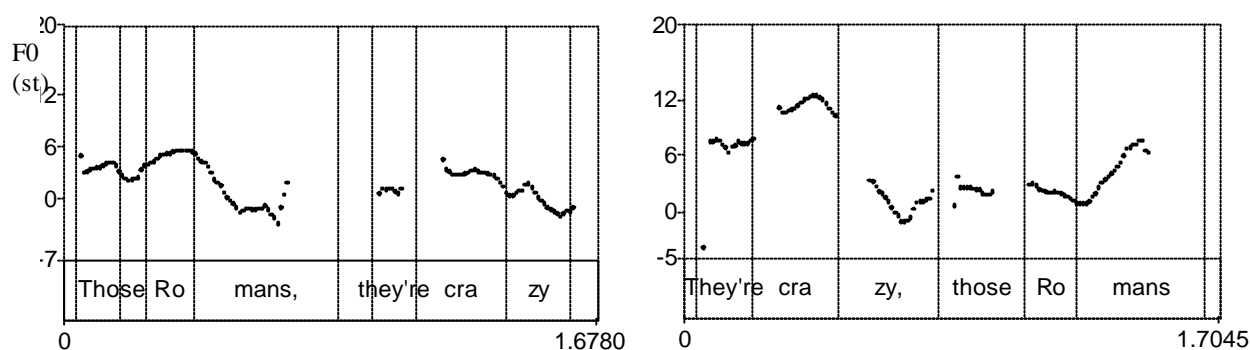
---

<sup>5</sup> A pause, in this pre-planned, read speech has been defined here as a period of silence (which is not caused by the presence of obstruents) of 100 ms or longer. Anything shorter than 100 ms has been considered as pre-boundary lengthening.

Consequently, ESEs fall into two groups according to the type of tonal subordination they display. Dislocated phrases, quotations markers, and epithets follow strategy (7a); that is, they show deaccentuation. Parentheses, non-restrictive relatives, and appositions follow strategy (7b): they show tonal reduplication at a lower level and with a compressed pitch range. Sentential adverbs and vocatives do not fit clearly into either (7a) or (7b).

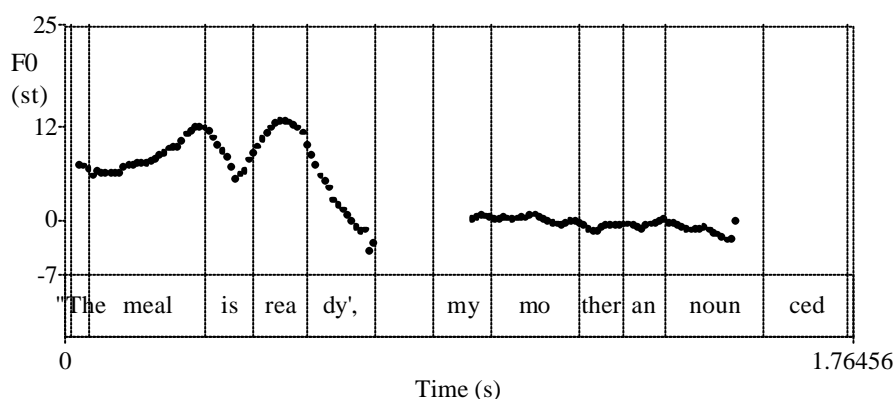
### 2.2.2 Dislocated phrases, quotations markers, and epithets

In English, dislocated phrases form independent phrases 88% of the time, both in initial and final position. 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 (but see also section 3), and they are nearly always deaccented (95%). Unlike in English, they are not followed by a final rise.



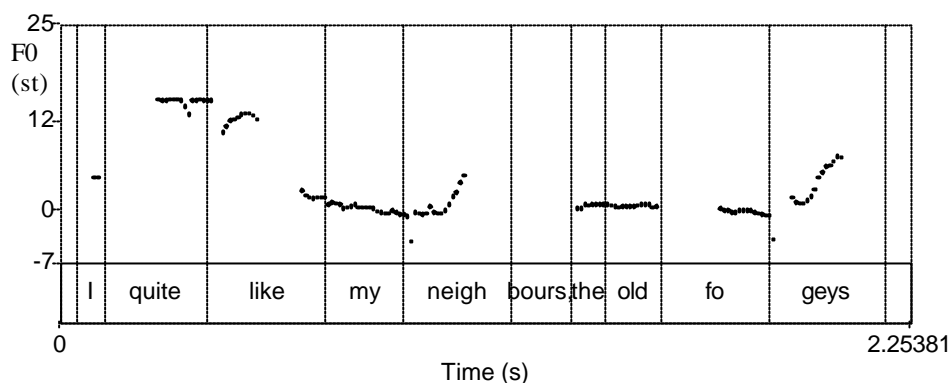
**Fig. 1** In the left panel, an example of left-dislocated subject: 'Those Romans, they're crazy'. In the right panel, an example of right-dislocated subject: 'They're crazy, those Romans'.

Direct speech markers, both in 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).



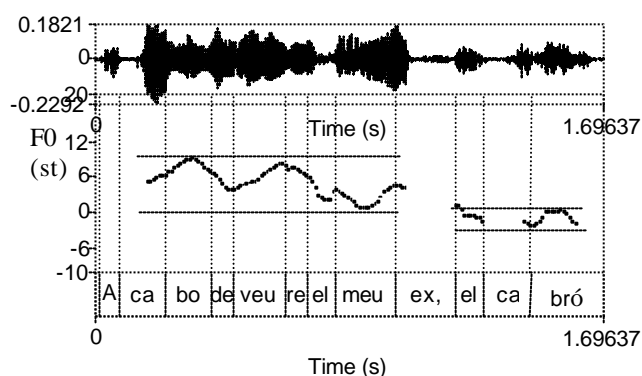
**Fig. 2** An example of direct speech marker in English

In English, epithets 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. They always form independent phrases. Fig. 3 shows the intonation that typically corresponds to English epithets:



**Fig. 3** An example of epithet in English

Unlike in English, epithets in Catalan<sup>6</sup> can be accented (50% of the time) and they never carry a final rise. The epithets in the corpus are at times produced so 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 accent had a pitch range subordinated to that of the main phrase, as can be seen in the pitch trace below:



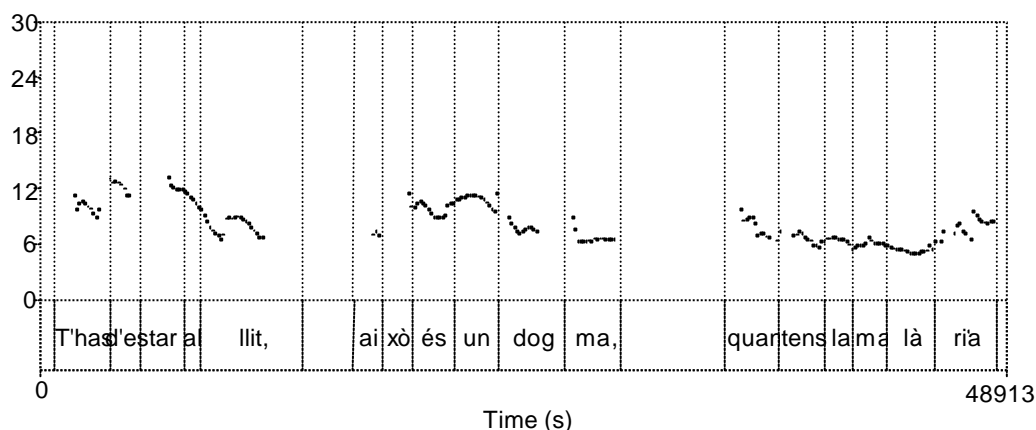
**Fig. 4** An example of epithet in Catalan: 'Acabo de veure el meu ex, *el cabró*' ('I've just seen my ex, *the bastard*')

Parallel lines have been fitted by hand to facilitate the comparison of the pitch range and pitch level in the main clause and in the 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.

### 2.2.3 Parentheses, non-restrictive relatives, appositions, and vocatives

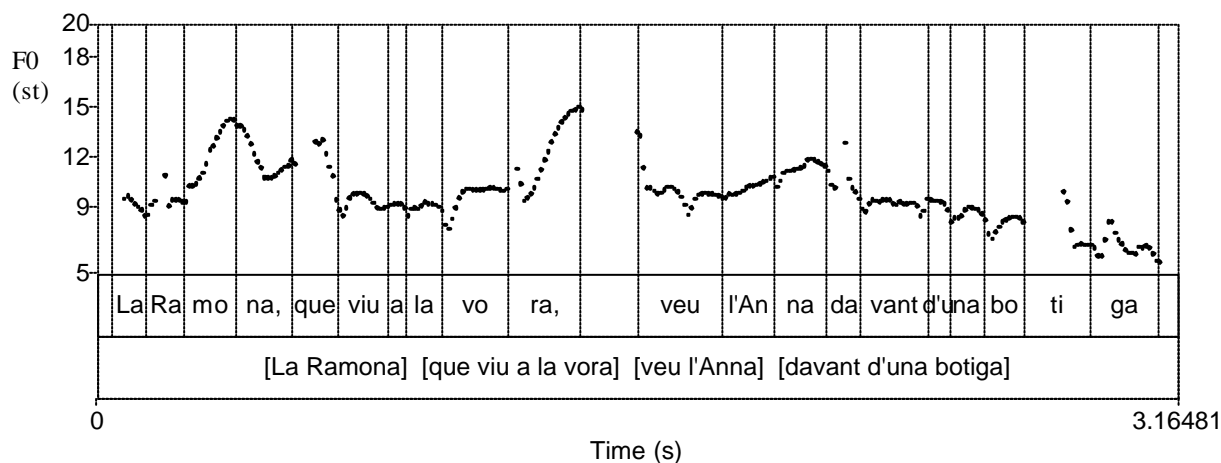
<sup>6</sup> Epithets however can carry an emotional load, in which case they are accented, irrespective of language. This seldom happens, and moreover, the few occurrences of accenting correspond to the same phrase, 'I've just caught a glimpse of my ex, the bastard' and to its interrogative counterpart 'Have you seen my ex, the bastard?.'

Parentheses, non-restrictive relatives, and appositions follow strategy (7b): 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 times), and are mostly accented (83%), often with a reduplicating contour (53% of times). In Catalan, parenthesis always form independent units which are mostly demarcated by pauses (58%). 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 with a compressed pitch range.



**Fig. 5** An example of parenthesis in Catalan: 'T'has d'estar al llit, *això és un dogma*, quan tens malària' ('You have to stay in bed, *this is a rule*, when you have malaria')

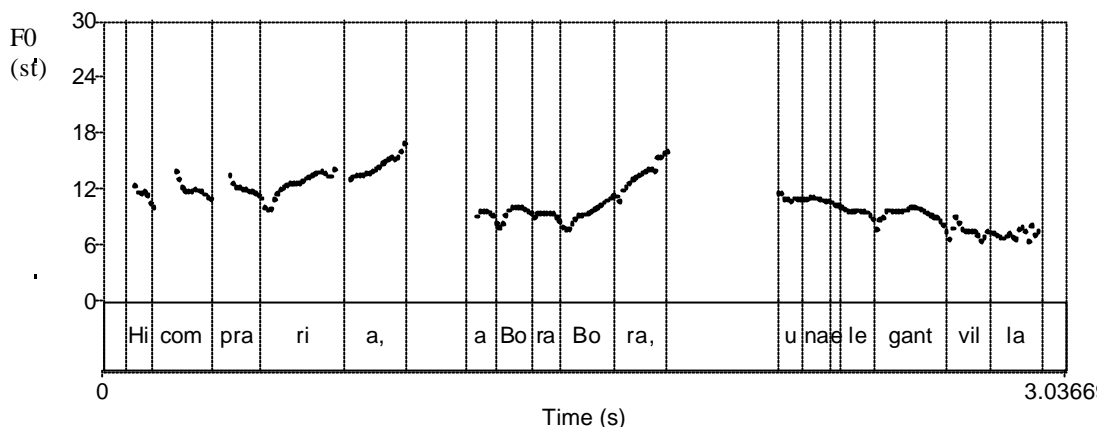
One of the first things that is observed about non-restrictive relatives is that they show more cohesion with the clause to their left than with that to their right. Both in English and Catalan, the non-restrictive relative clause reduplicates the tonal pattern of the left-hand clause (60% of the time in Catalan, 53% in English). 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 relatives 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 in gender.



**Fig. 6** An example of non-restrictive relative in Catalan: 'La Ramona, *que viu a la vora*, veu a l'Anna davant d'una botiga' ('Ramona, *who lives nearby*, saw Anna next to shop window')

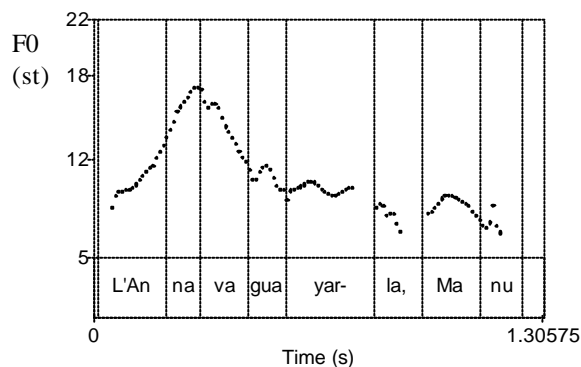


Similar to the case of non-restrictive relatives, appositions, both in English (88% of the time) and in Catalan (83%) tend to form independent intonational phrases that reduplicate the contour of their anchor.<sup>7</sup>



**Fig. 7** 'Hi compraria, a Bora Bora, una elegant vil·la' ('I would buy an elegant villa there, in Bora Bora')

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. 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).



**Fig. 8** An example of Catalan vocative in final position: 'L'Anna va guanyar-la, Manu' ('Anna won it, Manu')

<sup>7</sup> However, in a very few cases the apposition does not receive an independent contour. After Huddleston and Pullum (2002: 1064, 447) such instances are interpreted as 'appositive modifiers', which provide identifying information about the referent, in a way akin to that of restrictive relatives. For instance, 'Norman' in 'This is my husband Norman' would be interpreted as an 'appositive modifier', that is, as a succinct 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, and much less relevant than the information presented first. 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 as regard to their accentuation, but they were very consistent as regards their phrasing: they formed independent sentences nearly always. 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 2005a, and Astruc and Nolan to appear <sup>8</sup>).

### 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 right-dislocated phrase is that of introducing background information in a position where a high informative content would normally be expected (Huddleston and 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 2002b, 2002c; 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.

As for their prosodic form, it was noted in the results of Experiment 1, that the differences between left- and right-dislocations generally correspond to those between initial and non-initial ESEs. Both types of dislocations receive prosodically detached tonal contours, but only left-dislocations get full tonal contours, while right-dislocations show complete deaccentuation. However, there were also some instances of ‘miniature accents’ following the focal accent; little ‘bumps’ with an excursion size of 1 semitone at most. Post-focal accents have been identified so far in narrow focus sentences in Catalan (Estebas-Vilaplana 2000), in Spanish (Zubizarreta 1998), in Italian (Grice 1995, D’Imperio 2002), and in Germanic languages such as Icelandic (Nolan and Jónsdóttir 2001). 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. Experiment 2 and Experiment 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.

---

<sup>8</sup> Sentential adverbs are studied in detail in Astruc (2005), and Astruc and Nolan (to appear). The core categories of sentential adverb (modality, evaluation, and speech-act adverbs) were compared to homonymous manner adverbs in final position in the sentence, in English and Catalan. The results showed that sentential adverbs nearly always form independent phrases in both languages, thus supporting claims in the literature. Accentuation is used far less consistently than phrasing for disambiguating sentential and phrasal adverbs. Furthermore, there are cross-linguistic difference since English speakers deaccent sentential adverbs more frequently than Catalan speakers do.

## 3.2 Experiment 1

### 3.2.2 *Experimental material*

The corpus is formed by 112 phrases read by four speakers, thus yielding 448 phrases in total. The material consists of seven minimal pairs (see Appendix), each containing an apposition and a dislocation, introduced by a short background text intended to clarify their meaning

- (6) a. Apposition: És Mona, *la dona* ('[She] is Mona, *the wife*')  
b. Dislocation: És mona, *la dona* ('[She] is cute, *the wife*')

Two sets of comparisons were planned: (1) comparison of the stressed syllables in the apposition and in the dislocation, that is the initial syllable 'mo' in *Mo*-na (3a, the given name), and in *mo*-na (3b, 'cute'); (2) comparison of the most prominent syllables (the penultimate ones) in the main phrase and in the dislocation, that is, the comparison of *mo*-na and *do*-na<sup>9</sup>. The pragmatic context is more controlled than in the comparative study described in section 2, though it still aims 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, there were the phrases used in Experiment 3, which will be described in section 3.3. The 14 target sentences were read twice under four experimental noise conditions by four speakers<sup>10</sup>, thus yielding 448 phrases in total.

### 3.2.3 *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 (e.g. Lane and Tranel 1971, Castellanos, Benedí and Casacuberta 1996, Junqua 1996, Junqua et al 1999).

A pilot test was conducted with a female Central Catalan speaker, who was asked to read a text containing the target structures three times while wearing a set of headphones through which she could hear masking noise. The masking noise was real traffic noise obtained in a busy street in central Cambridge at peak hour. The first time, she read the text without background noise (condition 1). The second time, she was hearing the street noise (condition 2). The third time, the noise was set at an even higher volume (condition 3). In this way, the same set of sentences was elicited at three different pitch ranges with a minimum of variation. The technique worked well, as the results showed a significant increase in pitch scaling of all the target syllables. In fact, all tonal targets went up, except the final low. The near constancy of the final low for this subject, across the different experimental conditions, is in line with Pierrehumbert's observations (e.g. Liberman and Pierrehumbert 1984, and elsewhere). On the other hand, 'speaking up' also entailed a rephrasing of most of the contours in different intonational units, which complicated the interpretation of the results. Another problem was that the traffic noise was discontinuous and, furthermore, the increases in loudness were not monitored. Arguably, variation in the results might be partly attributed to such methodological deficiencies.

---

<sup>9</sup> The planned comparisons were not carried out finally because a better method for testing the accentuation of the right-dislocations were devised and applied in the following experiment, as reported in section 3.4.

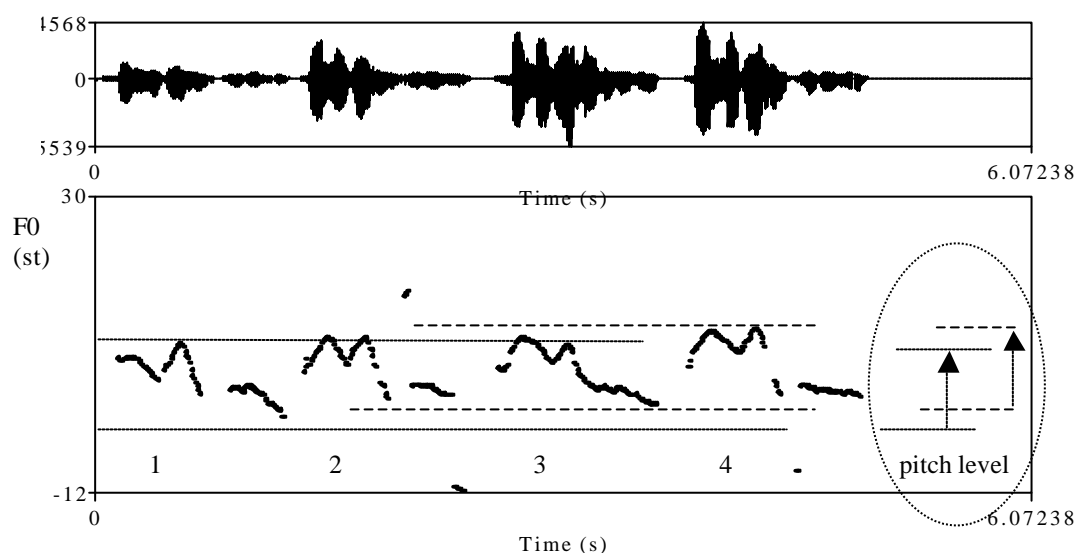
<sup>10</sup> The productions of one speaker were discarded, as the speaker refused to raise her voice, alleging that she had a sore throat.

The masking noise finally used in Experiment 1 is the type of noise regularly used in perception experiments<sup>11</sup>. It is continuous and very similar to the long term average spectrum of the typical speaker (Byrne et al 1994).

The equipment was the same used in the pilot test: a pair of open-ear headphones (Sennheiser HD 570) and two portable Sony TCD-D3 DAT, one for playing back the noise and the other for recording. The participants, four Central Catalan speakers, two men and two women, were asked to wear the headphones from the start. The first time, they read the text without any background noise. The second time, they read the text while masking noise was played back from the second DAT recorder through the headphones<sup>12</sup> at 75.8 dB. The third time they read the text while the volume was increased to 78.8 dB. And for the fourth and final reading, the volume went up to 83.4 dB. In this way four experimental conditions were obtained: silence, and three increasing levels of noise of a constant quality.

### 3.2.4 Results: effects on the speaker's tonal space

A preliminary inspection of the data confirms the impression that there has been a noticeable change in the speaker's tonal space. Fig. 9 shows the wave forms and the pitch traces corresponding to the four experimental conditions for the right-dislocated sentence 'Vol la vela, la vella' ('She wants the sail, the old lady'):



**Fig. 9** Wave forms and pitch traces corresponding to the right-dislocated phrase 'Vol la vela, la vella' ('She wants the sail, the old lady') under four noise conditions

The four experimental conditions are, from left to right: no noise (condition 1), the first level of noise (condition 2), the second level of noise (condition 3), and the third level of noise (condition 4). Over the pitch traces, lines have been fitted by hand to signal the upper and the

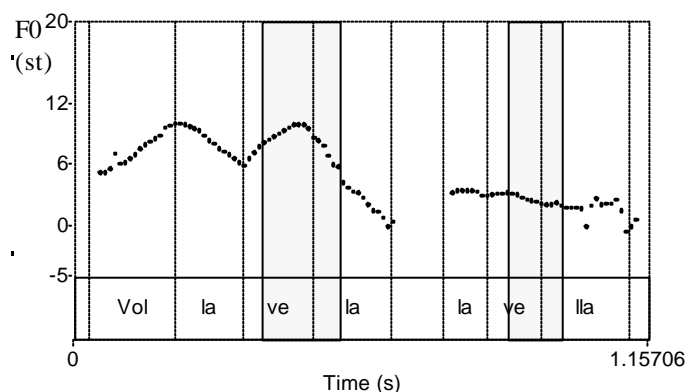
<sup>11</sup> Dr Tom Baer, from the department of Experimental Psychology, University of Cambridge, kindly facilitated producing 60 minutes of masking noise on a digital audiotape.

<sup>12</sup> Geoff Potter, technician of the Phonetics laboratory of the University of Cambridge, kindly assisted in calculating the output SPL (sound pressure level). This was calculated indirectly by measuring the average voltage level at the input of the headphone, after the DAT recorder used to play the noise was loaded with the pair of headphones. The dial was calibrated so that the noise could be increased in controlled intervals by shifting up the digital dial from point 10 (75.8 dB) to point 12 (78.8 dB), and finally to point 14 (83.4 dB). In this way the increases in the volume of the noise can be precisely monitored.

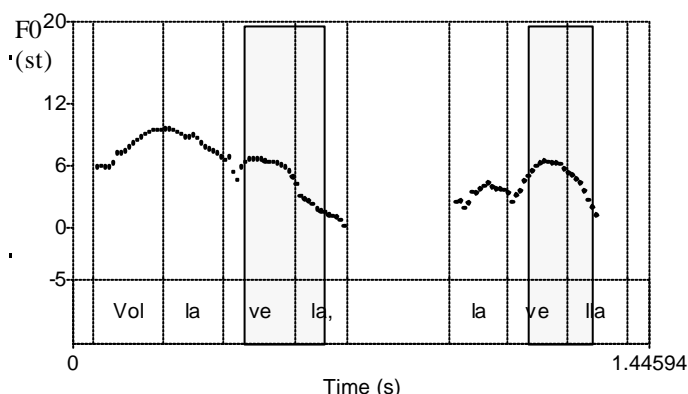
lower limits of the pitch range under condition 1 (the solid lines) and condition 4 (the dashed lines), the other two conditions triggering changes too tiny for such a crude comparison. In addition, the drawing on the right hand side of Fig. 9 schematically indicates that the pitch level has been raised. 48 sentences were measured at 6 data points and paired samples t-tests comparing condition 1 to condition 2, condition 2 to condition 3, and so forth to confirm that there is a statistically significant effect of condition 1 ( $p < 0.05$ ), that is, the first noise level, but there is no significant effect either of condition 3 or of condition 4 ( $p > 0.05$  in both cases). The difference between condition 1 and condition 4 is, as expected, significant ( $p < 0.05$ ). This indicates the presence of saturation effects conflated with the general raising of the pitch level. This outcome, the general raising in pitch level, runs contrary to expectations and does not contribute to answering the research question that drive this experiment, whether right-dislocations are deaccented. Although a more thorough analysis of all available data is needed to reach a firm conclusion, the evidence so far suggests a change in methodology, which is undertaken in Experiment 2.

The first author also carried out an acoustical and instrumental analysis of the data, 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 (version 4.1.21). This analysis was repeated twice more, with 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.

Figs 10 and 11 below show, respectively, an example of a right-dislocation and an example of an apposition:



**Fig. 10** Example of right-dislocation ‘Vol la vela, la vella’ (‘[She] wants the sail, the old lady’)



**Fig. 11** Example of apposition ‘Vol la vela, la vella’ (‘[He] wants the sail, the old sail’)

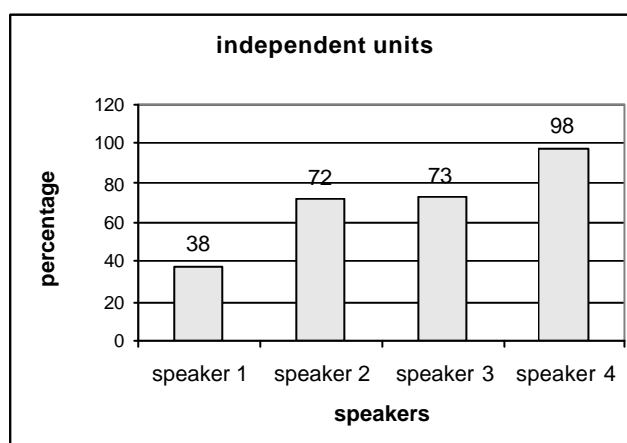
Fig. 10 presents the phrase ‘Vol la vela, la vella’ (‘She wants the sail, the old lady’), with a dislocated NP subject. This interpretation was elicited by a sentence about Mary, an old lady who has a ship in a bottle that is missing a sail, followed by the question ‘What does the old

lady want?’ The answer is the dislocated subject interpretation in Fig. 10 ‘[She] wants the sail, *the old lady*’. Fig. 11 presents the appositive interpretation of the same segmental string, ‘Vol la vela, la vella’. This is introduced by a phrase saying that a fisherman has been fixing his boat and now he is going to work on the sail, which is old. ‘What does the fisherman want?’ The answer is: ‘[He] wants the sail, *the old sail*’.

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.

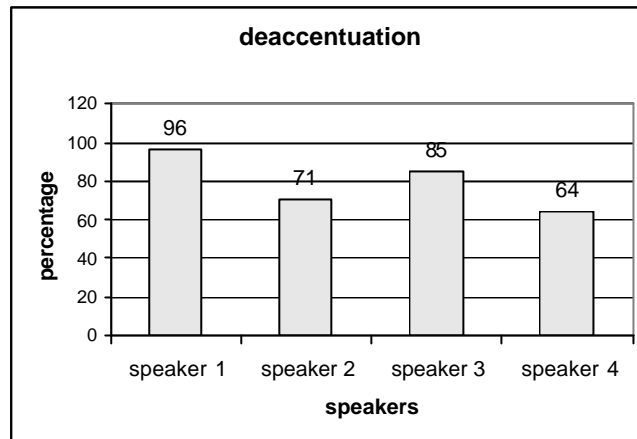
### 3.2.5 Right-dislocations: phrasing and intonation

In this section I will present an overview of the phrasing and the intonation of the 220 right-dislocated phrases. The analysis is based on acoustic and instrumental observations made by careful listening and by the examination of the pitch traces. It was decided that a right-dislocation was deaccented if it sounded less loud than the main phrase and without any perceivable pitch movement. The 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 form independent units 70% of the time. This percentage includes some sizeable degree of inter-speaker variation, as can be seen in the bar graph presented in Fig. 12:



**Fig. 12** Percentages of right-dislocated phrases forming separate units (on the x-axis) for each speaker (on the y-axis)

In Fig. 12, 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 the productions of the last speaker fall into separate units (98% of the time). 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.



**Fig. 13** Percentages of deaccented right-dislocated phrases (on the x-axis) for each speaker (on the y-axis)

Fig. 13 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% of deaccented tokens respectively.

### 3.2.6 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 speaker's 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 as 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 less prosodic breaks and deaccent more frequently, as is the case with speaker 1, the fastest reader. 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 2

### 3.3.1 Methodology

In Experiment 2, instead of using masking noise to elicit variations in pitch range, the levels of prosodic prominence were manipulated.

The experimental material (which can be found in the Appendix) was highly controlled. The target syllables were the initial syllable in disyllabic words with stress on the second syllable (as in *Vi<sup>1</sup>la*, a family name), and with stress on the first syllable (as in *<sup>1</sup>Vila*, short name name for 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 *<sup>1</sup>Vila<sup>1</sup>bella*, a place name and also the name of a club). In this way, the target syllables have identical

segmental composition but different degrees of stress – unstressed, primary stress, and secondary stress<sup>13</sup>. Therefore, the 3 stress conditions are:

- (7) *Stress conditions*
- a. Stress 0: *Vi'la* (family name)
  - b. Stress 1: *'Vila* (short name for a football club)
  - c. Stress 2: *ᵛVila'bella* (a football club)

There were twelve such words in total embedded in three right-dislocated subject phrases and in three right-dislocated object phrases.

- (8)
- a. *el Vilà, el Vila, el Vilabella*, as above
  - b. *la llimona* (lemon), *la llima* (lime), *la llimonada* (lemonade)
  - c. *la mel* (honey), *el meló* (watermelon), *la melonada* (melon juice)
  - d. *la mà* (mom), *la mare* (mother), *la Mamabona* ((imaginary) football club)

The words in (3.5a) and in (3.5b) appeared twice each, once with each type of right-dislocated sentence. The material was balanced to level out vowel-specific pitch differences (Lehiste 1970, Lehiste and Peterson 1972), 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:

- (9)
- a. *Va guanyar la lliga, el Vilabella?* (Did they win the league, *the Vilabella*)
  - b. *Va guanyar-la, el Vila* (They won it, *the 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 2003b, 2004).

It was expected that stressed syllables, including those with primary and those with secondary stress, 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.

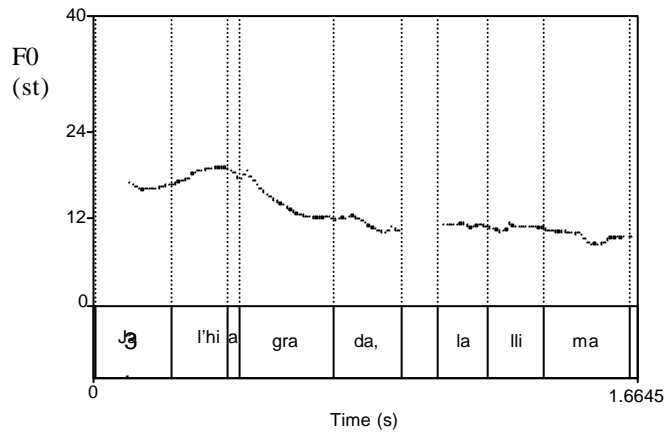
### 3.2.2 Results

With regards to phrasing, it was found that dislocations, as in Experiment 1, tend to form separate tonal units. This can be observed in the two pitch traces in Fig. 14 and 15 below, corresponding respectively to the sentences ‘*Ja l’hi agrada, la llima*’ (‘S/he likes it, lime’) and ‘*Ja l’hi agrada, la llimonada*’ (‘S/he likes it, lemonade’):

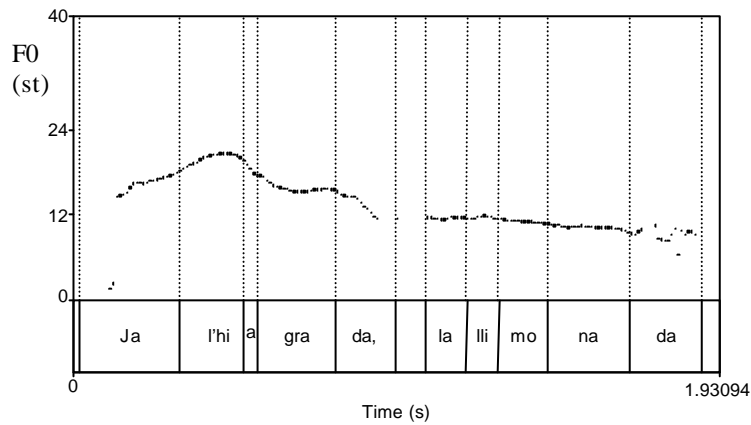
---

<sup>13</sup> According to Mascaró (2002), compound names such as ‘*Vilabella*’ and ‘*Mamabona*’ lose the stress of the first component (*'Vila-*, *'Mama-*) instead of keeping it as a secondary accent. However, Oliva and Serra (2002) propose that three or more unstressed syllables, as would be the case of ‘*el-Vi-la-'be-lla*’), are not allowed in Catalan. In order to avoid this situation, a support stress has to be re-introduced (‘*el ᵛVila'bella*’).



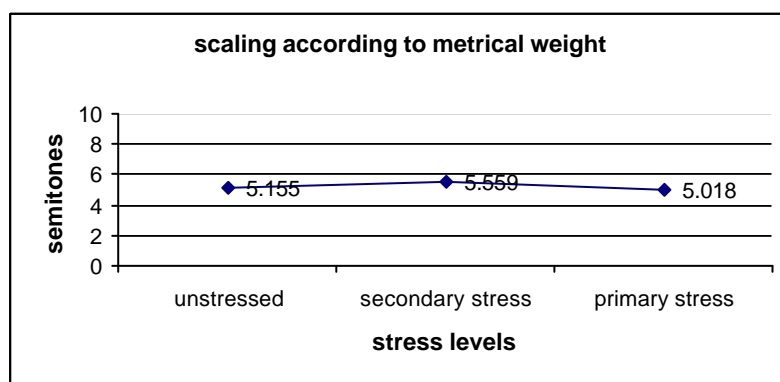


**Fig. 14** 'Ja l'hi a'grada, la 'llima' ('S/he likes it, *lime*')



**Fig. 15** 'Ja l'hi a'grada, la 'llimo'nada' ('S/he likes it, *lemonade*')

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 Fig. 16 below.



**Fig. 16** Scaling of the target syllables. Fundamental frequency in semitones (on the x-axis) and three levels of stress (on the y-axis)

As Fig. 16 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 words (*lli'mona*) and words with primary stress (*'llima*) are shorter than words with secondary stress (*llimo'nada*) and so the measurement point is earlier in the overall pitch downtrend.

### 3.2.3 Discussion

Both the auditory and the instrumental analyses of the corpus of right-dislocated phrases confirm that these structures are deaccented in Catalan. 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 rather unexpected outcome 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 that ESEs form syntactically and prosodically independent units. However, as argued in section 1.2, they do not form a homogeneous grammatical category although they have some 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?

The purpose of Experiment 1 was to establish whether ESEs always form independent tonal units and are always deaccented, or rather they 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 of signalling tonal subordination (appositions, with reduplication; right-dislocations, with deaccenting). Unfortunately, the masking noise technique used did not yield quantitative confirmation of the analysis.

Experiment 3 followed a much stricter methodology, which involved measuring the F0 excursion of syllables with identical segmental composition appearing in initial position in words such as *'llima*, *lli'mona*, and *llimo'nada*, so that the target syllables had 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.

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 belong to the same prosodic unit, and there is hence scope for ambiguity. This behaviour hints at the existence of a trade-off between rhythm and melody in the prosodic marking of ESEs, as has been observed with Catalan right-dislocations. It was also observed in 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. This trend was not confirmed statistically, perhaps owing to a certain degree of inter-speaker variation. A tightly controlled methodology is needed to find statistical confirmation.

The possibility that the prosodic behaviour of dislocations and extra-sentential elements in general might not be amenable to the influence of syntax, but rather to general principles of information structure and textual organization, was suggested by Liberman (1975: 185), and 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). However, 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.

## 5. References

- Allerton, D. J. and Cruttenden, A. 1974 English sentence adverbials: their syntax and their intonation in British English. *Lingua* 34, 1-30.
- Allerton, D. J. and Cruttenden, A. 1976. The intonation of medial and final sentence adverbials in British English. *Archivum Linguisticum* 7, 1, 29-59.
- Allerton, D. J. and Cruttenden, A. 1978 Syntactic, illocutionary, thematic, and attitudinal factors in the intonation of adverbials. *Journal of Pragmatics* 2, 155-188.
- Astruc, Ll. and Nolan, F. in press, Variation in the intonation of sentential adverbs. In T. Riad and C. Gussenhoven (eds.) *Tone and intonation*. Berlin/New York: Mouton de Gruyter.
- Astruc, Ll. 2003a Sentence-external elements in Catalan. *Catalan Journal of Linguistics. Romance intonation* 2, 15-31.
- Astruc, Ll. 2003b Right-dislocations in Catalan. *Proceedings of the 15th International Congress of the Phonetic Sciences (ICPhS)*, Barcelona, 1265-1268.
- Astruc, Ll. 2004 Right-dislocations: influence of information structure on prosodic phrasing and intonation. *Cambridge Occasional Papers in Linguistics*. Cambridge: Department of Linguistics, University of Cambridge, 1-14.
- Astruc, L. 2005a The intonation of extra-sentential elements in Catalan and English. Doctoral dissertation, University of Cambridge (available from [www.astruc.info](http://www.astruc.info))
- Astruc, L. 2005b The form and function of extra-sentential elements. *Cambridge Occasional Papers in Linguistics*, 2. Cambridge: Department of Linguistics, University of Cambridge.
- Beckman, M. and Pierrehumbert, J. 1986 *Intonational structure in Japanese and English*. *Phonology Yearbook* 3, 225-310.
- Bing, J. M. 1984 A discourse domain identified by intonation. In D. Gibbon and Richter, H. (eds.) *Intonation, accent and rhythm studies in discourse phonology*, 429-451 Berlin/New York: Mouton de Gruyter.

- Bing, J. M. 1985 *Aspects of English prosody*. New York/ London: Garland Publishing.
- Bolinger, D.W. 1954 English prosodic stress and Spanish sentence order. *Hispanica* 3, 152-156.
- Bonet, E. 1984 *Aproximació a l'entonació del català central*. Master's dissertation, Universitat Autònoma de Barcelona.
- Byrne, D., H. Dillon, K. Tran, S. Arlinger, K. Wilbraham, R. Cox, B. Hagerman, R. Hetu, J. Kei, C. Lui, J. Kiessling, M.N Kotby, N.H.A. Nasser, W.A.H El Kholy, Y. Nakanishi, H. Oyer, R. Powell, D. Stephens, R. Meredith, T. Sirimanna, G. Tavartkiladze, G.I. Frolenkov, S. Westermann and C. Ludvigsen 1994 An international comparison of long-term average speech spectra. *Journal of the Acoustical Society of America* 4, 2108-2120.
- Castellanos, A., J. M. Benedí and F. Casacuberta 1996 An analysis of general acoustic-phonetic features for Spanish speech produced with the Lombard Effect. *Speech Communication* 20, 23-25.
- Cooper, W. E; and Paccia-Cooper, J. 1980 *Syntax and speech*. Cambridge, Massachusetts/ London: Harvard UP.
- Cruttenden, A. 1997, *Intonation*. (2<sup>nd</sup> edition) Cambridge: Cambridge University Press.
- Oxford English Dictionary 1989, 2<sup>nd</sup> edition on-line.
- Delattre, P. 1972 The distinctive function of intonation. In: D. Bolinger (ed) *Intonation: selected readings*. Hamondsworth, Middlessex, England/ Baltimore, USA/Ringwood, Victoria, Australia: Penguin, 159-174.
- D'Imperio, M. P. 2002 Italian intonation: an overview and some questions. *Probus* 14, 37-69.
- Emonds, J. E. 1976 *A transformational approach to English syntax*. New York/San Francisco/London: Academic Press.
- Emonds, J. E. 1979 Appositive relatives have no properties. *Linguistic Inquiry* 102, 211-243.
- Estebas-Vilaplana, E. 2000 The Use and Realisation of Accentual Focus in Central Catalan with a Comparison to English. Doctoral dissertation, University College London.
- Fabb, N. 1990 The difference between English restrictive and non-restrictive relative clauses. *Journal of Linguistics* 26, 57-77.
- Fagyal, Z. 2002a Tonal template for background information: the scaling of pitch in utterance-medial parentheticals in French. In B. Bel and I. Marlien (eds.) *Proceedings of the 1st International Conference on on Speech Prosody*, Aix en Provence, france 279-282.
- Fagyal, Z. 2002b Prosodic boundaries in the vicinity of utterance-medial parentheticals in French. *Probus* 14, 93-111.
- Firbas, J. 1964 On defining the theme in functional sentence analysis. *Travaux linguistics de Prague 2. L'école de Prague d'aujourd'hui*. Prague: Éditions de l'Académie Tchèque des Sciences, 267-280.
- Geluykens, R. 1992 *From discourse process to grammatical construction: on left-dislocation in English*. Amsterdam: John Benjamins.
- Geluykens, R. 1994 *The pragmatics of discourse anaphora in English: evidence from conversational repair*. Berlin: Mouton de Gruyter.

- Grice, M. 1995 *The intonation of interrogation in Palermo Italian: implications for intonation theory*. Niemeyer: L.A. series.
- Gussenhoven, C. 1985 The intonation of George and Mildred: post-nuclear generalizations. In Johns-Lewis, C. ed *Intonation in discourse* pp. 77-123 San Diego: College Hill.
- Gussenhoven, C. 1993 *On the grammar and semantics of sentence adverbs*. Dordrecht: Foris.
- Gussenhoven, C. 2004 *The phonology of tone and intonation*. Cambridge: Cambridge University Press.
- Hansen, G. F. (2002) A preliminary investigation of intonation in Danish sentences with parenthetical phrases. *Proceedings from Fonetik 2002*, KTH Stockholm, in *Speech Music and Hearing, Quarterly progress and status report*, TMH-QPSR 44, 169-172.
- Hoye, L. 1997 *Adverbs and modality in English*. English language series. London/New York: Longman.
- Huddleston, R. and Pullum, G. K. 2002) *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
- Junqua, J. C. 1996 The influence of acoustics on speech production: a noise-induced stress phenomenon known as the Lombard reflex. *Speech Communication* 20: 13-22.
- Junqua, J.C., S. Finke, and K. Field 1999 The Lombard effect: a reflex to better communicate with others in noise. *Proceedings International Conference on Acoustics, Speech and Signal Processing* 1999.
- Kempson, R., Meyer-Viol, W., and Gabbay, D. 2000 *Dynamic syntax: the flow of language understanding*. Oxford: Blackwell.
- Ladd, D. R. 1980 *The structure of intonational meaning: evidence from English*. Bloomington/London: Indiana University Press.
- Ladd, D. R. 1996 *Intonational phonology*. Cambridge: Cambridge University Press.
- Lambrecht, K. 1981 *Topic, antitopic, and verb agreement in non-standard French*. Amsterdam: J. Benjamins.
- Lambrecht, K. 1994 *Information structure and sentence form: topic, focus, and the mental representation of discourse referents*. Cambridge Studies in Linguistics. Cambridge: Cambridge University Press.
- Lane, H. and B. Tranel 1971 The Lombard Sign and the role of hearing in speech. *Journal of Speech and Hearing* 14: 677-709.
- Lehiste, I. 1970 *Suprasegmentals*. Cambridge, MA/London, England: The MIT Press.
- Lehiste, I. and G.E. Peterson 1972 Some basic considerations in the analysis of intonation. In D. Bolinger (ed) *Intonation: selected readings*. Harmondsworth, Middlessex, England/ Baltimore, USA/Ringwood, Victoria, Australia: Penguin Books, 367-384.
- Liberman, M. 1975 *The intonational system of English*. Bloomington, Indiana: Indiana University Linguistics Club.
- López, A. and Morant, R. 2002 L'adverbi. In Solà, J., Lloret, M. R., Mascaró, J. and Pérez Saldanya, M. eds) *Gramàtica del català contemporani* pp. 1797-1852 Barcelona: Editorial Empúries.
- Martin, P. 1987 Prosodic and rhythmic structures in French. *Linguistic Inquiry* 8, 249- 33.

- Mascaró, J. 2002 Reduïció vocàlica. In J. Solà, M. R. Lloret, J. Mascaró and M. Pérez Saldanya (eds) *Gramàtica del català contemporani*. Barcelona: Editorial Empúries, 89-123.
- Nespor, M. 1993 *Fonologia: le strutture del linguaggio*. Bologna: il Mulino. Publications.
- Nespor, M. and Vogel, I. 1986 *Prosodic phonology*. Dordrecht: Foris Publications.
- Nolan, F. and H. K. Jónsdóttir 2001 Accentual patterns in Icelandic. In van Dommelen, W. A. and T. Fretheim (eds) *Nordic Prosody. Proceedings of the VIII Conference, Trondheim 2000*, 187-198. Frankfurt am Main: Peter Lang.
- Oliva, S. and P. Serra 2002 Accent. In J. Solà, M. R. Lloret, J. Mascaró and M. Pérez Saldanya (eds) *Gramàtica del català contemporani*. Barcelona: Editorial Empúries, 345-391.
- Payà, M. 2002 Incidental clauses in spoken Catalan: prosodic characterization and pragmatic function. In B. Bel and I. Marlien, (eds). *Proceedings 1<sup>st</sup> International Conference on Speech Prosody*, Aix-en-Provence, France, 559-562.
- Payà, M. 2003 Prosody and pragmatics in parenthetical insertions in Catalan. *Catalan Journal of Linguistics 2. Special issue on Romance intonation*, 207-227.
- Pierrehumbert, J. 1980 The phonetics and phonology of English intonation. Doctoral dissertation, MIT.
- Pierrehumbert, J. and Hirschberg, J. 1990 The meaning of intonational contours in the interpretation of discourse. In P.R. Cohen, J. L. Morgan and M. E. Pollack (eds) *Intentions in communication*. Cambridge, Massachusetts/London, England: The MIT Press, 271-311.
- Prieto, P. 2002a *Entonació: models, teoria, mètodes*. Editorial Ariel: Barcelona.
- Prieto, P. 2002b Entonació. In J. Solà, M., Lloret, R., Mascaró, J. and Pérez Saldanya, M. (eds.) *Gramàtica del català contemporani* pp. 393-462 Barcelona: Editorial Empúries.
- Recasens, D. 1993 *Fonètica i fonologia*. Barcelona: Enciclopèdia Catalana.
- Safir, K. 1986 Relative clauses in a theory of binding and levels. *Linguistic Inquiry* 17, 663-683.
- Sag, I. A. 1997 English relative clause constructions. *Journal of Linguistics* 33, 431-483.
- Selkirk, E. 1984 *Phonology and syntax: the relation between sound and structure*. Cambridge, Massachusetts: The MIT Press.
- Truckenbrodt, H. 1995 *Phonological phrases: their relation to syntax, focus, and prominence*. Doctoral dissertation, MIT.
- Truckenbrodt, H. 1999 On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry* 30 (2), 219-255.
- Trim, J.L. 1959 Major and minor tone groups in English. *Le Maître Phonétique* 112, 26-29.
- Vallduví, E. 1990 The role of plasticity in the association of focus and prominence. *Proceeding of the Eastern States Conference on Linguistics (ESCOL)* 7, 295-306.
- Vallduví, E. 1994 Detachment in Catalan and information packaging. *Journal of Pragmatics* 22, 573-601.
- Wunderli, P. 1987 *L'intonation des sequences extraposées en français*. Tübingen: Gunter Narr Verlag.

## Appendix

### Experiment 2

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

#### Appositions

Mira, allí hi ha el Joan. I qui és la senyora de vermell que està al seu costat?

- És la Mona, *la dona*. Sí, es diu així: Mona.

(Look, there's John. Who's the lady in red that is next to him? She is Mona, *the wife*. Yes, that's her name, Mona.)

La mama va veure els nuvis, abans de la boda?

- Va veure la Núria, *la núvia*.

(Did Mum see the bride and the groom before the wedding? She saw Núria, *the bride*)

Saps quin poble és aquest?

- És Mora, *la Nova*. No l'altra Mora, la Mora d'Ebre.

(Do you know which town this is? This is Mora, *la Nova*)

Fa temps que vull anar a veure un drama, 'La Mama', al teatre. No sé si ja el fan o encara fan la comèdia tonta aquella...

-Fan el drama, '*La Mama*'.

(I've been looking forward to going to the theatre to see a drama, 'The Mum'. Is it on or are they still that silly comedy? They are showing the drama, '*The Mum*'.)

Els veïns s'han separat. I tenen una filla, la Lena. El mare ha portat el pare a judici. Què vol la mare?

- Vol la nena, *la Lena*.

(The neighbours have split up, and they have a little girl, Lena. The mother has brought the father to court. What does the mother want? She wants the little girl, *Lena*.)

El pescador està treballant a la barca i ara ha d'arreglar la vela, que és vella. Què vol el pescador?

- Vol la vela, *la vella*.

(The fisherman has been fixing his boat. Now he's going to do some work on the sail, which is old. What does he want? He wants the sail, *the old one*.)

No sabien si posar una comèdia o un drama, al cine club de la escola. Què han posat finalment?

- Han posat Dallas, *el drama*.

#### Dislocations

Com és la dona de l'Emili?

- És mona, *la dona*. Però no sembla pas gaire agradable.

(What is Emili's wife like? She's cute, *the wife*. But she doesn't seem very nice.)

Qui és l'amiga que va anar a veure la núvia?

- Va veure la Núria, *la núvia*.

(Who's the friend that the bride went to see? She saw Núria, *the bride*)

Has vist la noia nova. Saps com és?

-És mora, *la nova*.

(Have you seen the new girl? What is she like? She's a moor, *the new one*.)

Què li passa a la mama, que xiscla tant? No li facis cas. No té res.

-Fa drama, *la mama*.

(What's the matter with mother, who's screaming so much? Don't pay any attention to her. She's fine. She's making a drama out of it, *mother*.)

La Lena està de 3 mesos. Encara no saben si és nen o nena. Què vol ella?

- Vol una nena, *la Lena*.

Lena is three months pregnant. They don't know yet whether it's a boy or a girl. What does she want? She wants a girl, *Lena*.)

La vella senyora Maria ha comprat un vaixell a dins d'una ampolla per al menjador de casa, però al vaixell li falta la vela. Què vol la vella?

- Vol la vela, *la vella*.

(Old Mary bought a ship in a bottle for her dinning room. But the new ship is missing the sail. What does she want? She wants the sail, *the old lady*.)

Recordes aquella història terrible? Sí, tot un drama: la dona que va assessorar l'home i després es va penjar? On va passar?

- Va passar a Dallas, *el drama*.

(At the cinema club of the school they weren't sure whether to show a comedy or a dram. What have they shown finally? They have shown Dallas, *the drama*.) (Do you remember that terrible story? It was a real drama: the wife murdered the husband and then she hanged herself. Where did it happen? It happened in Dallas, *the drama*.)

### Experiment 3

Right-dislocated phrases in Catalan with translations in English

Ja l'hi agrada, la mel (S/he likes it, *honey*)

Ja l'hi agrada, el meló (S/he likes it, *watermelon*)

Ja l'hi agrada, la melonada (S/he likes it, *watermelon juice*)

Ja en vol, de mel (S/he wants it, *honey*)

Ja en vol, de meló (S/he wants it, *watermelon*)

Ja en vol, de melonada (S/he wants it, *melon juice*)

Va guanyar-la, la mare (She won it, *the mother*)

Va guanyar-la, la mamà (S/he won it, *the mamà*)

Va guanyar-la, la Mamabona (They won it, *the Mamabona*)

Ja l'hi agrada, la llima (S/he likes it, *lime*)

Ja l'hi agrada, la llimona (S/he likes it, *lemon*)

Ja l'hi 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*)

Va guanyar-la, el Vilà (He won it, *Vilà*)

Va guanyar-la, el Vilabella (They won it, *the Vilabella*)