

***THE USE AND REALISATION OF ACCENTUAL FOCUS  
IN CENTRAL CATALAN  
WITH A COMPARISON TO ENGLISH***

**Eva Estebas-Vilaplana**

Submitted for the degree of PhD in Phonetics

**University College London**

2000

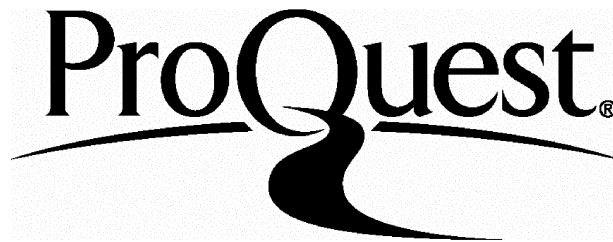
ProQuest Number: 10013937

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10013937

Published by ProQuest LLC(2016). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code.  
Microform Edition © ProQuest LLC.

ProQuest LLC  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106-1346

*To my parents, Toni and Dolors  
To my grandma, Coloma*

*To Rafa*

## ***Abstract***

The aim of this thesis is threefold: 1) to investigate to what extent intonational (accentual) strategies are used in Central Catalan to signal the information structure of sentences as compared to syntactic devices, 2) to describe the phonetic and phonological properties of broad and narrow focus sentences in Central Catalan, and 3) to compare the results of Central Catalan with those of a language that mainly uses accentual focus, such as English.

The first question is tested by means of a production test, a perception test and an acceptability test. Sentences produced with a broad focus and a narrow focus reading are analysed. Within narrow focus, two kinds of contexts are included: focus triggered by contrast and by identification. The results of the three tests show that Central Catalan speakers can both produce and recognise focus conveyed by intonational means and that the use of accentual devices is perceived as natural.

The analysis and interpretation of the production data is done within the Autosegmental-Metrical framework of intonational analysis for both broad and narrow focus sentences. The results show that Central Catalan associates tonal primitives to metrically strong syllables and to the edges of three higher levels of prosodic structure: the accented word, the intermediate phrase and the intonation phrase. The accentual cues to signal narrow focus involve the location of a pitch accent on the stressed syllable of the focal element and the introduction of an intermediate phrase boundary at the end of the focal domain. The deaccentuation of postfocal material is optional and varies in form (from reduction to absence of pitch accent). Broad and narrow focus sentences use the same nuclear accent type. Scaling differences are not incorporated into the phonological system but are a matter of choice on the speaker's part (i.e. emphasis).

Finally, the expression of focus in Central Catalan is similar to English in that the two languages use the same nuclear tone for both broad and narrow focus. However, differences in phrasing and in deaccenting are detected between the two languages. Whereas Central Catalan introduces a prosodic boundary after the focussed element and deaccenting is optional, no boundary is observed in English and deaccentuation is compulsory.

## *Acknowledgements*

I would like to thank my supervisor John Maidment for his guidance and support throughout these years, and for giving me the opportunity to work in the Department of Phonetics and Linguistics of UCL, which allowed me to carry on with this thesis.

I am also highly indebted to Pilar Prieto for her generous and inspiring discussions, for her comments on early drafts of this thesis and for making her work available to me so quickly. I also want to express my gratitude to several people who allowed me to discuss some parts of my work with them: Jill House, Mercedes Cabrera-Abreu, Bob Ladd, Maria Black and Mitsuhiro Nakamura. I also owe many thanks to Maria Josep Solé, who first introduced me to phonetics and who has always encouraged me with my work.

Special thanks go to Steve Nevard for his kind, prompt and cheerful help with technical and computer assistance and for his invaluable patience whenever I got stuck. Also many thanks to Judith Crompton and Molly Bennett for their encouragement at all times.

I am very grateful to many people in Ripoll, my village, who kindly offered me their time to be my informants. Unfortunately, I cannot mention them all due to the amount of people involved. However, special thanks go to Cristina Amils and Núria Gili for allowing me to do some of the recordings in their houses. I also thank my English subject, Karen Froud, for her quick response in doing the recordings.

The work presented in this thesis has been partially supported by two grants: Beca Batista-Roca from la Generalitat de Catalunya, and British Academy Postgraduate Studentship.

Finally, I thank my parents, Toni and Dolors, and my grandmother, Coloma, for their invaluable and ceaseless support during this work and throughout all my life, and Rafa Ávila, without whom I would have never reached the end.

## ***Table of contents***

<i>List of Figures</i>	10
<i>List of Tables</i>	17
<b>CHAPTER 1. INTRODUCTION</b>	<b>20</b>
1.1 Introduction	20
1.2 Focus	21
1.2.1 Semantic/pragmatic approach	21
1.2.2 Phonetic/phonological approach	23
1.2.3 Focus and normal stress	25
1.2.4 Ways of expressing focus: phonological vs syntactic devices	28
1.3 Modelling intonation	31
1.3.1 Traditional analyses	32
1.3.1.1 The British tradition	32
1.3.1.2 The American tradition	34
1.3.2 Autosegmental-Metrical Phonology	35
1.3.2.1 Background	35
1.3.2.2 Phonological units	39
1.3.2.3 Levels of prosodic structure	44
1.3.2.4 Pitch downtrends	45
1.4 Catalan	47
1.4.1 Stress language	48
1.4.2 Intonation language	49
1.4.2.1 Neutral declarative intonation	51
1.4.2.2 Marked declarative intonation	55
1.5 Conclusion and outline	56
<b>CHAPTER 2: ACCENTUAL FOCUS IN CENTRAL CATALAN</b>	<b>58</b>
2.1 Introduction	58
2.2 Accentual focus cross-linguistically	59
2.3 Production test	62
2.3.1 Experimental design	63

2.3.1.1	Materials	63
2.3.1.2	Data collection procedures	67
2.3.1.3	Informants	68
2.3.1.4	Instrumentation and recording	69
2.3.2	Analysis of the data	70
2.3.2.1	Auditory analysis	71
2.3.2.2	Acoustic analysis	73
2.3.3	Results	77
2.3.3.1	Reading test	77
2.3.3.2	Semi-spontaneous test	80
2.4	Perception test	84
2.4.1	Experimental design	84
2.4.1.1	Materials	84
2.4.1.2	Editing the stimuli	85
2.4.1.3	Informants	86
2.4.1.4	Testing	86
2.4.2	Results	87
2.5	Acceptability test	90
2.5.1	Experimental design	91
2.5.1.1	Materials	91
2.5.1.2	Stimuli	92
2.5.1.3	Informants	93
2.5.1.4	Testing	93
2.5.2	Results	94
2.6	Conclusion	103
 <b>CHAPTER 3. BROAD FOCUS</b>		<b>105</b>
3.1	Introduction	105
3.2	Framework	106
3.2.1	Phonological structure	107
3.2.2	Phonetic variability	110
3.3	Data	112
3.4	Results	113

3.4.1 General properties	114
3.4.2 Acoustic analysis	118
3.4.2.1 Phrasing	118
3.4.2.2 Non-final accents (prenuclear accents)	124
3.4.2.2.1 Start of the rise	127
3.4.2.2.2 Location of the peak	129
3.4.2.3 Sentence-final accent (nuclear accent)	143
3.5 Discussion	151
3.6 Conclusion	157
<b>CHAPTER 4. NARROW FOCUS</b>	<b>159</b>
4.1 Introduction	159
4.2 Phonetic and phonological cues of narrow focus	160
4.3 Data	164
4.4 Results	166
4.4.1 Accentual focus	166
4.4.1.1 Early narrow focus	166
4.4.1.1.1 The nuclear accent	169
4.4.1.1.2 Prenuclear accents	179
4.4.1.2 Late narrow focus	180
4.4.1.3 Levels of phrasing	185
4.4.1.4 Postfocal accents	190
4.4.2 Syntactic focus	192
4.5 Discussion	194
4.6 Conclusion	197
<b>CHAPTER 5. F0 DOWNTRENDS</b>	<b>198</b>
5.1 Introduction	198
5.2 Global and local interpretations of F0 downtrends	199
5.3 The scaling of peaks	205
5.3.1 Broad focus sentences	205
5.3.1.1 Data	205



5.3.1.2 Results	206
5.3.1.2.1 Utterance initial and final F0 values and pitch range	208
5.3.1.2.2 Effects of temporal distance between peaks	210
5.3.1.2.3 Predicting F0 values	216
5.3.1.2.4 H with final lowering or L target?	219
5.3.1.3 Discussion	222
5.3.2 Narrow focus sentences	223
5.3.2.1 Data	224
5.3.2.2 Results	225
5.3.2.2.1 Effects of narrow focus on downstep	228
5.3.2.2.2 Predicting F0 values	231
5.3.2.3 Discussion	235
5.4 The scaling of valleys in broad focus and in narrow focus sentences	236
5.4.1 Data	237
5.4.2 Results	238
5.4.3 Discussion	239
5.5 Conclusion	239
<b>CHAPTER 6. COMPARISON TO ENGLISH</b>	<b>242</b>
6.1 Introduction	242
6.2 Accentual focus in English and in Central Catalan	242
6.3 Experimental design	245
6.3.1 Materials	245
6.3.2 Informant	247
6.3.3 Data collection procedures	247
6.3.4 Instrumentation and recording	247
6.3.5 Acoustic analysis	248
6.4 Results	248
6.4.1 Phonetic and phonological properties	248
6.4.1.1 Broad focus	248
6.4.1.1.1 Prenuclear accents	250
6.4.1.1.2 Nuclear accents	258
6.4.1.2 Narrow focus	261

6.4.2 Downtrends	269
6.4.2.1 Broad focus	271
6.4.2.2 Narrow focus	274
6.5. Discussion	276
6.6 Conclusion	280
<b>CHAPTER 7. SUMMARY AND CONCLUSIONS</b>	<b>282</b>
7.1 Introduction	282
7.2 Summary	284
7.3 Implications	288
7.4 Further research	289
<b>REFERENCES</b>	<b>291</b>
<b>APPENDIX 1</b>	<b>305</b>
<b>APPENDIX 2</b>	<b>310</b>
<b>APPENDIX 3</b>	<b>313</b>
<b>APPENDIX 4</b>	<b>316</b>
<b>APPENDIX 5</b>	<b>321</b>

## *List of Figures*

2.1. SFS display screen showing 1) the speech waveform, 2) the Lx waveform, 3) the excitation period measurements and 4) the F0 contour for the sentence <i>l' au venia de l'illa</i> produced with broad focus.....	72
2.2. Speech waveform and F0 trace for the sentence <i>L'AU venia de l'illa</i> with narrow focus on the subject.....	72
2.3. Speech waveform and F0 trace for the sentence <i>la Neus ANIMA els alumnes</i> with narrow focus on the verb.....	73
2.4. Speech waveform and F0 trace for the sentence <i>la Neus ANIMA els alumnes</i> with narrow focus on the verb and a postnuclear accent.....	73
2.5 Table and pie chart showing the number of responses produced with accentual narrow focus and with broad focus for the whole reading test.....	75
2.6. Table and histograms showing the number of accentual focus responses and broad focus responses for the variables subject focus vs verb focus and focus triggered by contrast vs focus triggered by identification obtained in the reading test .....	76
2.7. Table and pie chart showing the distribution of responses obtained in the semi-spontaneous test.....	79
2.8. Table and histograms showing the number of accentual focus responses and broad focus responses for the variables subject focus vs verb focus and focus triggered by contrast vs focus triggered by identification obtained in the semi-spontaneous test.....	80
2.9. Table and pie chart showing the number of accentual focus responses, broad focus responses and unexpected responses for the whole perception test.....	85
2.10. Table and histograms showing the number of accentual focus, broad focus and unexpected responses in the perception test for the variables subject focus vs verb focus and contrastive focus vs identification focus.....	86
2.11. Table and histograms showing the number of responses for the whole acceptability test according to three categories: A = narrow focus conveyed	

intonationally, B = narrow focus conveyed syntactically, and C = broad focus. The acceptability scale ranges from 4 (perfectly acceptable) to 0 (totally unacceptable)..... 92

2.12. Tables and histograms showing the number of acceptable/unacceptable responses for sentences with narrow focus on the subject and on the verb..... 94

2.13. Tables and histograms showing the number of acceptable/unacceptable responses for sentences with narrow focus triggered by contrast and by identification..... 96

3.1. Mean frequency value (in Hz) of peak 1 (p1) and peak 2 (p2) in double accented subject domains for four speakers..... 116

3.2. Mean duration (in seconds) of the first (non-preboundary) and second (preboundary) accented syllables in double stressed subject domains for four speakers..... 117

3.3. Mean F0 values at the onset (of) and offset (of) of prenuclear accented syllables and at the peak (p) for all speakers. The lines below and above the squares indicate the mean values less and plus the standard deviation respectively..... 124

3.4. Distance between the onset of the accented syllable and the peak (on-peak) against distance between the onset and the offset of the accented syllable (on-of) in seconds for all prenuclear accents of all speakers..... 127

3.5. Distance between the onset of the accented syllable and the peak (on-peak) against distance between the onset of the accented syllable and the end of the word (on-end) in seconds for all speakers in paroxytones..... 130

3.6. Distance between the onset of the accented syllable and the peak (on-peak) against distance between the onset of the accented syllable and the end of the word (on-end) in seconds for all speakers in oxytones..... 131

3.7. Distance between the onset of the accented syllable and the peak (on-peak) against distance between the onset of the accented syllable and the end of the word (on-end) in seconds for speaker DV in proparoxytones..... 133

3.8. Distance between the onset of the accented syllable and the F0 valley (on-valley) against distance between the onset of the accented syllable and the end of the word (on-end) for the first accent of a double stressed subject produced with a falling pitch..... 138

4.1. Distance between the onset of the accented syllable and the valley (on-valley) against distance between the onset of the accented syllable and the end of the word (on-end) in seconds for focal oxytones in subject and verb focal domains for all speakers.....	169
4.2. Distance between the onset of the accented syllable and the valley (on-valley) against distance between the onset of the accented syllable and the end of the word (on-end) in seconds for focal paroxytones in subject and verb focal domains for all speakers.....	170
4.3. Distance between the onset of the accented syllable and the valley (on-valley) against distance between the onset of the accented syllable and the end of the word (on-end) in seconds for focal proparoxytones in subject and verb focal domains for speaker DV.....	171
4.4. Mean F0 values of the valley between two peaks separated by a different number of intervening syllables (1-2 or 3-4 syllables) for all speakers.....	173
4.5. Distance between the onset of the accented syllable and the peak (on-peak) against distance between the onset of the accented syllable and the end of the word (on-end) for prenuclear oxytones and paroxytones in sentences with a narrow focus reading produced by speakers CP and NG.....	175
4.6. Mean duration of syllable 1 (non-preboundary) and syllable 2 (preboundary) in seconds for all speakers.....	181
5.1. Mean schematised F0 contours for neutral declarative sentences with different number of potential stresses and actual accents (in brackets) for four speakers.....	202
5.2. Mean values (in Hz) of utterance-initial and utterance-final F0 values of sentences of different lengths (3 to 4 stressed syllables).....	204
5.3. Mean F0 values for H1 and H2 corresponding to an increase in the number of intervening syllables (1-2 or 3-5) and in different sentence conditions for speakers NG and DV.....	209
5.4. Mean F0 values for H1 and H2 corresponding to an increase in the number of intervening syllables (3-5 or 6-9) and in different sentence conditions for speakers CP and MC.....	211
5.5. Predicted F0 values of a given peak (H2 or H3) as opposed to the	

observed F0 values of the same peak for speakers NG and DV. The line indicates $x = y$ .....	214
5.6. Mean F0 values of the last H (measured as the F0 height at the onset of the accented syllable) and the last L of a neutral declarative contour.....	215
5.7. Predicted F0 values for the last H of an utterance as opposed to the real F0 values observed in the data for speakers NG and DV. The line indicates $x = y$ .....	217
5.8. Schematised F0 contours for sentences produced with focus in medial position triggered by contrast and identification.....	221
5.9. Schematised F0 contours for sentences produced with focus in final position triggered by contrast and identification.....	222
5.10. Mean H1 and H2 for medial narrow focus declaratives triggered by a contrastive and an identification prompt.....	224
5.11. Mean H1 and H2 for final narrow focus declaratives triggered by a contrastive and an identification prompt.....	225
5.12. Observed against predicted prenuclear H2 values in verb narrow focus sentences for speaker NG. The line indicates $x = y$ .....	227
5.13. Observed against predicted prenuclear peak values for H2 and H3 in sentences with narrow focus on the object for speaker NG. The line indicates $x = y$ .....	228
5.14. Observed against predicted values for H in sentences with narrow focus in medial position produced with no emphasis (speaker NG) and with emphasis (speaker DV). The line indicates $x = y$ .....	229
5.15. Prediction with final lowering ratio of peak on focus on the object for speakers NG (both contrast and identification) and DV (only identification sentences).....	229
5.16. Mean F0 values of the valley between two peaks separated by a different number of intervening syllables (1-2 and 3-5) for speakers DV and NG in broad focus sentences.....	233
6.1. Mean values and standard deviations for the scaling of L as a function of the number of unaccented syllables between accents.....	247
6.2. Mean values for the duration between the onset of the first accent and L (on1-valley) and for the duration between L and the onset of the second	

accent (valley-on2) as a function of the number of unaccented syllables between accents.....	249
6.3. Distance between the onset of the accented syllable and the peak (on-peak) against distance between the onset and the offset of the accented syllable (on-of) in seconds for prenuclear accents in different kinds of words (oxytones, paroxytones and proparoxytones).....	251
6.4. Mean values and standard deviations for the scaling of L as a function of the number of unaccented syllables between accents in sentences with narrow focus on the second element.....	259
6.5. Mean values for the duration between the onset of the first accent and L (on1-valley) and for the duration between L and the onset of the second accent (valley-on2) as a function of the number of unaccented syllables between accents in narrow focus.....	260
6.6. Distance between the onset of the accented syllable and the valley (on-valley) against distance between the onset of the accented syllable and the end of the word (on-end) in seconds for nuclear accents in different kinds of words (oxytones, paroxytones and proparoxytones).....	262
6.7. Mean schematised F0 contours for neutral declarative sentences with different number of potential stresses and actual accents (in brackets).....	267
6.8. Mean F0 values for H1 and H2 corresponding to an increase in the number of intervening syllables (1-2, 3-4 or 5-8).....	268
6.9. Mean F0 values for H1 and H2 in sentences with 3 stresses but 2 accents (produced with a different number of unaccented syllables between accents) and mean F0 values for H1, H2 and H3 in sentences with 4 stresses but 3 accents.....	269
6.10. Schematised F0 contours for sentences produced with focus in medial position and in final position triggered by contrast and by identification.....	270
6.11. Mean H1 and H2 for medial and final narrow focus declaratives triggered by a contrastive and an identification prompt.....	271
A.1. Figure for example (3.4). <i>La <u>Mila</u> <u>nega</u> la <u>maionesa</u></i> (speaker CP).....	317
A.2. Figure for example (3.5). <i>La <u>Mila</u> <u>nega</u> la <u>maionesa</u></i> (speaker NG).....	318
A.3. Figure for example (3.6). <i>Les <u>meves</u> <u>amigues</u> <u>eren</u> <u>bones</u> <u>nenes</u></i> (speaker NG).....	318

A.4. Figure for example (3.7). <i>Menjaven</i> (speaker CP).....	319
A.5. Figure for example (3.13a). <i>L'home venia llimones madures</i> (speaker ER).....	319
A.6. Figure for example (3.13b). <i>L'home venia llimones madures</i> (speaker CP).....	320
A.7. Figure for example (3.14). <i>Les noies vénen de veure l'examen</i> (speaker ER).....	320
A.8. Figure for example (3.19). <i>En Ramon mira les novel·les</i> (speaker DV)....	321
A.9. Figure for example (3.20). <i>La Mila mira les novel·les</i> (speaker DV).....	321
A.10. Figure for example (3.21). <i>La Melanie mira les novel·les</i> (speaker DV)..	322
A.11. Figure for example (3.23a). <i>La nena morena ...</i> (speaker NG).....	322
A.12. Figure for example (3.23b). <i>La nena morena ...</i> (speaker NM).....	323
A.13. Figure for example (4.6). <i>L'EVA guarda monedes romanes</i> (speaker NG).....	323
A.14. Figure for example (4.7). <i>En Joan DOMINA l'àlgebra</i> (speaker NG).....	324
A.15. Figure for examples (4.8) and (4.25). <i>ELL anava a Girona</i> (speaker DV).....	324
A.16. Figure for examples (4.9) and (4.26). <i>L'Elena mira de MOURE el mobiliari</i> (speaker DV).....	325
A.17. Figure for example (4.10). <i>En Juli hauria de DEMANAR l'hora</i> (speaker CP).....	325
A.18. Figure for example (4.11). <i>L'Elena volia MUNYIR l'ovella</i> (speaker NG).....	326
A.19. Figure for example (4.12). <i>L'Elena volia MUNYIR l'ovella</i> (speaker DV).....	326
A.20. Figure for example (4.14). <i>NOU miraven el mar</i> (speaker DV).....	327
A.21. Figure for example (4.15). <i>La MARE menja arengades</i> (speaker DV)....	327
A.22. Figure for example (4.16). <i>El NÒMADA menjava al carrer</i> (speaker DV).....	328
A.23. Figure for example (4.17). <i>La Glòria ve de VILANOVA</i> (speaker CP)....	328
A.24. Figure for example (4.18). <i>La Rosa llegia LLIBRES</i> (speaker CP).....	329
A.25. Figure for example (4.19). <i>La Glòria ve de VILANOVA</i> (speaker CP)....	329
A.26. Figure for example (4.28). <i>És l'ISIDRE, qui du una gavardina vermella</i>	



(speaker DV).....	330
A.27. Figure for example (4.29). <i>De robes negres, en duia l'AMÈLIA</i>	
(speaker DV).....	330
A.28. Figure for example (4.30). <i>Hi bullia l'OU, a l'olla</i> (speaker DV).....	331
A.29. Figure for example (5.1). <i>Les nenes haurien de donar l'enhora bona</i>	
(speaker NG).....	331
A.30. Figure for example (6.4). <i>Amelia married a marine</i> (speaker KF).....	332
A.31. Figure for example (6.5). <i>Jo relies on my money</i> (speaker KF).....	332
A.32. Figure for example (6.6). <i>The marmalade</i> (speaker KF).....	333
A.33. Figure for example (6.12). <i>MILLIE lives in Ireland</i> (speaker KF).....	333
A.34. Figure for example (6.13). <i>The neighbour MARRIED the model</i>	
(speaker KF).....	334
A.35. Figure for example (6.14). <i>John married MELANIE</i> (speaker KF).....	334
A.36. Figure for example (6.15). <i>Gary borrowed the MONEY</i> (speaker KF)....	335

## *List of Tables*

2.1. Number of target sentences, distractors and trial sentences recorded for each speaker.....	63
2.2. Overall mean acceptability value for accentual and syntactic focus and results of the t-tests for the whole test and for the different variables.....	98
2.3. Overall mean acceptability value and results of the t-tests comparing the acceptability of accentual, syntactic and broad focus in subject and verb focus domains.....	99
2.4. Overall mean acceptability value and results of the t-tests comparing the acceptability of accentual, syntactic and broad focus in contrast and in identification focus.....	99
3.1. Difference between the mean distance “on-end” and the mean distance “on-peak” in seconds.....	134
3.2. Mean values of F0 peak position (expressed as % of the foot duration), foot duration, and distance between the onset of the accented syllable and the peak (on-p) in seconds.....	136
3.3. Mean values of F0 peak position (expressed as % of the syllable duration), syllable duration (on-of) and distance between the onset of the accented syllable and the peak (on-p) in seconds for words produced in isolation.....	140
3.4. Mean values of F0 peak position (expressed as % of the syllable duration), syllable duration (on-of) and distance between the onset of the accented syllable and the peak (on-p) in seconds for paroxytones produced emphatically.....	145
4.1. Results of the t-tests comparing the F0 values of valleys with a different number on intervening syllables between peaks for all speakers.....	174
4.2. Results of the t-test comparing the mean duration of syllable 1 (non-preboundary) and syllable 2 (preboundary) for all speakers.....	181
5.1 Mean F0 range in utterances with three and four stresses and results of t-tests comparing the F0 range values for all speakers.....	205
5.2. Mean durational distance (in seconds) between peaks with 1-2 intervening syllables and with 3-5 intervening syllables in different kinds of	

accented utterances.....	207
5.3. Mean F0 values of H2 preceded by 1-2 or 3-5 unaccented syllables for the two speakers in different sentence conditions.....	208
5.4. Mean F0 difference between H1 and H2 with 1-2 and 3-5 intervening syllables for the two speakers and in different sentence conditions.....	209
5.5. Mean duration (in seconds) between peaks with 3-5 intervening syllables and with 6-9 intervening syllables in sentences with three and four stresses but two pitch accents.....	210
5.6. Mean F0 values of H2 preceded by 3-5 or 6-9 unaccented syllables for the two speakers.....	210
5.7. Mean F0 difference between H1 and H2 with 1-2 and 3-5 intervening syllables for the two speakers and in different sentence conditions.....	210
5.8. Results of the t-test comparing the height of H1 and H2 in medial position for narrow focus triggered by contrast and by identification.....	224
5.9. Results of the t-test comparing the height of H1 and H2 in final position for narrow focus triggered by contrast and by identification.....	225
5.10. Mean F0 value of L with 1-2 and 3-5 intervening syllables between H tones for the two speakers and in different sentence conditions and results of the t-tests.....	233
6.1. Number of English sentences recorded for broad focus and narrow focus. In narrow focus, sentences are divided according to the focal domains (subject, verb, object), the triggering question (contrast, identification) and the stress pattern of the word (oxytone, paroxytone, proparoxytone).....	241
6.2. Mean duration of various numbers of unaccented syllables between accents.....	247
6.3. Mean values of F0 peak position (expressed as % of the duration of the accented syllable), the accented syllable duration (on-of) and the distance between the onset of the accented syllable and the peak (on-p) in seconds for prenuclear accents.....	252
6.4. Mean values of F0 peak position (expressed as % of the foot duration), foot duration, and distance between the onset of the accented syllable and the peak (on-p) in seconds.....	253
6.5. Mean values of F0 peak position (expressed as % of the duration of the	

accented syllable), the accented syllable duration (on-of) and the distance between the onset of the accented syllable and the peak (on-p) in seconds in isolated words.....	254
6.6. Mean values of F0 peak position (expressed as % of the duration of the accented syllable), the accented syllable duration (on-of) and the distance between the onset of the accented syllable and the peak (on-p) in seconds for the nuclear accent of multi-stressed sentences.....	255
6.7. Mean values of F0 peak position (expressed as % of the duration of the accented syllable), the accented syllable duration (on-of) and the distance between the onset of the accented syllable and the peak (on-p) in seconds for the nuclear accent of a subject focal domain.....	263
6.8. Mean values of F0 peak position (expressed as % of the duration of the accented syllable), the accented syllable duration (on-of) and the distance between the onset of the accented syllable and the peak (on-p) in seconds for the nuclear accent of a verb focal domain.....	263
6.9. Mean values of F0 peak position (expressed as % of the duration of the accented syllable), the accented syllable duration (on-of) and the distance between the onset of the accented syllable and the peak (on-p) in seconds for the nuclear accent of an object focal domain.....	264
6.10. Mean F0 range in utterances with three and four stresses and results of t-tests comparing the F0 range values.....	267
6.11. Mean F0 difference between H1 and H2 with 1-2 and 3-5 intervening syllables for the two speakers and in different sentence conditions.....	268
6.12. Results of the t-test comparing the height of H1 and H2 in medial and in final position for narrow focus triggered by contrast and by identification.....	271

## ***Chapter 1. Introduction***

### **1.1 Introduction**

This dissertation deals with the expression of focus by intonational means in Central Catalan with a comparison to Southern British English. Traditionally, Catalan and English have been categorised as languages that signal the information structure of sentences by different means. Whereas English alters the intonational pattern of the sentence to express focus, Catalan is supposed to shift the syntactic structure, while the accentual pattern remains the same. In the present study, this categorical cross-linguistic classification is thrown into doubt by analysing several samples of Central Catalan data and comparing them to some English sentences. The aim of this dissertation is threefold:

- 1) to examine to what extent Central Catalan speakers can use (produce and recognise) focus conveyed by intonational means, and to investigate how natural the use of accentual strategies is felt to be in this language, as compared to syntactic strategies;
- 2) to describe the phonetic and phonological properties of Central Catalan sentences produced both with a neutral reading (or broad focus) and a marked interpretation (or narrow focus) mainly conveyed by an accentual modification;
- 3) to detect differences and similarities between Central Catalan and Southern British English accentual devices.

This chapter consists of a brief and very general introduction to some aspects of focus, intonational phonology and Central Catalan intonation relevant for this dissertation. Section 1.2 deals with the notion of focus from different perspectives and presents various ways of expressing focus cross-linguistically. Section 1.3 introduces the Autosegmental-Metrical (AM) framework of intonational analysis, which will be used in this dissertation to describe the Central Catalan and the English utterances. It also provides some references to previous models of intonational representation and the reasons for choosing the AM approach to analyse the data. Section 1.4 presents a brief

description of Central Catalan, with particular attention to the work pursued on intonation. Finally, in section 1.5 an outline of the whole dissertation is provided.

## 1.2 Focus

The notion of *focus* is subject to a variety of meanings and interpretations, depending on the perspective of study: semantic, pragmatic, syntactic, phonetic or phonological. This dissertation is mainly concerned with the phonological/phonetic side of focal phenomena, although a few references to the concept of focus from other perspectives will be provided in this introduction.

### 1.2.1 Semantic/pragmatic approach

From a semantic/pragmatic point of view, focus has to do with the conveying of new information in discourse. Within this perspective, the focussing activity is related to contextual factors as well as to the speaker's intentions (Bolinger 1972, Schmerling 1976, Halliday 1970, House 1990, Maidment 1990, Lambrecht 1994). Most of the studies within the pragmatic approach claim that the information conveyed in a sentence is segmentable into "old" (or presupposed) and "new". Old information involves the information in the utterance "that is assumed by the speaker to be shared by him and the hearer" (Jackendoff 1972:230). New information, on the other hand, is what is at the centre or focus of the communicative act<sup>1</sup>. For example, in the dialogue presented in (1.1), the fact that one of the speakers found a diary is old information since it is shared knowledge by both interlocutors. The person who this diary belongs to (*Mary's*) is new for speaker A. Thus, speaker B highlights or focusses this part of the message, since it is the carrier of new information<sup>2</sup>.

(1.1) A -Which diary did you find?

B -MARY's diary

---

<sup>1</sup> Several terms with different theoretical implications have been used to account for the "old-new" portions of information in sentences: given-new (Halliday 1970), theme-rheme (Firbas 1964, Contreras 1976), topic-focus (Hajičová 1984, Lambrecht 1994, King 1995), topic-prominence (Kiss 1995), topic-comment (Gundel 1988), background-focus (Dahl 1974, Chafe 1976) and presupposition-focus (Chomsky 1971, Jackendoff 1972, Selkirk 1984), among others.

<sup>2</sup> In this dissertation any time an element is focussed will be in capitals, unless it appears in phonetic transcription where it will be underlined.

Sometimes, however, the intended focus does not fall on a particular item but on the whole phrase or sentence. In this case, there is no shared knowledge between the two speakers and the whole utterance is new or under focus. All approaches dealing with focal phenomena distinguish between these two focus types, which according to Ladd (1980) are referred to as *broad* focus and *narrow* focus. Broad focus involves focus on the whole sentence, not just on individual words. Narrow focus, on the other hand, involves focus on only a particular part of the sentence. According to Leusen and Kálmán (1993), as cited in Frota (1998), each of these focus types can only occur felicitously in specific contexts like the following: for broad focus, in narratives, list of structures, answers to “what happens” or out-of-the blue sentences; for narrow focus, in answers or replies. Examples of these two kinds of focus are provided in (1.2) and (1.3) below, along with the types of questions that trigger the different focal structures<sup>3</sup>.

(1.2) Broad focus:           What happened?  
                                  Peter brought the books

(1.3) Narrow focus: a.       Who brought the books?  
                                  PETER brought the books  
                                  b.       Did John bring the books?  
                                  No, PETER brought the books

Narrow focus may cover several meanings. Two of the most common ones are *identification* and *contrast*. (1.3a) presents an example of identification narrow focus. In this context, the speaker wants to identify or retrieve a particular piece of information. (1.3b), on the other hand, illustrates a contrast scenario, where a given item is wrongly assumed by one of the interlocutors and subsequently corrected by the other speaker. Sometimes identification narrow focus has been named “new information” (García-Lecumberri 1995). However, since several studies (Rochemont and Culicover 1990, King 1995 and Kiss 1995, among others) have sometimes referred to broad focus as new information focus, this term is not used here to avoid confusion. In this dissertation, we will use “identification” and “contrast” to refer to

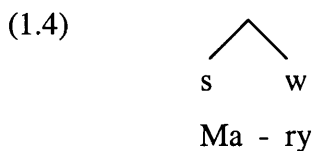
---

<sup>3</sup> In some cases, two focal points can be found in the same sentence. For example, *did John bring the books? No, MARY brought the MAGAZINES*. Double focus patterns are beyond the scope of this study.

the kinds of sentences illustrated in (1.3a) and (1.3b) respectively. Whenever the term “narrow focus” is used, both senses (identification and contrast) are involved.

### 1.2.2 Phonetic/phonological approach

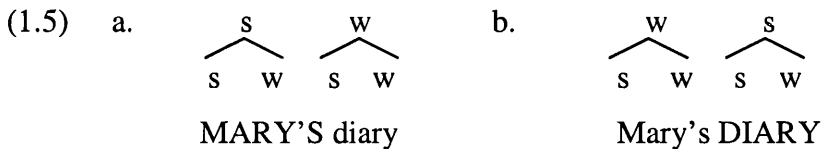
Given the assumption that focus depends on contextual factors and on the communicative intentions of the speaker, the literature agrees that the expression of focus always has a phonetic/phonological correlate (e.g. stress, pitch accent). From a phonological point of view, focus has to do with the assignment of *relative prominence* at the sentence level. Sentence-level prominence involves the identification of the *stress pattern* of a sentence and its further association to the *pitch* or *intonational pattern* (Ladd 1996). *Stress* is a relational or syntagmatic notion that applies at different linguistic levels. At the word level, syllables can be either stressed or unstressed. Stressed syllables are more acoustically and rhythmically salient than unstressed syllables. Liberman and Prince (1977) characterised the stress patterns of words by means of a tree structure consisting of two nodes: *s* (strong) and *w* (weak). This tree structure indicates that the node marked *s* is more prominent than the one marked *w*. Thus, for example, (1.4) shows the prominence structure of the word *Mary*, which has stress on the first syllable. This prominence relation is in fact a linguistic abstraction that may have several phonetic manifestations, such as, duration (Cooper et al 1985), loudness (Wells 1986, Toledo 1986) and pitch (Fry 1958, Gay 1978, also see Ladd 1996:47). Thus, stressed syllables are usually louder and longer than unstressed syllables and may show a pitch or intonational movement.



At the sentence level, the stress pattern reflects the intended focus or the information structure of the utterance. The stress pattern derives from the speaker’s intention to highlight or make more noticeable a particular item. Thus, in the phrase *Mary’s diary*, two stress patterns can be defined, depending on which part of the utterance the speaker considers more informative. If the speaker wants to highlight *Mary’s*, as in the context presented in (1.1), *Mary’s* is more prominent than *diary*. Within the tree-



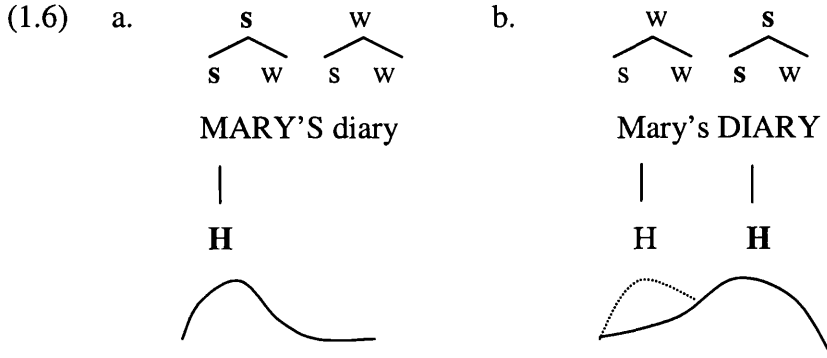
structure analysis, an *s* node will then be assigned to *Mary's* and a *w* node to *diary*. This is illustrated in (1.5a). If the speaker wants to highlight *diary*, the reverse pattern applies, as in (1.5b). The phrase in (1.5b) has two possible readings: either narrow focus on the last word or broad focus. The tree structures in (1.5) define a strongest point of prominence for the whole phrase. This most prominent point corresponds to the element that is only dominated by *s* labels. In Liberman and Prince's terminology, this element is called the *Designated Terminal Element* (henceforth DTE). The assignment of focus has to do with the location of the DTE in the tree structure (see section 1.3.2.1 for more details in utterance-level prominence patterns).



According to Liberman and Prince (1977) and Ladd (1980, 1996), among others, the stress pattern of a sentence is used as the input to the *pitch* or *intonational pattern*. That is, once the prominence relations between the items of a sentence have been decided, the intonational events are distributed. These events are known as *pitch accents* (Bolinger 1958, Pierrehumbert 1980). A pitch accent usually involves some audible (or visible) pitch change associated to a stressed or metrically strong syllable. Even though all stressed syllables are potentially accentable, not every stress is necessarily associated to an accent and the final decision of which syllables become accented is a matter of choice on the speaker's part. Although sentence accentuation is flexible, there is always one syllable that has to become accented, namely, the syllable dominated by all *s* nodes in the tree structure (or DTE). The accent associated to the DTE is usually referred to as the *nuclear accent*. In sentences with a broad focus reading, the nuclear accent tends to occur on the stressed syllable of last lexical word<sup>4</sup>. Usually, there are no accents after the nuclear one. In narrow focus utterances, the nuclear accent falls on the stressed syllable of the focussed item. This is illustrated in (1.6) which presents two possible pitch contours for the prominence patterns in (1.5). In the two structures, the nuclear accent is in bold. In both cases, a peak in the pitch contour (characterised by a H(igh) pitch accent) is observed over the focussed word.

<sup>4</sup> In languages such as English, there are several exceptions to this generalisation. Some of them are discussed in section 1.2.3.

In (1.6b) an optional peak (dashed contour) appears on the first item. No pitch movements are observed after the nuclear accent in (1.6a).



In short, the assignment of focus has to do with the prominence relations of the words making up a sentence and one way to manifest these prominence relations is by means of the stress distribution and the subsequent accent location. Thus, in this dissertation we assume that sentence stress is defined independently from intonational structure (Lieberman and Prince 1977, Ladd 1980, 1996, Pierrehumbert 1980) and that pitch accents are distributed after the prominence relations between the elements of a sentence have been established<sup>5</sup>.

### 1.2.3 Focus and normal stress

The notion of broad focus is closely linked to the idea of “normal stress” developed within the framework of Generative Grammar (Chomsky and Halle 1968). Normal stress involves the assignment of an unmarked stress pattern to sentences by means of phonological rules, which apply at the output of the syntactic component of the grammar, i.e. at the level of surface syntactic structure. The idea of normal stress is formalised in the *Nuclear Stress Rule* (henceforth NSR), which assigns primary stress or sentence stress to the rightmost lexical word of an utterance. The idea of normal stress can be linked to the notion of broad focus in the sense that normal stress rules

<sup>5</sup> This view (sometimes referred to as the “stress-first” view) differs from other approaches (basically Selkirk 1984) who claim that sentence stress plays no role in determining the assignment of pitch accents and propose that the relation between focus and sentence stress is mediated by intonation (“accent-first” view). It is beyond the scope of this thesis to determine which of the two traditions better accounts for the accent-focus relationship. In this dissertation, we assume that sentence stress controls accentuation since this is the basis of the AM framework of intonational analysis, which will be used to describe the Central Catalan data (see section 1.3.2 for an overview on the AM approach).

can predict the stress/accent location when focus is broad (Ladd 1980, 1996). Thus, in the sentence illustrated in (1.7), produced with a broad focus reading, the nuclear stress rule predicts primary stress on the stressed syllable of the last lexical word (i.e. *directions*). Within this approach, narrow focus sentences would be related to the idea of “contrastive stress” which is seen as unpredictable and beyond the scope of the NSR (see section 1.3.2.1 for more details on the application of this rule).

(1.7) George follows DIRECTIONS

The tenets of the NSR have been criticised in two directions: the “syntactic” and the “semantic” views (terms suggested by Ladd 1980). The syntactic approach, mainly supported by Bresnan (1971) and critics (Lakoff 1972, Berman and Szamosi 1972), provides examples of sentences where the NSR fails to assign normal stress and proposes the assignment of stress at the level of deep syntactic structure. For example, the NSR cannot account for the “normal stress” pattern of the sentence in (1.8) (discussed in Bresnan 1971), since it predicts sentence stress on the last constituent (i.e. *follow*). To be able to handle cases like that, Bresnan (1971: 259) suggested that “the NSR is ordered after all the syntactic transformations on each transformational cycle”. Thus, *directions* receives primary stress at the level of deep structure, when it is the object of the embedded sentence (i.e. *George follows directions*) and before it is deleted from its original position. Although Bresnan’s modification of the NSR can account for the “normal stress” pattern of a sentence such as (1.8), there are still several utterances which cannot be explained within this proposal, as argued by the followers of the semantic view.

(1.8) Helen left DIRECTIONS for George to follow  
(meaning: George should follow the directions)

The semantic view follows two main lines of research, which in Ladd’s (1996) terms are the “highlighting-based approach” and the “structure-based approach”. The highlighting-based approach has its basis in the works of Bolinger (1972) and Schmerling (1976). They claim that accent is not predictable from syntax but depends on semantic and contextual factors, such as “the point of information focus” in the utterance or on contrastive or highlighted items. For example, in the sentence

presented in (1.9a), the last lexical word is not accented because it is semantically empty. In (1.9b), on the other hand, the last lexical word receives primary stress since it is semantically richer.

- (1.9) a. I have a POINT to make  
b. I have a point to EMPHASISE

Finally, the structure-based approach (Ladd 1980, Gussenhoven 1983a, 1984, 1994) tries to reconcile the syntactic and the semantic views by assuming that once focus is defined on a semantic basis (according to the speaker's intention and the context), pitch accent distribution is determined by structural factors. In brief, Gussenhoven proposes that accent is assigned once focus domains have been formed within each sentence. Thus, if a constituent has already been introduced into the discourse, then it is marked [-focus] and it does not get an accent. On the other hand, new information is marked as [+focus] and it is subsequently accented. Similarly, Ladd (1980:76) integrates the semantic and the syntactic approaches by defining normal stress in terms of the semantics of focus. Thus, he reformulates Bolinger's idea that "accent goes to the point of information focus" by adding that this is so, "unless the focus is unmarked in which case, accent goes on the location determined by syntax". He assumes that "if the focus constituent is the whole sentence, then we get normal stress; if not, we get narrow focus on the constituent identified by the placement of accent". According to this, normal stress is the accent placement that provides a broad focus interpretation of the utterance.

Following Ladd (1980, 1996) and Gussenhoven (1983a, 1984, 1994), the interpretation of focus assumed in this study means that once the semantic/pragmatic decisions have been taken to signal the focal structure of an utterance, the focal item must be associated to an accent<sup>6</sup>.

---

<sup>6</sup> Ladd (1996:223ff) presents some examples of focus without accent and accent without focus. These cases are not treated in this study.

### 1.2.4 Ways of expressing focus: phonological vs syntactic devices

Although most languages express focus by means of accentuation, not all of them attain the focus/accent association by the same means. Literature on the relation between focus and accentuation (Vallduví 1990, 1995, Vallduví and Zacharski 1994, Kiss 1995, Ladd 1996, among others) shows that whereas some languages (mainly Germanic) can modify their intonational pattern to achieve the focus/accent alliance, other languages (Romance) need other devices, such as word order movement, to attain the same link.

Example (1.10) (which reproduces the sentences in (1.2) and (1.3a)) and example (1.11) illustrate how English and Catalan use different strategies to signal focus. (1.10a) and (1.11a) show a S(ubject) V(erb) O(bject) simple declarative sentence produced with a neutral reading or broad focus in the two languages. Both in English and in Catalan the pronunciation of these utterances involves a nuclear accent on the last lexical word, i.e. *books/llobres*. Alternatively, (1.10b) and (1.11b) show a sentence where a particular item (*PETER/EN PERE*) is highlighted with respect to the others. In English, the strategies used to focus this item involve a reorganisation of the intonational pattern of the sentence (the nuclear accent is no longer on the last lexical item but on the focussed element), with no changes at the syntactic level. In Catalan, on the other hand, a syntactic shift rather than an accentual alteration occurs. In this case, the focussed word (*EN PERE*) moves into the nuclear accent position to get prominence and the item that used to occupy this site (*llibres*) is dislocated to the beginning of the sentence. Thus, whereas English prefers to mark focus by altering the phonological structure of the sentence, Catalan tends to retain the accentual pattern and modifies the syntactic structure.

- (1.10) a. Peter brought the books  
b. Who brought the books?  
PETER brought the books

- (1.11) a. En Pere va portar els llibres
- b. Qui va portar els llibres?  
Els llibres, els va portar EN PERE  
“The books, they were brought by PETER”  
(lit. The books, them brought PETER)

The strategies illustrated in (1.10) and (1.11) for English and Catalan are schematised in (1.12) below. The syntactic and accentual structures of sentences produced with a neutral declarative intonation are the same in both languages. The syntactic structure involves a SVO order and the accentual structure consists of a pitch accent (pa) associated to the stressed syllable of each lexical word<sup>7</sup>. The last pitch accent (shaded area) is the nuclear accent. However, when the sentences are produced with focus on the subject, different strategies are used to achieve the subject/nuclear-accent association. In English, the syntactic structure remains the same (SVO) but the nuclear accent is now located over the new focal position. No more pitch accents are realised after the focal material. In Catalan, the phonological structure is unchanged (the nuclear accent remains on the last position) but there is a syntactic shift (the focussed element moves into an intonationally prominent position, i.e. the last position of the sentence)<sup>8</sup>.

(1.12)	ENGLISH	CATALAN						
<i>Neutral declarative</i>	Peter brought the books	En Pere va portar els llibres						
<i>Syntactic structure</i>	<table border="1"><tr><td>S</td><td>V</td><td>O</td></tr></table>	S	V	O	<table border="1"><tr><td>S</td><td>V</td><td>O</td></tr></table>	S	V	O
S	V	O						
S	V	O						
<i>Accentual structure</i>	<table border="1"><tr><td>pa</td><td>pa</td><td>pa</td></tr></table>	pa	pa	pa	<table border="1"><tr><td>pa</td><td>pa</td><td>pa</td></tr></table>	pa	pa	pa
pa	pa	pa						
pa	pa	pa						
<i>Subject focus</i>	PETER brought the books	Els llibres, els va portar EN PERE						
<i>Syntactic structure</i>	<table border="1"><tr><td>S</td><td>V</td><td>O</td></tr></table>	S	V	O	<table border="1"><tr><td>O</td><td>V</td><td>S</td></tr></table>	O	V	S
S	V	O						
O	V	S						
<i>Accentual structure</i>	<table border="1"><tr><td>pa</td><td>-</td><td>-</td></tr></table>	pa	-	-	<table border="1"><tr><td>pa</td><td>pa</td><td>pa</td></tr></table>	pa	pa	pa
pa	-	-						
pa	pa	pa						

<sup>7</sup> Sometimes not all stresses become accented but here for the purpose of clarity we will assume that every lexically stressed syllable gets a pitch accent.

<sup>8</sup> The syntactic reorganisation used in Catalan to signal focus is more complicated than a simple change in word order. It mainly involves a left (or right) dislocation of the object, which retains its link to the sentence by the presence of a pronoun. The pronoun can be realised before the verb or after the verb. In this example it occurs before the verb (i.e. *els*). The subject is moved to the end of the sentence. If we assume that every time an element moves, it leaves an empty category (*e*) on its original site, and if we associate empty categories with moved elements by means of subindices, the syntactic reorganisation of the sentence is as follows. *En Pere* (subject) leaves an empty slot in its original position (index *i*) and *els llibres* (object) leaves another empty slot (index *j*) and its presence in the sentence is recovered by the pronoun also indexed with the letter *j*. That is, *els llibres<sub>j</sub>, e<sub>i</sub> els<sub>j</sub> va portar e<sub>j</sub> en Pere<sub>i</sub>*.

Vallduví (1990, 1991, 1994a/b) and Vallduví and Zacharski (1994) claim that the cross-linguistic variation observed in Catalan and in English is the reflex of a *Plasticity Parameter* of intonation relative to focus. In [+plastic] languages, such as English, the relation between focus and prominence is achieved by accentual means. In [-plastic] languages, such as Catalan, the relation between prominence and focus is achieved through syntax. This proposal is related to Kiss's (1995) claim that some languages (such as Catalan, Hungarian, Greek and Basque, among others) can be categorised as *discourse-configurational* languages. In these languages, phonological phenomena (such as stress or pitch accent distribution) do not suffice to express focus, as in languages such as English, and morphosyntactic strategies are needed.

However, other studies on the ways of expressing focus in Romance languages, namely, Spanish (García-Lecumberri 1995, Zubizarreta 1998), French (Zubizarreta 1998) and European Portuguese (Frota 1998), show that focus-signalling means are not so language-dependent and that a given language may use more than one strategy. Similarly, different strategies can be used more or less appropriately in distinct contexts. For example, Zubizarreta (1998) shows that whereas in French the sentence in (1.13a) can be the answer of a wh-question, in Spanish the equivalent sentence (1.13b) can only have a contrastive (or emphatic) interpretation.

- (1.13) a. JEAN a téléphoné  
Jean has phoned  
(Who phoned?)
- b. JUAN llamó (no Pedro)  
(Juan phoned, not Pedro)

Experimental work on focus in Peninsular Spanish (García-Lecumberri 1995 and García-Lecumberri et al 1997) partially disagrees with Zubizarreta's claim, since sentences such as (13b) are proved to be used in Spanish as responses to wh-questions, although the contrastive meaning is preferred. Finally, the results of Frota (1997, 1998) for focus in European Portuguese show that sentences where focus is signalled by intonational movements are equally favoured with sentences where phrasal prominence determines the focal structure. One of the goals of this

dissertation is to analyse whether intonational strategies can also be used in Central Catalan to signal the information structure of sentences.

Finally, given the assumption that any time an element is focussed, it is also accented, the terms “accentual focus” or “intonational focus” seem redundant and need clarification. In this dissertation, whenever the terms “accentual focus” or “intonational focus” are used, they mean focus conveyed by some alteration of the intonational structure of the sentence, as opposed to focus signalled by a syntactic modification (or “syntactic focus”).

### 1.3 Modelling intonation

*Intonation* refers to the use of *pitch* patterns to convey *non-lexical* or *sentence-level* meanings. Pitch is the perceptual sensation of *fundamental frequency* (F0), which is the acoustic correlate of the repetition rate of vocal fold vibration<sup>9</sup>. Intonation variations apply at the sentence level and convey meanings that affect the utterance as a whole. Thus, two utterances with the same segmental structure but different intonational patterns can differ in meaning. In some languages, known as *tone languages* (for tone languages see Goldsmith 1976), pitch variations are not only linguistically significant at the sentence level but also at the lexical level. In these languages, pitch is lexically controlled and conveys differences in meaning between words.

This dissertation deals with the intentionally controlled use of pitch patterns at the sentence level. Thus, it examines pitch variation as part of intonation but not as part of lexical tone. At the sentence level, intonational patterns are analysed as a string of categorically distinct entities. The number of distinctive units is limited in each language. Different approaches of intonational analysis differ in the characterisation of these units. However, studies (sometimes far apart in time and from different traditions, such as Halliday 1967 and Hirschberg 1999) agree in that to describe intonation four main parameters have to be taken into account: 1) the selection of pitch accents (or *tone* in Halliday's terminology), 2) the distribution of pitch accents

---

<sup>9</sup> In this dissertation, the terms “pitch” and “F0” are used interchangeably.



across sentences (*tonicity*), 3) the division of chunks of speech or levels of phrasing (*tonality*), and 4) the key or pitch range with which utterances are produced. In the present study, all these parameters will be examined within the *Autosegmental-Metrical* (henceforth, AM) approach of intonational analysis. Before presenting the main tenets of the AM model, a few references to previous works will be first made so as to have an overview on the different lines of research on intonation.

### 1.3.1 Traditional analyses

Before the AM tradition of intonational analysis, the study of intonation has followed two main lines of research, which in Ladd's (1996) terms are known as the *instrumental* or *phonetic* tradition and the *impressionistic* or *proto-phonological* tradition. The instrumental approach mainly aimed at characterising the acoustic cues of intonation but no attention was paid to identifying the phonological events that made up pitch contours. The impressionistic view, on the other hand, was more phonologically-oriented. Their work was mainly based on perceptual and auditory analyses of the data but little experimental work was carried out. One of the advantages of the AM model is that it combines an experimental approach to intonation with a theoretical insight.

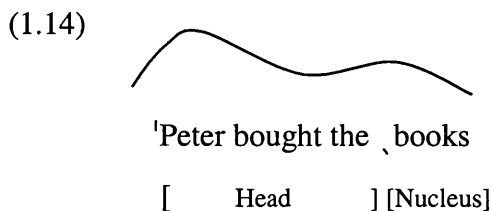
The impressionistic view covers the work of two main traditions: the British and the American schools (see Ladd 1980 and Cruttenden 1986 for reviews on these frameworks). In brief, the two schools constitute the two exponents of a long debate in the study on intonation, namely, the question whether pitch phenomena should be described in terms of *levels* or *configurations* (terms suggested by Bolinger 1951).

#### 1.3.1.1 The British tradition

One of the main goals of the British tradition of intonational analysis was to describe intonation for didactic purposes, basically for teaching pronunciation. The tenets of the British approach, although designed for the transcription of English intonation, have been used in many other languages (including a few works on Catalan). There are two basic aspects that characterise the work pursued in this tradition, namely, the division of pitch contours into several component parts, and the configurational

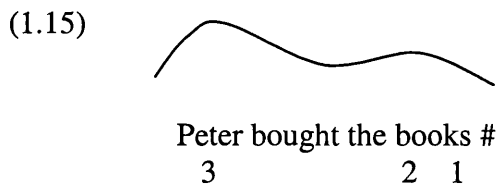
analysis of tone. Palmer (1922) was the first to divide the pitch contour into a nuclear constituent (the *nucleus*) preceded by a prenuclear portion (the *head*) and followed by a *tail*. Kingdon (1958) added a further subdivision of the prenuclear part of the contour into the head and a *prehead*. Thus, a tone unit (stretch of utterance) consists of an obligatory element, the nucleus, and three optional ones, the prehead, the head and the tail. Crystal (1969) described these four components in terms of rhythmic prominence, that is, by distinguishing between stressed, unstressed and accented syllables. The prehead consists of all syllables before the onset of the first accented syllable. The head comprises the string of syllables from the first accented syllable up to the syllable before the nucleus. Finally, the tail is made up of all syllables after the nuclear one. The compositional analysis of tune is followed by most scholars of the British tradition (O'Connor and Arnold 1973, Ashby 1978, Gimson 1980, Couper-Kuhlen 1986, Cruttenden 1986 and Tench 1996, among others).

The analysis of intonation pursued by the British School was mainly configurational, that is, it was based on pitch movement rather than on pitch levels (as opposed to the American tradition). For them, the phonological primitives involved categories such as *fall*, *rise*, *fall-rise*, *rise-fall*, *high* or *low* among others. A relevant premise of this approach is that certain entities can only be attached to specific structural constituents. For example, *high* and *low* are events related to the head constituent only, whereas *fall-rise* or *rise-fall* can only appear in the nuclear component. (1.14) illustrates the tonal transcription for the sentence *Peter brought the books* (produced with a broad focus reading) according to O'Connor and Arnold's analysis. The sentence consists of a high head (marked [']) followed by a low fall nuclear accent (marked [,]).



### 1.3.1.2 The American tradition

The aim of the American structuralist tradition is to describe intonation as part of a phonemic theory. This approach is mainly represented by the works of Pike (1945), Wells (1945) and Trager and Smith (1951), and subsequently developed in Liberman (1975) and in the autosegmental work of Leben (1976) and Goldsmith (1976). Pitch contours are described as a series of pitch level phonemes. In Pike, the speaker's pitch range is divided into four relative phonemic pitch levels, /1, 2, 3 and 4/, where /1/ means high and /4/ low. In Trager and Smith, the reverse order is proposed (/1/ for low and /4/ for high). Pitch phonemes are located at the beginning of an utterance, before the primary stress and at the end, although they may appear in other sentence positions as well. In Trager and Smith, three phonemes of terminal juncture are added into the system. They are characterised by pitch movement: fall /#/ , rise /| | / and level /| / . In this view, a pitch contour consists of a series of pitch phonemes and a terminal juncture. For example, the sentence *Peter bought the books* would be represented as in (1.15), where 3 stands for a mid level, 2 and 1 for lower pitch levels and # for a fall terminal juncture.



Later developments of the level analysis of intonation within the autosegmental approach propose to reduce the number of pitch accents to three levels and the idea of phonemes at the terminal junctures develops into the notion of boundary tone, i.e. an underlying phoneme manifested phonetically by pitch movement at an intonational boundary (Liberman 1975).

The AM approach of intonational analysis inherits several aspects from the American School, in particular, the level analysis of pitch (although reduced to two categories) and the notion of boundary tone.

### 1.3.2 Autosegmental-Metrical Phonology

#### 1.3.2.1 Background

The AM approach of intonation has its roots in two theoretical frameworks developed within the tradition of Non-linear Phonology, namely, the *Autosegmental* approach and the *Metrical* approach. These two frameworks, along with other approaches (Government Phonology, Lexical Phonology, Dependency Phonology, among others) emerged as a reaction to the linear approach of phonological analysis proposed by Chomsky and Halle (1968) in *the Sound Pattern of English* (henceforth SPE)<sup>10</sup>. In the SPE, phonological events were described by means of phonological rules that operate on a linear arrangement of sound-segments (composed of simultaneously occurring features). Subsequent studies have shown that there are several phonological aspects (e.g. nasalisation, tone or stress) that cannot easily be described in terms of a linear representation of features but need from a more complex (non-linear) organisation.

The Autosegmental approach stems from the analysis of tone languages, where tone features cannot be convincingly analysed within a segment-sized matrix as in the SPE but should be allowed to span over other domains or levels of phonological analysis. For example, the Bantu language Bakwiri (example taken from Durand 1990:243) has two tones, H(igh) and L(ow). In this language, there is a game in which two syllables of disyllabic words are transposed, as illustrated in (1.16).

(1.16)      k w e l i (dead)      →      l i k w e  
                   H    L                                      H    L

However, while segments making up the two syllables have changed, the tones have stayed in their original position. This contradicts the expectations of the SPE, where tones were part of the matrix of tone bearing units (basically vowels), and hence the movement of a particular segment would involve the movement of all its constituting features. The fact that in tone languages tones do not necessarily behave in unison with segments seems to indicate that tones and segments are located on different tiers of phonological representation.

<sup>10</sup> For a clear review of these models, see Durand (1990) and Carr (1993).

Thus, based on the tone evidence, Goldsmith (1976) places melodic and segmental units on separate tiers, the “tonal tier” for tones and the “CV tier” for segments. Elements on the tonal tier are related to elements on the CV tier by means of association lines, as in (1.17). Within this proposal, it is assumed that the transposition phenomenon illustrated in (1.16) only applies to items on the segmental tier and leaves the tonal tier unaffected. Thus, tone features are treated as autosegments, which are independent from the features that constitute a segment. The analysis of intonation as independent from (but associated to) the segmental string is adopted in the AM model.

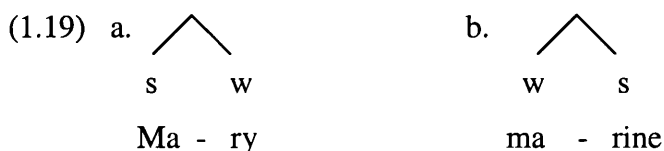
(1.17)            k w e l i  
                       | |  
                       H L

The other theory from which the AM framework departs is Metrical Phonology (Lieberman 1975, Lieberman and Prince 1977). Metrical Phonology appeared as a reaction to the treatment of stress proposed in the SPE. Similar to tone, the SPE analysed stress as a feature specified as part of the vocalic segment of the stressed syllable. Stress rules predict the stress patterns on the basis of the segmental, morphological and syntactic properties of constituents and apply cyclically from the smallest (word level) to the largest constituent (sentence level). At the word level, the *Main Stress Rule* (MSR) assigns primary stress or principal prominence ([1 stress]) to the most salient syllable of the word. At the phrase/sentence level, the assignment of stress is determined by the *Nuclear Stress Rule* (NSR). As stated in section 1.2.3, the NSR predicts the unmarked or neutral distribution of stresses (“normal stress”) by assigning primary stress to the rightmost lexical word of the phrase/sentence with subsequent downgrading of the other stress/es in the domain by one degree. The NSR first applies to phrases and then to sentences. An example is provided in (1.18).

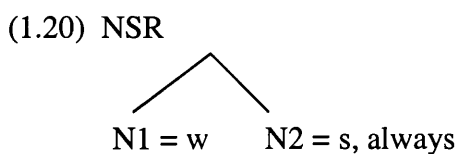
(1.18) [[Mary’s [sister]] [married [Albert]]]  
           1        1        1        1            MSR  
           2        1        2        1            NSR  
           3        2        3        1            NSR

One of the main problems derived from the SPE analysis of stress is that the prominence relationship between two constituents of a phrase varies depending on the degrees of prominence assigned to that phrase in larger domains. For example, if the phrase *married Albert* is said in isolation, two levels of prominence are assigned, [2 1]. However, if the phrase appears within a sentence, such as *Mary's sister married Albert*, the relation of prominence between the two words becomes [3 1], as illustrated in (1.18) above. This situation is rather undesirable since the prominence relation between two constituents of a phrase is entirely dependent on the stress configuration of larger domains. Thus, a formalism is needed whereby the degrees of prominence between two constituents of a phrase remains unaffected.

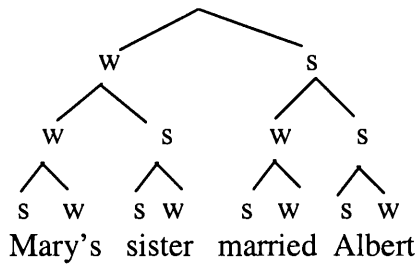
As stated in section 1.2.2, Liberman and Prince's (1977) alternative proposal involves interpreting relative prominence as a structural relation between two constituents rather than as a property of individual segments. Thus, stress (or prominence) is defined as a binary relation between a *w* (weak) element and a *s* (strong) element represented on a tree structure, known as the *metrical tree*. This is illustrated in (1.19) below ((1.19a) reproduces the example in (1.4)). The labels *w* and *s* acquire a meaning in relation to the other member of the pair: *s* is strong in relation to *w* and *w* is weak in relation to *s*.



As already pointed out, Metrical Phonology not only deals with the stress patterns of words but also of larger constituents. At the phrase/sentence level, the assignment of *s* and *w* nodes is governed by the NSR, which is reformulated as follows: for any pair of nodes [N1 N2], N2 is strong. Liberman and Prince's NSR is presented in (1.20). The tree structure derived from the application of the NSR is schematised in (1.21) below for the sentence *Mary's sister married Albert*.



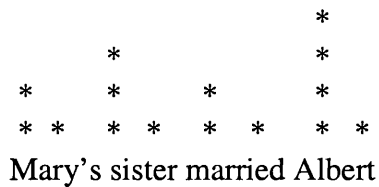
(1.21)



The advantage of using this formalism over the SPE-type of analysis is that relative prominence between two constituents is not affected by the stress configuration of larger domains. Thus, the relative prominence between the two words of the phrase *married Albert* remains the same (*w-s*), irrespective of whether the phrase is said in isolation or is part of a larger domain.

Apart from the metrical tree, Metrical Phonology introduces another device, the *metrical grid*, which defines relative prominence of successive stressable units. For example, in (1.21), the first syllable in *Albert* is more prominent than any of the preceding syllables but the relation of these syllables to each other is not specified. This relation is defined in the metrical grid. Metrical grids consist of rows and columns of asterisks, each of which corresponds to the nodes of the tree. Metrical grids are built by assigning an asterisk to each syllable, and then to any syllable labelled *s*, starting from the most embedded constituent to larger ones. This is illustrated in (1.22), where the different degrees of prominence for all the constituents of the sentence constituents are specified.

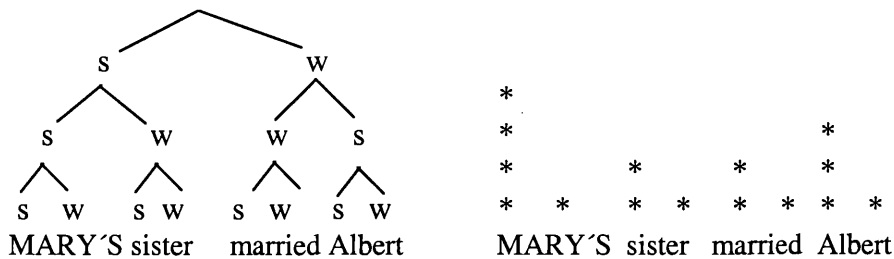
(1.22)



The choice of a grid for the metrical description of a sentence depends on which part of the sentence the speaker wants to focus. So, the labelling of any level of the tree may be reversed (*w-s* > *s-w*) to highlight particular information of the utterance. Thus, for example, if the sentence *MARY'S sister married Albert* is produced with narrow focus on the first element, then the canonical broad focus pattern (*w-s*) is altered into a

narrow focus pattern (*s-w*). The tree and grid representations for this sentence will be as follows.

(1.23)



These prominence patterns constitute the basis for the assignment of pitch accents in the AM framework of intonational analysis.

### 1.3.2.2 Phonological units

The point of departure for the AM model of intonational representation is the work of Pierrehumbert (1980). Since then, several studies have emerged elaborating and modifying the main ideas of the AM framework. These include Ladd (1983a, 1990, 1992, 1996), Liberman and Pierrehumbert (1984), Gussenhoven (1984), Beckman and Pierrehumbert (1986), Pierrehumbert and Beckman (1988) and Grice (1995a, b), among many others. Recently, the tenets of the AM model have developed into a system for prosodic labelling known as the ToBI system<sup>11</sup> (Silverman et al 1992, Beckman and Ayers 1994, Beckman and Hirschberg 1994). One of its main goals is to provide a potentially universal descriptive framework for intonational analysis<sup>12</sup>.

Within the AM approach, pitch patterns consist of a sequence of categorically distinctive entities, associated to elements of the segmental string and to the edges of prosodic domains. The association of these entities to the text is governed by principles of prosodic organisation, according to which either prominent syllables or

<sup>11</sup> ToBI stands for "Tones and Break Indices".

<sup>12</sup> The tenets of the AM model and the ToBI system have been applied to several languages, such as, English (Pierrehumbert 1980), Japanese (Beckman and Pierrehumbert 1986, Pierrehumbert and Beckman 1988, Venditti, 1995, 1999), Catalan (Prieto 1995), Palermo Italian (Grice 1995a), Neopolitan Italian (D'Imperio 1999b), Bari Italian (Savino 1999), German (Féry 1993, Grice and Benz Müller 1995, Grabe 1998a), Bengali (Hayes and Lahiri 1991), Greek (Arvaniti and Baltazani 1999), American Spanish (Sosa 1991, 1999), Korean (Jun 1996), Portuguese (Frota 1993, 1998) and French (Post 2000), among others.

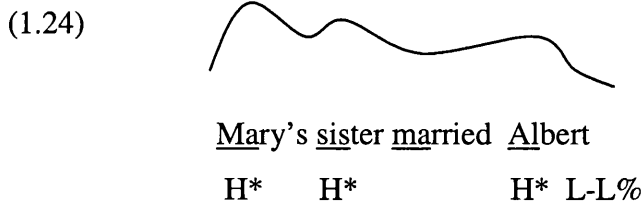


prosodic edges are possible anchors for the phonological events. For prominent syllables, the model assumes that for the generation of intonation patterns, the generation of patterns of sentence stress is necessary, as claimed in Liberman (1975) and in Liberman and Prince (1977). For prosodic edges, the model is based on the work of Bruce (1977) who postulated some tonal events after the last pitch accent and on the notions of phrasal constituency provided by Selkirk (1984) and Nespor and Vogel (1986). Thus, each phonological unit not only contains information about the *tonal* properties of the intonation pattern but also about the *rhythmic* structure.

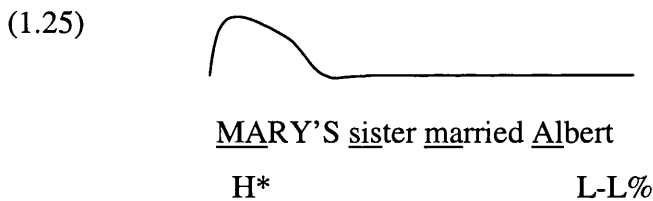
There are two kinds of phonological events: the *pitch accents* and the *edge tones*. The latter are further divided into *phrase accents* and *boundary tones*. The pitch accents consist of local F0 movements that accompany prominent syllables (stressed syllables). The phrase accents and the boundary tones represent the F0 movements at the edges of the prosodic domains. The boundary tone signals the end of an intonational phrase. The phrase accent has acquired two interpretations depending on which version of the model is used. In the initial version (Pierrehumbert 1980), the phrase accent is a floating tone that accounts for the F0 contour between the last pitch accent and the boundary tone. In later versions (Beckman and Pierrehumbert 1986, Pierrehumbert and Beckman 1988, Beckman and Hirschberg 1994, among many others), the phrase accent marks the end of an intermediate phrase. The latter interpretation of the pitch accent is adopted in this study. See section 1.3.2.3 for more details on the levels of prosodic structure.

The tonal properties of pitch accents, phrase accent and boundary tones are transcribed as H and L, which stand for relatively high and relatively low F0 targets. The rhythmic properties are indicated by assigning an asterisk [\*] following the tone that marks a pitch accent, a hyphen [-] after the tone for the phrase accent, and [%] after the tone for the boundary tone. An example is provided in (1.24) for the sentence *Mary's sister married Albert* produced with a broad focus intonation, as in (1.21). This sentence consists of three peaks in the F0 contour followed by a final falling pitch. Each peak is described in terms of an H\* pitch accent associated to a metrically strong syllable (underlined). Note that although all stressed syllables are potential pitch accent bearers, they do not necessarily have to become accented, as illustrated in

*married* which does not show any pitch movement. The final fall is described by means of the L-L% edge tones following the last H\* pitch accent<sup>13</sup>.



Although all stressed syllables can be associated to a pitch accent, Pierrehumbert (1980) points out that no pitch accents occur after the nuclear accent. Thus, for example, in the sentence *MARY'S sister married Albert*, with narrow focus on the first element of the utterance, the stresses after the nuclear stress are not accented. Thus, the pitch pattern of this sentence is illustrated in (1.25), where only the first (focussed) word receives a pitch accent. The L% boundary tone marks the end of the intonation phrase. The low F0 contour between the last pitch accent and the boundary tone is described by means of the L- phrase accent.



Within the AM framework, the notion of the nuclear accent is different from that proposed within the British tradition. For Pierrehumbert (1980), the nuclear accent is simply the last pitch accent of a series of pitch accents. This means that the nuclear accent is a more confined part of the contour than in other approaches, which have viewed the tonal material belonging to the final pitch accent, the phrase accent and the boundary tone as constituting a single tonal entity. For example, a *fall-rise* nuclear accent in the British approach is described in Pierrehumbert's framework as an H\* (nuclear accent) followed by an L- phrase accent and an H% boundary tone. This implies that the same pitch accent H\* can be used both in nuclear and prenuclear positions, as illustrated (1.24).

<sup>13</sup> The left edge has no boundary tone because the F0 contour is low (or mid) and Pierrehumbert (1980) suggests that only a high F0 start should be tonally specified (H%).

Pierrehumbert (1980) proposes the following tone inventory to describe American English F0 contours: pitch accents (H\*, L\*, H\*+L, L\*+H, H+L\*, L+H\*, H\*+H), phrase accents (H-, L-) and boundary tones (H%, L%). Whereas phrase accents and boundary tones are only monotonal, pitch accents can be both monotonal and bitonal. Bitonal accents can be left-headed or right-headed depending on whether the first or the second element of the accent is associated to the metrically strong syllable. As in monotonal pitch accents, the tone associated with the accented syllable is marked with an asterisk and the tones preceding (leading) or following (trailing) the starred one are left with no mark. The two tones are joined with the plus symbol<sup>14</sup>. The phonetic properties of the different tones are as follows:

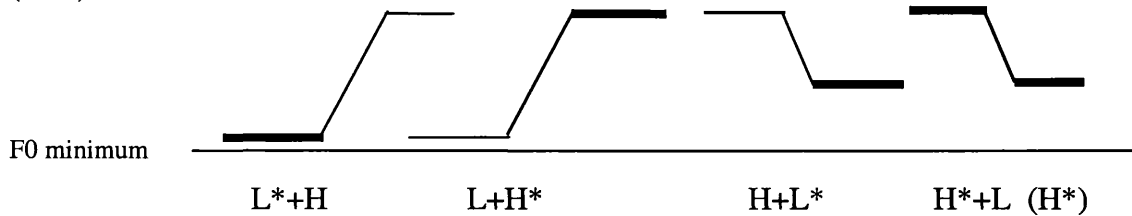
*1. Pitch accents.* H\* is a local F0 peak aligned in the vicinity of the accented syllable. L\* involves a local F0 valley. The categorical contrast between L+H\* and L\*+H is manifested phonetically by means of the timing association of the LH gesture to the accented syllable (see Pierrehumbert and Steele 1989). Whereas L+H\* is realised as a rising F0 movement over the accented syllable, L\*+H involves a low F0 target on the accented syllable, followed by a rising trajectory that starts at the end of the accented syllable. The relationship between H\*+L and H+L\* is different from that between L+H\* and L\*+H. H+L\* involves an F0 drop from a preceding syllable. However, the L is not a local valley or F0 minimum but is scaled at the level of a lowered (or downstepped) high tone. H\*+L, in its turn, indicates a high accent that triggers downstep in the following H tone (see section 1.3.2.4 for details on downstep). Thus, in H\*+L, the L has no manifestation in the F0 contour. The phonetic properties of the four bitonal accents are schematised in (1.26). The line in bold indicates the limits of the accented syllable. Finally, H\*+H was used to represent a high pitch followed by a sustained high level transition to a following high accent, as opposed to H\* which is followed by a sagging transition before another H\*. In Beckman and Pierrehumbert (1986), however, this pitch accent was eliminated from the tonal inventory since the

---

<sup>14</sup> In Pierrehumbert's original system, leading and trailing tones were marked with a hyphen (e.g. L\*+H-, L-+H\*). In some systems based on this framework (e.g. van den Berg et al 1992, Féry 1993 and Grice 1995a), the hyphen and the plus sign were no longer used (e.g. L\*H, LH\*). In this dissertation, we will keep the plus sign but omit the hyphen as in Ladd (1996) and in the ToBI notational system.

sustained transition between Hs was reanalysed as the result of an elevated pitch range<sup>15</sup>.

(1.26)



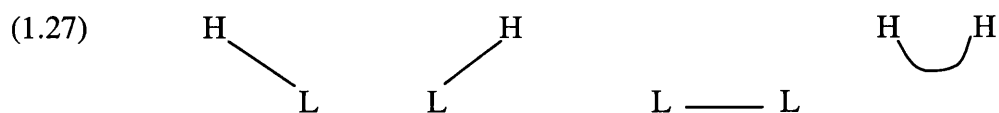
In more recent versions of the AM framework, such as the ToBI system (Beckman and Hirschberg 1994), the number of pitch accents for American English was reduced to H\*, L\*, L\*+H, L+H\* and H+!H\*. In this inventory, a lowered (or downstepped) high tone is marked by the symbol ! preceding the H value. Thus, H+!H\* stands for the former H+L\* and any time an H\* is lowered (i.e. !H\*) it does not have to be preceded by H\*+L. This version of the model will be taken in this dissertation.

2. *Phrase accents.* H- stands for a high pitch following the last pitch accent of the phrase and L- for a low pitch after the last pitch accent.

3. *Boundary tones.* H% indicates a final rise. L% can have two interpretations depending on the preceding tone. After an L- phrase accent, L% indicates a fall to the bottom of the speaking range. After an H- phrase accent, L% stands for a sustained level pitch from the previous tone.

Between the tonal events, the contour is phonologically unspecified (Pierrehumbert and Beckman 1988). However, the overall phonetic contour can be accounted for by phonetic interpolation rules that specify the transition between two tonal values (Pierrehumbert 1980, Ladd 1983b, Grice 1995b). The interpolation between H and L (or L and H) and two L tones is linear, whereas the interpolation between two H tones shows a sag in the F0 contour, as schematised in (1.27) below.

<sup>15</sup> For a discussion on this issue see Ladd (1996:273ff).



### 1.3.2.3 Levels of prosodic structure

In the first version of the AM framework (Pierrehumbert 1980), only one level of prosodic structure (the *intonation phrase*) was postulated. The intonation phrase was a prosodic domain similar to the notions of *tone unit* (Crystal 1969) or *tone group* (Ashby 1978). The intonation phrase consisted of at least one pitch accent and its right edge was delimited by the presence of a boundary tone (L% or H%), usually followed by an F<sub>0</sub> pause. In this early version, phrase accents were not associated to any kind of structure (neither metrical nor prosodic) but were treated as floating tones that accounted for the F<sub>0</sub> contour between the last pitch accent of an intonation phrase and the boundary tone. Later, Beckman and Pierrehumbert (1986) distinguished two levels of prosodic hierarchy in English, namely, the *intonation phrase*, which is a major domain of prosodic structure and the *intermediate phrase*, which is a smaller domain. Intonation phrases were now composed of one or more intermediate phrases. With the introduction of this level of prosodic structure, phrase accents were used to signal the end of this intermediate domain. Thus, an intermediate phrase was described by the presence of at least one pitch accent followed by a phrase accent.

The idea of a hierarchical organisation of prosodic structure in the AM framework is closely linked to the theory of *Prosodic Phonology* or *Prosodic Hierarchy* developed by Selkirk (1984), Nespor and Vogel (1986) and Hayes (1989). However, Prosodic Phonology and Intonational Phonology (or the AM approach) differ in the way phonological constituents are defined. Whereas Prosodic Phonology defines the phonological domains based on the syntactic structure of the sentence, Intonational Phonology defines them based on the intonation of the sentence. For example, Pierrehumbert and Beckman (1988) suggest that intermediate phrases are delimited by the phrase accent and constitute the domain of downstep (see section 1.3.2.4 and also chapter 5 for more details). In this line, Ladd (1986) claims that major levels of prosodic structure are demarcated by means of audible prosodic breaks and minor

levels on the basis of tonal structure (i.e. the presence of a nuclear pitch accent). Selkirk and Nespore and Vogel, in their turn, also postulate an Intonational Phrase of prosodic structure and a lower domain, which they called the Phonological Phrase. However, the identification of these prosodic constituents is determined by the application of junctural (or sandhi) phenomena, since it is argued that sandhi processes have their domains of application.

Despite differences in the definition of prosodic domains, both Prosodic Phonology and Intonational Phonology agree in that a hierarchy of prosodic structure has to conform to the *Strict Layer Hypothesis* (henceforth SLH, Selkirk 1984). In short, the SLH states that higher nodes of prosodic structure must always contain lower nodes in the hierarchy. Thus, an Intonational Phrase can only contain Intermediate Phrases. Below the intermediate phrase, minor domains such as the phonological word, the foot and the syllable have been postulated for several languages. However, as will be argued in chapter 3, the number and kinds of prosodic domains are not universal and vary cross-linguistically. Furthermore, prosodic structure defined on a syntactic basis does not necessarily coincide with prosodic structure determined by intonation, although some studies (Hayes and Lahiri 1991 and Frota 1998) have shown that a syntactic and intonational integrated approach may work for some languages. In this dissertation, prosodic structure will be mainly defined on an intonational basis.

#### 1.3.2.4 Pitch downtrends

As stated in section 1.3.2.2, the AM framework describes intonation by means of only two tones (H and L). This is possible because there is no preassumption that H and L correspond to a particular F<sub>0</sub> value. Thus, H and L are phonological abstractions that can be realised in different ways. For example, the H\* pitch accents illustrated in (1.24) are scaled at different F<sub>0</sub> levels, yet they have the same phonological interpretation. One of the reasons why within the AM approach F<sub>0</sub> contours can be described with only two tones derives from the treatment of pitch downtrend as a local phonologically controlled modification of the pitch range.

It is a well-known characteristic of speech that pitch tends to gradually decline over the course of an utterance (see Ladd 1984). Traditionally, this *declination* has been

analysed as a global phonetic effect, that is, it was believed to be triggered by an automatic physiological mechanism (Fujisaki 1983, 1988 for Japanese; Thorsen 1980 for Danish; Bruce 1977 for Swedish; Lieberman 1967, Liberman 1975, Cooper and Sorensen 1981 for English; among others). Pierrehumbert (1980) and Liberman and Pierrehumbert (1984) discovered that the value of each F0 peak in a descending contour could be expressed as a constant proportion of the previous peak. They found that time-dependent lowering was almost absent in their data and that pitch descent could be characterised as an accent-by-accent decay, which they termed *downstep* (or *catathesis* in Pierrehumbert and Beckman 1988), eliminating the need for a global declination component. Additionally, they found out that the final pitch accent in a sentence undergoes a more drastic lowering in F0 than expected by the application of the downstep rule and proposed to account for it by using a lowering constant. The existence of downstep and *final lowering* has been observed in many other languages apart from English (Japanese: Beckman and Pierrehumbert 1986, Poser 1984; Mexican Spanish: Prieto et al 1996; German: Möbius 1993, Grabe 1998a, to appear). The interpretation of F0 downtrends as the result of the repeated localised occurrence of downstep allows to describe the F0 contours by means of two tones, since differences in the scaling of tones are explained by the downstep (and final lowering) rules.

The phonological characterisation of downstepping contours varies in the literature. Pierrehumbert's (1980) proposes that downstep is triggered by the presence of certain sequences of tones which involved an L flanked by two Hs. Thus, for example, in the sequence H\*+L H\*, the second H\* will be downstepped due to the preceding L tone. This means that the L has no manifestation in the F0 contour but it only has the effect of lowering the following H tone. Later, Beckman and Pierrehumbert (1986) proposed that rather than the L being the downstep trigger, it is the occurrence of a preceding bitonal pitch accent that causes the downstepping effect. Thus, in L\*+H H\*, the second H\* is downstepped (due to the presence of L\*+H), even though there is no intervening L between the two H tones.

The idea that downstep is triggered by bitonality is criticised by Ladd (1983), among others. Ladd argues that downstep is an independent intonational choice which has nothing to do with the tonal sequence. Ladd accounts for downstep by means of a

feature [ $\pm$ downstep], which is a property of the tone that undergoes the lowering. Thus, the sequence  $H^*+L H^*$  is transcribed  $H^* !H^*$ , where ! indicates [+ downstep]. This proposal has been criticised in that it can generate unlikely pitch patterns, such as the presence of a sentence initial downstepped  $H^*$ . However, the ! diacritic has been adopted by the ToBI system (Beckman and Hirschberg 1994, Pitrelli et al 1994) to mark downstepped H tones. In the ToBI annotation conventions, the authors make explicit that this diacritic can never apply to the first H tone in a phrase<sup>16</sup>. Downtrends in Central Catalan will be analysed in chapter 5.

## **1.4 Catalan**

The final part of this introductory chapter presents a few notes on Catalan and its intonational characteristics. Catalan is a Romance language spoken by over seven million people in Spain, Andorra, the southeast of France and the city of Alguer in Italy. In Spain, Catalan is spoken in four autonomous communities: Catalunya, València, Aragón and the Balearic Islands. Catalan has several dialects which are commonly divided into two large groups: Western and Eastern Catalan (Veny 1993). The main phonological/phonetic grounds upon which this division has been made lie in the weak vowel system. Western Catalan has a weak vowel system of five vowels [i, e, a, o, u], whereas Eastern Catalan has a weak vowel system of three vowels [i, ə, u]. The dialect that will be analysed in this study belongs to the Eastern group and is known as Central Catalan. Although there is no well-defined standard Catalan pronunciation, Central Catalan is the best candidate to be treated as "standard", since it is the most widely spoken dialect of the whole Catalan linguistic domain. It is used by more than 80% of the population in Catalunya and is representative of an educated Catalan speaker of the provinces of Barcelona, Girona and part of Tarragona. Within Central Catalan, the subdialect that will be analysed in this dissertation is Northern Central Catalan, spoken in the northeast of Catalunya, near the French border.

---

<sup>16</sup> For other proposals in the modelling of downstep see Ladd (1990, 1993), Grice (1995a, b) and Cabrera-Abreu (1996).



### 1.4.1 Stress language

Catalan, along with other Romance languages, is a stress language. Stress is phonetically manifested by a greater relative intensity of the stressed syllable (Cerdà 1970), a greater duration (Cerdà 1972), as well as by the quality of the vowel (Wheeler 1979). In Central Catalan, stressed vowels have one of the following qualities [i, e, ε, a, ɔ, o, u], whereas unstressed vowels are [i, ə, u]<sup>17</sup>. The close relation between vowel quality and stress in Catalan is illustrated in (1.28). The vowel in *mà* is stressed and with a full vowel quality (i.e. [i<sup>ma</sup>]). When a suffix is added to the word, the location of the stress shifts to the following syllable (i.e. [-<sup>1</sup>nɛ-]) and the vowel in *ma-* is weakened into [ə].

(1.28)	<i>mà</i> (hand)	<i>maneta</i> (little hand)
	i <sup>ma</sup>	mə <sup>1</sup> nɛtə

Catalan has variable word stress, which can fall on the ultimate (e.g. *conill* [ku-<sup>1</sup>niʎ], “rabbit”), penultimate (e.g. *cabra* [kə-βrə], “goat”) or antepenultimate syllable (e.g. *àguila* [a-γi-lə], “eagle”) of a word. Usually there is only one stressed syllable per word, although some compounds show two stresses (e.g. *gratacels* [gratə<sup>1</sup>sɛls], “skyscraper”) and a small number of monosyllabic function words have no stress at all (e.g. articles (*en* [ən], *el* [əl], *la* [lə], “the”), some pronouns (*nos* [nɔs], “us”) and prepositions (*a* [ə], “to”; *en* [ən], “in”; *per* [pər], “for”; *amb* [əm], “with”).

Some studies (Wheeler 1974, 1979) have argued that Catalan stress can be predicted by phonological rules in terms of the segmental properties of strings. Wheeler proposes two rules to assign stress in Catalan: the *Major Stress Rule* (MSR) and the *Minor Stress Rule* (MiSR). The MSR assigns stress to the last non-word-final vowel of a lexical item. Thus, for example [ku-<sup>1</sup>niʎ] is stressed on the last vowel because the word ends with a consonant and hence the last vowel does not appear in word-final

<sup>17</sup> There are a few exceptions to this generalisation (see Wheeler 1979 for more details).

position. ['ka-βrə], on the other hand, is stressed on the first vowel because the last vowel is word-final. The MiSR assigns stress one syllable further back and accounts for those words with stress on the antepenultimate syllable, such as ['a-γi-lə]. Wheeler's arguments were criticised by Mascaró (1978) who found several cases of non-predictable stress in the application of the aforementioned rules and argued that some aspects of Catalan stress have to be specified at the lexicon.

Once stress is assigned, its location tends to remain fixed. However, a few readjustments can be made depending on the rhythmic pattern of the phrase or sentence (Oliva 1992). For example, when *disset* ("seventeen") appears in isolation, the stress falls on the last syllable ([di'sɛt]). If it is followed by a word with stress on the first syllable (e.g. *anys* ([ˈaɲs], "years"), two consecutive stresses appear, resulting in a stress clash. This stress clash triggers a reorganisation of the stress pattern of the phrase: the first stress is displaced to the right and absorbed by the next stress (*Beat Absorption Rule*, BAR, Oliva 1992). Sometimes a stress of "support" may appear on the first syllable of *disset*.

- (1.29) di'sɛt  
           di'sɛt ˈaɲs       (stress clash)  
           →  
           disɛt ˈaɲs       (BAR)  
           (ˈ)disɛt ˈaɲs     (optional stress of support)

As in other Romance languages, the stress distribution in Catalan is the input of intonational phenomena, since stressed syllables are potentially accentable units. In the next section, the main characteristics of Central Catalan intonation will be provided with particular attention to declarative intonation.

#### 1.4.2 Intonation language

Intonation in Catalan applies at the sentence level, that is, tone is not used to distinguish meanings between words but to convey information that affects the whole utterance. Two of the main linguistic functions of intonation in Catalan are those of

highlighting and delimiting chunks and phrases. Thus, intonation is used both to signal the information structure or focus of an utterance and to indicate the division and organisation of discourse.

Work on Catalan intonation has not been very prolific. The first study that refers to intonation in Catalan dates back to 1916. In this work, Barnils stressed the importance of intonational studies to describe languages. This paper mainly attempts to characterise intonation in musical terms but no actual description of Catalan pitch contours is provided. Since this early work, a gap of more than 50 years occurred in the study of Catalan intonation. In the 70s, only two studies appeared and it was not until the 80s and 90s that most studies were produced. These studies can be grouped into four main areas of research:

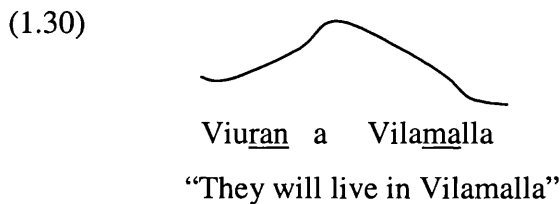
- 1) descriptive studies on Central Catalan (Virgili i Blanquet 1971; Recasens 1977, 1993; Bonet 1984, 1986; Badia i Margarit 1994; Prieto 1995, 1997, to appear b, c; Estebas-Vilaplana 1996; Estebas-Vilaplana and Maidment 1999a, b);
- 2) descriptive studies on dialectal variation (Mascaró 1986, 1987; Prieto to appear a);
- 3) experimental studies (Salcioli 1988a, b);
- 4) comparative studies between Catalan, Spanish and English intonation (Gutiérrez 1983, 1984; Estebas-Vilaplana 1997).

Most of the early work on Catalan intonation (Virgili i Blanquet, Recasens, Bonet, Gutiérrez) was based on a perceptual analysis of sentences. In most cases, no “established” framework of intonational analysis was used and the data was mainly described on a rather subjective basis. The first attempt to accommodate the pitch contours of Catalan into a theory of intonational representation was made by Prieto (1995) who described the data within the AM model of intonational analysis. In the following section, the basic intonational characteristics of declarative sentences produced with an unmarked, neutral intonation (or broad focus) and a marked intonation (or narrow focus) are provided. This description will be based on those

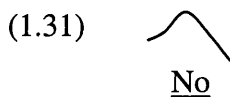
studies that deal with declarative intonation<sup>18</sup>. Particular attention will be paid to the works by Prieto and Bonet, since these studies are the most exhaustive descriptions on Central Catalan intonation available so far.

#### 1.4.2.1 Neutral declarative intonation

Most studies on Catalan intonation (Bonet 1984, Recasens 1993, Badia i Margarit 1994, Prieto 1995, to appear b, c) agree in that the neutral or unmarked intonation of declarative sentences involves two aspects: 1) a final fall in pitch which starts after the penultimate accented syllable and progressively descends till the end of the contour, and 2) a rise during the first accented syllable. This is illustrated in (1.30) reproducing an example from Prieto (to appear b). The stressed syllables are underlined.



According to Prieto, the contour starts at a mid pitch level which is maintained during the prestressed syllable(s). Then, a rise starts on the first accented syllable and expands over the next poststressed syllable. After that, there is a progressive lowering in pitch, which undergoes its major fall over the last accented syllable. When the sentence consists of only one accented syllable, then the rising-falling movement is compressed within this syllable. An example, also provided in Prieto (to appear b), is reproduced below.



Finally, when the utterance has more than two stressed syllables, the accent distribution is variable. There can be as many accents as stressed syllables, but the final number of accents is a matter of choice on the speaker's part. All non-final

---

<sup>18</sup> The works by Salcioli and Mascaró are not reviewed here. Salcioli's work mainly deals with the intonation of interrogations and Mascaró studies other Catalan dialects rather than Central Catalan.

accented syllables show a rise, followed by the fall on the last accented syllable. Two examples from Prieto are presented in (1.32) below. In the two pitch contours there are four accents. The main difference lies in the distribution of the accents. In (1.32a) there is no accent in *ahir*, whereas in (1.32b) *arribar* does not have an accent.

(1.32) a.



La meva cosina va arribar ahir a la nit

“My cousin arrived yesterday night”

b.



La meva cosina va arribar ahir a la nit

Sometimes declarative sentences are produced with more than one tone unit. This might be due to: 1) differences in style (reading speech vs casual speech), 2) the length of the constituents (long constituents tend to be followed by a prosodic boundary) and 3) the necessity to disambiguate semantically ambiguous sentences. These aspects are discussed in more detail by Prieto (to appear c). (1.33) presents a sentence with a prosodic boundary after a long subject (example from Prieto to appear b) and (1.34) shows two segmentally equivalent sentences, which acquire different meanings depending on the location of the prosodic boundary (example from Bonet 1984).

(1.33)



La meva cosina de Girona / arriba demà

“My cousin from Girona / comes tomorrow”

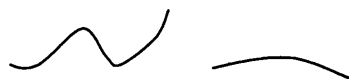
(1.34) a.



Una jove / veu l’amenaça

“A young girl / sees the threat”

b.

Una jove veu / l'amenaça

“A young voice / threatens him”

Prieto (1997, to appear b) describes the intonation of sentences with more than one tone unit as follows. The second tone unit behaves as if it was an isolated prosodic phrase with declarative intonation. The first tone unit, on the other hand, shows a significant change in pitch at the end of the phrase. This change in pitch can be characterised by a continuous rising pitch, as in the examples above, or can show a falling-rising movement.

Following the AM approach, Prieto (1995) interpreted the pitch contours with a declarative intonation as follows. When a sentence consists of only one accented syllable, as in (1.31), this is associated to an H\* pitch accent followed by L-L% edge tones. When the pitch contour has more than one accented syllable, as in (1.30) and in (1.32), the last accent is transcribed as L\* and all prenuclear accents as H\*. As before, the end of the intonation phrase is marked L-L%. The tonal transcriptions for the sentences in (1.30-1.32) are illustrated in (1.35) below.

- (1.35) a. No  
H\* L-L%
- b. Viuran a Vilamalla  
H\* L\* L-L%
- c. La meva cosina va arribar ahir a la nit  
H\* H\* (H\*) (H\*) L\* L-L%

As far as the transcription of the sentences produced with two tone units is concerned, Prieto (1997) proposes an intermediate phrase boundary (marked with an H- phrase accent) after the first tone unit. Thus, her interpretation of (1.33) and (1.34) would be as in (1.36) and (1.37) respectively. In none of her papers does she propose a notation for the falling-rising contour at the end of the first tone unit.

(1.36) La meva cosina de Girona / arriba demà  
           H\*    H\*            H\* H- H\*    L\* L-L%

(1.37) a. Una jove / yeu l'amenaça  
           H\* H- (H\*)    L\* L-L%

      b. Una jove yeu / l'amenaça  
           H\* H\* H-    L\* L-L%

Prieto's phonological interpretation of the declarative F0 traces differs in some ways from Bonet (1984). Bonet's interpretation is mainly based on her own perceptual judgements. Bonet's proposal is briefly presented below since her analysis will be relevant for the interpretation of our data. Bonet describes the intonation of accented syllables but not the pitch movement at the edges of the contour.

Bonet claims that the intonation of Central Catalan declaratives could be accounted for by means of three principles, which are related to the number of prosodic constituents (or tone units) within a sentence:

1. all accented syllables in the last prosodic constituent are low;
2. the last accented syllable of a prosodic constituent is low;
3. accented syllables (not in the last constituent and not the last syllable of its own constituent) are high.

According to these rules, sentences produced with a tone unit (examples (1.30)-(1.32)) have a low tone on every accented syllable. On the other hand, sentences with two tone units, such as (1.33) and (1.34), will present the following patterns:

(1.38) La meva cosina de Girona / arriba demà  
           H    H            L    L    L

(1.39) a. Una jove / yeu l'amenaça  
           L    L    L

- b. Una jove veu / l'amenaça  
H L L

Two basic differences can be observed between Bonet's and Prieto's tonal descriptions. First, whereas Prieto characterises the last pitch accent of the first tone unit as H\* (followed by H-), Bonet assumes that this pitch accent is low. Second, whereas in Prieto's analysis all prenuclear accents are H\*, in Bonet's analysis can be either L or H depending on whether they appear on the final prosodic constituent or not. This different interpretation raises the issue of which is the best way (perceptual or acoustic) to decide the tonal categories. Bonet treats some of the prenuclear syllables as low since she perceives them with a low pitch in comparison to the surrounding syllables. This contrasts with the acoustic analysis of Prieto who, by observing the phonetic details in the pitch contours, analyses the same pitch accents as high. In the third chapter of this dissertation, this issue will be revisited and both perceptual and acoustic evidence will be provided to elucidate this problem of interpretation.

#### 1.4.2.2 Marked declarative intonation

Only three studies on Catalan intonation (Bonet 1984 and Prieto 1995, to appear b) briefly deal with declaratives produced with a marked intonation or narrow focus. All these studies agree in that the non-neutral declarative intonation involves a pitch accent on the focussed element which has a higher pitch than the equivalent accent in neutral declaratives, and that the stretch of words after the focussed item is produced with a low and flat intonation. Prieto (to appear b) describes the pitch accent on the focussed material as consisting of a rising pitch over the accented syllable that falls drastically after the syllable down to the baseline of the speaker's range. She also claims that focus triggers a prosodic segmentation of the intonation phrase, which occurs right after the focussed material. Prieto's (to appear b) example of a pitch contour with narrow focus on the subject is illustrated below.



(1.40)



LA MARINA / va demanar-li

“MARINA / asked him”

Both in Bonet and in Prieto the phonological interpretation of narrow focus sentences is associated to some kind of high accent on the focussed element. Prieto (1995) proposes four tones that can be linked to the focussed word and that have different nuances: !H\* (for emphasis), L+H\* (for surprise and contrast), L+H\*+H (for contradiction) and H+!H\* (with a didactic meaning). The phonological interpretation of postfocal material is not clear in any work. In this dissertation, the modelling of narrow focus contours will be examined in more detail in chapter 4.

## 1.5 Conclusion and outline

In this introductory chapter the aims of this dissertation have been specified and a few works on focal and intonational issues relevant for the present study have been reviewed. In particular, works on three main areas of research have been covered. First, a brief introduction to the notion of focus from several perspectives and to the ways focus is expressed cross-linguistically has been provided. Second, the main tenets of the AM framework have been presented and along with previous models of intonational analysis. Finally, a brief account of Central Catalan and its intonational properties has been given. In the following chapters, the three research aims of this dissertation will be examined in the following way. Chapter 2 is concerned with the first question of this study, that is, it examines to what extent Central Catalan speakers use accentual focus and how natural the use of this strategy is felt to be in this language as opposed to a syntactic modification. This is investigated by means of three tests (a production test, a perception test and an acceptability test). Chapters 3-5 deal with the second aim of this thesis, that is, to describe the phonetic and phonological properties of Central Catalan sentences produced with broad and narrow focus. Chapter 3 analyses the main characteristics of a neutral or broad focus declarative intonation in Central Catalan. Chapter 4 examines the phonetic and phonological properties of sentences with a narrow focus reading and compares the

results to those of sentences with a broad focus interpretation. In these two chapters, differences between phonological association and phonetic alignment are analysed in detail. Chapter 5 investigates F0 downtrends in utterances with a broad and a narrow focus reading in Central Catalan. Chapter 6 is concerned with the last aim of this study, namely, to compare the intonational properties of Central Catalan narrow focus sentences produced with intonational means to those of a language that mainly uses accentual strategies to express focus, i.e. English. Finally, in chapter 7 the conclusions of this thesis and possible consequences and implications are presented.

## ***Chapter 2: Accentual focus in Central Catalan***

### **2.1 Introduction**

This chapter analyses to what extent Central Catalan, a language that has been classified as [-plastic] or discourse-configurational, can use accentual strategies alone to signal focus and how acceptable accentual strategies are felt to be as opposed to syntactic devices. As pointed out in the last chapter, the literature (Ladd 1980, 1996, Gussenhoven 1983a, 1984 and Vallduví 1990, among many others) agrees that there is a close relationship between focus and accent, that is, every time an element is focussed, it is also accented. However, the way this focus/accent relation is attained has been observed to be different cross-linguistically. The literature seems to divide languages into two groups: 1) languages that need some kind of syntactic reorganisation to attain the focus/accent association and 2) languages that do not need it. This two-way classification of languages has received different names. Languages that use syntactic devices have been called [-plastic] (Vallduví 1990, 1991, 1994a/b and Vallduví and Zacharski 1994), discourse-configurational (Kiss 1995) or [+syntactic] languages (Frota 1993). On the other hand, languages where syntactic strategies are used less frequently have been classified as [+plastic], non-discourse-configurational or [-syntactic].

Most studies on the expression of focus in Catalan (Vallduví 1990, 1991, 1994a/b) have shown that Catalan is a language that needs a syntactic shift to attain the focus accent relationship. As illustrated in the introduction (section 1.2.4), the focussed element moves into a prominent position (usually the last position of the sentence) to become accented. However, some studies on other Romance languages have demonstrated that this cross-linguistic variation is not so clear-cut and that sometimes languages that belong to the [+syntactic] group, such as Spanish or French, do use intonational devices to convey focus. Studies have claimed that factors, such as the kind of triggering question for narrow focus and the position of the focussed element within the sentence, might have an influence on the focal devices. For example, García-Lecumberri (1995) for Peninsular Spanish and Zubizarreta (1998) for French have shown that these languages prefer to use accentual devices in contrastive narrow focus rather than in identification narrow focus.

Similarly, they have proposed that accentual focus is more frequently used in sentence-initial position, such as subject position, than in other positions, namely, medial (verb) or final (object).

In this chapter, the possibility of using accentual focus in Central Catalan, a language that has been formerly categorised as [-plastic] or discourse-configurational, will be examined. In particular, two aspects will be covered:

1. the *extent* of usage of accentual focus, and
2. the *acceptability* of accentual strategies as opposed to syntactic strategies to convey focus.

The first aspect is investigated by means of a *production test* and a *perception test*. With these tests the claim that Central Catalan speakers can both produce and recognise focus triggered by accentual means is investigated. For the second aspect, an *acceptability test* is performed to examine how natural accentual strategies are felt to be, as compared to syntactic devices. In all tests, two variables will be taken into account: 1) the sentence position of the focussed element (initial vs non-initial) and 2) the triggering question, i.e. contrastive or identification narrow focus. The predictions concerning these two variables are that, as observed in other Romance languages, Central Catalan favors accentual focus in sentence-initial position and in contrastive prompts.

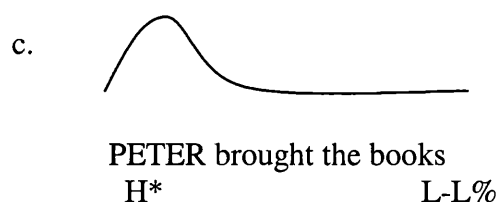
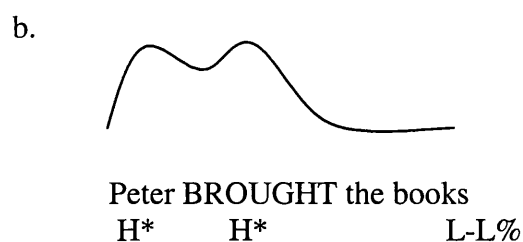
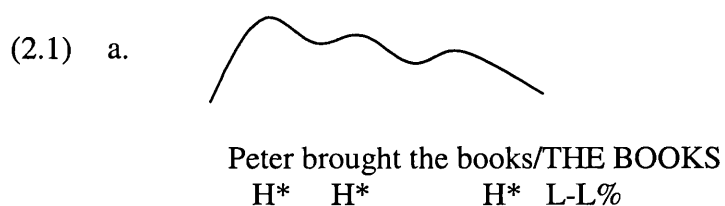
The organisation of this chapter is as follows. Section 2.2 presents a brief review of the ways accentual focus can be manifested cross-linguistically. Then, sections 2.3, 2.4 and 2.5 describe the experimental procedures and the results of the production test, the perception test and the acceptability test respectively. Finally, the conclusions of this chapter are provided in section 2.6.

## **2.2 Accentual focus cross-linguistically**

As will be discussed in more detail in chapters 4 and 6, the expression of focus by intonational means is manifested in three different ways across languages: 1) by locating

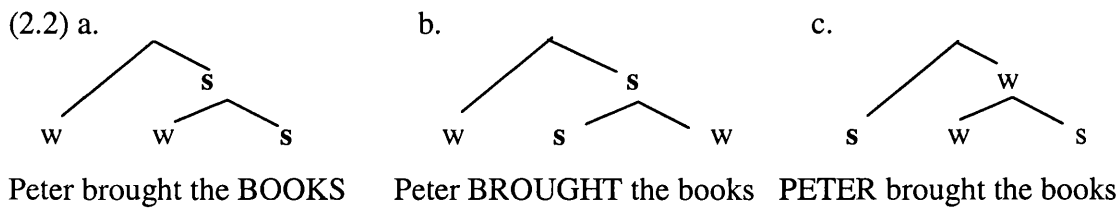
the nuclear accent on the focal constituent and deleting postfocal accents, 2) by choosing a different accent type and 3) by changing the phrasing structure of the sentence.

In English (Ladd 1996), German (Féry 1993) or Dutch (Gussenhoven 1994), focus is signalled by the placement of the nuclear accent on the highlighted item with deaccenting of postfocal material. (2.1) illustrates the intonational patterns for the sentence “Peter brought the books” produced with different readings. (2.1a) has two interpretations, either broad focus or narrow focus on the object. (2.1b) shows narrow focus on the verb and (2.1c) narrow focus on the subject. In all cases, a pitch accent, marked as H\*, is associated to the focussed element (in capitals). All prefocal stressed syllables also get a pitch accent, as in (2.1a) and in (2.1b). However, postfocal stresses do not receive any pitch accent. This is illustrated in (2.1b) and in (2.1c), where after the nuclear accent the contour remains low and flat (marked with L-L%).



The lack of pitch accents after the focal material is known as *deaccenting* (Ladd 1980). Ladd (1996:175) points out that deaccenting occurs when “a word that we might expect to be accented fails to be accented in a context where it has recently been used or where

the entity to which it refers has recently been mentioned". Ladd accounts for deaccenting not as a result of accent shift or displacement but as a "reversal of relative strength in a metrical tree" (Ladd 1996:229). Thus, when a word is deaccented according to pragmatic reasons, another word must be accented due to the relational nature of prominence. The prominence patterns of (2.1) above are illustrated in (2.2) below.



Studies on intonational focus across languages show that changes from a neutral prominence pattern (*w-s*) into a narrow focus pattern (*s-w*) do not always involve deaccenting of the postfocal material. Whereas in English, German and Dutch deaccenting after the focal pitch accent is compulsory, in other languages a postnuclear pitch accent seems to be accepted (e.g. Peninsular Spanish: García-Lecumberri 1995; Palermo Italian: Grice 1995a; Maltese: Vella 1995; Bari Italian: Grice and Savino 1997 and Neopolitan Italian: D’Imperio 1997a, b). All these studies agree in that the postnuclear accent has two properties: it coincides with the last stressed syllable of the utterance and it has a compressed pitch range.

Another way to express focus phonologically is by choosing a different tune type or pitch accent from that used in broad focus utterances. Whereas languages such as English (Ladd 1996), Dutch (Gussenhoven 1983b, 1994, 2000) and German (Féry 1992, 1993) use the same accent type for the nucleus of broad and narrow focus sentences, other languages, such as Bengali (Hayes and Lahiri, 1991), Palermo Italian (Grice 1995a), European Portuguese (Frota 1993, 1998) and Greek (Arvaniti and Baltazani, 1999) use a different tune type for the marking of broad and narrow focus. In European Portuguese, for example,  $H+L^*$  is used as the nuclear accent of broad focus and  $H^*+L$  is used for narrow focus.

Finally, the third way to mark focus intonationally is by means of phrasing. In languages such as Korean (Jun 1996) and Japanese (Venditti et al 1996), it has been argued that whenever an element is focussed, it forms a separate phrase. However, when it is not focussed the boundaries are lost, i.e. it is *dephrased*. This is illustrated below reproducing an example from Jun (1996), cited in Ladd (1996). When [iri:mi] “name” is the point of the question, it is produced in a separate phrase, as in (2.3a). When it is old information (2.3b), it is dephrased.

- (2.3) a. [sat]<sup>h</sup>un-ənni] [iri:mi][mwəni]  
“What is cousin’s name?” (lit. cousin name what)
- b. [sat]<sup>h</sup>un-ənni iri:mi][suni-dzi]  
“Cousin’s name is Suni” (lit. cousin name Suni)

In the present chapter, any time we refer to the use of accentual strategies in Central Catalan, we will be basically looking for a reorganisation in the location of the nuclear accent followed by postfocal deaccenting *and* no alteration of the syntactic structure. The possibility that the expression of narrow focus by accentual means in Central Catalan may also involve a different accent type or a reorganisation in the levels of phrasing (along with accent location and deaccenting) will not be dealt in this chapter but will be analysed in chapter 4.

### **2.3 Production test**

The production test and the perception test were designed to answer the first question posed in this study, concerning the extent of use of intonational focus in Central Catalan. The production test examines the claim that Central Catalan speakers can signal narrow focus by intonational means alone (such as accent location and deaccenting) with no need for a syntactic reorganisation.

## 2.3.1 Experimental design

### 2.3.1.1 Materials

The selection of the corpus for the production test followed some of the criteria proposed by García-Lecumberri (1995) for the analysis of Spanish and English intonational focus. In line with this study, only SVO declarative affirmative sentences are analysed according to two aspects: *focus domain* (or sentence position) and *focus trigger*.

Two focal domains are covered: focus on the *subject* and focus on the *verb*. Sentences with these focal domains consist of a marked, non-neutral stress/accent pattern of the kind *s-w*. These two domains are included to analyse the effects of sentence position on the realisation of accentual focus. Some studies on Romance languages (García-Lecumberri 1995, Zubizarreta 1998) have shown that accentual focus on the subject is more likely to occur than on the verb. In line with these studies, it is expected that Central Catalan speakers prefer to use intonational focus on initial position (subject) than on other positions (verb). In this test, focus on the object is not covered due to its potential ambiguity with broad focus structures, as observed in English, German or Dutch<sup>1</sup>. Furthermore, the analysis of subject and verb focal domains alone was considered sufficient to prove the claim that accentual strategies can be used in Central Catalan to express focus. Examples of subject and verb focus domains are provided in (2.4a) and (2.4b) respectively. The questions that triggered the two kinds of focus structures are also included. Overall, the data consisted of 32 sentences with narrow focus on the subject and 32 with narrow focus on the verb. See Appendix 1 for the whole corpus of sentences.

- (2.4) a. En Jordi llegia llibres? LA ROSA llegia llibres  
“Did Jordi read books? ROSA read books”

---

<sup>1</sup> Although in this test object focus was not investigated, for the phonological/phonetic analysis of the data (see chapter 4) a few sentences with narrow focus on the object were collected to detect possible differences in the accent type between broad and narrow focus sentences, and hence examine whether Central Catalan presents ambiguity (or not) between a broad focus and a late narrow focus reading.



- b. La Rosa talla els geranis? La Rosa REGA els geranis  
“Does Rosa cut the geraniums? Rosa WATERS the geraniums”

The second variable analysed in this study relates to the question that triggers narrow focus, namely, focus triggered by *contrast*, which basically involves the correction of an erroneous assumption, and focus triggered by *identification*, which consists in retrieving a new piece of information (see section 1.2.1 in chapter 1 for more details). Although the semantics of narrow focus may cover other meanings (see Kiss 1995), the literature agrees (García-Lecumberri 1995, Kiss 1995, Frota 1998) that contrast and identification are two of the commonest readings expressed by narrow focus. The contrast/identification variable was included to see whether intonational focus in Central Catalan covers these two functions or is restricted to one of them. Zubizarreta (1998) points out that Romance languages differ in the contexts where narrow focus can be signalled by intonational means. According to Zubizarreta, accentual focus is mainly used with a contrastive reading in Spanish. In French, on the other hand, it is used in both contrastive and identification contexts. Zubizarreta’s claims for Spanish partially agree with the findings of García-Lecumberri (1995), who showed that Spanish speakers do use accentual strategies in identification contexts but to a lesser extent than in contrastive contexts. In line with these statements, it is expected that, in Catalan, accentual focus is merely used with a contrastive reading, and that in identification contexts accentual focus is less used or not used at all. Overall, the data includes 32 sentences triggered by a contrastive prompt and 32 triggered by an identification question. The same number of contrastive and identification sentences are examined in the two focal domains, i.e. 16 contrastive sentences and 16 identification sentences for both subject and verb focus. Examples of both kinds of functions are exhibited in (2.5). (2.5a) illustrates contrast in verb position and (2.5b) identification in subject position. See Appendix 1 for the list of sentences along with the triggering questions.

- (2.5) a. La Rosa talla els geranis? La Rosa REGA els geranis  
“Does Rosa cut the geraniums? Rosa WATERS the geraniums”

- b. Qui vivia a Vilavella? L'ANNA vivia a Vilavella  
“Who lived in Vilavella? ANNA lived in Vilavella”

The corpus of sentences consist of simple, short structures in order to avoid speakers dividing their utterances into more than one intonation group. The average number of syllables per sentence varies between eight and ten. The shortest structure, however, has six syllables and the longest one fifteen syllables. The target sentences selected for this study includes the greatest number of voiced sounds possible so as to obtain uninterrupted F0 contours.

Each focal domain consist of either one or two lexically stressed (and hence potentially accentable) words. These two stress distributions were mainly included for further analysis of the data concerning aspects such as downstep, phrasing or peak alignment among others (see chapters 3-5). 16 sentences with a single stressed lexical item and 16 sentences with two lexically stressed words were elicited for focus on the subject and on the verb respectively. Overall, 32 single stressed domains and 32 double stressed domains were included. Double stressed subject structures comprise a pre or postmodified noun, namely, a noun preceded or followed by an adjective. Double stressed verb structures include periphrastic combinations of the kind "modal + infinitive". In double stressed domains, the nuclear accent was expected to be on the second item. Examples of single and double stressed structures in subject position are shown in (2.6). The stressed syllable is underlined.

- (2.6) a. *Single stressed structure*

L'EMÍLIA vol amanida  
“EMILIA wants salad”

- b. *Double stressed structures*

LES MEVES AMIGUES eren bones nenes  
“MY FRIENDS were good girls”

Each of the focus domains contains two kinds of stress distributions: 1) domains where the potentially accented syllable is not preceded by any other, stressed or unstressed, syllables and 2) domains where the potentially accented syllable is preceded by unstressed syllables, either in the same word or in different words. Again this was included for further phonological and phonetic analyses. Examples of type 1 are *L'Anna* (Anna) for a subject domain and *rega* (he/she waters) for a verb domain. Monosyllables, such as *jo* (I) or *beu* (he/she drinks), for subject and verb domains respectively, were also included. Examples of type 2 are *en Jaume* (James) for a subject domain and *anava* (he/she went) for a verb domain.

For each speaker, 136 sentences were elicited which comprised 8 trial sentences, 64 distractors and 64 target sentences. The number of target sentences according to the variables (focal domain and focus trigger) are the following: 32 sentences with expected narrow focus on the subject (16 triggered by contrast and 16 by identification) and 32 sentences with expected narrow focus on the verb (16 triggered by contrast and 16 by identification). This is summarised in Table 2.1.

Target sentences				Distractors	Trial sentences
Subject narrow focus		Verb narrow focus		64	8
Contrast	Identification	Contrast	Identification		
16	16	16	16		

Table 2.1. Number of target sentences, distractors and trial sentences recorded for each speaker.

Trial sentences were included for the informants to get used to the task procedure. These sentences, however, were not taken into account for analytical purposes. Additionally, sentences were included to distract the speakers from the focussing activity and avoid a routine intonation in the elicited data. Distractors mainly consisted of sentences with either broad focus or narrow focus on the object or on the final complement. The two kinds of distractors are illustrated in (2.7a) and (2.7b) respectively. Distractors and trial sentences are listed in sections 1.2 and 1.3 of Appendix 1.

- (2.7) a. Quin temps fa? Està plovent  
“What's the weather like? It's raining”
- b. On aniràs de vacances? De vacances aniré a Cuba  
“Where will you go for holidays? For holidays I'll go to Cuba”

### **2.3.1.2 Data collection procedures**

The data used in this study were gathered by means of two tasks: a *reading activity* and a *response to visual cue task*. The reading task was included to guarantee the acquisition of sentences with the major number of voiced sounds possible, so as to be able to perform a later acoustic analysis of the F0 traces. The inclusion of a response to visual cue task was aimed at obtaining more natural and (semi-)spontaneous speech. This task was also planned to give the informants more freedom to choose the kinds of strategies they prefer to signal the information structure of Central Catalan utterances, namely, intonational or syntactic structures. In both the reading and the semi-spontaneous task, speakers were asked to answer some questions posed by the researcher. 68 sentences were elicited in each task (4 trial sentences, 32 distractors and 32 target sentences).

The eliciting procedures used for the reading task were designed following García-Lecumberri (1995). For the reading activity the speakers were given nine cards. The first card consisted of four trial sentences. The other cards had eight sentences each: four distractors and four target sentences. Each sentence was the answer to a question. The order of the sentences, however, was randomised and did not correspond to the order in which the questions were posed. This had the purpose of distracting the speakers from being too concerned with the "right" pronunciation of the utterances and at obtaining the highest naturalness possible in their answers. For this task, the prompting questions were previously recorded by the researcher so that the stimuli were the same for all the informants. Between each question a gap of 10 seconds was left for the speakers to answer.

For the response to visual cue task (or semi-spontaneous activity), the informants were shown three pictures and were asked some questions related to them. The pictures contained several sketches of people doing different activities. Most of the characters had a label at the top, so as the informants could easily detect the target scene for each question. These labels were included in order to obtain a large number of similar, and hence comparable, responses. An example of an sketch for a sentence with expected narrow focus on the subject triggered by identification and another one with expected narrow focus on the verb triggered by contrast are illustrated in Appendix 1 (section 1.4). Similar to the reading activity, the order of the questions and that of the different sketches did not coincide so that speakers were involved in the task, rather than in the accuracy of their productions, and hence their speech was the most natural possible. Unlike the reading task, the triggering questions for the response to visual cues were not recorded but were produced live by the researcher in order to prompt more informal and spontaneous speech.

In Appendix 1, sentences designed for the reading activity and those designed for the semi-spontaneous task are differentiated by means of the letter font. Whereas sentences in standard letters are those gathered for the reading task, sentences in italics are the expected responses of the semi-spontaneous activity (if the speaker decided to use accentual strategies to convey focus).

### **2.3.1.3 Informants**

For the production test, thirty-four subjects (seventeen male and seventeen female adult native speakers of Central Catalan) were recorded. All the speakers had similar characteristics as far as age, geographical origin, social status and education are concerned. The informants were between 24 and 32 years old at the time of the recordings. Each of the subjects was born and lived in the same town, Ripoll, situated in the province of Girona, 100 km from Barcelona. They were brought up in middle class Central Catalan-speaking families. Like all Catalan speakers, all the subjects were bilingual and spoke Spanish as a second language. However, Catalan was the language they always used at home, work and in normal daily conversations. Some of them also

know a foreign language (English or French), mainly used for reading purposes. All the speakers had finished secondary school and most of them had a university degree or were involved in completing one. None of the speakers had lived in countries where other languages are spoken. The variety of Catalan they used is Central Catalan.

#### **2.3.1.4 Instrumentation and recording**

The recordings included two simultaneous but separate signals: *speech* and *laryngeal* (Lx) signals. For the speech signal, the instrumental tools used to obtain the data were a Marantz Superscope/CD 330 tape recorder and a Beyer Dynamic microphone on a stand. For the laryngeal signal, a portable laryngograph LX 12 with two electrodes was used. The laryngograph is an apparatus used to record vocal fold vibration. It operates by passing a weak electric current between a pair of electrodes placed in contact with the skin on both sides of the larynx. The variation of electrical impedance produced with the vibration of the vocal folds is measured by the device. Thus, the output of the device is a signal analogous to the degree of vocal fold contact at a given instant. This waveform of impedance against time is known as Lx. Lx may be further processed to derive information about the duration of the period of vocal fold vibration (Tx) and to measure instantaneous fundamental frequency (F0)<sup>2</sup>. A Thandar portable DRO 26 oscilloscope was also used to check the activity of the vocal folds. The data were recorded onto Fuji Chrome DR-II Type 2 tapes.

The recordings were conducted in a soundproof room of a private recording studio in Ripoll. Each subject performed the task individually. For the reading task, the speakers wore headphones to listen to the prompt questions played on a Philips A21101 cassette recorder. For the semi-spontaneous task, the researcher posed the questions. The microphone for the recordings of the speech signal was set at about twenty centimetres from the speakers. To get the Lx signal a pair of electrodes was set on the speakers' neck by means of a collar band. The electrodes were connected to the portable laryngograph. Before the actual recording, subjects were asked to read a short passage to adjust the gain controls of the tape recorder and the laryngograph.

---

<sup>2</sup> Refer to section 2.3.2.2 to see how F0 traces were obtained from the Lx signal.

### 2.3.2 Analysis of the data

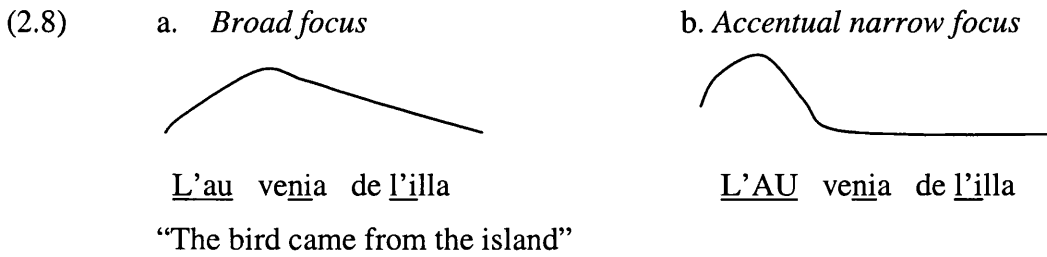
The overall number of sentences obtained in the two production tests consisted of 4352 responses (136 sentences x 34 speakers). Out of this amount, sentences were divided in the following way: 2176 target sentences, 2176 distractors and 272 trial sentences. Finally, within the target structures, 1088 were gathered by means of the reading test and 1088 as responses to the visual cue task.

In order to investigate the first question posed in this study, namely, whether accentual strategies are used in Central Catalan to signal the information structure of sentences, *auditory* and *acoustic* analyses of the target sentences obtained in the whole production test (reading task and semi-spontaneous activity) were performed.

Since for the reading activity the informants had no chance of using any focussing device other than accentual strategies, it was assumed that if Central Catalan speakers do not signal focus by intonational means, then sentences would be produced with a broad focus intonation. On the other hand, if they use accentual devices, this would involve the location of the nuclear accent on the focussed word and possible deaccenting of postfocal material. As reported in the literature (Bonet 1984, Recasens 1993, Badia i Margarit 1994, Prieto 1995, to appear b, and also chapter 3 of the present study), if sentences were performed with a broad focus intonation, we would expect a rising pitch over the first accented syllable (or over any prenuclear accented syllables) followed by a progressively falling contour. This is illustrated in (2.8a) for the sentence *l'au venia de l'illa*. On the other hand, if utterances were produced with the expected narrow focus on the subject or on the verb (signalled by a prosodic shift), a high pitch is expected over the accented syllable of the focal element followed by an immediate fall and postfocal deaccenting. Prefocal stressed items might or might not be accented. If so, a rise is expected to occur over the accented syllable. This is illustrated in (2.8b) for the same utterance as in (2.8a) but with narrow focus on the subject<sup>3</sup>.

---

<sup>3</sup> Since in this chapter we are mainly interested in the *use* of accentual focus, for more details in its phonetic and phonological properties, see chapter 4.



For the semi-spontaneous task, before carrying out the auditory and acoustic analyses of the target sentences, the kinds of responses provided were annotated. Basically, four types of utterances were obtained: 1) sentences produced with accentual narrow focus, 2) with broad focus, 3) with syntactic narrow focus and 4) other. For more details on the last two types of responses, see section 2.3.3.2.

Before presenting the results of the reading task and the semi-spontaneous activity concerning the use of accentual focus in Central Catalan, a few notes are provided about the procedures used to conduct the auditory and the acoustic analyses of the data.

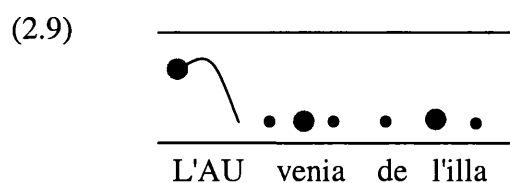
### 2.3.2.1 Auditory analysis

For the auditory analysis, the researcher listened to the target sentences as many times as necessary and outlined the pitch movement perceived in each utterance. This was done on an interlinear graph representation or tadpole diagram (O'Connor and Arnold 1973). A dot was placed at the top of each syllable. Big dots indicated stressed syllables and small dots unstressed syllables. The perceived pitch of each syllable was marked according to the height of the dots. In monosyllabic words, the pitch movement was drawn by means of a line. Although this analysis might not be very precise and sometimes difficult to assess, it still gave us a first impression of the intonational tendencies obtained in the data. Sentences were also listened by two other Central Catalan native speakers, who confirmed the auditory judgements of the researcher.

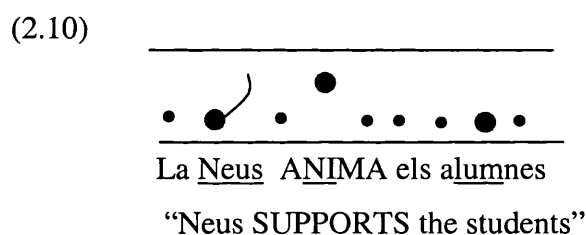
An example of an auditory trace for an utterance with narrow focus on the subject, signalled by accentual means, is provided in (2.9). A high pitch was perceived on the



accented syllable of the focussed word. This high pitch was followed by an immediate fall. The intonation for the rest of the sentence was heard as low and level.



For a sentence with narrow focus on the verb, such as *la Neus ANIMA els alumnes*, a rising contour was perceived over the first accented syllable. As before, a high tone was observed over the focussed item followed by an immediate fall. Postfocal material tended not to be accented. This is illustrated in (2.10).



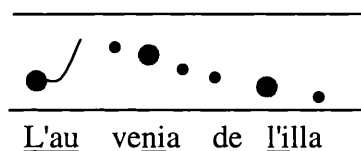
Although in most sentences postfocal material was perceived without any pitch change, in a few utterances a postnuclear accent was detected over the last lexical word of the sentence. This tended to be perceived as a weaker accent than the nuclear accent. An example is provided in (2.11) for the sentence *la Neus ANIMA els alumnes* with focus on the verb. Postnuclear pitch accents have been observed in other Romance languages (e.g. Peninsular Spanish: García-Lecumberri 1995; Palermo Italian: Grice 1995a; Maltese: Vella 1995; Bari Italian: Grice and Savino 1997; and Neopolitan Italian: D’Imperio 1997a, b). In this study, more details on postnuclear accents in Central Catalan are provided in chapter 4 (section 4.4.1.4).

(2.11)



Finally, (2.12) provides the auditory trace of an utterance where the speaker failed to use accentual focus but produced a broad focus intonation as a response to a question which was supposed to trigger narrow focus. In these cases, prenuclear accented syllables were perceived with a rising pitch followed by a progressive fall till the end of the utterance.

(2.12)



### 2.3.2.2 Acoustic analysis

In order to confirm the intonational judgements obtained in the auditory analysis, an acoustic analysis of some of the utterances was carried out by means of the Speech Filing System (SFS) program. This program allows for a simultaneous inspection of several signals: in this case, speech waveform and F0 trace. To transfer the speech and laryngeal signals obtained in the recordings into the computer, the two signals were played on a Dennon DN-770R tape recorder connected into a Sun Sparc-10 computer running SFS. Acquisition of the signals was done at 16 KHz sampling rate, following the routines of the program.

In order to obtain the F0 traces from the laryngeal signal, the VTX and FX programs were used. VTX converted the laryngeal waveform (Lx) into excitation period measurements (Tx). From VTX excitation period measurements, F0 traces were obtained by means of the FX program, which converts a Tx item into a F0 item. Within this program the option -s (smoothing) was used in order to get a pitch contour with less abrupt changes. The smoothing was done with a 5-point median filter.

For each speaker, a separate directory was created, which contained some of the target productions stored on individual files. However, not all target sentences were introduced for all speakers due to the large amount of data (overall 2176 target sentences and 2176 distractors). Basically, only a few examples (around 5 or 6) were stored for each speaker. All sentences that were doubtful in the auditory analysis were also included. At this stage of research, the aim of the acoustic analysis was just to backup the auditory judgements. A more detailed acoustic analysis of all utterances for 6 speakers is carried out at a later stage of study (see chapters 3-5).

Figure 2.1 illustrates the SFS display screen for the sentence *l'au venia de l'illa* produced with a *w-s* or broad focus intonation, as in (2.12). The order of the different signals observed in the screens is: 1) speech waveform, 2) Lx waveform, 3) excitation period measurements and 4) F0 contour. The last screen is used to annotate relevant points in the utterance. In this example and in the following ones, the beginning of each syllable is marked with a vertical line. Once F0 traces were obtained for each sentence, the screens corresponding to the Lx and the Tx items were deleted. This is illustrated in Figures 2.2-2.4. The F0 traces presented in these Figures correspond to the utterances presented in (2.9-2.11). As expected, the F0 traces agree with the auditory judgements illustrated above. Figure 2.2 shows the utterance *L'AU venia de l'illa* produced with narrow focus on the subject. Figure 2.3 exhibits *la Neus ANIMA els alumnes* produced with narrow focus on the verb. In both Figures, an F0 peak can be observed over the focussed element. Postfocal material is clearly deaccented. Figure 2.4 shows *la Neus ANIMA els alumnes* produced with narrow focus on the subject. In this example, a postnuclear accent is observed on the last lexical word.

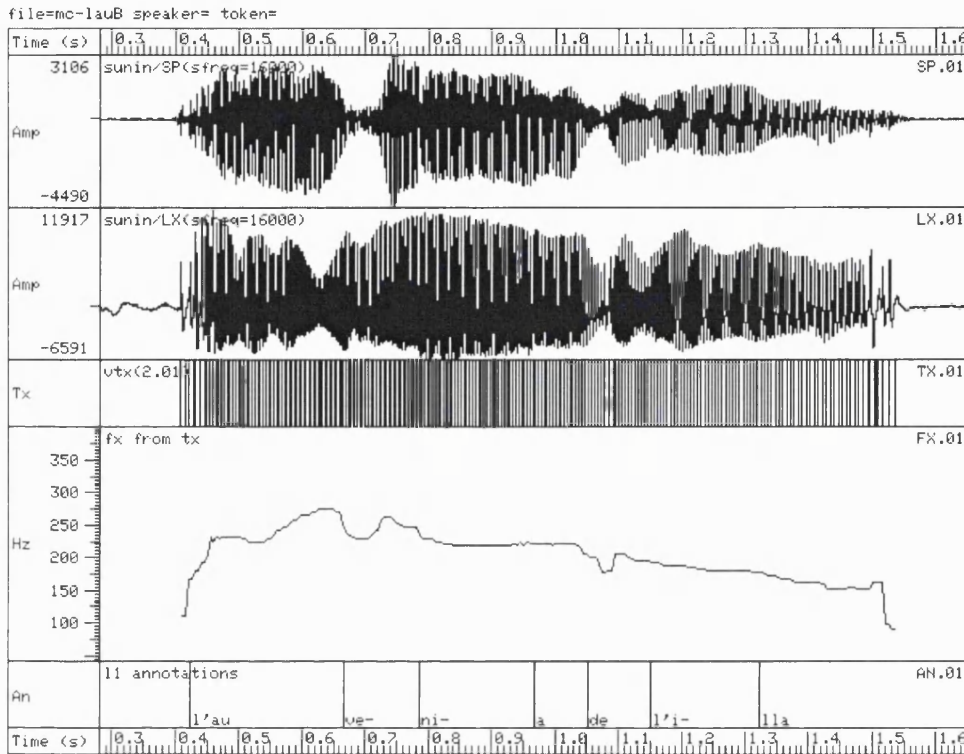


Figure 2.1. SFS display screen showing 1) the speech waveform, 2) the Lx waveform, 3) the excitation period measurements and 4) the F0 contour for the sentence *l' au venia de l'illa* produced with broad focus.

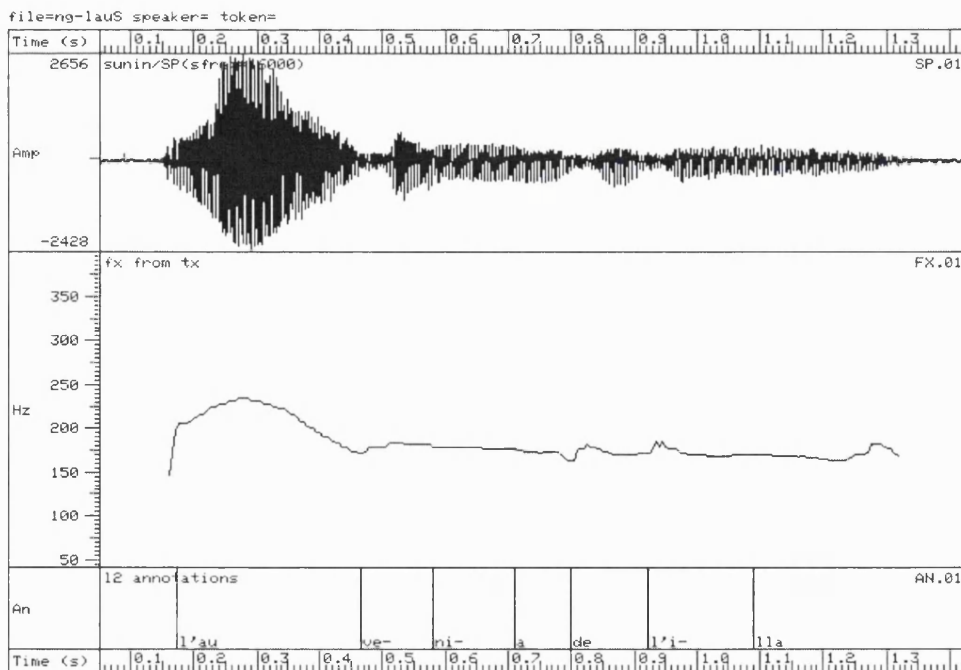


Figure 2.2. Speech waveform and F0 trace for the sentence *L'AU venia de l'illa* with narrow focus on the subject.

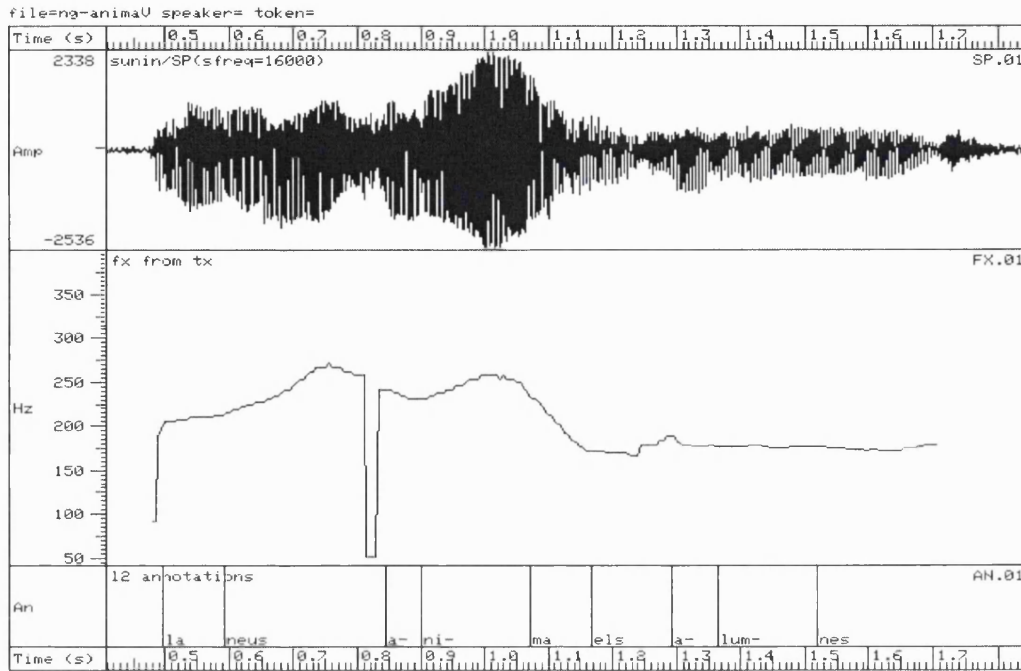


Figure 2.3. Speech waveform and F0 trace for the sentence *la Neus ANIMA els alumnes* with narrow focus on the verb.

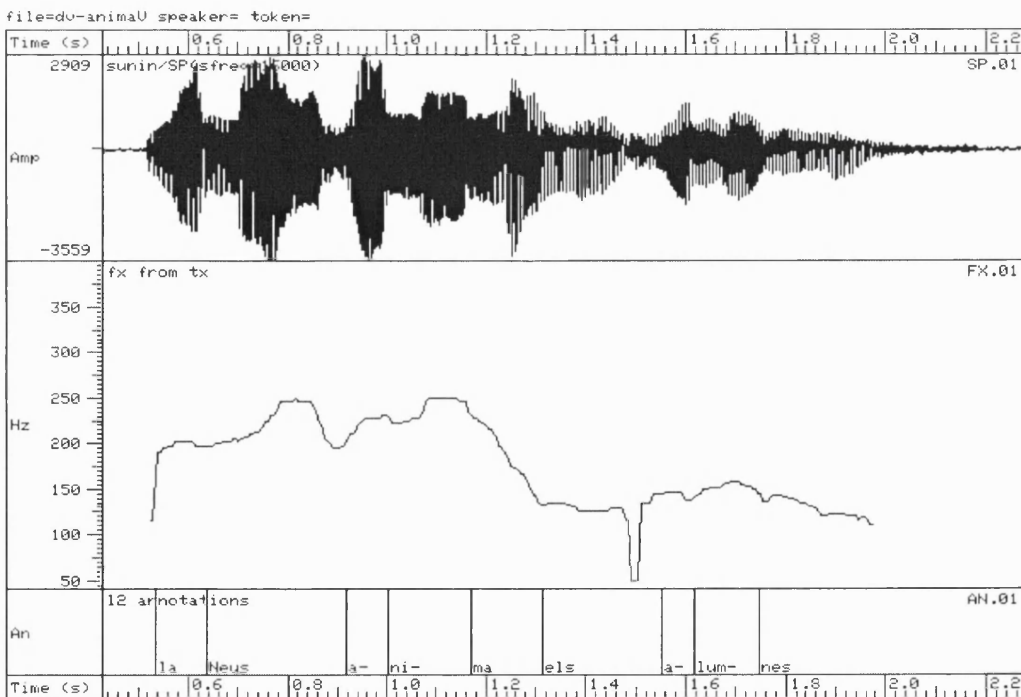


Figure 2.4. Speech waveform and F0 trace for the sentence *la Neus ANIMA els alumnes* with narrow focus on the verb and a postnuclear accent.

### **2.3.3 Results**

The results of the production test are presented in two sections: 1) results of the reading task and 2) results of the semi-spontaneous activity. This division is due to the different kinds of responses obtained in the two activities.

#### **2.3.3.1 Reading test**

As presented in section 2.3.2, the target sentences of the reading task were divided into two categories: 1) responses produced with accentual narrow focus and 2) responses produced with a broad focus intonational pattern. Responses with accentual focus involved two kinds of utterances: a) utterances with a nuclear accent on the focussed item and deaccenting of postfocal information, as illustrated in Figures 2.2 and 2.3 above, and b) utterances with a nuclear accent on the focussed item but no deaccenting of postfocal material, as shown in Figure 2.4. The decision to include these two kinds of utterances within the accentual focus responses derives from the fact that Central Catalan speakers perceive no semantic differences in sentences produced with and without a postnuclear accent. This is examined in more detail in chapter 4, where the results of a perception test show that sentences with and without a postfocal accent are interpreted in the same way by two Central Catalan speakers (i.e. both kinds of utterances were judged as responding to an *s-w* prominence pattern). Finally, responses where accentual focus was not observed consisted of sentences uttered with a broad focus intonation, that is, sentences where the *w-s* pattern was kept and hence the nuclear accent was on the last lexical word rather than on the expected focussed element. The *w-s* pattern was illustrated in Figure 2.1.

The overall number of accentual narrow focus and broad focus productions for the reading activity is presented in Figure 2.5. The results show that the signalling of focus by accentual means is possible in Central Catalan since the total number of accentual focus productions for the whole test (853 responses out of 1088, 78.4%) is much higher than the total number of broad focus productions (235 responses out of 1088, 21.6%). A higher number of accentual focus responses was consistently found across all speakers.

Number of target responses	Number and percentage of accentual focus responses	Number and percentage of broad focus responses
32 (x34 speakers) = 1088	853 (78.4%)	235 (21.6%)

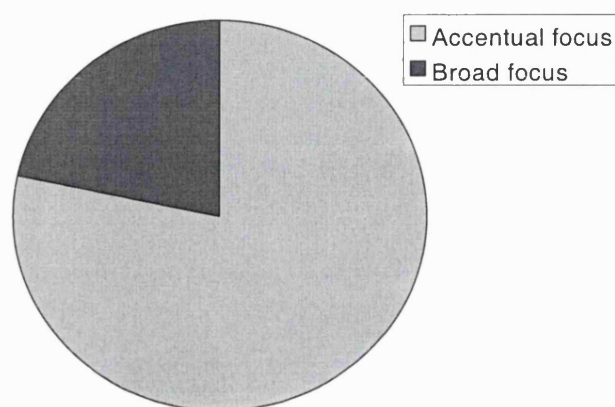


Figure 2.5. Table and pie chart showing the number of responses produced with accentual narrow focus and with broad focus for the whole reading test.

The number of accentual focus and broad focus responses for the variables *subject* focus vs *verb* focus, and narrow focus triggered by a *contrastive* prompt vs narrow focus triggered by an *identification* prompt are displayed in Figure 2.6. The results of Figure 2.6 show that for all variables the number of accentual focus responses is much higher than the number of broad focus responses. This indicates that accentual focus can be used both in subject and in verb focus domains and with a contrastive as well as with an identification reading.

	Number of target responses	Number and percentage of accentual focus responses	Number and percentage of broad focus responses
Subject	16 (x34 speakers) = 544	436 (80.2%)	108 (19.8%)
Verb	16 (x34 speakers) = 544	417 (76.6%)	127 (23.4%)
Contrast	16 (x34 speakers) = 544	476 (87.5%)	68 (12.5%)
Identification	16 (x34 speakers) = 544	377 (69.3%)	167 (30.7%)

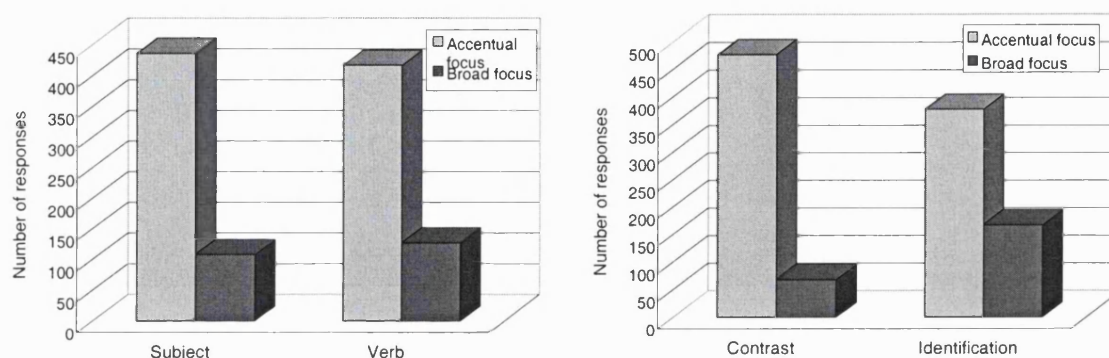


Figure 2.6. Table and histograms showing the number of accentual focus responses and broad focus responses for the variables subject focus vs verb focus and focus triggered by contrast vs focus triggered by identification in the reading test.

In order to see whether the focal domains (subject vs verb) and the triggering questions (contrast vs identification) had an effect on the realisation of accentual focus, t-tests comparing the number of accentual focus responses for the two variable groups were performed. The significance level was considered to be 1 per cent ( $p < 0.01$ ). The critical value of  $t$  for this level is 2.45 with 33 degrees of freedom.

Although for the subject/verb variable the number of accentual focus responses in verb focus (417 responses, 76.6%) is slightly lower than the number of accentual focus responses in subject focus (436 responses, 80.2%), differences between the two domains are not significant ( $p > 0.01$ ,  $t = 1.37$ ). This indicates that use of intonation to convey focus in Catalan is equally favoured in subject and verb domains. These findings are different from the results reported in García-Lecumberri (1995) for Peninsular Spanish where accentual focus was more favoured in subject than in verb domains. According to our results, accentual devices are equally used in subject and in verb domains, indicating that



sentence position does not have an effect on the production of accentual focus in Central Catalan.

For the variable contrast vs identification focus, the results showed that contrastive questions elicited more intonational focus productions (476 responses, 87.5%) than identification questions (377 responses, 69.3%). The difference between both categories is significant ( $p < 0.01$ ,  $t = 6.15$ ) and supports Zubizarreta's (1998) claim that the use of accentual focus in Romance languages is more compatible with a contrastive interpretation than a non-contrastive one. However, as she points out, not all Romance languages behave in exactly the same way. French seems to allow more identification productions than Spanish. The results observed in our data show that Central Catalan speakers can also use accentual focus as a response to an identification question, although the contrastive meaning seems to be preferred. This suggests that the usage of intonational strategies to signal narrow focus does not behave categorically among languages but has different degrees of application.

### **2.3.3.2 Semi-spontaneous test**

For the semi-spontaneous task, responses were divided into four main categories: 1) sentences with accentual focus, 2) sentences with syntactic focus, 3) broad focus responses and 4) other. Sentences with accentual focus consisted of utterances where the focus/accent association was carried out by intonational strategies, mainly by means of nuclear accent location on the focussed item with (or without) deaccenting of postfocal elements. Sentences with syntactic focus involved utterances where the focus/accent association was attained by means of a syntactic shift. In these sentences, the focussed element was moved to a prominent position (or accent-bearing location). Two kinds of syntactic reorganisations were observed. The first one involved the displacement of the focussed element to the sentence-final position and dislocation of the object, which is linked to the main sentence by the presence of a pronoun. The second consisted of using a cleft structure: the focussed element is extracted from its position and forms a relative clause with the unfocussed material. Both structures are illustrated in (2.13) below. As in the reading test, broad focus responses consisted of sentences produced with a broad

focus intonation. Finally, "other" consists of unforeseen responses such as "I don't know", "maybe", among others.

- (2.13) a. La faldilla, la rentava LA NÚVIA  
"The skirt, it was washed by THE BRIDE"  
(lit. The skirt, it washed THE BRIDE)
- b. És L'ISIDRE qui du una gavardina vermella  
"It is ISIDRE who wears a red waterproof"

Figure 2.7 shows the results obtained in the semi-spontaneous test. The results of this test are very similar to the ones obtained in the reading test. The percentage of responses with accentual focus (74.7%) is much higher than the percentage of syntactic responses (20.1%), broad focus responses (2.4%) and other kinds of responses (2.8%). This confirms the idea that intonational strategies are used in Central Catalan to signal the information structure of sentences. Rather unexpectedly, a very small number of responses with syntactic focus (218 out of 1088) was obtained. One possible reason to explain this behaviour seems to derive from the semi-spontaneous task itself. Even though speakers were free to choose the focussing strategies, the semi-spontaneous task was performed after the reading activity and maybe speakers were already inclined to use accentual strategies. Another possibility could be that speakers do prefer accentual strategies rather than syntactic strategies to convey narrow focus. However, this situation is very unlikely because the usage of syntactic focus in Catalan is well accepted (as will be corroborated with the acceptability test in section 2.5). Finally, the small usage of syntactic strategies in the semi-spontaneous activity is partly explained by the effects of the triggering questions. The data in Figure 2.7 includes both responses to contrastive and identification questions. As will be observed in Figure 2.8, whereas identification narrow focus utterances are produced both with syntactic and intonational devices, contrastive sentences mainly prefer accentual strategies.

Number of target responses	Number and percentage of accentual focus responses	Number and percentage of syntactic focus responses	Number and percentage of broad focus responses	Number and percentage of other responses
32 (x34 speakers) = 1088	813 (74.7%)	219 (20.1%)	26 (2.4%)	30 (2.8%)

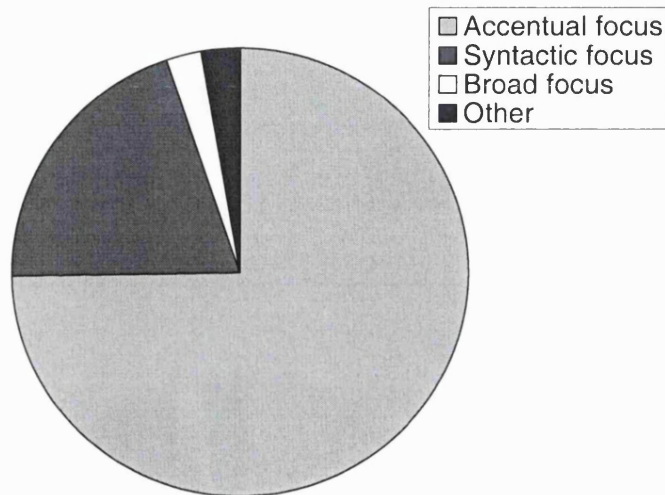


Figure 2.7. Table and pie chart showing the distribution of responses obtained in the semi-spontaneous test.

The number of responses according to the variables subject/verb and contrast/identification is provided in Figure 2.8. As in the reading activity, the results of Figure 2.8 show that for all variables the number of accentual focus responses is much higher than the number of syntactic focus, broad focus or other kinds of responses.

	Number of target responses	Number and percentage of accentual focus responses	Number and percentage of syntactic focus responses	Number and percentage of broad focus responses	Number and percentage of other responses
Subject	16 (x34) = 544	409 (75.2%)	91 (16.7%)	26 (4.8%)	18 (3.3%)
Verb	16 (x34) = 544	404 (74.3%)	128 (23.5%)	0	12 (2.2%)
Contrast	16 (x34) = 544	490 (90%)	28 (5.2%)	13 (2.4%)	13 (2.4%)
Identification	16 (x34) = 544	323 (59.4%)	191 (35.1%)	13 (2.4%)	17 (3.1%)

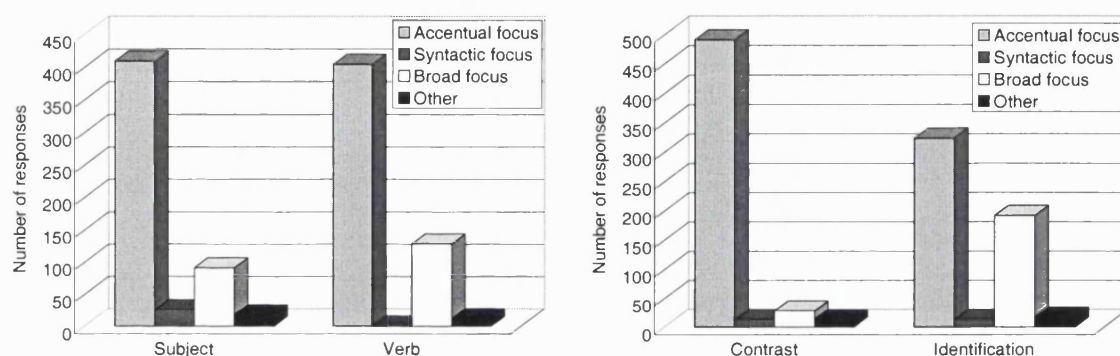


Figure 2.8. Table and histograms showing the number of accentual focus responses and broad focus responses for the variables subject focus vs verb focus and focus triggered by contrast vs focus triggered by identification obtained in the semi-spontaneous test.

As with the reading activity, t-tests comparing the effect of sentence position on the realisation of accentual focus show that there is no significant effect of subject/verb position in the assignment of accentual focus in Central Catalan ( $p > 0.01$ ,  $t = 0.29$ ). Thus, both subject and verb positions equally favour accentual focus. For the variable contrast/identification, effects similar to the ones observed in the reading activity are found. As expected, the use of intonational strategies is significantly higher with a contrastive than with an identification prompt ( $p < 0.01$ ,  $t = 7.33$ ). However, the fact that a high number of identification utterances was also produced with accentual devices may indicate that the possibility of conveying identification with accentual strategies is not at all disfavoured. Further research including more spontaneous data is needed.

Although the use of syntactic strategies is very small still two tentative generalisations can be drawn from the results. Syntactic focus is more frequently used in verb domains

(23.5%) than in subject domains (16.7%). As expected, syntactic focus is more favoured when triggered by identification (35.1%) than by contrast (5.2%).

Finally, comparisons between the production of intonational focus in reading (853 responses) and in semi-spontaneous speech (813 responses) exhibited no significant differences ( $p > 0.01$ ,  $t = 1.5$ ). This suggests that accentual focus is a common device in Central Catalan for signalling the information structure of sentences and that the usage of this strategy in reading speech seems not to be forced by the reading activity.

## **2.4 Perception test**

The perception test was designed to investigate whether Catalan speakers are able to identify focus conveyed through accentual means. This test mainly involved the reverse process of the production test. Listeners had to find out whether sentences produced with accentual narrow focus could be associated to a marked, non-neutral reading.

### **2.4.1 Experimental design**

#### **2.4.1.1 Materials**

The stimuli used for the perception test were the utterances produced by one of the speakers in the production test (including the data of both the reading activity and the response to visual cue task). The speaker chosen was one that provided a high number of accentual focus productions of narrow focus in the recordings (32 accentual focus responses out of 32 targets in the reading task and 30 out of 32 in the semi-spontaneous task). As pointed out in section 2.3.3.1, accentual focus involved accent location on the focal material with (or without) deaccenting of postfocal information. The speaker's productions were assessed auditorily by the researcher and two other native Catalan speakers. An acoustic analysis of the responses was also performed so as to confirm the validity of the auditory judgements (see section 2.3.2.2 for details on the acoustic analysis). In order to have a comparable number of sentences for each of the variables in both the production and the perception tests, the two broad focus responses obtained in

the semi-spontaneous test were performed again (with accentual narrow focus) by the same speaker in a second recording session.

The variables examined in the perception test are the same ones as for the production test, namely, 1) focus on the subject vs focus on the verb, and 2) focus triggered by contrast vs focus triggered by identification (see section 2.3.1.1 for further details). Given the results obtained in the production test, no significant differences are expected to be found in the perception of accentual focus in subject and verb domains. On the other hand, accentual focus is expected to be better identified in contrastive contexts than in identification contexts.

The perception test included stimuli obtained both in the reading and in the semi-spontaneous activities since no major intonational differences were encountered in the accentual focus responses obtained in the two tasks (this is discussed in more detail in chapter 4).

Overall, 136 stimuli were presented: 64 target sentences (sentences with accentual narrow focus), 8 trial sentences and 64 distractors. Trial sentences were included for the listeners to get used to the task procedure (see section 2.4.1.4). Distractors had the aim of preventing identification of the purpose of the test. As in the production test, distractors were stimuli with a broad/object focus intonation and hence the appropriate triggering question was the one that prompted the *w-s* pattern. The list of target sentences, distractors and trial sentences for the perception test is presented in Appendix 1.

#### **2.4.1.2 Editing the stimuli**

Before running the perception test, the questions that triggered the answers in the production test (stimuli for the perception test) were removed from the initial recording. This was done by means of the SFS program (see section 2.3.2.2 for more details). The productions of the selected speaker along with the triggering questions were introduced into a Sun Sparc-10 computer in which SFS ran. SFS allowed us to break the signal into different speech chunks and exclude the researcher's interventions. 6 seconds of silence

were added between stimuli by means of the E MAX program so as to leave some time between utterances. Then, the data was transferred to a 60 ES DAT recorder and finally recorded onto a UHER CR 160 tape.

#### **2.4.1.3 Informants**

The perception test was carried out by twenty-eight subjects (fourteen male and fourteen female adult native speakers of Central Catalan). Their background characteristics were similar to those of the speakers recorded for the production test. Their age was between 25 and 35 years old. Each of the subjects was born and lived in Ripoll within a middle class Catalan-speaking family. All the subjects were Catalan-Spanish bilinguals, although Catalan was quite exclusively used in their daily conversations. The mastery of foreign languages (English or French) was not very high. Similar to the speakers selected for the production test, all the listeners had finished secondary school and some of them had a university degree. None of the listeners had lived in countries where foreign languages are spoken.

#### **2.4.1.4 Testing**

For the perception test, subjects were asked to listen to the stimuli (answers of the two production activities) and find out the type of question that could trigger each answer in a multiple choice kind of test. The test was designed in the form of three potential questions for each answer. Only one of the questions was relevant to each stimulus. For each stimulus, the following three question types were presented: 1) a question that triggered narrow focus on the subject, 2) another one that triggered narrow focus on the verb and 3) a last one that triggered a broad focus reading or narrow focus on the object. The hypothesis was that if Central Catalan speakers cannot perceive narrow focus signalled by intonational means, then stimuli produced with accentual narrow focus (on the subject or on the verb) will be perceived as answers to a broad/object focus type of question. For each stimulus, one of the three possible questions was a distractor. Specifically, the question that triggered narrow focus on the subject was the distractor for a verb focus stimulus, and the reverse, a verb focus question was the distractor for a

subject focus stimulus. For each stimulus, the order of the questions was randomised. An example of two stimuli and the three question types is shown in (2.14). Both stimuli were produced with intonational focus on the subject. Stimulus A was the result of a contrastive prompt and stimulus B of an identification prompt. In these two examples, listeners are expected to mark question 1 as the prompt that triggers the stimuli.

(2.14) *Stimulus A*

LA ROSA llegia llibres  
"ROSA read books"

*Questions*

1. En Jordi llegia llibres?  
"Did Jordi read books?"
2. La Rosa comprava llibres?  
"Did Rosa buy books?"
3. Què tenia la Rosa?  
"What was the matter with Rosa?"

*Stimulus B*

LA NÚVIA rentava la faldilla  
"THE BRIDE was washing the skirt"

*Questions*

1. Qui rentava la faldilla?  
"Who was washing the skirt?"
2. Què feia la núvia amb la faldilla?  
"What was the bride doing with the skirt?"
3. Què feia la núvia?  
"What was the bride doing?"

In order to carry out the perception test, listeners sat in a quiet room of a private house. The procedures of the test were explained by the researcher. Listeners were asked to find the question for each stimulus. A minimum of one and a maximum of three subjects carried out the test at the same time. After each stimulus, the researcher stopped the tape so that subjects were not worried about time constraints. Stimuli could be replayed a second time if listeners wished.

## 2.4.2 Results

The overall number of judgements obtained in the perception test is as follows: 1792 target responses (64 stimuli x 28 speakers), 1792 distractors and 224 trial sentences (8 x 28 speakers). The target responses obtained in the perception test were divided into three categories corresponding to the three question types presented in the multiple choice test: 1) items where accentual focus was perceived in the appropriate word (either subject or



verb), 2) items where accentual focus was perceived in an unexpected word (either on the verb in subject focus sentences or on the subject in verb focus sentences), and 3) items where the listener interpreted a broad/object focus reading. The results of the perception test are presented in Figure 2.9, which shows the number of accentual focus responses, broad focus responses and unexpected accentual focus responses for the whole perception test.

Number of target responses	Number and percentage of accentual focus responses	Number and percentage of broad focus responses	Number and percentage of unexpected accentual focus responses
64 (x28 listeners) = 1792	1512 (84.4%)	233 (13%)	47 (2.6%)

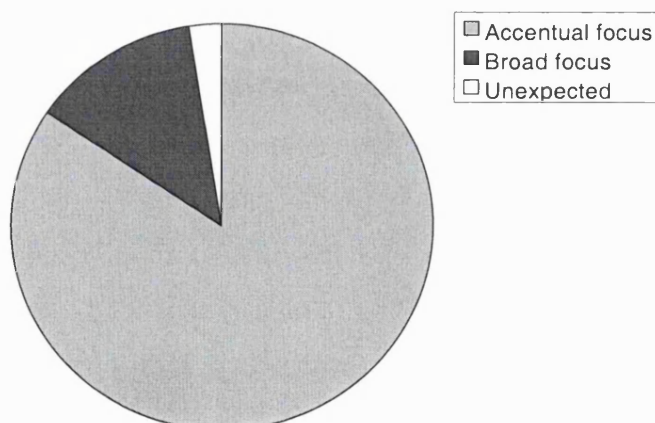


Figure 2.9. Table and pie chart showing the number of accentual focus responses, broad focus responses and unexpected responses for the whole perception test.

The results of the perception test show that the signalling of information structure by means of intonational focus is robustly perceived by all listeners. The percentage of responses where accentual focus was appropriately recognised for the whole perception test (84.4%) is much higher than the percentage of responses where an inappropriate broad focus interpretation was provided (13%) or where accentual focus was perceived on an unexpected item (2.6%). This confirms the idea that Catalan is a language that can

use intonation as a means of conveying focus. A higher number of accental focus responses over other kinds of responses was consistently found across all listeners.

The effects of the variables subject/verb and contrast/identification on the perception of accental focus were also analysed. The results are shown in Figure 2.10. In line with the production test, t-tests were performed to see whether differences in the perception of accental focus in subject/verb domains and in contrast/identification contexts were significant. The significance level used was  $p < 0.01$ , which for 27 degrees of freedom gives a value of  $t$  of 2.47.

	Number of target responses	Number and percentage of accental focus responses	Number and percentage of broad focus responses	Number and percentage of unexpected accental focus responses
Subject	32 (x28 listeners) = 896	859 (95.9%)	30 (3.3%)	7 (0.8 %)
Verb	32 (x28 listeners) = 896	653 (72.9%)	203 (22.6%)	40 (4.46%)
Contrast	32 (x28 listeners) = 896	763 (85.1%)	123 (13.7%)	10 (1.12%)
Identification	32 (x28 listeners) = 896	749 (83.6%)	110 (12.3%)	37 (4.13%)

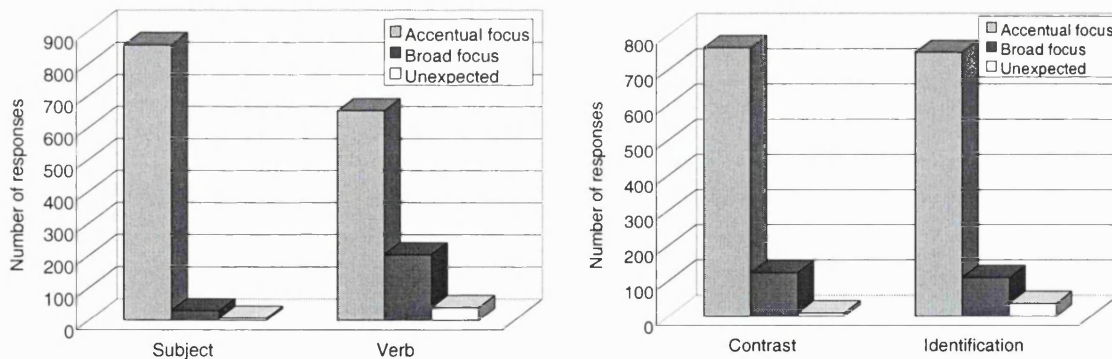


Figure 2.10. Table and histograms showing the number of accental focus, broad focus and unexpected responses in the perception test for the variables subject focus vs verb focus and contrastive focus vs identification focus.

The results of Figure 2.10 show that for all variables the number of responses where accental focus was perceived is much higher than the number of responses attributed to broad focus or unexpected accental focus. This indicates that accental focus is perceived both in subject and in verb focus domains and in contrast as well as in identification prompts.

The results of the t-test comparing the effects of sentence position on the recognition of accentual focus show that the percentage of perceptual judgements for intonational focus on the subject is significantly higher ( $p < 0.01$ ,  $t = 8.8$ ) than for intonational focus on the verb: 95.87% vs 72.88% respectively. This indicates that Catalan listeners can identify accentual focus on the subject much more easily than on the verb, although verb accentual focus is still well perceived. This tendency was also observed in García-Lecumberri (1995) for the analysis of Spanish focus. These results, however, differ from the ones obtained in the production test, which exhibited no significant differences in the use of accentual strategies in subject and verb focal domains in Catalan. Thus, whereas at the production level accentual strategies are equally favoured in the two sentence positions, at the perception level accentual focus in sentence-initial position is more easily recognised<sup>4</sup>.

No significant differences ( $p > 0.01$ ,  $t = 0.7$ ) were observed for the perception of accentual focus in contrastive vs identification contexts. These results disagree with the findings of the production test, where contrastive focus exhibited a significantly higher number of accentual responses than identification focus. This behaviour seems to indicate that even though listeners are able to perceive intonational focus quite satisfactorily in identification sentences, speakers might prefer to use different strategies for signalling non-contrastive focus.

## **2.5 Acceptability test**

So far, the results of the production test and the perception test have shown that Central Catalan speakers can both produce and recognise narrow focus conveyed through intonational means and that the use of accentual strategies is quite high (more than 70% of accentual focus responses were obtained in both tests). These results challenge the

---

<sup>4</sup> A possible explanation for that might derive from the fact that in verb focus the pitch height of the focal accent can be lowered (downstepped) with respect to the previous accent (Estebas-Vilaplana and Maidment 1999b). This means that the pitch height of the focal accent in medial position is lower than that of the focal accent in initial position. This pitch lowering might have an effect on perception and thus verb focus can be more difficult to perceive than subject focus. Further research is needed on this hypothesis. For more information on the relation between focus and downstep in Central Catalan see chapter 5.

categorical classification of Central Catalan as a [-plastic] language (Vallduví 1991, 1994 and Vallduví and Zacharski 1994) and suggest that accentual strategies should not be disregarded as a way of signalling focus in Central Catalan. See chapter 4 for more discussion on this issue.

One of the criticisms that may arise from the data is that the high use of accentual focus may be derived from the experimental procedures designed to elicit the utterances. For example, in the reading activity the speakers had no choice other than using accentual devices to convey narrow focus. In the semi-spontaneous task, although the speakers were free to choose the kinds of strategies, those might not have been as natural as in normal conversation. In order to overcome the problems deriving from the experimental design, an acceptability test was conceived to investigate the naturalness and acceptability of accentual narrow focus in Central Catalan. This test also includes narrow focus sentences conveyed by syntactic means, so as to explore differences and similarities in the acceptability of syntactic focus and intonational focus in Central Catalan.

In principle, we expect to find that both accentual strategies and syntactic strategies are recognised as natural. However, some variation might be found with respect to the focal domains (subject vs verb) or the triggering questions (identification vs contrast).

## **2.5.1 Experimental design**

### **2.5.1.1 Materials**

The acceptability test included 64 sentences where narrow focus was produced with accentual strategies, 64 sentences where narrow focus was produced with syntactic strategies and 64 broad focus sentences. Sentences with accentual focus were the same ones as for the production test and for the perception test (see Appendix 1). Thus, the variables 1) focal domain (focus on subject vs focus on verb) and 2) triggering question (contrast and identification) were also tested in terms of acceptability.

Sentences with syntactic focus were the exact counterparts of the accentual focus sentences (same meaning and possible answers to the same questions) but with focus

conveyed by means of a syntactic reorganisation. The syntactic strategies involved the placement of the focussed element (subject or verb) in sentence-final position, the dislocation of the object outside the sentence and the appearance of a preverbal pronoun linked to the displaced object. An example is provided in (2.15) below for focus on a subject domain. (2.15a) includes focus signalled by intonational means and (2.15b) focus signalled syntactically. In (2.15b), the subject (*la núvia*) is moved to the final position of the sentence to get accentual prominence. The object (*la faldilla*) is dislocated, while leaving a pronoun (*la*) before the verb. In both cases, the question that triggered these answers is an identification prompt.

- (2.15) a. Qui rentava la faldilla? LA NÚVIA rentava la faldilla  
"Who washed the skirt? THE BRIDE washed the skirt"
- b. Qui rentava la faldilla? La faldilla, la rentava LA NÚVIA  
"Who washed the skirt? The skirt, it was washed by THE BRIDE"  
(lit. The skirt, it washed THE BRIDE)

Finally, broad focus sentences consisted of utterances with the same segmental structure as accentual focus sentences but produced with an unmarked intonation. These sentences were expected to be perceived as unnatural responses to a narrow focus question. However, they were included to confirm that the speaker's judgements were based on intonation rather than on the segmental structure.

### **2.5.1.2 Stimuli**

For the acceptability test, each stimulus consisted of both the triggering question and the answer, as illustrated in (2.15) above. The question was included because the naturalness or acceptability of focussed sentences could only be fairly judged in the appropriate context. For every context, the sentence with accentual focus, the one with syntactic focus and the broad focus sentence were produced in this order.

The stimuli were recorded in an anechoic room in the Department of Phonetics and Linguistics of University College London. The triggering questions were read by a 23

year old Central Catalan speaker. The answers were provided by the researcher herself, since we wanted to ensure to record the “right” structures and with the “right” intonation<sup>5</sup>. For all the answers both speech and laryngeal signals were obtained. Although the laryngeal signal was not relevant for the acceptability test, it was included for the possible checking of any unexpected results in the test. Speech was recorded to a B&K sound level meter of the type 2231. The sound level meter was fitted with a 4165 microphone, which was placed at about twenty centimetres from the two speakers. The laryngeal signal was obtained by means of a laryngograph processor. The two signals were recorded on a Sony 1000 ES DAT recorder. Between each of the stimuli, a gap of eight seconds was left for the listeners to provide their judgements. The recordings on the audio digital tape were later transferred on chrome tape (Maxell XL II) to be played on a Philips A21101 tape recorder. Mistakes in the original recording were removed when editing the final tape.

### **2.5.1.3 Informants**

Twenty-eight adult native speakers of Central Catalan (fourteen male and fourteen female) were involved in the acceptability test. These informants were different from those that performed the production and the perception tests, although their background characteristics were similar. All subjects were born and lived in Ripoll within a middle class Catalan-speaking family and were between 25 and 38 years old. Subjects were Catalan-Spanish bilinguals, although Catalan was the language they predominantly used in their daily conversations. All informants had finished secondary school and some of them had a university degree.

### **2.5.1.4 Testing**

A written transcript including the different kinds of sentences and the triggering questions was provided to the informants. Instructions of how to perform the test were written on

---

<sup>5</sup> Sentences obtained in the production test were not used as stimuli in the acceptability test for two reasons. First, we wanted to have a similar number of sentences with accentual focus and syntactic focus and in the production test the two samples were very different. Second, we wanted to avoid speaker variability.

the paper and were repeated orally before running the tape. Listeners were asked to judge the acceptability of the stimuli within a scale of 4 to 0, where 4 was “perfectly acceptable”, 3 “quite acceptable”, 2 “uncertain”, 1 “rather unacceptable” and 0 “totally unacceptable”. The decision to use a 5-point acceptability scale was to give the listeners some flexibility in their judgements. A box with the scale numbers as the one exhibited in (2.16) was provided. Listeners were asked not to leave any stimuli unassessed and to tick only one box in each row. The tape was played with no interruptions since it was considered that there was enough time between the stimuli to give an answer.

(2.16) Qui rentava la faldilla?

- A. La núvia rentava la faldilla
- B. La faldilla, la rentava la núvia
- C. La núvia rentava la faldilla

	4	3	2	1	0
A					
B					
C					

### 2.5.2 Results

The overall results of the acceptability test are shown in Figure 2.11. On the vertical axis the number of responses for categories A, B and C is presented. A stands for accentual focus, B for syntactic focus and C for broad focus. The acceptability scale ranges from 4 to 0.

	Overall number of responses	Number and percentage of responses according to the acceptability scale				
		4	3	2	1	0
Accentual focus (A)	64 (x28 speakers) = 1792	1382 (77.1%)	393 (21.9%)	16 (0.9%)	1 (0.006%)	0
Syntactic focus (B)	64 (x28 speakers) = 1792	1268 (70.8%)	416 (23.2%)	87 (4.9%)	3 (0.2%)	0
Broad focus (C)	64 (x28 speakers) = 1792	2 (0.1%)	68 (3.8%)	207 (11.5%)	737 (41.1%)	778 (43.4%)

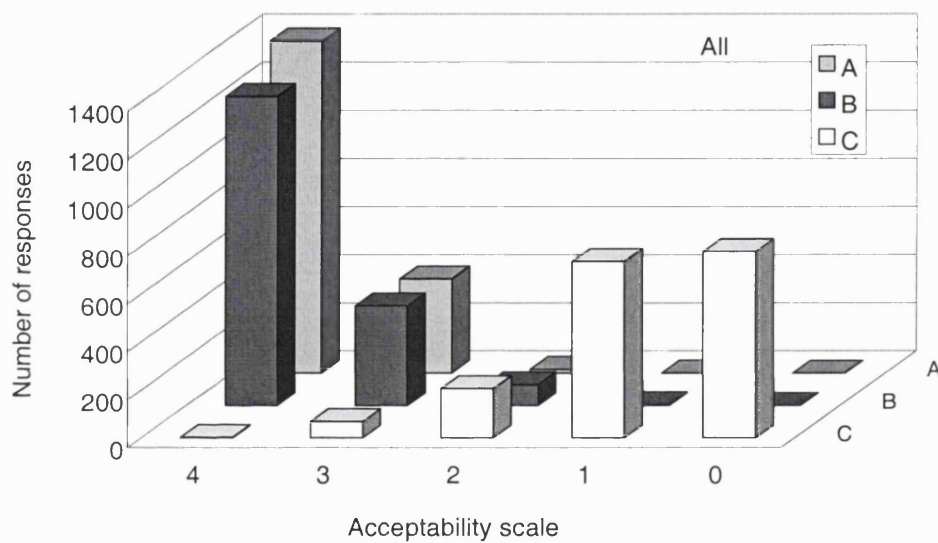


Figure 2.11. Table and histograms showing the number of responses for the whole acceptability test according to three categories: A = narrow focus conveyed intonationally, B = narrow focus conveyed syntactically, and C = broad focus. The acceptability scale ranges from 4 (perfectly acceptable) to 0 (totally unacceptable).

The overall results of the acceptability test presented in Figure 2.11 show that the acceptability of utterances with narrow focus signalled by accentual means (category A) is very high, since most of the answers scored the highest acceptability values, i.e. 4 (77.1% of the responses) and 3 (21.9% of the responses). Similarly, narrow focus signalled by syntactic devices (category B) also presented high levels of acceptability, i.e. 4 (70.8%) and 3 (23.2%). This suggests that Catalan speakers judge accentual and syntactic strategies quite similarly in terms of acceptability for the signalling of narrow



focus. As expected, broad focus structures (category C) scored the lowest levels of acceptability as answers to narrow focus questions, i.e. 0 (43.4%) and 1 (41.1%).

The results of the acceptability test in relation to the different variables: subject vs verb focus domains, and contrast vs identification triggering questions are presented in Figures 2.12-2.13.

	Overall number of responses	Number and percentage of responses according to the acceptability scale in Subject focus				
		4	3	2	1	0
Accentual focus (A)	32 (x28 speakers) = 896	690 (77%)	196 (21.9%)	9 (1%)	1 (0.006%)	0
Syntactic focus (B)	32 (x28 speakers) = 896	614 (68.5%)	228 (25.4%)	53 (5.9%)	1 (0.006%)	0
Broad focus (C)	32 (x28 speakers) = 896	0	23 (2.6%)	99 (11%)	340 (38%)	434 (48.4%)

	Overall number of responses	Number and percentage of responses according to the acceptability scale in Verb focus				
		4	3	2	1	0
Accentual focus (A)	32 (x28 speakers) = 896	692 (77.2%)	197 (21.9%)	7 (0.8%)	0	0
Syntactic focus (B)	32 (x28 speakers) = 896	672 (75%)	188 (21%)	34 (3.8%)	2 (0.2%)	0
Broad focus (C)	32 (x28 speakers) = 896	2 (0.2%)	45 (5%)	108 (12%)	397 (44%)	344 (38.4%)

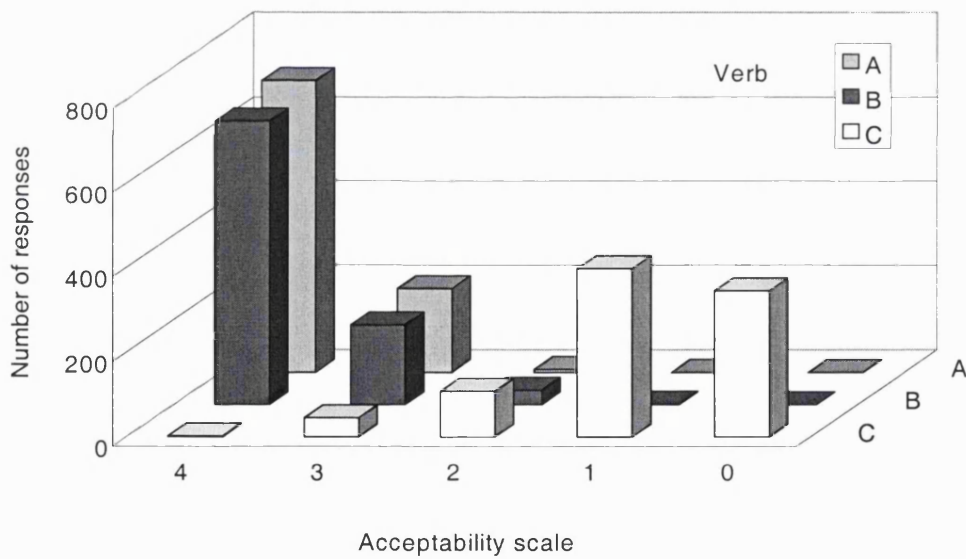
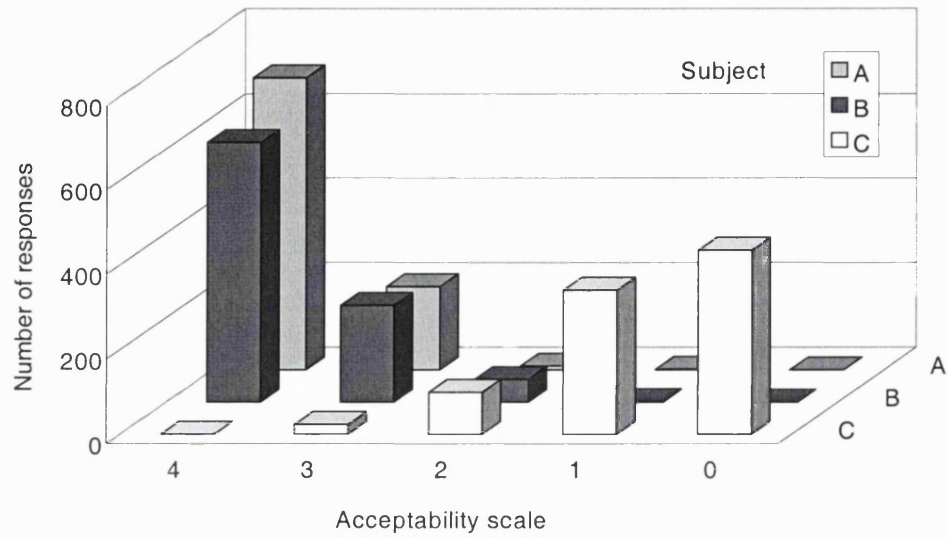


Figure 2.12. Tables and histograms showing the number of acceptable/unacceptable responses for sentences with narrow focus on the subject and on the verb according to three categories: A = narrow focus conveyed intonationally, B = narrow focus conveyed syntactically, and C = broad focus. The acceptability scale ranges from 4 (perfectly acceptable) to 0 (totally unacceptable).

The results of the acceptability test for narrow focus on the subject and on the verb provided in Figure 2.12 showed that narrow focus triggered by accentual means (category A) and by syntactic means (category B) is highly acceptable in the two focal domains.

For both strategies most of the responses scored 4 or 3. For narrow focus on the subject, 77% of the responses corresponded to 4 and 21% to 3 in focus conveyed through intonational means, and 68% to 4 and 25% to 3 in focus conveyed by syntactic means. In narrow focus on the verb, 77% of the responses scored 4 and 22% 3 in intonational focus, and 75% scored 4 and 21% 3 in syntactic focus. As expected, broad focus structures (category C) were felt to be unacceptable as answers to subject or verb focus triggering questions, scoring mainly 0 or 1 in both cases.

Figures 2.13 exhibit the results of the acceptability test according to the triggering question: contrast vs identification.

	Overall number of responses	Number and percentage of responses according to the acceptability scale in Contrastive focus				
		4	3	2	1	0
Accentual focus (A)	32 (x28 speakers) = 896	694 (77.4%)	194 (21.7%)	8 (0.9%)	0	0
Syntactic focus (B)	32 (x28 speakers) = 896	664 (74%)	187 (20.9%)	42 (4.7%)	3 (0.3%)	0
Broad focus (C)	32 (x28 speakers) = 896	0	4 (0.4%)	61 (6.8%)	352 (39.3%)	479 (53.5%)

	Overall number of responses	Number and percentage of responses according to the acceptability scale in Identification focus				
		4	3	2	1	0
Accentual focus (A)	32 (x28 speakers) = 896	688 (76.8%)	199 (22.2%)	8 (0.9%)	1 (0.1%)	0
Syntactic focus (B)	32 (x28 speakers) = 896	622 (69.4%)	229 (25.5%)	45 (5%)	0	0
Broad focus (C)	32 (x28 speakers) = 896	2 (0.2%)	64 (7.1%)	146 (16.3%)	385 (43%)	299 (33.4%)

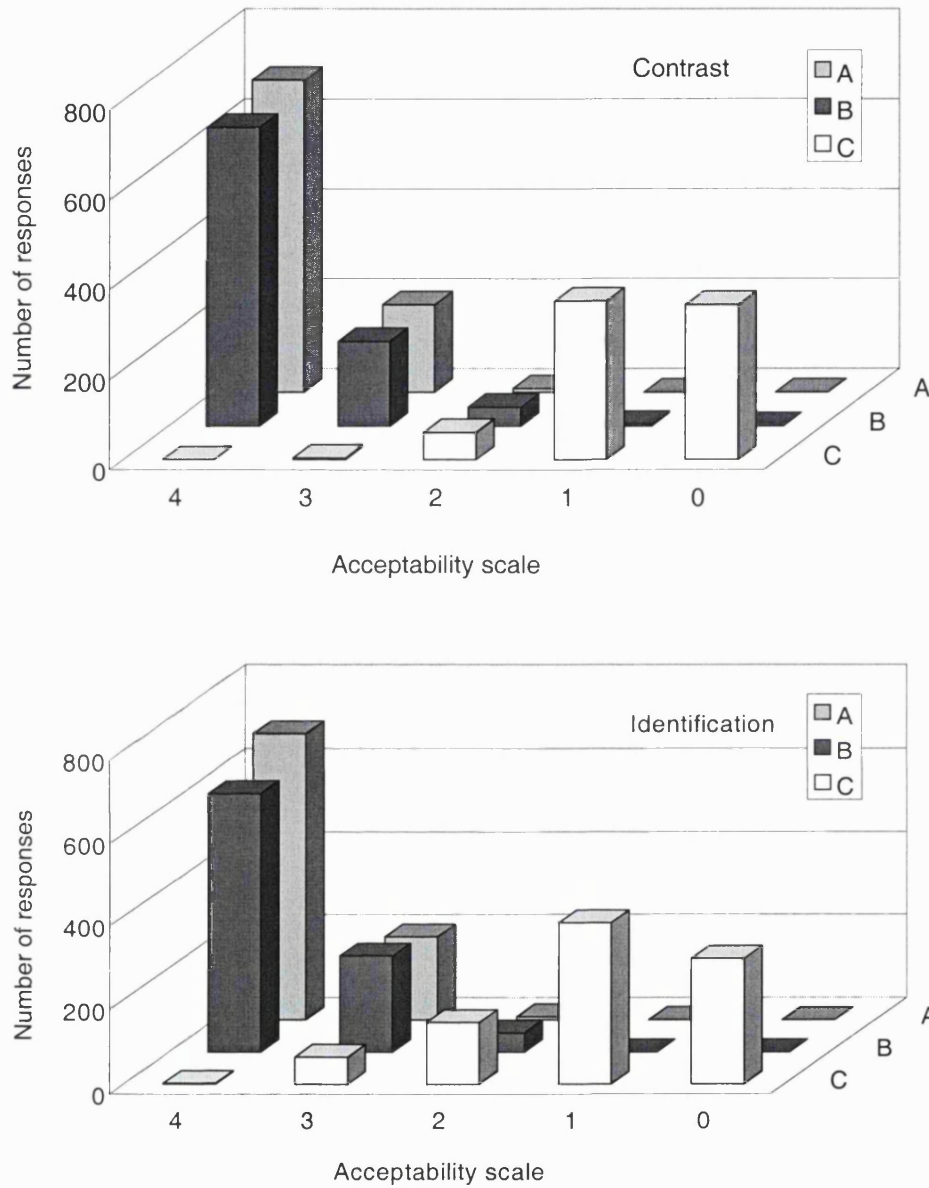


Figure 2.13. Tables and histograms showing the number of acceptable/unacceptable responses for sentences with narrow focus triggered by contrast and by identification according to three categories: A = narrow focus conveyed intonationally, B = narrow focus conveyed syntactically, and C = broad focus. The acceptability scale ranges from 4 (perfectly acceptable) to 0 (totally unacceptable).

The results of the acceptability test for narrow focus triggered by a contrastive prompt and by an identification prompt show that Central Catalan speakers found narrow focus triggered by accentual means (category A) and by syntactic means (category B) highly acceptable in the two conditions. For contrastive focus, 77% of the responses scored 4

and 21% 3 in focus conveyed through intonational means, and 74% scored 4 and 21% 3 in focus conveyed by syntactic means. For narrow focus triggered by an identification prompt, 77% of the responses corresponded to 4 and 22% to 3 in intonational focus and 64% corresponded to 4 and 26% to 3 in syntactic focus. Structures with a broad focus reading (category C) were clearly felt unacceptable as replies to a narrow focus question in a contrastive context where most responses scored 0 or 1. However, in an identification context a few more instances of higher acceptability values (such as 2 and 3) were observed for broad focus. These observations are consistent with the results obtained in the production test and support the idea reported in other studies on Romance languages (Zubizarreta 1998, García-Lecumberri 1995) that accentual focus on non-contrastive contexts is less natural. However, although a broad focus intonation was sometimes judged as slightly acceptable for a narrow focus context triggered by identification, still intonational strategies are felt to be more acceptable than unacceptable as responses to a non-contrastive question.

So far, the results of the acceptability test have shown that both intonational and syntactic strategies are felt to be highly acceptable to convey narrow focus. This is true for all variables, i.e. focal domain and triggering question. Moreover, the results show that in all cases accentual focus seems to be slightly more acceptable than syntactic focus. In order to see whether this difference is significant, a t-test was performed which compared the *mean acceptability value* of accentual focus to that of syntactic focus for all informants. The mean acceptability value for each case was calculated in the following way. For each participant the number of responses in each acceptability level was summed and multiplied by the acceptability number. Then, all values were added and divided by the overall number of items (64 for the whole test or 32 for the variable groups). This is illustrated in (2.16) for one informant in the category of accentual focus for the whole test.

(2.16)	54 responses scoring 4 (x4) =	216
	7 responses scoring 3 (x3) =	21
	2 responses scoring 2 (x2) =	4
	1 responses scoring 1 (x1) =	1
	0 responses scoring 0 (x0) =	0
	OVERALL SUM	242 : 64 items = 3.78 mean acceptability value

Once the mean acceptability value was obtained for all speakers both for the whole test and for the variable groups, t-tests were performed comparing the mean acceptability values for accentual and syntactic focus. As in the perception test, the significance level was considered to be 1 per cent ( $p < 0.01$ ), which gives a value of  $t$  of 2.47 for 27 degrees of freedom. The *overall mean acceptability value* (average of the mean acceptability values for all subjects) for accentual and syntactic focus and the results of the t-tests are presented in Table 2.2. The results of the t-tests show that in all cases differences in the acceptability of accentual and syntactic strategies are non-significant ( $p > 0.01$ ). This indicates that Catalan speakers judge accentual and syntactic strategies as equally acceptable to signal focus.

	Overall mean acceptability value		p	t value
	Accentual focus	Syntactic focus		
All	3.76	3.66	>0.01	1.8
Subject	3.75	3.62	>0.01	2.3
Verb	3.78	3.70	>0.01	1.5
Contrast	3.76	3.68	>0.01	1.5
Identification	3.75	3.68	>0.01	1.2

Table 2.2. Overall mean acceptability value for accentual and syntactic focus and results of the t-tests for the whole test and for the different variables.

Finally, in order to see whether there are any differences in the acceptability of the different strategies within the two domains (subject vs verb and contrast vs identification), t-tests comparing the mean acceptability values for each category (accentual, syntactic and broad focus) were performed. The results of the t-tests are presented in Table 2.3 for the subject/verb variable.

	Overall mean acceptability value		p	t value
	Subject	Verb		
Accentual	3.75	3.78	>0.01	-1.7
Syntactic	3.62	3.70	>0.01	-1.8
Broad	0.67	0.79	>0.01	-1.9

Table 2.3. Overall mean acceptability value and results of the t-tests comparing the acceptability of accentual, syntactic and broad focus in subject and verb focus domains.

The results show that (initial or medial) sentence position does not have an effect on the acceptability degree of the different strategies. Both for accentual focus and syntactic focus no significant differences are observed with respect to the subject and verb focal domains. Similarly, broad focus is also equally unacceptable as an answer to a narrow focus question in the two domains.

Finally, the results of t-tests comparing the mean acceptability value according to each category (accentual, syntactic and broad focus) in contrast and in identification contexts are presented in Table 2.4.

	Overall mean acceptability value		p	t value
	Contrast	Identification		
Accentual	3.76	3.75	>0.01	1.7
Syntactic	3.68	3.68	>0.01	1.7
Broad	0.47	1	<0.01	-3.78

Table 2.4. Overall mean acceptability value and results of the t-tests comparing the acceptability of accentual, syntactic and broad focus in contrast and in identification focus.

For narrow focus conveyed through both intonational and syntactic means, no significant differences are observed between contrast and identification prompts ( $p > 0.01$ ). This shows that both strategies are felt equally acceptable in the two contexts. However, significant differences are observed with the judgements of broad focus as a response to a narrow focus question ( $p < 0.01$ ). Although both in contrast and in identification contexts broad focus scores were quite low, more hesitation is observed in identification prompts (overall mean acceptability value=1). This seems to indicate that whereas in narrow focus

triggered by contrast the participants clearly identified a broad focus intonation as unacceptable, in narrow focus triggered by identification, a broad focus intonation was not so clearly rejected. Once more, this agrees with those studies that claim that in Romance languages intonational strategies are more clearly connected to a contrastive reading than a non-contrastive one.

Overall, the results of the acceptability test show that Central Catalan speakers seem to accept intonational and syntactic strategies to convey narrow focus on a quite similar basis for all the variables under study (subject vs verb focal domains and contrast vs identification triggering questions). As expected, broad focus sentences as responses to narrow focus triggering questions tend to be unacceptable in most of the cases. Only in non-contrastive contexts, sentences with a broad focus intonation were felt to be slightly more acceptable.

## **2.6 Conclusion**

In this chapter, the extent of usage and acceptability of accentual focus in Central Catalan has been tested. A production test and a perception test were carried out so as to investigate to what extent Central Catalan speakers could produce and recognise focus conveyed by accentual means. In addition, an acceptability test was performed to analyse how natural accentual focus is felt to be in Central Catalan and to examine similarities and differences in the acceptability of accentual vs syntactic strategies as means to convey narrow focus. In all tests, two variables were examined, namely, the effect of sentence position or focal domain (subject vs verb) and the triggering question (contrast vs identification). The results of the production test and the perception test showed that accentual focus is easily produced and perceived by Central Catalan speakers. The acceptability test demonstrated that Central Catalan speakers accept equally syntactic strategies and intonational strategies to signal the information structure of sentences. The effects of subject vs verb focal domains were only found in the perception test where listeners identified focus on subject better than on verb. Contrastive prompts elicited more intonational focus than identification prompts. Similarly, judgements in the acceptability of accentual focus in identification prompts were more indecisive than in



contrastive questions. Overall, the results of these experiments throw into doubt Vallduví's (1990, 1991, 1994a/b) and Vallduví and Zacharski's (1994) categorical idea that Catalan can only attain the focus/accent relation by means of syntactic strategies. The results of the three tests demonstrated that intonational strategies are frequently used and accepted as well. In the following chapters of this study, the phonetic and phonological properties of broad and narrow focus will be examined.

## ***Chapter 3: Broad focus***

### **3.1 Introduction**

The production, perception and acceptability tests presented in Chapter 2 showed 1) that the use of accentual strategies to signal focus is possible in Central Catalan (preferably with a contrastive meaning), and 2) that both accentual and syntactic devices are felt to be highly acceptable to convey focus. These results suggest that there is more flexibility in the choice of focal strategies in Central Catalan than was claimed in Vallduví (1990, 1991, 1994a, b) where the focus/accent association could only be mediated by syntax. Thus, the categorical classification of Central Catalan as a [-plastic] language is questioned in this study since intonational devices alone have proved to be used to attain the focus/accent alliance in a variety of contexts.

Given the evidence that narrow focus can be signalled by intonational means in Central Catalan, the next aim of this study is to describe the phonetic and phonological properties of accentual focus. However, before analysing in detail the intonational characteristics of narrow focus utterances, it is considered necessary to explore the intonational properties of sentences with a broad focus reading, in order to be able to identify differences and similarities between marked and unmarked intonational patterns. Thus, this chapter examines Central Catalan declaratives produced with a broad focus intonation. The phonological and phonetic properties of accentual focus in Central Catalan will be investigated in chapter 4. The data will be analysed within the AM framework of intonational representation.

This chapter consists of the following sections. Section 3.2 presents a brief recapitulation of the main tenets of the AM model with particular attention to the differences between phonological structure and phonetic variability. Then, section 3.3 introduces the data used for the analysis of broad focus structures. Section 3.4 explores the phonological and phonetic properties of Central Catalan neutral declaratives. In this section, issues such as 1) the kinds of tonal categories associated to relevant metrical primitives, 2) the levels of prosodic structure, and 3) differences

in the phonetic realisation of the tonal entities will be covered. Finally, sections 3.5 and 3.6 present the discussion of the results and the conclusions respectively.

### 3.2 Framework

The framework used to describe the data is the AM approach of intonational analysis. This framework started with Pierrehumbert (1980) and was subsequently developed in Beckman and Pierrehumbert (1986), Pierrehumbert and Beckman (1988), Ladd (1996) and in the different ToBI systems (see Beckman and Hirschberg 1994 for American English), among many others. Since Pierrehumbert's initial work, the framework has been subject to several modifications. The version we adopt in this thesis differs from Pierrehumbert (1980) in two aspects: 1) downstep is not triggered by a sequence of bitonal tones but is a property of the lowered tone (as in Beckman and Hirschberg 1994 or Ladd 1996), and 2) the phrase accent is no longer a floating tone but is associated to an intermediate level of prosodic structure (as in Beckman and Pierrehumbert 1986 or Beckman and Hirschberg 1994).

As presented in chapter 1, the AM framework decomposes intonational tunes into sequences of H and L tones, which are linked either to metrically strong syllables or to the edges of the prosodic units. Tones marking the prosodic edges are divided into *boundary tones*, which accompany a major phrase boundary, such as the *intonation phrase* (IP), and *phrase accents*, which are associated to a minor level of prosodic structure, namely, the *intermediate phrase* (ip). The phonological events accompanying the stressed syllables are called *pitch accents*. Not all syllables are specified as having a tone. Tonally unspecified syllables get their surface tone by interpolation between two consecutive tonal targets. The rhythmic properties of these tonal primitives are indicated as follows: [%] marks a boundary tone, [-] a phrase accent and [\*] a pitch accent. Whereas tones signalling the end of a prosodic domain can only be monotonal (i.e. H% and L% for a boundary tone, and H- and L- for a phrase accent), pitch accents can be both monotonal and bitonal (e.g. H\*, L\*, H\*+L, L\*+H, etc.)<sup>1</sup>.

---

<sup>1</sup> For a more detailed account of the tonal primitives proposed within the AM model see Ladd (1996) and also section 1.3.2 of the present study.

At the phonological level, the relationship between tonal events and metrically relevant items, such as strong syllables or phrase edges, is mediated by autosegmental *association*. Association involves a link between two items on separate phonological tiers. Thus, pitch accents are associated to metrically strong syllables, and edge tones to the ends of higher prosodic domains. However, at the phonetic level, phonologically equivalent units may be subject to realisational differences. In recent research (Ladd 1996, Grabe 1998a, b, Nolan 1999), several sources of phonetic variability have been identified, such as the *compression* and *truncation* of accents or tunes, and differences in the *alignment* of peaks associated to H\* pitch accents. The levels of phonological structure observed in different languages and the main sources of phonetic variability are briefly discussed in the next section.

### 3.2.1 Phonological structure

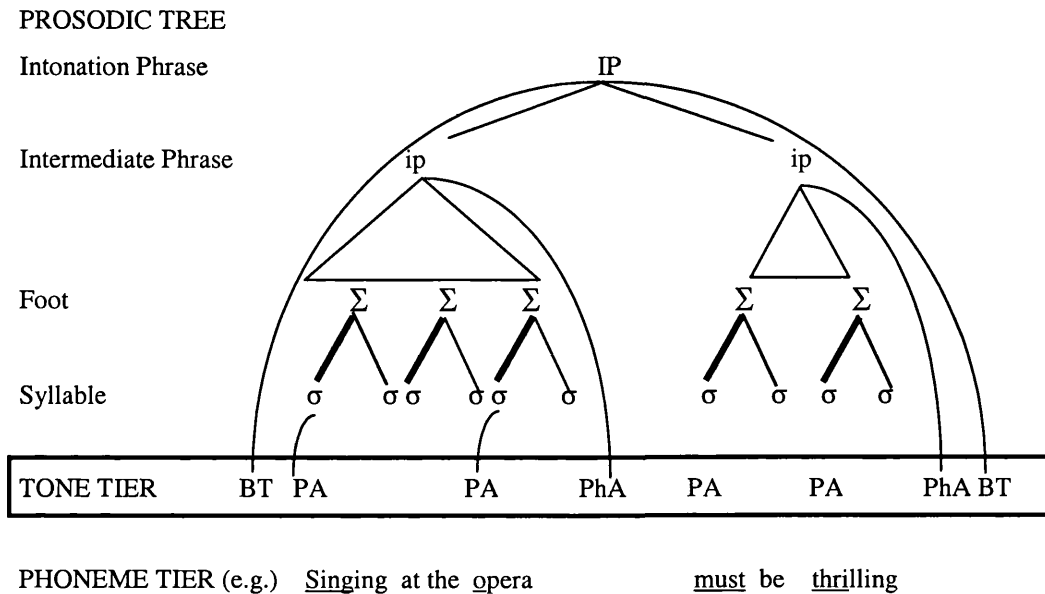
The first detailed account on how tonal primitives are phonologically related to metrical primitives is provided in Pierrehumbert and Beckman (1988) for their analysis of Japanese tone structure. In this study, metrical information is represented by means of a *prosodic tree*, which consists of a hierarchy of phonologically defined units. The idea of prosodic trees has been applied to several languages. However, the number and kinds of levels of phonological structure within the tree varies among these languages. In English, for example, four prosodic levels have been posited, namely, the IP, the ip, the foot and the syllable. The IP is the highest level of prosodic structure whose limits are demarcated by a boundary tone, either L% or H% at the right edge, and no tone or H% at the left edge<sup>2</sup>. The IP is the prosodic domain within which pitch range is specified and where a large disjuncture, such as a major pause or break, is perceived at the boundary. The ip is a lower constituent, which must include at least one pitch accent. The ip is tonally demarcated by the presence of a phrase accent, either L- or H- at the right edge. It has a medium degree of disjuncture (i.e. no large breaks). In English, the ip is defined as the domain of downstep (Beckman and Pierrehumbert 1986). However, this varies cross-linguistically since in other languages downstep has been shown to occur at the IP level (e.g. Greek: Arvaniti and

---

<sup>2</sup> The default initial boundary is supposed to have a mid or low pitch range and is left unmarked for transcription (Pierrehumbert 1980, Beckman and Hirschberg 1994). For criticisms on this aspect, see Cabrera-Abreu (1996) who proposes a model with no L tones.

Baltazani 1999). A simplified adaptation of Pierrehumbert and Beckman's tree for English is presented in (3.1) below<sup>3</sup>.

(3.1)



As observed in (3.1), the highest level of the prosodic tree is the IP, which dominates lower daughter nodes representing relevant entities of phonological structure. Domination is schematised by straight lines. Curved lines represent the association between elements at the prosodic tier and at the tone tier. Association between tonal entities and metrical entities operates both at a low-level node, where syllables are linked to pitch accents (PA) and at higher level nodes, where edge tones (PhA and BT) are associated to the edges of prosodic phrases. The association between syllables and pitch accents is not arbitrary but governed by metrical strength (see chapter 1 for more details). Pitch accents are linked to the strong branch of the foot, marked with a thicker black line.

The prosodic tree varies cross-linguistically. For example, languages such as Japanese (Venditti 1999) and Korean (Jun 1996) have no ip level of phonological structure.

<sup>3</sup> There are several aspects related to this hierarchical representation that have not been shown since they are not relevant for the purposes of this study. For example, pitch accents might have an internal structure (bitonality) as opposed to edge tones, which are monotonal (see Grice 1995a, b for more details).

Instead, the domain below the IP is the Accentual Phrase<sup>4</sup>. Alternatively, Jun and Fougeron (to appear) claim that French has both an ip and an Accentual Phrase. The Accentual Phrase is defined by a specific tonal pattern, which demarcates the right edge of the phrase. In Korean and French, the delimitative tones are LH. In Japanese, there is a rise associated to the second mora and a subsequent fall to a low tone at the end of the phrase. No Accentual Phrase level of phonological structure has been proposed for English since there are no delimiting edge tones for such a constituent (Beckman and Pierrehumbert 1986).

Cross-linguistic differences not only apply in the levels of phonological hierarchy but also in the nature, function and distribution of pitch accents. For example, whereas in English the association between pitch accents and metrically strong syllables occurs at a postlexical level (Pierrehumbert 1980), in Japanese the location of pitch accents is specified at the lexical level (Beckman and Pierrehumbert 1986, Pierrehumbert and Beckman 1988 and Venditti 1999). This means that the distribution of pitch accents in Japanese is fairly fixed as opposed to English where the distribution of accents is more flexible and is closely linked to a focussing function.

The intonational structure described so far for English (with two levels of phrase prosodic structure: the IP and the ip) and the postlexical association of pitch accents as a reflex of focussing have been argued in other languages, such as Palermo Italian (Grice 1995a), European Portuguese (Frota 1998), Bengali (Hayes and Lahiri 1991)<sup>5</sup>, Greek (Arvaniti and Baltazani 1999), German (Benzmüller and Grice 1999) and Central Catalan (Prieto 1995, 1997), among others. In this study, a similar prosodic representation to the one sketched in (3.1) for English will apply to Central Catalan. Thus, the two levels of phrasing proposed by Prieto will be confirmed by the data as well as the postlexical association of tonal primitives to metrical relevant units. However, a smaller prosodic domain positioned below the ip will also be justified on the basis of the data presented.

---

<sup>4</sup> Pierrehumbert and Beckman (1988) proposed three levels above the word in the prosodic hierarchy of Japanese: the Accentual Phrase, the ip and the Utterance. In the J\_ToBI model (Venditti 1999), the ip and the Utterance have merged into a single level of phrasing, the IP.

<sup>5</sup> The levels of prosodic structure proposed for European Portuguese and Bengali are defined both on intonational grounds and on the application of junctural phenomena. In both studies, the level below the IP is the Phonological Phrase (as in Selkirk 1980 or Nespors and Vogel 1986).

### 3.2.2 Phonetic variability

As in segmental phonetics, research on intonation has shown that the phonological units used to describe a particular pitch contour are subject to variability. This means that functionally equivalent entities can be realised differently in different environments. Nolan (1999) identifies three processes that cause variation at the tone level, namely, the alignment of F0 peaks in H\* pitch accents, compression and truncation.

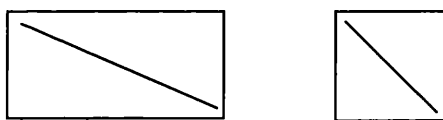
Several studies (Silverman and Pierrehumbert 1990, Prieto et al 1995, among others) have shown that at the phonetic level the synchronisation between an F0 feature and a prominent syllable does not always coincide. Sometimes the F0 peak corresponding to an H\* pitch accent is not realised within the accented syllable but aligned considerably later. The distinction between *association* and *alignment* was formalised in Ladd (1983, 1996), who claimed that whereas association is an abstract structural property, alignment is a phonetic property of the timing of F0 events and segmental events. Studies on the alignment of F0 contours in several languages have argued that the location of peaks in H\* pitch accents depends on a number of factors (see House and Wichmann 1996 for a review). In brief, some of the factors that have an effect on the timing of peaks are the following: 1) segmental duration, 2) right-hand prosodic environment (e.g. word or phrasal boundaries), 3) left-hand prosodic contexts (e.g. presence of sentence-initial unstressed syllables or anacrusis), 4) rhythmical organisation (e.g. stress clash) and 5) discourse-structure.

Steele and Altom (1986) and Silverman and Pierrehumbert (1990) showed for American English that F0 peaks tend to shift rightwards as the segmental duration of the rhyme increases and that upcoming prosodic events, such as word or phrasal boundaries and following accented syllables, have retracting effects on peak location. Similar results were found in Prieto et al (1994, 1995) for Mexican Spanish. F0 peaks are more delayed as the duration of the accented syllable increases and retracted when adjacent to a word, an ip or an IP boundaries and in stress clash contexts. Nolan and Farrar (1999) showed that in British English peak delay is more frequent when there is no anacrusis (i.e. when the first accent of the sentence is not preceded by unstressed

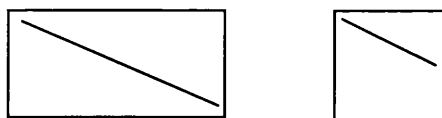
syllables). Finally, House and Wichmann (1996) and Wichmann et al (1999) demonstrated that in British English the timing of F0 peaks is determined by discourse position: paragraph-initial peaks are more delayed than peaks in sentence-final position.

Recent experimental work (Grabe 1998a, b) has shown that the realisation of pitch accents may vary depending on the length of the segmental stretch and on the availability of voiced segmental material. Languages appear to adopt one of two strategies to accommodate accents to short or poorly voiced sequences, namely, truncation or compression of accents. In truncation, part of the contour is not realised in shorter sequences or in segmental structures with reduced voiced material. Compression, on the other hand, consists of shrinking the tonal sequence to fit into a smaller segmental space. For example, Grabe shows that the realisation of falls (H\*+L) in English and in German presents cross-linguistic differences. In English, the fall compresses, as sonorant segmental material is shorter. In comparable German structures, the fall is truncated and emerges as an apparently high accent. This is illustrated in (3.2) (adapted from Nolan 1999). Thus, in both languages, the falling pattern involves the same underlying phonological category (H\*+L), subject to realisational differences depending on the segmental environment.

(3.2) COMPRESSION of English H\*+L



TRUNCATION of German H\*+L



Studies on truncation and compression have shown that languages and dialects vary in the use of such strategies. For instance, Danish (Grønnum 1991), Palermo Italian (Grice 1995a) and Hungarian (Ladd 1996) seem to prefer truncation rather than compression to accommodate the F0 patterns in short or poorly voiced segmental sequences. On the other hand, English (Ladd 1996) and Central Catalan (Prieto to appear d) tend to favour the compression of the F0 movements in comparable



structures. Grabe et al (to appear), as mentioned in Nolan (1999:7), show dialectal differences in the usage of truncation and compression in several varieties of English. Whereas Standard Southern British and Newcastle compress both rises and falls in short voiced material, Leeds and Belfast truncate.

In the next sections of this chapter, both the phonological and phonetic properties of Central Catalan declaratives are examined. Particular attention is devoted to identifying the types of phonological entities associated to metrical primitives as well as to investigating differences in their phonetic realisation.

### **3.3 Data**

The broad focus declaratives analysed in this chapter consisted in exactly the same sentences as the ones designed for narrow focus on the subject and on the verb but produced with an unmarked intonation (see section 1.1 in Appendix 1 for the whole list of sentences). These sentences were used since they had the advantage of having the same segmental configuration as narrow focus utterances and hence the only source of variation was expected to be intonational. The broad focus utterances were gathered at a second recording session<sup>6</sup>. To obtain utterances with a broad focus pattern, the informants were asked to read a list of sentences as if they were answers to a "what happens?" question-type. The informants for this test were six female speakers who also participated in the narrow focus production test and whose productions were selected for the analysis of narrow focus (see section 4.3 in chapter 4 for the criteria used to select these six speakers out of the thirty-four informants). The six speakers are identified as CP, DV, ER, MC, NG and NM. The instrumentation and recording procedures used in this second production test were exactly the same ones as for the first test (see section 2.3.1.4 in chapter 2 for details).

The broad focus sentences gathered in this test comprised a minimum of 3 and a maximum of 5 lexically stressed words per utterance. Examples of sentences with

---

<sup>6</sup> Although some broad focus sentences were already acquired as distractors in the first production test, those sentences were not used for the intonational analysis since their segmental structure was very different from that of narrow focus utterances.

different number of stresses are provided in (3.3), where stressed syllables are underlined. Finally, a few structures consisting of only one stressed word were also included in the materials to compare the intonation of single-stressed utterances with that of multi-stressed utterances. Single-stressed structures consisted of both nouns and verbs. The list of structures with only one lexical word is displayed in Appendix 2 (section 2.1). Overall, the number of multi-stressed broad focus sentences gathered in this test is 384 (64 utterances for each speaker). The overall number of single stressed utterances is 216 (36 items for each speaker).

- (3.3) a. La Mila nega la maionesa (3 stresses)  
“Mila spoils the mayonnaise”
- b. L'home venia llimones madures (4 stresses)  
“The man sold ripe lemons”
- c. Les meves amigues eren bones nenes (5 stresses)  
“My friends were good girls”

### 3.4 Results

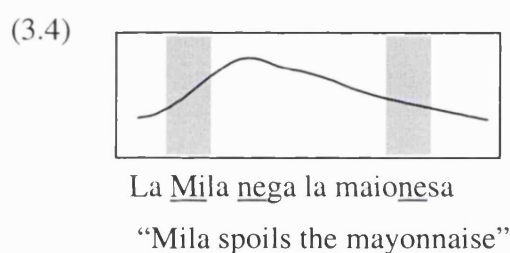
The results of Central Catalan declaratives produced with a broad focus intonation are presented in two parts. The first part includes a general overview of the basic intonation characteristics observed in the F0 traces of neutral declaratives. In this section, the phonological representation provided in Bonet (1984) and in Prieto (1995) is reviewed and applied to the data. The second part consists of a more detailed acoustic analysis of the utterances. In this section, some of the ideas proposed in the former phonological interpretations are questioned and modified on the basis of our results. In this second part, three main issues are examined in detail: 1) the levels of phrasing, 2) the intonational properties of prenuclear (non-final) accents, and 3) the intonational properties of the nuclear (final) accent.

From now onwards, in order to present the data in a more flexible way, F0 traces will be sketched after the original trace and incorporated within the text. In these reproductions, shaded areas correspond to the limits of accented syllables. The

original F0 traces and speech waveforms for each sketched trace are presented in Appendix 5 for further reference.

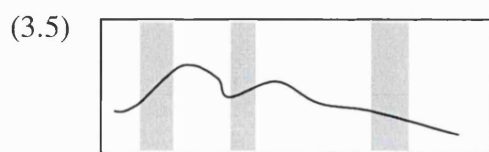
### 3.4.1 General properties

The majority of F0 traces for broad focus declaratives obtained in the recordings showed an F0 rise in the vicinity of the first stressed syllable followed by a progressively falling contour, which reached its F0 minimum at the end of the utterance. An example is provided in (3.4) which sketches the F0 trace for the sentence *la Mila nega la maionesa* produced by speaker CP. Even though there is no obvious F0 movement within or near the limits of the last stressed syllable, we still assume that it is accented (see section 3.4.3.3 for a justification on this claim).



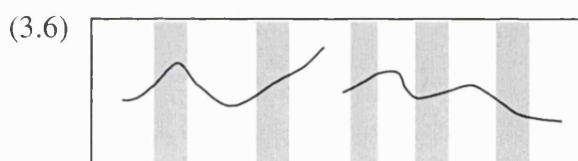
Although in most cases the stressed syllable/s occurring between the first and the final accents did not become accented, as illustrated in (3.4) above, inter- and intra-speaker variability was observed in the accentuation of medial stresses. For example, whereas speaker NG consistently produced as many accents as stressed syllables available, speakers CP and MC tended only to accent the first and last stressed syllables of all their utterances. Finally, speakers DV, ER and NM showed some fluctuation between the accenting and non-accenting of medial stresses. Usually, the higher the number of medial stressed syllables, the higher the number of accents<sup>7</sup>. Whenever realised, non-final or prenuclear accents were characterised by a rising pitch around the stressed syllable. This is illustrated in (3.5) for the same sentence as in (3.4) produced by speaker NG.

<sup>7</sup> Sentences where most or all medial stresses become accented were considered more expressive than sentences where only the first and last accents were realised. This impression was independently confirmed by two Catalan speakers.



La Mila nega la maionesa

Most sentences were uttered with no evident prosodic breaks. However, a few speakers (mainly ER and NM) tended to produce some kind of intonational boundary between the subject and the predicate. The presence of this boundary was consistent for most speakers (except for speaker MC) in sentences with a double stressed subject. Phonetically, the boundary was signalled by an abrupt F0 change in the pitch contour, which consisted of a drastic rise. This is illustrated in (3.6) for an utterance with a double-stressed subject. Sometimes speaker ER produced a fall-rise movement at the end of the subject domain. This pattern will be analysed in more detail in section 3.4.2.1.



Les meves amigues eren bones nenes

"My friends were good girls"

Finally, utterances consisting of only one stress were produced with an F0 peak over the accented syllable. The fall occurred on the poststressed syllable/s. This is illustrated in (3.7). If there were no poststressed syllables, the whole rise-fall movement was compressed on the accented syllable.



Mejaven

"They ate"

The intonational characteristics of Central Catalan neutral declaratives observed so far in the data are consistent with the results reported in other descriptive studies, both

based on acoustic data (Prieto 1995, 1997, Estebas-Vilaplana and Maidment 1999a) and on auditory data (Prieto to appear b, Bonet 1984, Recasens 1977, 1993, Badia i Margarit 1994). In all cases, the neutral declarative intonation is described as consisting of an initial rise in the F0 trace followed by a fall, with possible differences in the number of prenuclear accents as well as in the phrasing.

Even though studies seem to agree on the phonetic properties of Central Catalan neutral declaratives, studies do not entirely coincide with the tonal interpretation proposed for the F0 contours. As presented in chapter 1 (section 1.4.2.1), Prieto's (1995) interpretation of single-accented structures involves an H\* associated with the accented syllable followed by L-L% edge tones, as reproduced in (3.8).

$$(3.8) \quad [ [ \text{me} \underline{\text{j}} \text{a} \text{v} \text{e} \text{n} ] ]$$

$$\begin{array}{c} | \quad | \\ \text{H}^* \quad \text{L-L}\% \end{array}$$

On the other hand, multi-stressed structures are characterised by an H\* associated to the prenuclear accents (as many as realised) and an L\* associated to the final accented syllable. As before, an L- phrase accent and an L% boundary tone mark the end of the contour. This is illustrated in (3.9).

$$(3.9) \quad [ [ \text{La} \underline{\text{M}} \text{i} \text{l} \text{a} \underline{\text{n}} \text{e} \text{g} \text{a} \text{ la} \underline{\text{m}} \text{a} \text{i} \text{o} \text{n} \text{e} \text{s} \text{a} ] ]$$

$$\begin{array}{c} | \quad | \quad | \quad | \\ \text{H}^* \quad (\text{H}^*) \quad \text{L}^* \quad \text{L-L}\% \end{array}$$

Finally, whenever an intonational boundary is present between the subject and the predicate, this is interpreted as an intermediate phrase boundary marked by the presence of H- (Prieto 1997). According to this, the tonal description of the contour in (3.6) would be as shown in (3.10).

$$(3.10) \quad [ [ \text{Les} \underline{\text{m}} \text{e} \text{v} \text{e} \text{s} \underline{\text{a}} \text{m} \text{i} \text{g} \text{u} \text{e} \text{s} ] [ \underline{\text{e}} \text{r} \text{e} \text{n} \underline{\text{b}} \text{o} \text{n} \text{e} \text{s} \underline{\text{n}} \text{e} \text{n} \text{e} \text{s} ] ]$$

$$\begin{array}{c} | \quad | \quad | \quad | \quad | \quad | \quad || \\ \text{H}^* \quad \text{H}^* \quad \text{H-} \quad \text{H}^* \quad \text{H}^* \quad \text{L}^* \quad \text{L-L}\% \end{array}$$

As already reported in chapter 1, Prieto's phonological interpretation has both differences and similarities with the analysis proposed in Bonet (1984) mostly based on the author's perceptual judgements. Both studies agree that the last accent of multi-stressed Central Catalan neutral declaratives is low<sup>8</sup>. However, they disagree in the interpretation of prenuclear accents. Bonet accounts for the tonal entities of non-final accents depending on how the sentence is subdivided into prosodic constituents. In sentences produced with one single prosodic phrase, as (3.4) or (3.5), all accented syllables would be low. This is illustrated in (3.11).

(3.11) [La Mila nega la maionesa]  
           |    |                    |  
           L   (L)                   L

Sentences produced with more than one prosodic constituent present the following characteristics: all accented syllables in the last constituent are low and the accented syllables in non-final constituents are high, except for the final syllable in a non-final constituent which is low. This is represented in (3.12).

(3.12) [[Les meves amigues] [eren bones nenes]]  
           |        |        |    |    |  
           H       L        L    L    L

In the next section, a detailed acoustic analysis of the data will be provided to supply more information about the phonetic and phonological properties of Central Catalan neutral declaratives. On the basis of this analysis, it will be argued that the data do not entirely agree with either Bonet's or Prieto's phonological interpretations and an alternative proposal will be provided. The acoustic analysis of the F0 contours will be carried out in three parts: 1) levels of phrasing, 2) non-final or prenuclear accents, and 3) the final or nuclear accent<sup>9</sup>.

<sup>8</sup> Bonet does not specify the accent type of single-stressed utterances.

<sup>9</sup> The reason for examining prenuclear and nuclear accents separately is because they show different pitch movements (rise vs fall respectively).

### **3.4.2 Acoustic analysis**

#### **3.4.2.1 Phrasing**

The data analysed in this study varied with respect to the levels of phrasing or prosodic domains observed in neutral declaratives. Although most sentences were produced with no prosodic breaks, sometimes some kind of intonational boundary, presumably an intermediate phrase (ip) boundary, was observed between the subject and the verb. No ip boundary was ever detected between the verb and the object, whose level of cohesion was much stronger than that between the subject and the verb. This behaviour is consistent with syntactic structure where the boundary between an NP and the following VP is stronger than the boundary between the constituents making a VP (namely, between the V and the NP). As reported in Prieto (to appear b, c) and Bonet (1984), the appearance of prosodic boundaries between the subject and the verb in Central Catalan neutral declaratives depends on several factors, such as the length of the constituents, speech style or the need to disambiguate otherwise ambiguous sentences.

Before analysing in more detail the factors that trigger the presence of this optional boundary in our data, it is necessary to confirm its phonological status. Does it correspond to an ip boundary or to a major level of prosodic phrasing, such as an IP boundary? Studies on the cues to identify major boundaries, such as those marking an IP, and minor boundaries, such as those showing an ip, claim that it is not always straightforward to differentiate between these two levels of phrasing. Ladd (1986) for English and the different ToBI systems (e.g. Arvaniti and Baltazani 1999 for Greek, Venditti 1999 for Japanese, Grice and Benz Müller 1995 for German, among others) claim that a major pause (F0 break) signals the presence of an IP. Other cues such as an abrupt F0 change or preboundary lengthening can be used to identify an ip boundary. These criteria have been followed in this study, whenever possible, to confirm the status of the boundary between the subject and the predicate, observed in some of the utterances.

The presence of three cues justified the classification of this prosodic break as an ip boundary: 1) a rather abrupt rise in the F0 contour, 2) the lengthening of the preboundary syllable and 3) the absence of a major F0 break. These three pieces of

evidence are clearly observed in sentences with a double accented subject. These sentences are examined below in relation to the data of four speakers (CP, DV, NG and NM). The productions of speaker MC are not analysed because she only accented one of the two potentially accentable syllables in double stressed subject domains. Speaker ER is analysed later because she produced a different F0 trajectory at the end of the ip. Whereas most speakers showed an abrupt F0 rise, she tended to produce a fall-rise pattern.

For speakers CP, DV, NG and NM, the F0 contours of double stressed subjects showed an F0 peak in the vicinity of the two accented syllables (see (3.6) above for an example). In all cases, the second peak was much higher than the first one, suggesting the presence of a boundary. This is illustrated in Figure 3.1 below, which displays the mean frequency value (in Hz) for the two peaks in double accented subject domains for the four speakers. In all cases, the mean F0 value of the second peak is much higher than that of the first peak. In order to see whether the differences in the F0 values were significant, t-tests comparing the two F0 samples were performed. The results of the t-tests showed that the F0 mean of the second peak is significantly higher ( $p < 0.001$ ,  $t = 1.15$ ,  $df = 15$ ) than the F0 mean of the first peak for all speakers. These results confirm the presence of an abrupt F0 rise at the end of the double accented subject, which may signal an ip boundary.

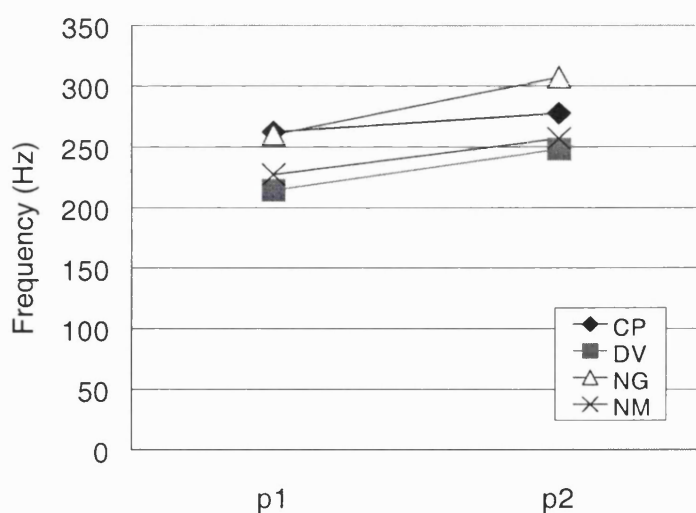


Figure 3.1. Mean frequency value (in Hz) of peak 1 (p1) and peak 2 (p2) in double accented subject domains for four speakers.



The presence of an ip prosodic boundary is further supported by the lengthening of the preboundary syllable. One of the cues that traditionally has been used to identify the presence of a boundary is the longer duration of the preboundary syllable as opposed to the duration of similar (non-preboundary) syllables in the same utterance (Cruttenden 1986). In our data, the length of the preboundary syllable was measured and compared to the length of other syllables, which contained the same number of segments and which were located in a similar position within the word but not preceding a boundary. Thus, for example, in the double stressed subject *les meves amigues* [ləz 'meβəz ə'miɣəz] the duration of the first underlined syllable (non-preboundary syllable) was compared to the duration of the second underlined syllable (preboundary syllable). The results are shown in Figure 3.2, which displays the mean duration of the first (non-preboundary) and the second (preboundary) syllables in double stressed subjects for the four informants.

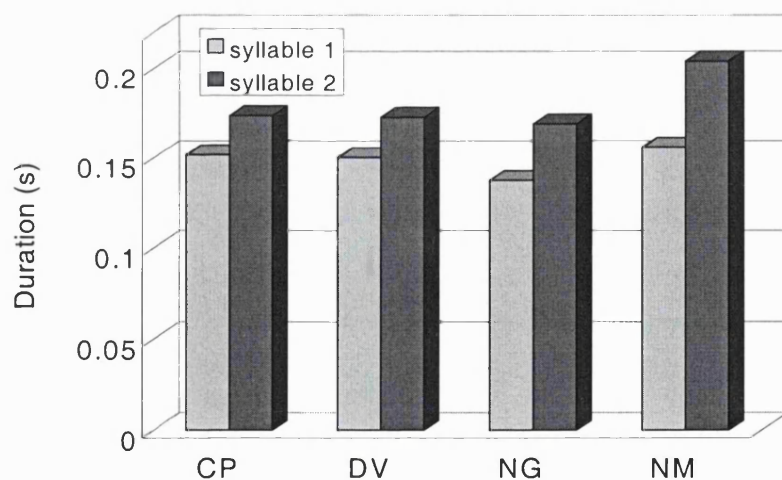


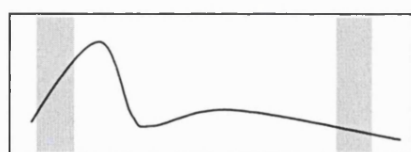
Figure 3.2. Mean duration (in seconds) of the first (non-preboundary) and second (preboundary) accented syllables in double stressed subject domains for four speakers.

In all cases,  $p < 0.001$  and  $t = 1.75$  for 15 df, indicating that the duration of the last syllable of the second word (syllable 2) is significantly longer than the duration of the last syllable of the first word (syllable 1). This provides further evidence for the presence of a boundary between the subject and the verb in structures with a double accented subject.

So far, cues such as the abrupt F0 rise and the lengthening of the preboundary syllable confirm the presence of a boundary between the subject and the verb in sentences with a double accented subject. Furthermore, since no major pause (i.e. no major F0 break) was observed between the subject and the verb, this boundary was interpreted as an ip boundary, marked by the presence of an H- phrase accent. This agrees with the results reported in Prieto (1997) and in Estebas-Vilaplana and Maidment (1999a).

The presence of an ip boundary between the subject and the predicate (marked H-) could also be observed in some of the utterances with a single stressed subject. Although in these cases the kind of phonetic evidence used to detect the presence of an ip boundary was not as clear as in utterances with a double stressed subject, a boundary could still be noticed in most of the productions of speakers NM and ER<sup>10</sup>. The piece of evidence that allowed us to determine the presence of an ip boundary in utterances with a single stressed subject for speakers NM and ER was found by comparing the F0 contours of these speakers with the F0 contours of the other informants. Whereas most of the pitch traces of the other speakers showed a progressively falling contour after the initial rise, the productions of speakers NM and ER did not exhibit a smoothly falling trajectory after the rise but a quite abrupt fall, most probably indicating the beginning of a new ip. This is illustrated in (3.13), which compares the same sentence produced with an ip break as in (3.13a) and without an ip break as in (3.13b) for speakers ER and CP respectively.

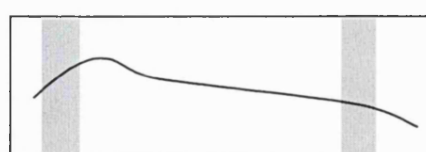
(3.13) a.



L'home venia llimones madures

“The man sold ripe lemons”

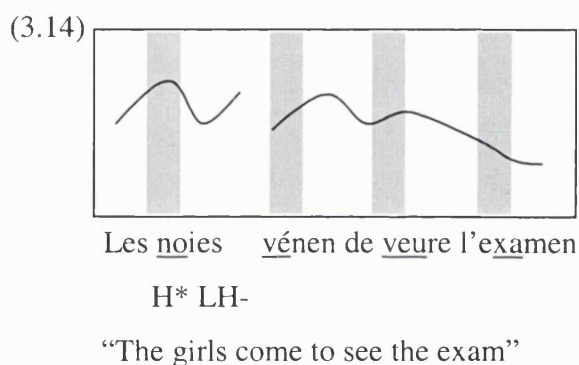
b.



L'home venia llimones madures

<sup>10</sup> In utterances with a single stressed subject, it was difficult to find a reference point to see whether the rising movement observed over the accented syllable of the subject was abrupt enough to be analysed as an ip marker. Whereas in sentences with a double accented subject, the height of the final rise could be compared with the height of the first accented syllable, in sentences with a single stressed subject only accented syllables after the subject were available. In these cases, it was difficult to determine whether the lower pitch of an accented syllable after the subject was due to the presence of a preceding boundary or was due to the presence of downstep over the lowered syllable itself. Second, with respect to the preboundary syllable lengthening, it was difficult to find entirely comparable syllables both in terms of segmental structure and in terms of stress/accent distribution.

Some of the productions of speaker ER were produced with a fall-rise rather than a rise between the subject and the verb. The appearance of a fall-rise F0 movement confirms the presence of an ip boundary, since this pattern is mainly used to signal the end of a phrasal domain. This fall-rise involved an H\* pitch accent associated to the accented syllable followed by a fall over the postaccent syllable/s and a subsequent rise whose peak was anchored at the end of the word. This is illustrated in (3.14) for an utterance with a single stressed subject. The use of a fall-rise pattern to mark the end of a prosodic domain in Central Catalan was also observed in Estebas-Vilaplana (1996) for the intonation of enumerations. Most of the elements making up a list ended with a fall-rise intonation.



Overall, the results show that Central Catalan neutral declaratives are mainly produced in one single tone group, although sometimes a boundary can be observed between the subject and the predicate. Phonetic evidence shows that this boundary tends to signal a minor prosodic domain (ip). The results of this study indicate that the presence of this boundary is optional and it can be realised as an abrupt F0 rise or as a fall-rise pattern. This agrees with former studies on Central Catalan intonation (Bonet 1984, Prieto 1997, to appear b, among others). No boundary is observed between the verb and the object, confirming a closer link between these two constituents, as in syntactic structure.

Although the appearance of an ip boundary between the subject and the predicate is optional, the presence/absence of this boundary seems to be conditioned by factors such as the length of the constituents, the number of accents in each constituent, and the speech style. The data showed that an ip boundary was more consistently observed in sentences with a double stressed subject than in those containing a single stressed

subject. This behaviour agrees with the idea that intonational boundaries are closely related to the “weight” of the constituents. In Ladd’s (1996: 234) terms, “the heavier a constituent is, the more likely it is to constitute its own intermediate phrase”. According to Ladd, the “weight” of constituents can be the result of several factors, such as the length of the speech chunks, the degree of semantic separateness, and the presence of contrastive focus. The data show that the length of constituents does have an effect on the number of prosodic phrases, since the longer the constituent in terms of words (e.g. subjects with two lexically stressed words), the more likely the presence of an ip boundary. According to the data, the number of pitch accents seems to be another aspect that might make a constituent heavier since the data show that whenever the two stresses of a double stressed subject become accented, the H-boundary was present. However, when only one of the two stresses was accented, no ip boundary was realised. This was observed in the data of speaker MC and occasionally for speaker CP. Finally, according to Prieto (to appear b) another factor that might influence the presence of boundaries is the rate of speech. The fact that the data gathered for this study was mainly reading data and produced with a quite slow tempo could also influence the presence of boundaries, especially in short sentences.

To sum up, the data show that Central Catalan declaratives tend to be produced with one single tone unit. However, an optional boundary (signalling the presence of an ip) can be observed between the subject and the predicate. The appearance of this boundary is influenced by aspects such as the length of the constituents, the number of pitch accents in each constituent and speech style or rate. In line with other studies (Ladd 1986, Arvaniti and Baltazani 1999, Grice 1995a, among others), the ip in Central Catalan is a level of prosodic structure situated below the IP and signalled by an abrupt F0 change at its right edge, indicating the presence of a phrase accent. So far, only evidence for an H- phrase accent has been observed. Evidence for an L-phrase accent will be provided in chapter 4. In Central Catalan, the ip must contain at least one pitch accent (and a word edge tone)<sup>11</sup>. The IP, on the other hand, is made up of at least one ip and is demarcated by an F0 pause.

---

<sup>11</sup> See section 3.4.2.2 for more details on word edge tones.

### 3.4.2.2 Non-final accents (prenuclear accents)

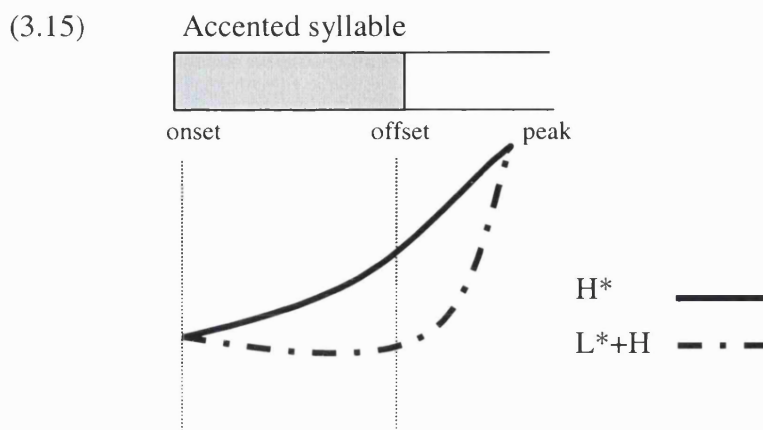
Several studies have shown that in many languages prenuclear accents in neutral declarative structures involve a rising movement, whose peak tends to be aligned after the accented syllable. However, the phonological interpretation of such rises varies cross-linguistically and sometimes studies on the same language maintain competing views for the phonological description of such rises. Prenuclear rises, for example, have been widely studied in Greek, yet no conclusion on the best way to categorise them has been reached. Arvaniti and Ladd (1995) and also the Greek version of the ToBI system (Arvaniti and Baltazani 1999) propose the L\*+H bitonal pitch accent to describe prenuclear rises. Arvaniti and Ladd argue for this accent due to the stability in the alignment and in the scaling of the starred L. However, further research in Greek (Arvaniti et al 1998, to appear) shows that even though to describe rising prenuclear movements in Greek a bitonal accent seems to be required (LH), there is no clear evidence of which of the two tones is aligned with the accented syllable and hence it is difficult to decide which is the starred tone. The L\*+H accent type is now ruled out because H does not exhibit the stable alignment predicted by bitonality, where the trailing tone is expected to follow the starred tone by a fixed interval (Grice 1995b), and also because in clash contexts, L\* is more robust than H. Similarly, the option L+H\* is discarded, since L does not precede H\* by a fixed interval. A further possibility is investigated, namely, the idea that H is an edge tone that marks the end of the accented word. However, this option is also discarded since in words with antepenultimate stress the H is located relative to the accented syllable and not at the word-edge. Finally, the possibility of using two starred tones (L\*H\*) is left open. Their proposal for two starred tones is based on the fact that both L and H are aligned relative to the stressed syllable despite being realised outside its boundaries.

Similar controversial views have been argued for Spanish prenuclear rises. Some studies on Spanish intonation (Prieto et al 1994, 1995, Prieto 1999 and Nibert 2000), have classified prenuclear rises as H\*. Phonetically, these accents present two characteristics: 1) a rise in F0 that takes place during the accented syllable and 2) a displaced peak where the F0 maximum of prenuclear H\*s is not located within the accented syllable but after it. Prieto et al have shown that the amount of peak delay is subject to factors such as segmental duration or an upcoming prosodic boundary. On

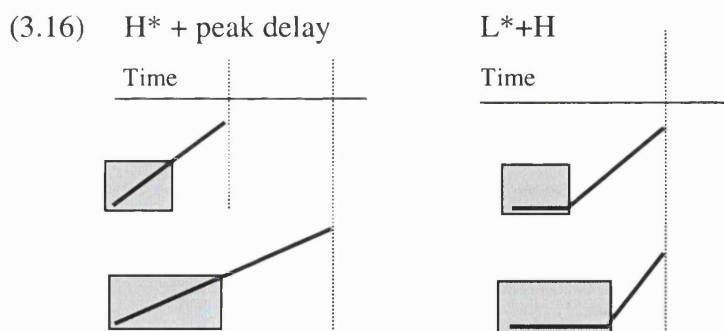
the other hand, Sosa (1999) and Face (to appear) have postulated a different interpretation for the rising movement of prenuclear accents in Spanish, namely, L\*+H. This accent involves a low F0 within the accented syllable and a sharp rise after the accented syllable. More recently, Hualde (2000) finds that neither H\* nor L\*+H describe prenuclear rises in Spanish satisfactorily. He claims that H\* fails to account for the fact that the tone is low at the onset of the stressed syllable. However, he discards L\*+H on the basis that it is not contrastive with L+H\*. He proposes, like Arvaniti et al (to appear) for Greek, a pitch accent where both tones are associated with the stressed syllable (L+H)\*.

Prenuclear accents in Central Catalan declaratives produced with a neutral intonation also show a rising movement. Previous studies on Central Catalan prenuclear rises have also shown heterogeneity of interpretations. For example, Prieto (1995, 1997) classifies those rises as H\* (with a delayed peak) on an acoustic basis, since the start of the rise takes place within the accented syllable. Bonet (1984), on the other hand, interprets prenuclear accents as low tones associated to the accented syllable since she perceives the peak of the rise after the offset of the stressed syllable. In the following pages, prenuclear rises in Central Catalan declaratives will be examined in more detail. Three possible analyses will be taken into consideration: 1) an H\* pitch accent with peak delay (as in Prieto 1995, 1997), 2) a bitonal accent (L\*+H), and 3) an L\* pitch accent followed by an H edge tone. The option L+H\* will not be examined because we assume that, with respect to the properties of the starred tone (i.e. with respect to the location of the peak and the onset of the rise), the same behaviour is expected for H\* and L+H\*. Also the possibility of using two starred tones (L\*H\*) is not considered because the data will provide evidence for one of the aforementioned analyses. First, the option L\*+H vs H\* is examined and finally the L\*H analysis is considered.

Theoretically, L\*+H differs from H\* with a peak delay in two aspects: 1) the location of the onset of the rise, and 2) the location of the peak. Whereas in H\* accents the rise starts at the onset of the accented syllable, in L\*+H accents the rising movement begins after the accented syllable. The two kinds of rising contours are illustrated in (3.15), where the straight line represents the expected contour for H\* with peak delay and the dotted line shows the contour for L\*+H.



The second aspect that differentiates H\* with peak delay and L\*+H is the location of the peak. In H\* the placement of the peak is partly predicted as a function of the segmental duration of the rhyme or syllable (Steele and Altom 1986, Silverman and Pierrehumbert 1990, Prieto et al 1995) among other factors. In the bitonal accent, on the other hand, the trailing tone is expected to be located at a fixed distance with respect to the starred tone (Grice 1995b, Arvaniti et al to appear). This is schematised in (3.16) below. For L\*+H, we assume that the starred tone is located at the onset of the accented syllable.



Our data suggested that: 1) there is variation in the location of the start of the rise, and 2) the location of the peak seems to be neither dependent on the syllable duration nor placed at a fixed distance with respect to the accented syllable. In fact, the peak seems to be located near the end of the word. These two aspects will be examined in more detail in the following sections.

### **3.4.2.2.1 Start of the rise**

For all observations, the only aspect that presents regularity among speakers is the low F0 at the onset of the accented syllable. However, the location of the rise with respect to this low onset shows inter- and intra-speaker variability. Whereas sometimes the rise takes place within the accented syllable, at other times the accented syllable is low and the rise occurs over the postaccent syllable. This is illustrated in Figure 3.3, which presents the mean F0 value at the onset and offset of the accented syllable and at the peak for all speakers. The graphs show that whereas for speakers CP and NM the rise starts after the syllable offset, for the other speakers the F0 starts to rise during the accented syllable.

The question that arises from these observations is whether the two phonetic patterns correspond to two phonologically contrastive units (as H\* and L\*+H respectively) or whether they can be interpreted as allotonic variations of the same category. In languages, such as English (Pierrehumbert 1980), Dutch (Ladd 1992) and Portuguese (Frota 1998) contrastive differences between prenuclear H\* and L\*+H have been observed.



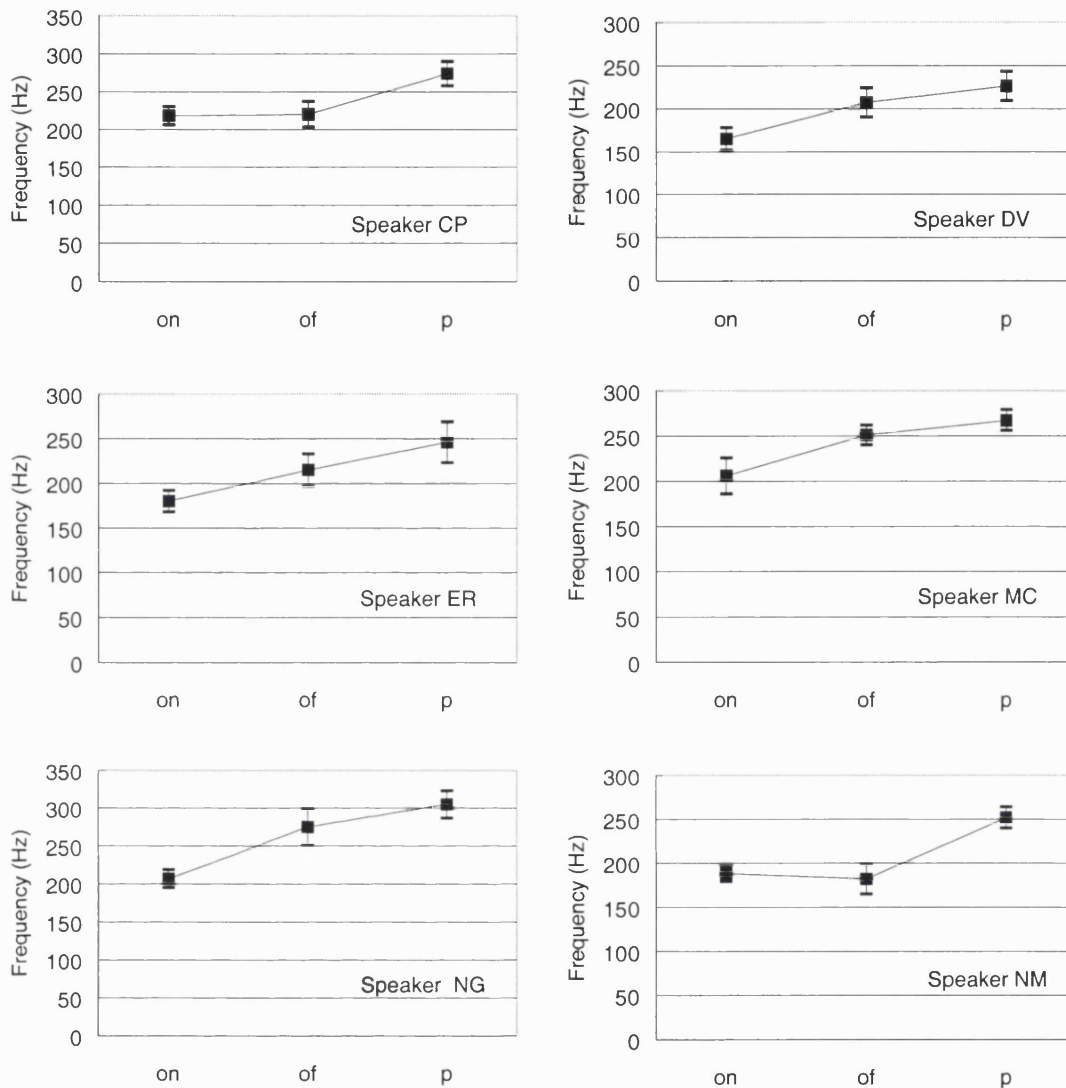


Figure 3.3. Mean F0 values at the onset (on) and offset (of) of prenuclear accented syllables and at the peak (p) for all speakers. The lines below and above the squares indicate the mean values less and plus the standard deviation respectively.

In order to see whether in Central Catalan prenuclear accents show a linguistic contrast depending on the location of the rise, a perception test was performed for two listeners. The test included sentences where the prenuclear rise started within the accented syllable, as in the sketched contour of (3.17a), as well as sentences where the rise occurred after the accented syllable, as in (3.17b). Each box represents a syllable. The shaded box is the accented syllable. Finally, a third group of structures was also included, where prenuclear rises were preceded by a high F0 before the low accented syllable, as in (3.17c). This pattern was only observed in some of the productions of speaker NM. However, it was included in the perception test since impressionistically

it was felt to be linguistically relevant. Overall, 30 sentences from the corpus of data were analysed (ten for each group). For each group, the same sentences were selected which differed only in the phonetic characteristics of the prenuclear rises but not in segmental structure. From the whole corpus, the 10 sentences of group (3.17a) were produced by speaker NG, the 10 of group (3.17b) by speaker CP and finally, the 10 of group (3.17c) by speaker NM. Sentences were presented in pairs where all possible combinations between groups were included. Two speakers listened to the tape and were asked to say whether they could perceive any difference in meaning or nuance in the two productions.

(3.17)



The results of the perception test show that the phonetic differences observed between (3.17a) and (3.17b) were not perceived as contrastive. Thus, both patterns can be classified as phonetic realisations of the same phonological entity. On the other hand, a linguistic difference was perceived between (a/b) and (c). The two listeners perceived pattern (c) as more marked and insistent. This indicates that in prenuclear position in Central Catalan there is no contrast between  $H^*$  and  $L^*+H$  but the contrast seems to be between  $L^*+H$  and something like  $H\% L^*+H$ . At this point, the classification of initial rises as  $L^*+H$  rather than  $H^*$  is due to the fact that both patterns (a) and (b) can be described as realisations of  $L^*+H$  but pattern (b) cannot be interpreted as a realisation of  $H^*$ . However, in the next section it will be argued that the  $L^*+H$  entity does not entirely account for Central Catalan prenuclear rises.

#### 3.4.2.2.2 Location of the peak

One of the characteristics of Central Catalan prenuclear accents is that the peak of the rise is neither placed at a fixed distance after the starred tone, as it would be in the case of an  $L^*+H$  pitch accent (Grice 1995b, Arvaniti et al to appear), nor dependent on the syllable duration, as expected in  $H^*$  (Steele and Altom 1986, Silverman and Pierrehumbert 1990, Prieto et 1995). In order to confirm these claims empirically, the

following measurements were obtained for all prenuclear accents of the six informants<sup>12</sup>.

1. Distance between the onset of the accented syllable and the location of the peak (on-peak).
2. Distance between the onset and the offset of the accented syllable, i.e. duration of the accented syllable (on-of).

If Central Catalan prenuclear accents corresponded to L\*+H, no correlation between the two variables would be expected, that is, L\*+H would involve a fairly fixed on-peak distance indicating that the location of the peak is unchanging, no matter what the duration of the syllable. Alternatively, if prenuclear accents were characterised as H\* with a delayed peak, a strong correlation between the two variables would be expected. This would indicate that the location of the peak could be explained as a function of syllable duration, that is, the longer the duration of the accented syllable, the longer the delay of the peak.

In order to observe the relation between the two variables, correlations were performed for the measurements of all speakers. The results are presented in Figure 3.4. The graphs plot the distance between the onset of the accented syllable and the location of the peak (on-peak) against the duration of the syllable (on-of) in seconds for all prenuclear accents of all speakers. The regression lines summarise the correlation between the two variables. Correlation coefficients are displayed in each graph.

---

<sup>12</sup> In this study, we assume that the starred tone is located at the onset of the accented syllable.

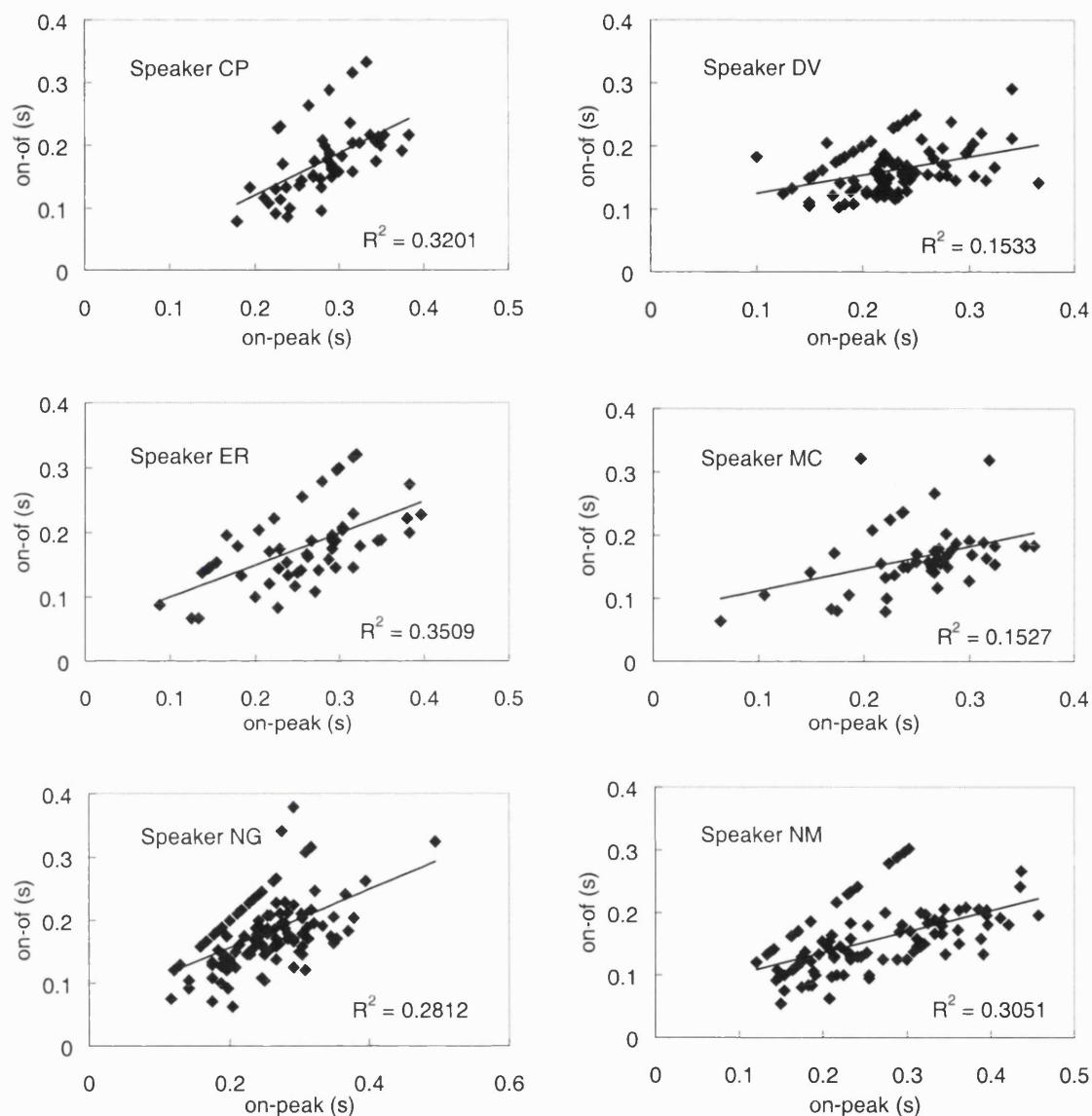


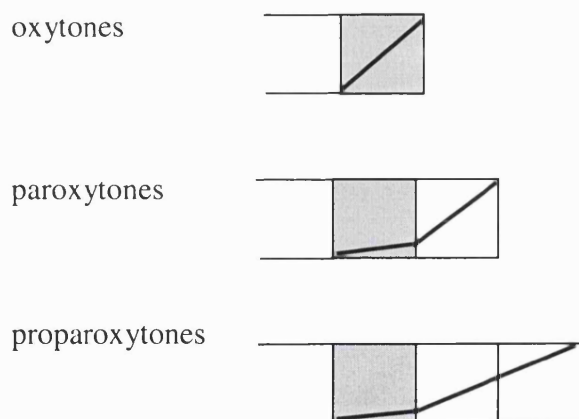
Figure 3.4. Distance between the onset of the accented syllable and the peak (on-peak) against distance between the onset and the offset of the accented syllable (on-of) in seconds for all prenuclear accents of all speakers.

The results displayed in Figure 3.4 favour neither the interpretation of Central Catalan prenuclear accents as  $L^*+H$ , nor that of  $H^*$  with peak delay. On the one hand, the data show a low correlation between the location of the peak with respect to the onset of the accented syllable and the duration of the accented syllable. In all cases,  $R^2$  values are very low, ranging from 0.15 to 0.35. This indicates that the placement of the peak cannot be explained as a function of syllable duration. Thus, the option of characterising prenuclear accents as  $H^*$  with peak delay is not supported by the evidence above. However, the alternative of describing prenuclear accents as  $L^*+H$  is

not entirely sustained either, since the on-peak distance is not as fixed as it would be expected by the bitonal accent. In Figure 3.4, two clusters of points can be spotted in each graph. At the top of each representation, a well-defined line of dots is observed, indicating that sometimes the distances “on-of” and “on-peak” are clearly correlated. Looking at the F0 contours in more detail, cases where a clear correlation between the two variables were observed corresponded to words with stress on the last syllable. This suggests that maybe the peak of the rise is anchored at a word boundary.

This possibility is considered in more detail in the next paragraphs with the following predictions: if the H signals the right edge of words, the peak will be anchored at the end of the word no matter how many unstressed syllables appear after the accented syllable. This is schematised in (3.18) below for words containing different stress distributions. In oxytones (words with stress on the last syllable), the peak coincides with the end of the accented syllable, which is the last syllable of the word. In paroxytones (words with stress on the penultimate syllable), the peak is located at the offset of the poststressed syllable, and in proparoxytones (words with stress on the antepenultimate syllable) at the offset of the second poststressed syllable.

(3.18)



The possibility that the H is located at the end of the word was tested empirically by taking the following measurements from all prenuclear accents.

1. Distance between the onset of the accented syllable and the location of the peak (on-peak).
2. Distance between the onset of the accented syllable and the end of the word (on-end).

For the analysis of the data, prenuclear accents followed by an ip boundary were separated from those *not* followed by an ip boundary. This was done to control boundary effects since the presence of an ip immediately triggers the alignment of the edge tone at the end of the word. Thus, in order to examine the location of the peak in prenuclear pitch accents *not* followed by an ip boundary, the following pitch accents were excluded from the data: 1) the second pitch accent of sentences with a double stressed subject (for all speakers except for MC), 2) the first pitch accent of sentences with a single stressed subject for speakers ER and NM, and 3) the first accent of double stressed subjects for all informants. This last group of accents was analysed separately because the intonational pattern of pitch accents presented fluctuation among speakers (i.e. some speakers produced a rise and others a fall). This is examined in more detail later on in this section.

Once the measurements for all prenuclear pitch accents were obtained, the results were divided into two groups according to the stress distribution of the words available in the initial corpus: i.e. oxytones and paroxytones. Since proparoxytones were not included in the corpus, a few sentences with proparoxytones were gathered later on for one speaker to confirm the findings on the other stress patterns. First, the results on paroxytones and oxytones will be presented for all speakers and then the results of proparoxytones will be analysed.

The results of the measurements are presented in Figures 3.5-3.6 for paroxytones and oxytones respectively. The graphs plot the distance between the onset of the accented syllable and the location of the peak (on-peak) against the distance between the onset of the accented syllable and the end of the word (on-end). The regression lines summarise the correlations observed between the two variables.  $R^2$  values are included in each graph.

## PAROXYTONES

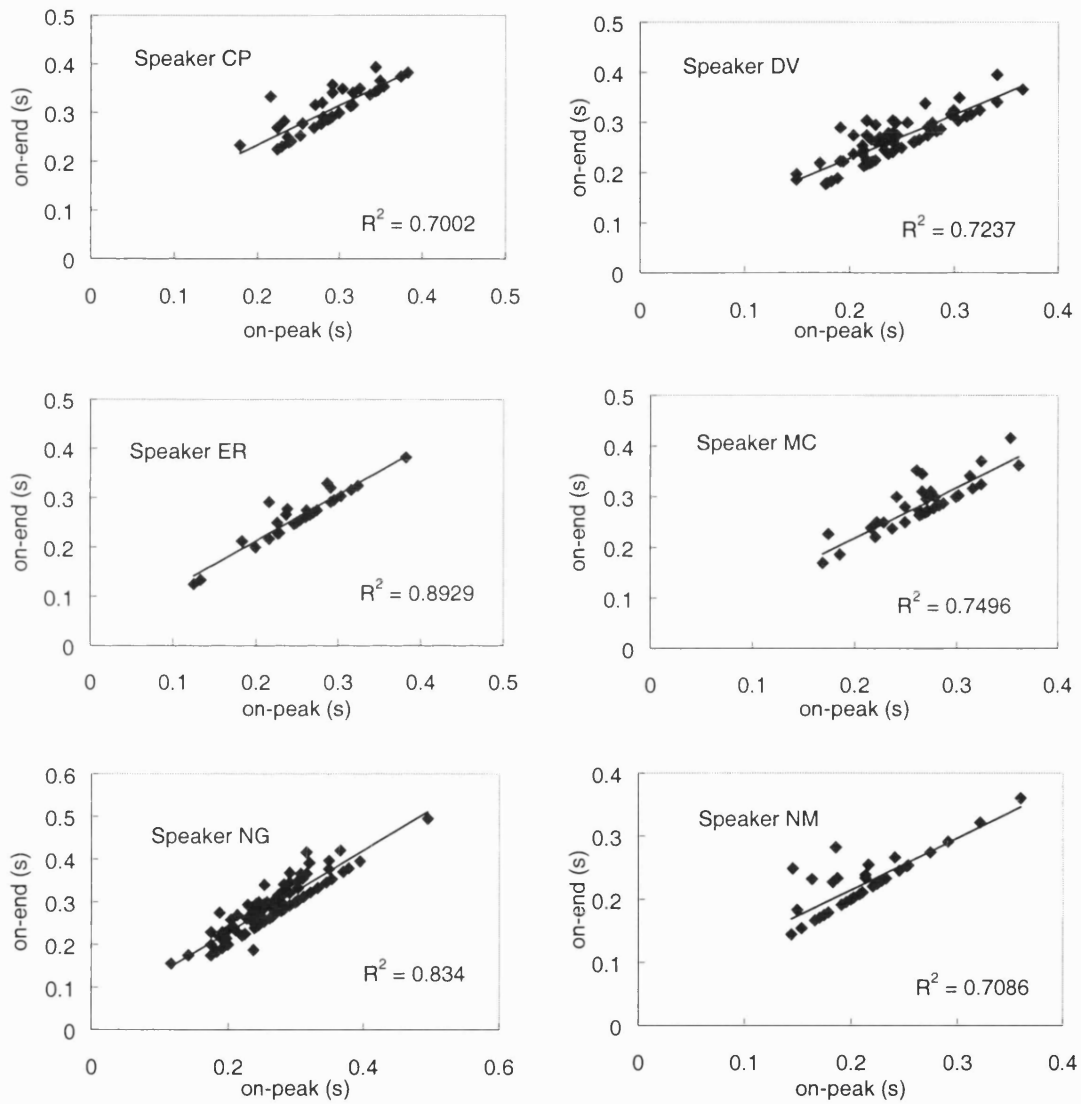


Figure 3.5. Distance between the onset of the accented syllable and the peak (on-peak) against distance between the onset of the accented syllable and the end of the word (on-end) in seconds for all speakers in paroxytones.

## OXYTONES

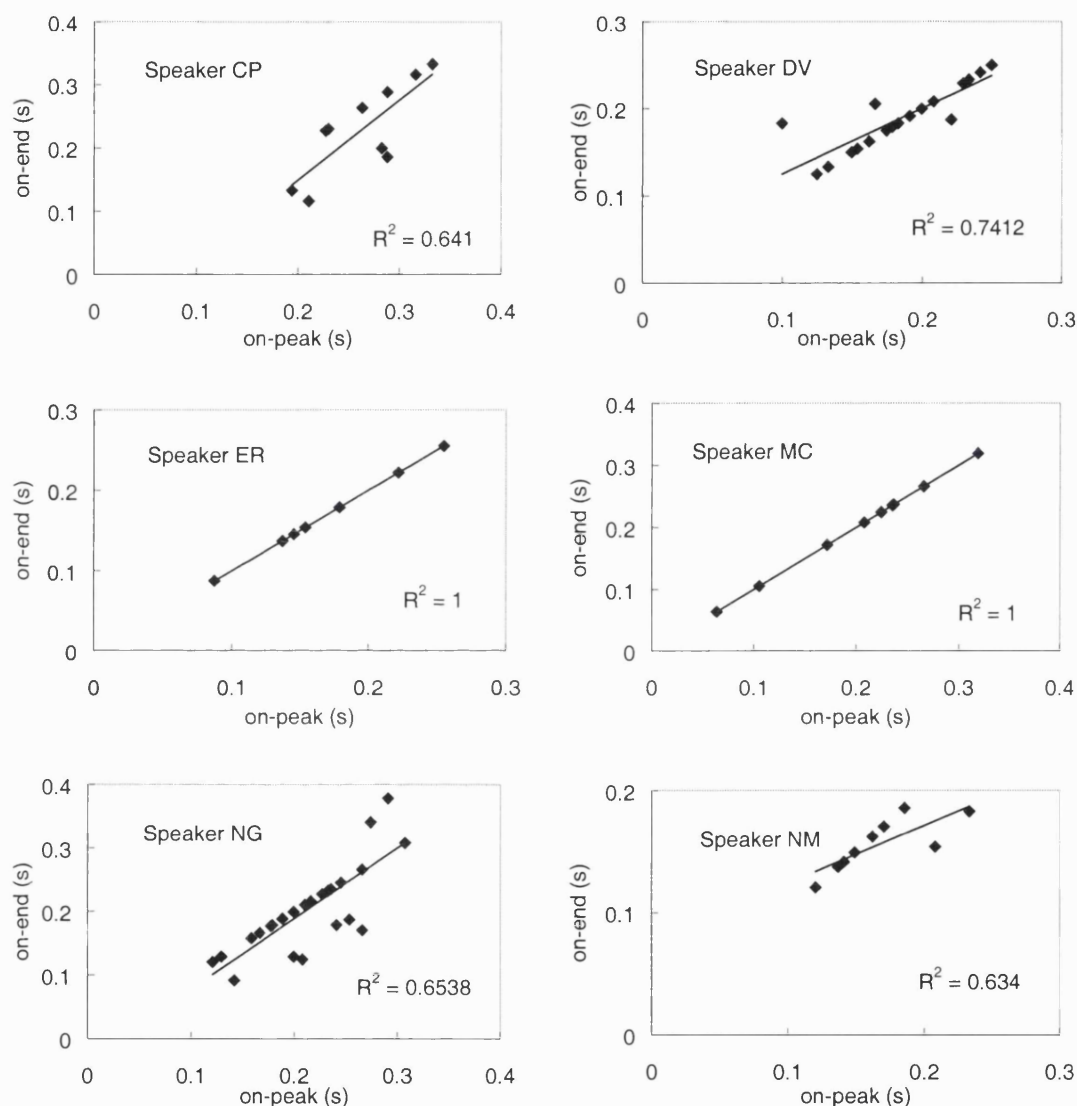


Figure 3.6. Distance between the onset of the accented syllable and the peak (on-peak) against distance between the onset of the accented syllable and the end of the word (on-end) in seconds for all speakers in oxytones.

The results of Central Catalan prenuclear accents in neutral declaratives (Figures 3.5-3.6) show that for all speakers there is a strong correlation between the location of the peak with respect to the onset of the accented syllable and the distance between this onset and the end of the word. This is true for both paroxytones ( $R^2$  values vary between 0.7 and 0.89) and oxytones ( $R^2$  values vary between 0.63 and 1). This indicates that the placement of peaks in Central Catalan prenuclear rises seems to be



anchored at the end of the word, irrespective of the number of poststressed syllables. These results show that in Central Catalan the location of the peak in prenuclear rises can be predicted as a function of word boundary.

Although the results presented so far for oxytones and paroxytones seem consistent enough to conclude that prenuclear peaks in Central Catalan are anchored at the end of the word, it was considered necessary to confirm these observations with the analysis of proparoxytones, that is, words which have two poststressed syllables. Arvaniti et al (to appear) found that proparoxytones behave differently from oxytones and paroxytones in Greek. In proparoxytones, the alignment of the H is relative to the accented syllable and not to the following word-edge. Thus, a few sentences were designed which included proparoxytones both in non-final and in final sentence positions (final proparoxytones were included for future analysis of final accents). Overall, 44 sentences were designed: 16 with proparoxytones in initial sentence position (subject), 12 in medial position (verb), and 16 in final position (object)<sup>13</sup>. Sentences are presented in Appendix 2 (section 2.2). Also a few utterances were included with exactly the same segmental material expect for the first word, which consisted of items with the different stress patterns: oxytones, paroxytones and proparoxytones. These items were included to have a few examples where the only source of variation was the stress pattern of the word. Sentences are listed in Appendix 2 (section 2.3). All sentences were produced by one speaker (DV), who was recorded following the same procedures as described in chapter 2 (section 2.3.1.4). This informant was chosen for two reasons: 1) the good quality of the F0 traces obtained in her previous recordings, and 2) the fact that she produced most of her neutral sentences with one ip.

The results are presented in Figure 3.7. As before, the graph plots the distance between the onset of the accented syllable and the peak (on-peak) against the distance between the onset of the accented syllable and the end of the word (on-end) for speaker DV in proparoxytones (both in subject and verb positions).

---

<sup>13</sup> Less sentences with verb proparoxytones were gathered because it is difficult to find verb forms with stress on the antepenultimate syllable in Catalan.

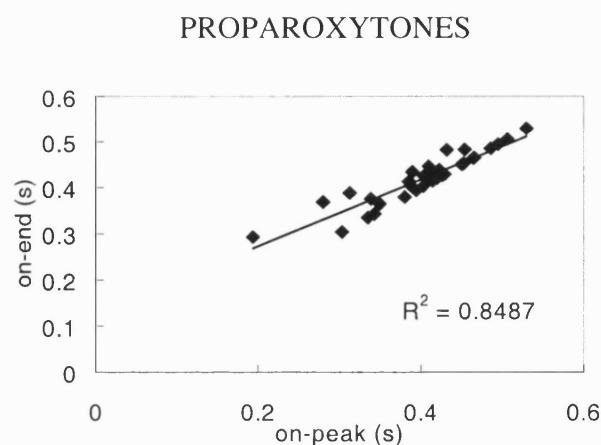
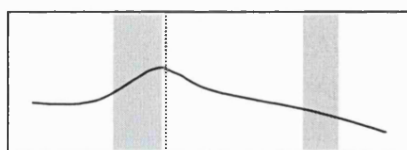


Figure 3.7. Distance between the onset of the accented syllable and the peak (on-peak) against distance between the onset of the accented syllable and the end of the word (on-end) in seconds for speaker DV in proparoxytones.

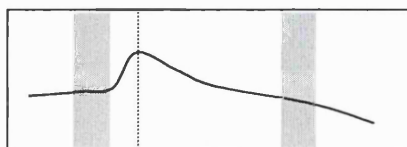
Again the results show a strong correlation ( $R^2 = 0.85$ ) between the location of the peak and the location of the word boundary in proparoxytones. These results confirm the initial observations that the peak of prenuclear rises is anchored towards the end of the word no matter how many syllables follow the accented one. An example of the consistent alignment of the H at the end of words is provided in (3.19-3.21) below for an oxytone, a paroxytone and a proparoxytone in sentence initial position. The end of the word is marked by the dotted line. See Appendix 5 for the original F0 traces.

(3.19)



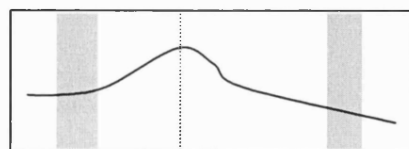
En Ramon mira les novel.les  
 “Ramon is looking at the novels”

(3.20)



La Mila mira les novel.les

(3.21)

La Melanie mira les novel.les

Finally, even though the peak of Central Catalan prenuclear rises seems to be anchored relative to the end of the word, sometimes its location is not precisely at the word edge but at a few milliseconds before or after it, as observed in Figures 3.5-3.7. In order to see where exactly the peak was anchored with respect to the word offset, the mean distance between “on-peak” was subtracted from the mean distance between “on-end” for all speakers in all kinds of words. The results are presented in Table 3.1.

	CP	DV	ER	MC	NG	NM
oxytones	-0.034	0.022	0	0	-0.010	-0.012
paroxytones	0.018	0.023	0.009	0.021	0.026	0.013
proparoxytones	-	0.018	-	-	-	-

Table 3.1. Difference between the mean distance “on-end” and the mean distance “on-peak” in seconds.

Except for the oxytones of speakers CP, NG and NM, the results show that the peak is either located exactly at the word boundary (0 seconds) or bit earlier (ranging from 0.009 to 0.26 seconds) but always within the last syllable of the word. In the oxytones of speakers CP, NG and NM, the peak is after the end of the word (ranging from -0.010 to -0.034). This unexpected behaviour might suggest that the peak is located with respect to a foot boundary rather than a word boundary. The location of the peak as a function of the foot has been explored in House et al (1999a, b) and Dankovičová and House (2000), who showed that the alignment of peaks in British English was better explained in terms of foot duration than in terms of syllable duration. That is, relating the peak location to the duration of the foot reduced variability in peak alignment. Given the results observed in Central Catalan where H values are anchored at the end of words, the idea of a foot related peak location is rather incompatible with the previous observations. However, since the data of speakers CP, NG and NM showed that in oxytones the peak could be anchored after the word offset, some sentences were analysed according to the peak/foot hypothesis for speaker NG. This

speaker was chosen because she presented the biggest sample of accented oxytones. The structures we analysed involved a few sentences with 2 syllable feet but different location of the word boundary. An example of both contexts is illustrated in (3.22). The two structures consist of a foot with two syllables but with different word boundary positions, one in the middle of the foot and the other at the end of it. Word boundaries are marked with the symbol #. *s* and *w* indicate stressed and unstressed syllables. These structures were chosen because it was expected to find similar foot durations in the two contexts, despite differences in word boundary location. Providing that this is the case, if peak alignment is expressed as a function of foot duration, it will be expected to find the F0 maximum at an equivalent point of the foot duration in the two contexts. In order to test this hypothesis, the duration of the foot and the duration from the onset of the accented syllable (beginning of the foot) up to the F0 peak were measured. Overall, 18 structures were analysed.

- (3.22) a.  $\overbrace{\text{Jo \# men-ja-va}}^{\text{Foot}} \dots$   
           s    w    s  
           “I ate ...”
- b.  $\overbrace{\text{La ma-re \# men-ja}}^{\text{Foot}} \dots$   
           s    w    s  
           “The mother eats ...”

The results are presented in Table 3.2, which shows the mean values of F0 peak position (expressed as % of the foot duration), the foot duration, and the distance between the onset of the accented syllable and the peak (on-p) in seconds. The results show that, as expected, the foot duration in the two contexts is not significantly different at a 1 per cent significance level ( $p=0.02$ ,  $t=2.25$ ). However, the percentage of peak position with respect to the foot duration is different in the two contexts. When the word boundary coincides with the foot offset ( $[s \ w] \# \ s$ ), the alignment of the peak is near the end of the foot/word. Otherwise, the peak is much earlier. This corroborates the idea that the location of the peak in Central Catalan prenuclear rises cannot be expressed as a function of foot duration since it is anchored at (or near) the end of the word.

	[s # w] s	[s w] # s
F0 peak position (%)	58.58	89.3
Foot duration (s)	0.425	0.376
on-p (s)	0.249	0.336

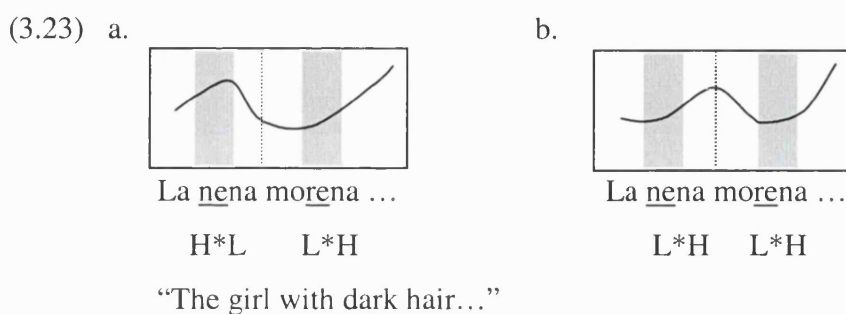
Table 3.2. Mean values of F0 peak position (expressed as % of the foot duration), foot duration, and distance between the onset of the accented syllable and the peak (on-p) in seconds.

These results mean that the postword peak observed in oxytones for speakers CP, NG and NM cannot be explained as a foot effect but requires an alternative interpretation. Looking at the results of these three speakers in more detail, it was observed that all cases where the peak was placed after the end of the oxytone, the prenuclear accent involved was always located immediately before the nuclear one. That is, the postword peak displacement was only observed on the last prenuclear accent of the utterance. Since the end of Central Catalan declaratives is characterised by a falling pitch, one possible explanation for this postword boundary peak in prenuclear oxytones is to make sure that there is a clear rise to enable the production of the nuclear fall. That is, maybe in proparoxytones and in paroxytones the final prenuclear rise can become steep enough within the limits of the word to ensure the final fall. However, in oxytones the space for the rise is much smaller and maybe has to exceed the word boundary to be high enough to prepare the final fall. This is only a speculation and further research is needed on this topic.

Apart from these cases, the data show that the peaks of prenuclear rises in Central Catalan are clearly anchored within the limits of the last syllable of the word, irrespective of the number of postaccent syllables. This behaviour seems to indicate that prenuclear rising accents in Central Catalan cannot be accounted for by means of a single phonological category (e.g. L\*+H) but need to be described as two phonological entities associated with different points in the prosodic structure, namely, 1) an L\* pitch accent associated with the accented syllable and 2) a high edge tone (H) associated with the end of the accented word.

The fact that in Central Catalan tonal events are located at the end of an accented word is confirmed by the behaviour observed in some of the first accents of a double stressed subject. The pitch patterns of these accents presented variability among

speakers. Although two speakers (NM and ER) produced a rising movement as the one described for the prenuclear accents analysed so far (i.e. L\* H), three speakers (CP, DV and NG) showed a falling movement. This trajectory involves a smooth rise within the accented syllable with the peak aligned at or near the end of this accented syllable followed by a fall afterwards (i.e. H\*L). As will be observed in the next section and also in the next chapter, this pattern is typical of Central Catalan falling accents. The two patterns are illustrated in (3.23a) and (3.23b) for speakers NG and NM respectively. The dotted line marks the end of the word.



In order to see whether the two pitch patterns presented in (3.23) were recognised as contrastive or were realisations of the same tone, two perception tests were carried out involving one Central Catalan subject. The first test consisted in playing 16 sets of sentences containing the same structure twice but produced with a different pitch pattern (rise vs fall) on the prenuclear accent of the double accented subject (sentences were produced by speaker NG for the fall and NM for the rise). The informant was asked to say whether she could perceive any differences in the meaning of the two utterances. The informant consistently reported that whenever the prenuclear accent was falling, she perceived a more neutral and less involved meaning. On the other hand, when the prenuclear accent was rising, the first element was perceived as being more insistent. The second test consisted in playing only the first phonological word of the double stressed subject (i.e. the first content word with any attached function word as in *la nena*) and ignoring the rest of the sentence. Again words with a rise were alternated with words with a fall. As with the first test, the speaker was asked to say whether she could perceive semantic differences between the two words. The informant claimed that whenever the word was produced with a fall, the meaning was the same as if the word was said in isolation (neutral declarative). When it was produced with a rise, the speaker said that a meaning of non-finality was involved and

some other words were expected to come. These judgements suggest that the two patterns illustrated in (3.23) are contrastive.

Given the fact that the prenuclear fall in double accented subject was perceived as a different tonal entity, the alignment properties of this falling accent were examined in more detail. The results showed that like the peak in prenuclear rises, the valley of the falling movement was anchored at the end of the word. This is illustrated in Figure 3.8 for the data of the three informants that produced prenuclear falls in the first accent of a double accented subject. Figure 3.8 shows the correlations between the onset of the accented word and the valley (“on-valley”) and the onset of the accented word and the end of the word (“on-end”). These data only consisted of paroxytones.

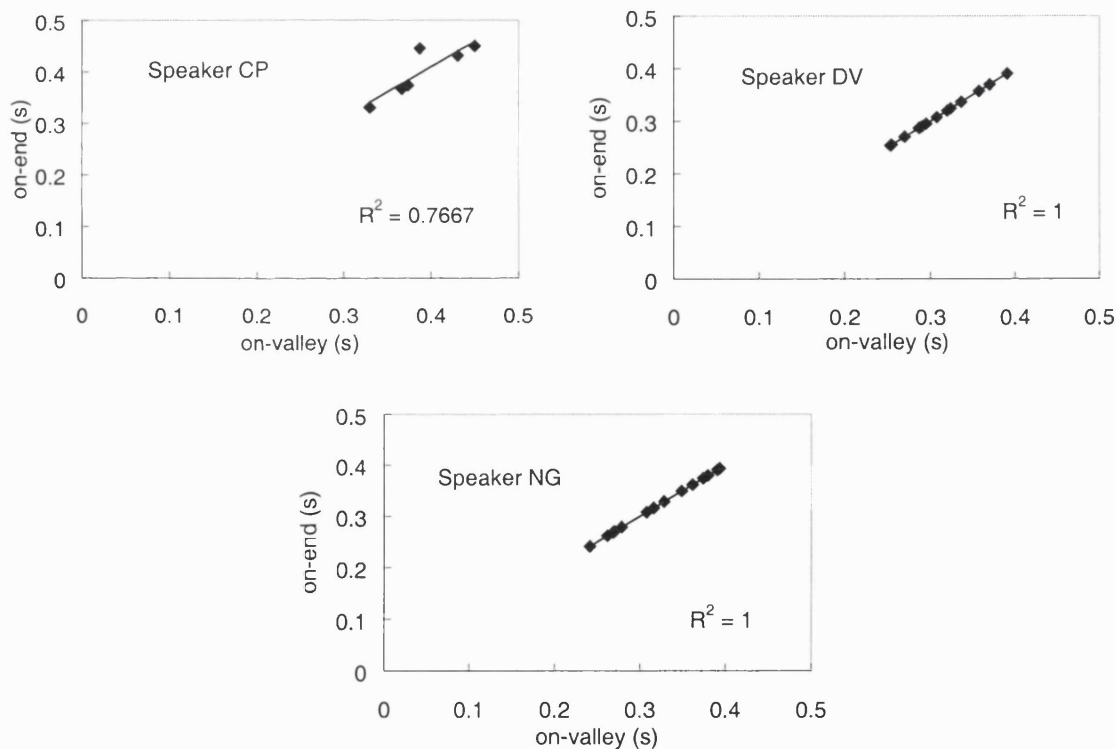


Figure 3.8. Distance between the onset of the accented syllable and the F0 valley (on-valley) against distance between the onset of the accented syllable and the end of the word (on-end) for the first accent of a double stressed subject produced with a falling pitch.

Even though the data only included paroxytones and thus words with other stress patterns need to be investigated, the results so far show a quite stable alignment of the valley at the end of the word. Thus, whereas rising accents were categorised as L\* H, falling accents seem to be described as H\* L. In both cases the starred tone is

associated with an accented syllable and the non-starred one is associated with the right edge of the accented word.

Given the consistent alignment of tones at the right edge of words observed in the data, there seems to be no need in Central Catalan for the existence of bitonal accents, at least for those that have a trailing tone. Thus, the pitch movements that in other Romance languages have been described as L\*+H and H\*+L, in Central Catalan are better interpreted as L\*H and H\*L as far as the alignment evidence is concerned. In these tone sequences, H and L stand for word edge tones, whose docking site is the right edge of the word.

The immediate implication for phonological structure is that a word level of prosodic hierarchy seems to be relevant for tonal anchoring in Central Catalan. Even though more research is needed on this topic, this prosodic domain seems to have similarities with the domain of the *prosodic word* proposed in Prosodic Phonology (Selkirk 1984, Nespor and Vogel 1986). In general terms, this domain is located below the Phonological Phrase and involves the presence of a content word plus left- and/or right-attached function words. In Prosodic Phonology, prosodic domains are basically defined in terms of the application of sandhi processes. On the basis of the data presented in this study, we might suggest that, in Central Catalan, intonation also plays an important role in defining prosodic constituents. The integration of a syntactic based approach and an intonation based approach to define prosodic structure has already been suggested for Bengali (Hayes and Lahiri 1991) and Portuguese (Frota 1998). In Central Catalan, three levels of prosodic structure seem to be defined on the basis of intonational features: 1) an IP (delimited by a boundary tone), 2) an ip (delimited by a phrase accent), and 3) a prosodic word domain (marked by the presence of a word edge tone). The correspondence between intonationally defined domains and domains identified on the basis of sandhi evidence is a matter of further research.

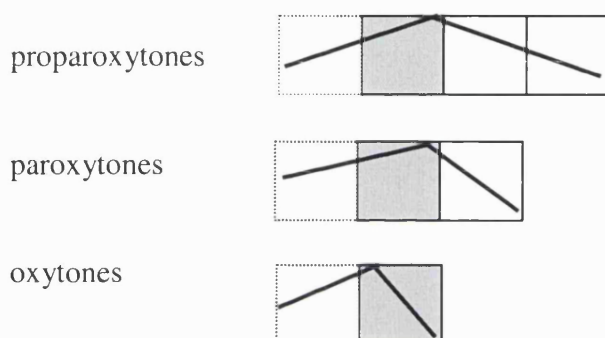
#### 3.4.2.3 Sentence-final accent (nuclear accent)

As observed in section 3.4.1, the phonetic properties of the final pitch accent in Central Catalan neutral declaratives differ depending on the number of accents in the



utterance. If the utterance consists of one stressed item, a smooth F0 rise is observed within the accented syllable followed by a fall. In isolated proparoxytones and paroxytones, the peak of the rise is aligned at or near the offset of the accented syllable and the fall takes place over the poststressed syllables. In oxytones, the rise/fall movement is compressed within the limits of the accented syllable. This is illustrated in (3.24). Each box represents a syllable. The shaded box is the accented syllable and the dotted box indicates that this syllable is optional.

## (3.24) Single-stressed utterances



In order to confirm these observations, the following measurements were taken for the single-stressed utterances of one informant (speaker CP) in the three kinds of words<sup>14</sup>:

1. Distance between the onset and the offset of the accented syllable (on-of), i.e. duration of the accented syllable.
2. Distance between the onset of the accented syllable and the peak (on-peak).

Table 3.3 shows the mean value of F0 peak position, expressed as % of the duration of the accented syllable, as well as the mean syllable duration (on-of) and the mean distance between the onset of the accented syllable and the peak in seconds.

<sup>14</sup> This speaker was chosen because she produced late narrow focus sentences with an emphatic accent, which will be compared to single-accented patterns later in this section.

	Oxytones	Paroxytones	Proparoxytones
F0 peak position (%)	50	87	88.9
on-of (s)	0.251	0.2	0.181
on-p (s)	0.127	0.174	0.161

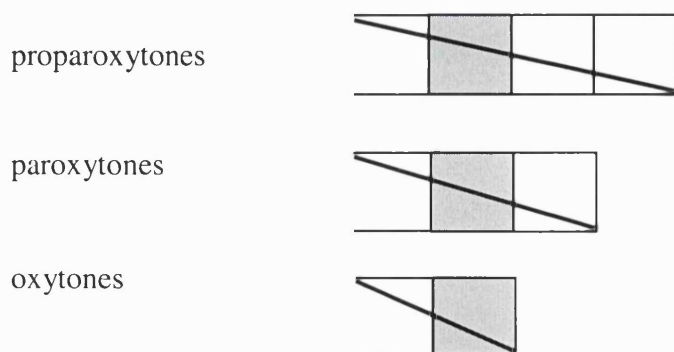
Table 3.3. Mean values of F0 peak position (expressed as % of the syllable duration), syllable duration (on-of) and distance between the onset of the accented syllable and the peak (on-p) in seconds for words produced in isolation.

The results provided in Table 3.3 show that the peak in single-stressed utterances tends to be located at the end of the accented syllable in proparoxytones (at 88.9% of the accented syllable) and in paroxytones (at 87% of the accented syllable). In oxytones, the peak is located in the middle of the accented syllable (50%) to be able to produce the fall. Thus, the whole rise/fall gesture compresses into one syllable.

In multi-stressed sentences, the phonetic characteristics of the final accent of neutral declaratives differ from single-stressed productions. The final accent of multi-stressed sentences exhibits a fall in the F0 contour. However, this fall not only occurs within the limits of the last accented syllable but the falling movement usually starts after the last prenuclear accent, and progressively falls till the end of the utterance. Thus, the fall seems to expand or accommodate within the space between the last non-final accent and the end of the sentence, no matter how many intervening syllables there are. This involves no drastic pitch movement within or near the limits of the accented syllable and thus the same pitch trajectory is observed in all kinds of stress distributions, as illustrated in (3.25). The open box indicates that there might be a variable number of preceding syllables.

(3.25)

Final accent in multi-stressed utterances



The lack of an obvious F0 prominence within or near the limits of the last stressed syllable in multi-stressed sentences poses the question of whether this syllable is actually accented. In fact, the falling F0 movement observed in neutral declaratives could be explained as an interpolation between the last H word edge tone and the L-L% phrase edge tones. However, this option is discarded on two grounds. First, the non-accentuation of the last lexical item would involve main prominence on a non-final lexical word and this would trigger a narrow focus reading (see chapter 4 for more details in narrow focus). Second, the lack of an F0 movement on the last stressed syllable of a neutral declarative occurs not only in utterances produced with one ip but also in utterances where an ip boundary is produced between the subject and the predicate. In this case, no obvious F0 movement is observed within the second ip, unless non-final stresses become accented. Despite the lack of F0 prominence, a pitch accent has to be postulated in theoretical terms since an ip domain is defined by the presence of at least a pitch accent (Beckman and Pierrehumbert 1986, Ladd 1996). Given a broad focus reading, this pitch accent must be associated to the stressed syllable of the last lexical word. As will be analysed in the following paragraphs and in chapter 5, the lack of pitch prominence in the last accent of neutral declaratives is the result of an F0 downtrend, namely, final lowering.

In the next sections, we will discuss the phonological interpretation of the last pitch accent in single and multi-stressed utterances. In previous works on Central Catalan intonation, differences in the F0 traces of single and multi-stressed utterances have been subject to different phonological representations. Prieto (1995) proposes an H\* associated to the accented syllable of single-stressed utterances and an L\* for the last accented syllable of multi-stressed neutral declaratives. Prieto's interpretation is acoustically based. She observes a low F0 in the final accented syllable of a multi-stressed declarative pitch trace, as opposed to the peak observed in single-stressed items. Similarly, Bonet (1984) also proposed a low tone in the last accent of multi-stressed declaratives. Her interpretation is perceptually based since she perceives the last accent of a declarative as much lower than the other accents. The same L\* has been postulated for Spanish final declarative pitch accents (Sosa 1999).

One of the problems with Prieto's interpretation is that two phonological entities (L\* and H\*) are used to describe the final accent of neutral declaratives, depending on the

number of stresses per sentence. Although the phonetic evidence does show a different pattern in the F0 contour of single and multi-stressed utterances, the two patterns do not seem to trigger a contrastive opposition. That is, in both cases the meaning conveyed is that of a neutral, unmarked declarative. Thus, the two patterns are functionally equivalent and their distribution seems to depend on whether the stressed word appears in isolation or not.

In this study, the same phonological unit will be used to represent the last pitch accent of both single and multi-stressed declarative sentences and the patterns observed in (3.24) and (3.25) above will be interpreted as realisational differences of the same category. We propose that the nuclear F0 movement of Central Catalan neutral declaratives should be described as an H\* pitch accent (followed by an L word edge tone), irrespective of whether there are prenuclear accents or not. Since the presence of word edge tones was quite clear in the analysis of prenuclear accents, we assume that in the nuclear position the end of the word is also marked with a word edge tone (L). However, no relevant data for this tone can be presented in final position due to the presence of other edge tones. In the next pages, we will provide evidence that H\* is the nuclear accent of both single and multi-stressed utterances.

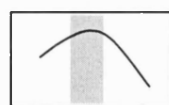
Whereas in single-stressed structures the H\* is realised with a peak in the F0 contour, in multi-stressed structures the H\* does not exhibit a peak on the final accented syllable. The lack of an F0 peak in multi-stressed structures can be explained by the general tendency to lower F0 targets over the course of an utterance. Pierrehumbert and Beckman (1988) discuss three sources of F0 downtrends at the sentence level: 1) declination (time-dependent decrease of F0 over the utterance), 2) downstep (lowering of accents beyond the effect of declination), and 3) final lowering (bigger lowering in the final peak of an intonation phrase than the lowering expected by downstep). These phenomena have been observed in many languages, such as English (Lieberman and Pierrehumbert 1984), Japanese (Pierrehumbert and Beckman 1988), German (Möbius 1993, Grabe 1998a, to appear) and Mexican Spanish (Prieto et al 1996), among others. In this study, we postulate that the last pitch accent of multi-stressed Central Catalan declaratives is a strongly lowered H\*. This allows us to use the same phonological entity (H\*) to describe the neutral intonation of single and multi-stressed utterances, while still keeping the perceptual impression that the last accent of multi-

stressed sentences is lower than the previous ones, as reported in Bonet (1984). See chapter 5 for more detailed analyses of pitch downtrends in Central Catalan.

At this stage, the proposal to use the same phonological primitive for both single and multi-stressed utterances is supported by 1) perceptual and 2) acoustic evidence and 3) comparative evidence showing that the last pitch accent of Central Catalan declaratives behaves differently from that of other Romance languages.

Perceptual evidence was based on the auditory judgements of two speakers comparing the non-lowered pitch accent observed in isolated words, as in (3.26a), with the last (strongly lowered) accent of a multi-stressed sentence, where the prenuclear portion was removed, as in (3.26b). The auditory impression of the two informants was that when heard in isolation lowered H\* accents sounded very similar to non-lowered H\* accents (produced with a low register). Similar observations are reported in Grabe (1998a) for German, where a downstepped H\*+L pitch accent appeared to be auditorily equivalent to a non-downstepped H\*+L. Thus, the distinction between non-lowered accents (in isolated words) and lowered accents (last accent of a multi-stressed sentence) in Central Catalan declaratives seems to be gradient rather than categorical. This observation throws into doubt Prieto's (1995) proposal for treating the two accents as phonologically distinct and supports the idea of analysing the two contours as gradient realisations of a single tone (H\*).

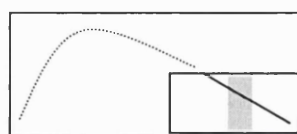
(3.26) a.



les nenes

H\*

b.



les nenes

H\*  
(+final lowering)

The second piece of evidence for classifying the final pitch accent of long declaratives as a strongly lowered H\* rather than L\* is acoustic or phonetic. The acoustic analysis of the final accent in the productions of all speakers shows no rising pitch excursion within the limits of the last accented syllable in multi-stressed utterances, as opposed to single stressed utterances. Instead the highest F0 point within the accented syllable is located right at the onset of the accented syllable as part of the progressively

descending contour. However, when the last pitch accent of multi-stressed declaratives was said with some emphasis, the same pattern as in single stressed utterances was observed<sup>15</sup>. That is, the peak was aligned near the offset of the accented syllable. If emphatic accents are treated as expanded versions of non-emphatic or neutral accents, this favours the idea that the final accent of multi-stressed declaratives is a lowered H\* which becomes non-lowered (i.e. with a peak similar to single stressed structures) when produced emphatically. This is illustrated in Table 3.4 for the productions of speaker CP said with emphasis on the final word. Only paroxytones were included in the data. As in Table 3.3, Table 3.4 presents the mean values of the following measurements: 1) F0 peak position (expressed as % of the duration of the accented syllable), 2) duration of the accented word (on-of), and 3) distance between the onset of the accented syllable and the peak (on-p). The results show that in final emphatic words the location of the peak in paroxytones is near the end of the accented syllable (at 92.3% of the accented syllable). This behaviour is very similar to the pattern observed in single-stressed utterances (see Table 3.3), where the peak of paroxytones was located at 87% of the accented syllable. This seems to corroborate the idea that the last pitch accent of a multi-stressed neutral declarative is a strongly lowered H\*.

	Paroxytones
F0 peak position (%)	92.3
on-of (s)	0.173
on-p (s)	0.157

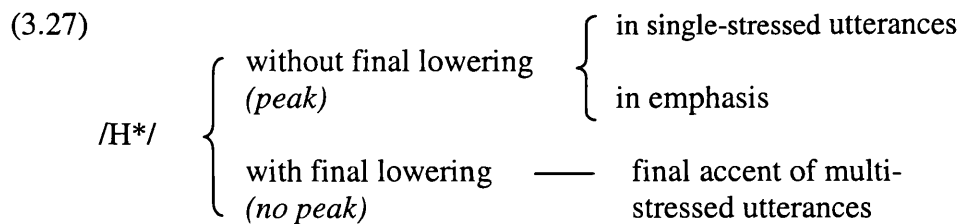
Table 3.4. Mean values of F0 peak position (expressed as % of the syllable duration), syllable duration (on-of) and distance between the onset of the accented syllable and the peak (on-p) in seconds for paroxytones produced emphatically.

Furthermore, anticipating the results obtained for narrow focus (see chapter 4), it will be observed that narrow focus is associated to an H\* pitch accent. In subject and in verb narrow focus, this accent is clearly realised as a peak aligned within the accented syllable (as in single stressed structures) and is subject to downstep effects (see chapter 5). In object narrow focus, however, the F0 trace corresponding to the H\* pitch accent presents the same behaviour as the final accent of multi-stressed broad

<sup>15</sup> Emphatic accents were obtained in some utterances with narrow focus on the object (see chapter 4).

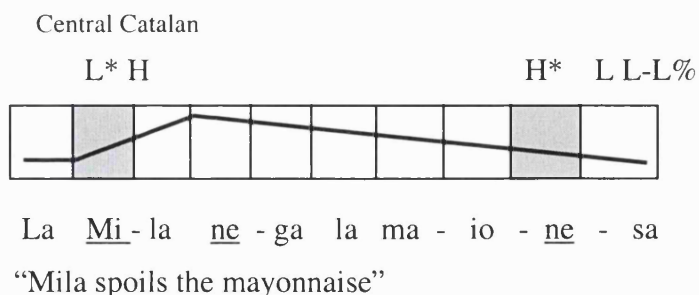
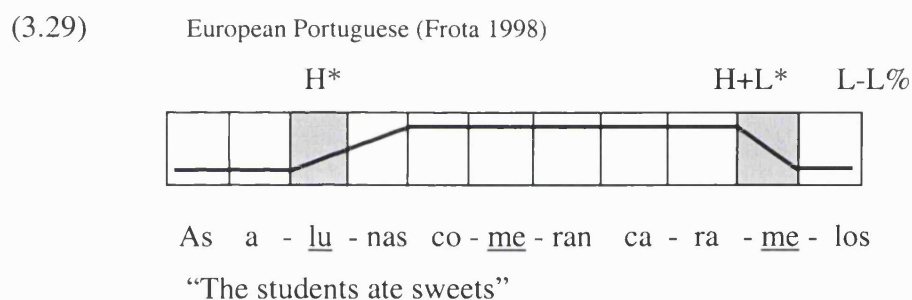
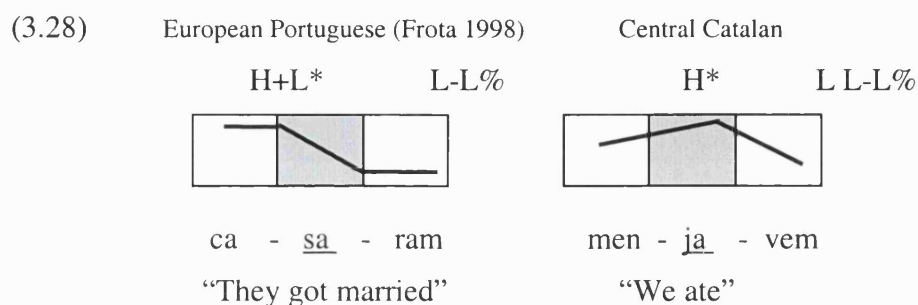
focus sentences, that is, no F0 peak is observed within the accented syllable. However, the F0 trajectory of the accented syllable is part of a progressively falling contour. Once more, the phonetic differences observed between initial or medial narrow focus and final narrow focus corroborate the idea that final H\*s are realised with final lowering, unless emphatically produced. The evidence confirms the idea that the last pitch accent of multi-stressed broad focus declaratives is a strongly lowered H\*.

Thus, rather than proposing two different phonological entities to account for the intonational behaviour of the final pitch accent in single and in multi-stressed structures, this study provides evidence for a single category (H\*) subject to realisational differences, as summarised in (3.27).



Finally, while trying to identify the best tonal category to describe the last pitch movement of Central Catalan declaratives, the possibility of using H+L\* was considered. H+L\* is the pitch accent proposed in several studies as the nuclear accent of declarative sentences in Romance languages such as Palermo Italian (Grice 1995a), Portuguese (Frota 1998) or American Spanish (Sosa 1991). This accent involves a high pitch before the accented syllable and a fall within the accented syllable. A comparison between the F0 contours presented in these studies and those obtained for Central Catalan in fairly similar sentences shows that the pitch patterns are considerably different both in single and in multi-stressed utterances. Whereas Central Catalan single-stressed utterances show a smooth rise within the accented syllable followed by a fall over the following syllables, in the aforementioned Romance languages the fall takes place within the accented syllable and is preceded by a high F0. This is illustrated in (3.28) where the schematised F0 patterns of single-stressed utterances in European Portuguese and in Central Catalan are presented. Similarly, in multi-stressed sentences, Romance languages show a drastic fall within the limits of

the accented syllable. In Central Catalan, on the other hand, the fall is not specifically on the last syllable but the F0 progressively falls from the penultimate accent onwards, as observed in (3.29). This clearly shows that the last pitch accent of neutral declaratives in Central Catalan is different from other Romance languages and the possibility of presenting the final movement as H+L\* is ruled out.



Given the evidence presented above the classification of the final pitch accent in Central Catalan declaratives is H\* irrespective of the number of accents per utterance.

### 3.5 Discussion

In this chapter, the phonetic and phonological properties of Central Catalan declaratives produced with a broad focus intonation have been examined. The results



of the acoustic analysis of the F0 traces suggested a different phonological interpretation from the representations proposed in earlier studies on Central Catalan intonation, namely, Prieto (1995, 1997) and Bonet (1984).

With respect to the *prenuclear* accents, Prieto's interpretation of non-final rises as H\* (with a peak delay) was thrown into doubt by two pieces of evidence: 1) the results of a perceptual test and 2) the alignment of F0 peaks. First, the perceptual test showed that no matter where the rising trajectory started (i.e. either within or at the offset the accented syllable), the two informants perceived no semantic differences between the two patterns. This disagrees with Prieto (1995, to appear e) who postulates a categorical distinction between H\* (with peak delay) and L\*+H in Catalan prenuclear declarative accents. In the perception test, differences were only perceived when an H% was located before the low tone. Thus, the categorical contrast was between H%L\*H and L\*H but not between H\* (with peak delay) and L\*H. Second, an acoustic analysis of the data showed that the peak of prenuclear rises was consistently anchored at the end of the word. This was confirmed for words with different stress distributions (i.e. oxytones, paroxytones and proparoxytones), where the peak of the rise was aligned at the right end of the word, irrespective of the number of unstressed syllables. These two factors suggested that the phonological representation of Central Catalan prenuclear rises involves an L\* pitch accent associated to the accented syllable and an H edge tone associated to the end of the word. The representation of prenuclear accents as L\* is consistent with the perceptual interpretation of Bonet (1984) who postulates a low tone in similar types of utterances.

The possibility of analysing prenuclear rises as L\*+H, as proposed in languages such as Greek (Arvaniti and Ladd 1995, Arvaniti and Baltazani 1999) or Spanish (Sosa 1999, Face to appear) was also taken into consideration. However, given the consistent location of H at the word boundary, the peak of prenuclear rises was considered to reflect some kind of prosodic domain rather than the trailing part of a bitonal accent. Furthermore, the peak was not aligned at a fixed position with respect to the starred tone, as it would have been predicted by bitonality. The consistent alignment of peaks at the right edge of accented words seems to strengthen the idea proposed in recent research (Arvaniti et al 1998, Ladd 1996) that turning points in the F0 contour, such as peaks or valleys, have to be relevant for phonological structure.

This means that prenuclear peaks either stand for the trailing tone of the bitonal L\*+H or represent the end of a prosodic domain, as observed in our data. This proposal questions the notion of peak delay, which involves the interpretation of F0 peaks as phonetic realisations of H\*s.

A few speakers produced a falling prenuclear accent in sentences with a double accented subject. Contrastive differences were identified between prenuclear rises, which were perceived with a more involved nuance, and falls, which were identified as more neutral. The data also showed that, like the peak of a prenuclear rise, the valley of a fall was anchored at the end of the word. Thus, falls seem to be characterised by means of the H\*L sequence, where H\* is associated with the accented syllable and L with the offset of the word that contains the pitch accent.

The phonological interpretation of the *nuclear* accent in neutral declaratives provided in this study differs from Bonet (1984) and from Prieto (1995). Both Bonet and Prieto claim that the last accent of a multi-stressed neutral declarative is low. Whereas Bonet's argument is based on perceptual evidence (i.e. the last accent is perceived as much lower than non-final accents), Prieto's conclusion is based on an acoustic analysis of the F0 contours, which show a very low F0 on the last accented syllable of neutral declaratives. However, based on the same acoustic evidence Prieto points out that if the utterance has only one stressed syllable, then the accent is high. Thus, according to Prieto, the tonal configuration H\*L-L% accounts for the final F0 movement of single-stressed neutral declaratives and the pattern L\*L-L% describes the final F0 movement of multi-stressed sentences. In this study, acoustic and perceptual analyses of the data provided evidence that the final contours of both single and multi-stressed declaratives can be interpreted in the same way, namely, as a fall (H\*L). Although it is true that the last pitch accent of multi-stressed utterances is both perceived and produced as a much lower accent than the preceding ones, this final lowering can be accounted for by a sentence level F0 downtrend. Studies on different languages (English: Liberman and Pierrehumbert 1984, Japanese: Pierrehumbert and Beckman 1988, German: Möbius 1993, and Mexican Spanish: Prieto et al 1996) have shown that the last accent of neutral declaratives tends to undergo a lowering much bigger than the one observed in non-final accents. This sentence phenomenon explains the low perceptual and acoustic properties of the last accent in multi-

accented sentences and allows to provide the same phonological interpretation for the nuclear accent of single and multi-stressed declarative sentences (H\*L). The phenomenon of final lowering and other F0 downtrends in Central Catalan are studied in more detail in chapter 5.

Finally, the data also presented evidence for two levels of prosodic phrasing above the word constituent, namely, a major domain or Intonation Phrase (IP), and a minor domain or intermediate phrase (ip). As in English (Pierrehumbert 1980), Palermo Italian (Grice 1995a) or Spanish (Sosa 1999), among many other languages, the IP in Central Catalan is defined as the highest level of prosodic structure. It is demarcated by a boundary tone (H% or L%) and also by the presence of a major pause or break at the boundary. In the data, the right boundary of an utterance produced with a neutral intonation was marked by the presence of L%, as also reported in Prieto (1995) or Estebas-Vilaplana and Maidment (1999a). No cases of final H% were found in the data since all sentences involved a declarative intonation. However, Prieto (1995) provides evidence for this tone in Central Catalan, at the end of interrogative sentences.

An ip level of prosodic structure was also observed in the productions. According to the results, the ip in Central Catalan is a level of prosodic phrase structure which is below the IP and contains a minimum of one pitch accent (followed by a word edge tone). The ip level is marked by the presence of a phrase accent (H- or L-). Evidence for this domain was found in sentences with a double stressed subject. Those utterances presented a much higher H tone at the end of the second accented word of the subject than that expected by a word edge tone. This H tone could not be accounted for by means of H% because it was not at the end of the whole IP and no major F0 break was observed after it. However, it could neither be an H word edge tone because it was not downstepped with respect to the preceding H, as observed in utterances with no double accented subject. Thus, the only way to account for this higher rise was by means of a phrase accent H-, indicating the end of an ip. The presence of this ip was corroborated by the longer duration of the syllables preceding it. Whenever present, the ip boundary was always located between the subject and the predicate. In longer subjects the ip boundary was present most of the time, confirming the idea that the weight of the constituents might influence prosodic boundaries (Ladd

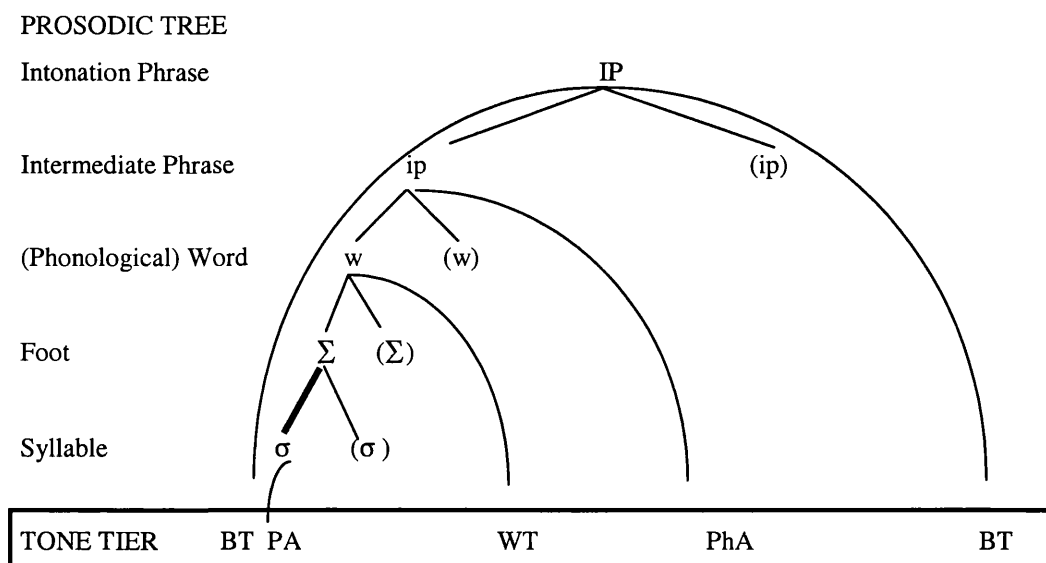
1996). The H- is also reported in Prieto (1997) and Estebas-Vilaplana and Maidment (1999a).

The presence of two levels of phrasing in intonation has been postulated several times in the literature (e.g. Trim 1959: major and minor tone groups, and O'Connor and Arnold 1973: single and double bar boundaries). Within the AM approach, an ip level of prosodic hierarchy has been postulated for English (Beckman and Pierrehumbert 1986, Pierrehumbert and Beckman 1988), Greek (Arvaniti and Baltazani 1999) or Palermo Italian (Grice 1995a). However, languages not always coincide in the levels of prosodic phrasing. In Japanese (Beckman and Pierrehumbert 1986, Pierrehumbert and Beckman 1988, Venditti 1999), in Korean (Jun 1996) and in French (Jun and Fougeron, to appear) a lower domain in the prosodic tree, called the Accentual Phrase, has been postulated<sup>16</sup>. The Accentual Phrase is a tonally defined domain. Thus, for example, in Japanese, the Accentual Phrase has at most one High tone and is delimited by a Low boundary tone. Furthermore, in Japanese and Korean, the notion of accent is different from English and Catalan. In English and Catalan a pitch accent is associated to the stressed syllable of the word. In Japanese and Korean, the “accent” is distributed among the Accentual Phrase and functions to delimit prosodic groupings of words rather than to mark prominent syllables. The possibility that the H and L tones observed at the end of accented words in Central Catalan signal an Accentual Phrase domain is ruled out on two grounds: 1) Central Catalan pitch accents are associated with stressed syllables, rather than distributed among an Accentual Phrase, and 2) there is no specific tonal pattern that can be attributed to a particular domain. Instead H and L are anchored at the end of a word and their presence is linked to the presence of a pitch accent. The consistent location of H and L at the end of accented words suggest that such tones might signal the end of a word domain, similar to the notion of *phonological word*, as proposed by Selkirk (1984) and Nespor and Vogel (1986) on phonological-syntactic grounds. Although more research is needed on this topic, a preliminary prosodic tree, based on intonational features, is proposed in (3.25) for Central Catalan. Contrary to English, Central Catalan may have an intonationally-defined phonological word, marked with the presence of a right-hand word edge tone (WT).

---

<sup>16</sup> The idea of an Accentual Phrase in French is not supported by recent work on French intonation within the AM framework (Post 2000).

(3.25)



Thus, according to the data, the phonological primitives used to describe Central Catalan intonation are:

- 1) pitch accents associated to metrically strong syllables at a postlexical level (so far L\* and H\* have been observed, although in the next chapter another pitch accent will be postulated, namely, H+!H\*)
- 2) boundary tones marking the presence of an IP (H% and L%)
- 3) phrase accents signalling an ip boundary (H- and L-)
- 4) word edge tones marking the end of an accented word (or phonological word) domain (H and L).

Given these generalisations, the phonological representation of utterances presented in (3.3-3.6) and (3.14) would be as follows<sup>17</sup>.

<sup>17</sup> If we look at the data from a purely phonetic point of view, the tonal transcriptions of (3.30-3.33) may prove problematic in the sense that at the end of sentences three tones with the same tone level (L) are used to indicate the fall. However, the three edge tones (word edge tone, phrase accent and boundary tone) are phonologically relevant in Central Catalan since they signal different levels of prosodic structure. Thus, even though this accumulation of L tones does not trigger any relevant phonetic information, we will still keep them because they are relevant for prosodic structure.

(3.30) [ [ [mejavem] ] ]

$\begin{array}{c} | \quad ||| \\ H^* \quad L L-L\% \end{array}$

(3.31) [ [ [La Mila] [nega] [la maionesa] ] ]

$\begin{array}{c} | \quad | \quad | \quad | \quad \quad | \quad ||| \\ L^* H \quad (L^* H) \quad \quad H^* L L-L\% \end{array}$

(3.32) [ [[Les meves] [amigues]] [[eren] [bones] [nenes]] ]

$\begin{array}{c} | \quad | \quad | \quad ||| \quad | \quad | \quad | \quad | \quad | \quad ||| \\ \left. \begin{array}{l} L^* \quad H \\ H^* \quad L \end{array} \right\} L^* \quad H H- \quad L^* H \quad L^* H \quad H^* L L-L\% \end{array}$

(3.33) [ [[Les noies]] [[vénen] [de veure] [l'examen]] ]

$\begin{array}{c} | \quad ||| \quad | \quad | \quad | \quad | \quad | \quad ||| \\ H^* \quad L H- L^* \quad H \quad L^* H \quad H^* \quad L L-L\% \end{array}$

### 3.6 Conclusion

In this chapter, the phonetic and phonological properties of Central Catalan declaratives produced with a broad focus intonation have been examined within the AM framework of intonational analysis. Particular attention has been devoted to three aspects: 1) the intonational behaviour of prenuclear (or non-final) accents, 2) the intonational behaviour of nuclear (or final) accents and 3) the levels of prosodic domains or phrasing within each sentence.

Similar to the descriptions provided in previous studies on Central Catalan intonation, the results obtained in this work agree in that the unmarked intonation of a declarative sentence involves a rising pitch in the vicinity of the prenuclear accented syllables followed by an F0 fall. This fall starts after the last prenuclear accent and ends at the end of the F0 contour. Not all stressed prenuclear syllables become accented, and in most cases only the first stressed syllable of the contour and the last one get an accent. In single-stressed words, the neutral declarative intonation involves a smooth F0 rise within the accented syllable followed by a fall over the poststressed syllables if any.

Even though the phonetic characteristics observed in the contours of Central Catalan neutral declaratives are pretty similar to those reported in earlier works, the phonological interpretation provided in this study differs considerably. Prenuclear accents usually show a rising movement whose peak is consistently aligned with the end of the word. This has been interpreted as L\*H, which involves an L\* pitch accent associated to the accented syllable and an H tone associated to the end of the accented word. The fall over the nuclear accent is characterised by the tonal sequence H\*L, both when there is a clear F0 peak (as in single-stressed utterances) or not (as in multi-stressed utterances). The absence of a peak in sentences with more than one accent is explained by a final F0 lowering (see chapter 5).

Above the accented word, two levels of prosodic phrase structure are observed in the data: a minor level (the ip) demarcated by the presence of a phrase accent (H- or L-) and a major level (the IP), marked by a boundary tone (H% or L%). Central Catalan neutral declaratives end with an L% boundary tone. Sometimes an ip boundary is observed between the subject and the predicate. This can be marked by a sharp rise (L\*HH-) or a fall rise (H\*LH-).

## Chapter 4: Narrow focus

### 4.1 Introduction

In the previous chapter, the phonological and phonetic properties of Central Catalan declaratives produced with a broad focus reading were examined. Neutral sentences involved an initial rise, described as L\*H, followed by a final fall, which was interpreted as H\* L. Differences in the number of prenuclear accents as well as in the number of tone units were perceived among speakers. Sometimes an ip boundary, marked with an H- phrase accent, was produced between the subject and the predicate. The end of the sentence was interpreted as L-L%. In this chapter, the F0 traces of Central Catalan declaratives produced with a narrow focus reading are analysed and compared to broad focus sentences. The narrow focus productions discussed in the following sections are of two kinds: 1) utterances where focus is expressed by intonational means, that is, utterances where the focus/accent association is attained through a reorganisation of the intonational pattern, whereas the syntactic structure remains the same, and 2) sentences where the focus/accent alliance is achieved after a syntactic alteration.

The organisation of this chapter is as follows. Section 4.2 consists of a brief introduction recalling the different ways languages convey narrow focus, with particular attention to accentual strategies. Section 4.3 describes the data used for the analysis of narrow focus in Central Catalan. Section 4.4 presents the results of sentences where focus is conveyed both by accentual means (4.4.1) and by syntactic means (4.4.2). In 4.4.1, sentences with *early* focus (focus on the subject and on the verb) are compared to sentences with *late* focus (focus on the object). In this section, issues such as the levels of phrasing and the phenomenon of postfocal accents are covered. In section 4.4.2, the intonational properties of utterances where the accent/focus alliance is attained by syntactic means are analysed and compared to the intonational patterns of sentences where focus is signalled by accentual means. The discussion of the results and the conclusions of the whole chapter are presented in 4.5 and 4.6 respectively.



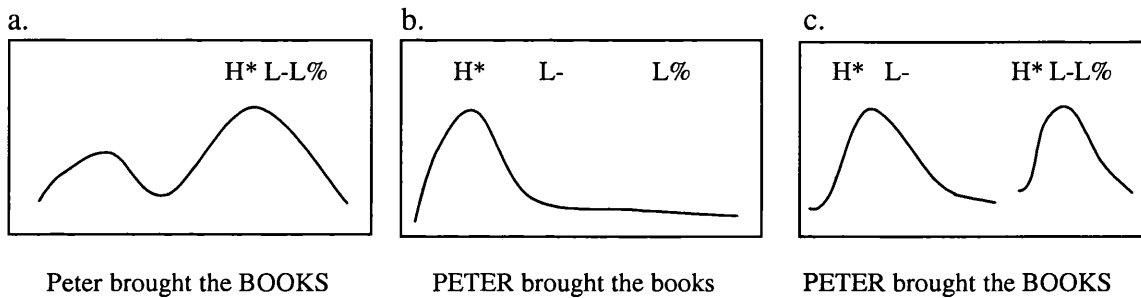
## **4.2 Phonetic and phonological cues of narrow focus**

As presented in chapter 1, any time a word is focussed, it immediately becomes more salient or more prominent than their neighbouring words. It has been observed cross-linguistically that in order to make an item more salient, both phonetic and phonological cues are used. Phonetically, a focussed item has a longer duration, a higher amplitude and a larger pitch range than a non-focussed item. For example, Cooper et al (1985) and Eady and Cooper (1986) for English, and Jun (1991) and Jun and Lee (1998) for Korean, found that a focussed word is longer than a neutral word. Sometimes focus also has an effect on the duration of postfocal material. In English (Erickson and Lehiste 1995) and in Korean (Jun and Lee 1998), the duration of the postfocus sequence is shorter than the equivalent sequence in neutral utterances. This phenomenon, however, seems to be language specific, since Jun and Fougeron (to appear) showed that in French the postfocal sequence does not differ in length from the same stretch in a neutral sentence. Similarly, the focussed word tends to be louder than the neighbouring non-focussed words (Wells 1986, Toledo 1986) and it always shows an F0 change (Fry 1958, Gay 1978, also see Ladd 1996:47). Studies on the perception of focus claim that, from all the suprasegmental features associated to focus, the cue that tends to be more relevant and robust for the perception of focus is the F0 pattern (see Fry 1958 and Beckman 1986 for English, and Solé 1984 for Spanish).

Studies on the phonological expression of focus across languages show that there are two kinds of effects induced by focal phenomena: 1) stress or accent effects, and 2) phrasing effects. The relevance of stress or accent distributions to express broad and narrow focus has been observed in many languages: English (Ladd 1980, 1990, 1996), Dutch (Gussenhoven 1983b, 1984), German (Féry 1992, 1993), Palermo Italian (Grice 1995a), Bari Italian (Grice and Savino 1997), Spanish (García-Lecumberri et al 1997), European Portuguese (Frota 1998) and French (Jun and Fougeron to appear), among others. In these languages, focus is mainly expressed by sentence stress and nuclear pitch accent location. However, accent effects can be manifested in two ways: 1) by the deletion of pitch accents or deaccenting, and 2) by the selection of pitch accents.

The deletion of postfocal accents is a common device to signal focus in most Germanic languages. In English, for instance, focus changes the tonal pattern of an utterance by placing the nuclear or last pitch accent over the most prominent word within the ip and deaccenting the postfocal material. Deaccenting involves no pitch accent after the focussed word and the appearance of a low plateau tone (assuming a fall on the nuclear accent). In Beckman and Pierrehumbert (1986), this plateau is analysed as a phrase accent (i.e. L-), covering the syllables between the focussed word and the phrase boundary. This is illustrated in (4.1) below.

(4.1)



Both (4.1a) and (4.1b) have one ip with one focussed word in it. In (4.1a), the focal word is the last word of the ip (i.e. late focus). In (4.1b), the focussed element is the first one of the ip (i.e. early focus) and hence all postfocal material is deaccented. In (4.1c), both words are focussed and each word is the nuclear accent of its own ip.

Whereas in English the same pitch accent is used to signal both broad and narrow focus (hence the sentence in (4.1a) has two readings), in other languages, on the other hand, specific pitch accents or tune types are used to convey the two meanings. As reported above, focussed elements are more prominent than non-focussed elements. One of the phonetic cues used to make items more prominent is to increase the pitch range. In some languages, this pitch range variation is merely phonetic and optional. In other languages, however, the boost has been incorporated into the phonological structure. Thus, a different pitch accent is used to convey broad focus and narrow focus. In Palermo Italian (Grice 1995a), in Bari Italian (Grice and Savino 1997) and in European Portuguese (Frota 1998), H+L\* is used with a broad focus reading and H\*+L with a narrow focus reading. In Spanish (Sosa 1999), L\*L% conveys a neutral intonation, whereas H\*L% and H+L\*L% a marked intonation. According to Sosa,

H\*L% has a nuance of surprise or amazement and H+L\*L% indicates reiteration. In Bengali (Hayes and Lahiri 1991), H\* L<sub>i</sub> conveys broad focus and L\* H<sub>p</sub> L<sub>i</sub> narrow focus<sup>1</sup>. In these languages, there is no ambiguity between a broad focus reading and a late narrow focus reading, as happens in English (Ladd 1996), Dutch (Gussenhoven 1983b, 1984) or German (Féry 1992, 1993). An example is provided in (4.2) for European Portuguese. In Portuguese, the two sentences in (4.2) are clearly different in meaning: (4.2a) conveys a neutral, broad focus reading, and (4.2b) a marked, narrow focus interpretation.

(4.2) a. *Neutral sentence*

As americanas ofereceram a enciclopèdia ao jornalista

H\*

H+L\* Li<sup>2</sup>

“The Americans-FEM gave the encyclopaedia to the journalist”

b. *Narrow focus sentence*

As americanas ofereceram a enciclopèdia ao JORNALISTA

H\*

H\*+L Li

“The Americans gave the encyclopaedia to the JOURNALIST”

Finally, the effects of focus on the phrasing structure of an utterance have been observed in languages, such as Bengali (Hayes and Lahiri 1991), Chichewa (Kanerva 1990), Korean (Jun 1998, Jun and Lee 1998 and Oh 1999) and Greek (Baltazani and Jun 1999). However, phrasing can be modified in different ways. Whereas sometimes an edge tone is introduced either before or after the focussed word, other times focus deletes rather than inserts a boundary. In Bengali and Chichewa, for example, the phrase tone is located after the focussed item. In Greek, focus deletes the boundary and deaccents the following words. Finally, in Korean, focus creates a boundary between the focussed item and the preceding word and dephrases postfocal material. A example from Korean (Jun 1998) is provided in (4.3). The focussed element (underlined) initiates an Accentual Phrase and dephrases the subsequent Accentual Phrases within an IP.

---

<sup>1</sup> In Hayes and Lahiri’s notational conventions, tones marked with the subindices P and I are phrase edge tones. P signals the end of a phonological phrase (similar to the ip ) and I the end of an IP.

<sup>2</sup> In Frota’s tonal transcription the subindex i marks a high level of prosodic phrase structure (i.e. IP).

(4.3) a. *Neutral sentence*

[kjəure] [dʒebəhan] [oiga] [masit'anɪnte]  
 winter-LOC grow-REL a cucumber-NOM delicious-they say  
 "They say a cucumber grown in winter is delicious"

b. *Narrow focus sentences*

(i) [kjəure dʒebəhan oiga masit'anɪnte]


"They say a cucumber grown in WINTER is delicious"

(ii) [kjəure dʒebəhan] [oiga masit'anɪnte]

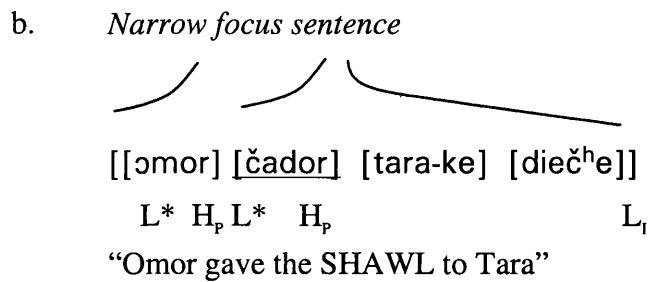
"They say a CUCUMBER grown in winter is delicious"

Languages may use one of these devices alone or may combine them. For example, in Bengali (Hayes and Lahiri 1991), phrasing, deletion of pitch accents and selection of pitch accents appear at the same time to signal focus. This is illustrated in the sentences of (4.4), which are adapted from Hayes and Lahiri's work. (4.4a) presents a sentence with a broad focus reading. The tonal structure of the utterance is described as a sequence of pitch accents (T\*), phrase accents (T<sub>P</sub>) and boundary tones (T<sub>I</sub>). The nuclear tone of a neutral-focus utterance involves an H\* pitch accent followed by an L<sub>I</sub> boundary tone. (4.4b) shows the same sentence as (4.4a) but with narrow focus on the word *shawl*. In (4.4b), the focussed element (underlined> is signalled by an accent type different from that of neutral focus, namely, L\*. Additionally, an H<sub>P</sub> tone marks the right edge of the focussed phrase. Finally, postfocal material is deaccented.

(4.4) a. *Neutral sentence*



[[ɔmor] [čador] [tara-ke] [dieč<sup>h</sup>e]]  
 L\* H<sub>p</sub> L\* H<sub>p</sub> L\* H<sub>p</sub> H\* L<sub>1</sub>  
 Omor the-shawl Tara-ACC gave  
 "Omor gave the shawl to Tara"



In this chapter, the accentual properties of narrow focus in Central Catalan declaratives are investigated. In particular, two issues are examined in detail: 1) the type of pitch accent associated with narrow focus and 2) the effects of focus on prosodic phrase structure.

### 4.3 Data

The narrow focus sentences analysed in this chapter consist of the utterances produced by six informants in the first production test (both in the reading task and in the semi-spontaneous activity)<sup>3</sup>. These speakers were selected according to two criteria: 1) the number of responses with accentual focus (all of them produced more than 85% of intonational focus responses in the production test), and 2) the quality of the Lx signal (speakers with bad F0 traces were rejected). Overall, 6 female speakers were selected: CP, DV, ER, MC, NG and NM. As reported in chapter 3, these informants were later recorded to obtain broad focus sentences.

The narrow focus utterances gathered in the first production test consisted of sentences with narrow focus on the subject and on the verb (see Appendix 1, section 1.1, for the list of sentences). Even though in the productions of the 6 speakers the focus/accent association was mainly attained by a reorganisation of the accentual pattern, a few sentences with a syntactic shift were also obtained. In this chapter, the two types of sentences are examined.

In addition to subject and verb focus utterances, a few sentences with narrow focus on the object were also recorded to examine differences and similarities in the *w-s* pattern

<sup>3</sup> See chapter 2 (section 2.3) for details.

when produced with a broad focus intonation vs an object narrow focus intonation. These sentences were recorded along with the broad focus utterances in the second recording session for the six speakers mentioned above. 16 sentences with narrow focus on the object were collected for each speaker. Those sentences consisted of some of the declarative sentences already designed for narrow focus on the subject or on the verb but with expected narrow focus on the object. In order to get object focus, both contrastive and identification triggering questions were used, as in subject and verb focus (see section 2.3.1.1 for details). Finally, the object focus data consisted of 8 sentences with a single stressed focal domain and 8 sentences with a double stressed focal domain. Examples of utterances with late narrow focus triggered by contrast and by identification are provided in (4.5) below. For the whole list of sentences with narrow focus on the object, see Appendix 3 (section 3.1).

(4.5) a. *Contrastive prompt*

La Rosa llegia revistes? La Rosa llegia LLIBRES  
“Did Rosa read magazines? Rosa read BOOKS”

b. *Identification prompt*

Què llegia la Rosa? La Rosa llegia LLIBRES  
“What did Rosa read? Rosa read BOOKS”

Sentences with narrow focus on the object were gathered by means of a reading task, similar to that designed to obtain utterances with narrow focus on the subject and on the verb. Thus, speakers were given a piece of paper where sentences with expected narrow focus on the object were included along with distractors (sentences with broad focus) in a randomised order. Informants were asked to find the appropriate answer to the questions recorded by the researcher. For late narrow focus, no semi-spontaneous task was carried out since at this stage the aim was no longer to analyse the use of accentual focus in Central Catalan but to obtain a few examples with focus on the object for a phonological interpretation of the data. The instrumentation and recording procedures used in this second production test were exactly the same ones as for the first test (section 2.3.1.4 for details).

## 4.4 Results

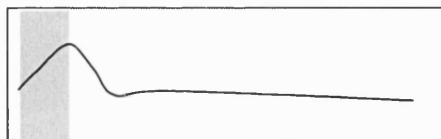
The results of Central Catalan declaratives produced with a narrow focus reading are presented in two parts. The first part analyses those sentences where the focus/accent alliance is attained by intonational means, without a reorganisation of the syntactic structure. Within this section four aspects are covered: 1) the intonational properties of sentences with *early* focus (i.e. focus on the subject or on the verb), 2) the intonational properties of sentences with *late* focus (i.e. focus on the object), 3) the levels of phrasing, and 4) postfocal accents. In the second part, the intonational characteristics of a few sentences where the focus/accent association is attained after a syntactic reorganisation are analysed and compared to sentences where the focus/accent association is not mediated by syntax.

### 4.4.1 Accentual focus

#### 4.4.1.1 Early narrow focus

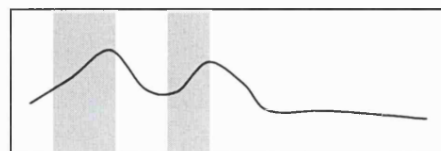
In Central Catalan, the F0 contours with an early narrow focus reading are characterised by a smooth rise over the stressed syllable of the focussed word followed by a fall in the poststressed syllables. The peak of the rise is aligned towards the end of the accented syllable, as in single-stressed broad focus utterances. In most cases, the stretch of words after the focussed element shows a low plateau until the end of the sentence. This behaviour was consistent in all narrow focus sentences irrespective of the kind of triggering question, i.e. contrastive prompt or identification prompt. This is illustrated in (4.6) for the sentence *L'EVA guarda monedes romanes* with focus on the subject and in (4.7) for the sentence *En Joan DOMINA l'àlgebra* with narrow focus on the verb. In the two cases, the pitch trace shows a clear peak over the accented syllable of the focussed element followed by an immediate fall and deaccenting of the remaining sequence. This behaviour agrees with the few brief references to narrow focus obtained in the literature (Bonet 1984 and Prieto 1995, to appear b). For the original F0 traces of (4.6) and (4.7) and the following sketched F0 contours, see Appendix 5.

(4.6)



L'EVA guarda monedes romanes  
 "EVA keeps Roman coins"

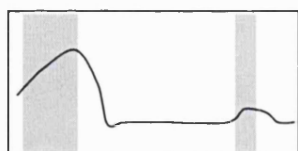
(4.7)



En Joan DOMINA l'àlgebra  
 "Joan MASTERS algebra"

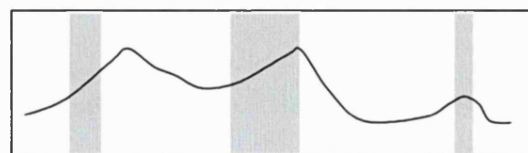
Although in most cases postfocal stresses did not become accented, sometimes a very compressed pitch accent was observed on the last stressed syllable of the whole IP. Postfocal accents appeared in both subject and verb focus sentences. Examples of postfocal accents in utterances with subject and verb focus are illustrated in (4.8) and (4.9) respectively. Postfocal accents will be analysed in more detail in section 4.4.1.4.

(4.8)



ELL anava a Girona  
 "HE went to Girona"

(4.9)

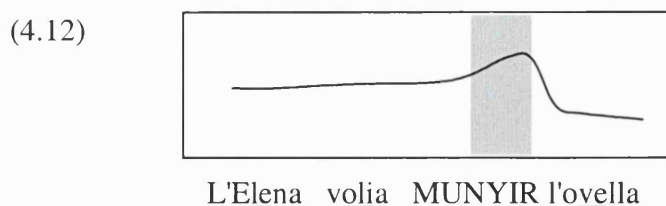
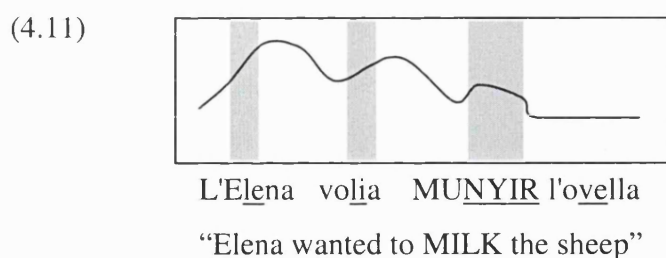
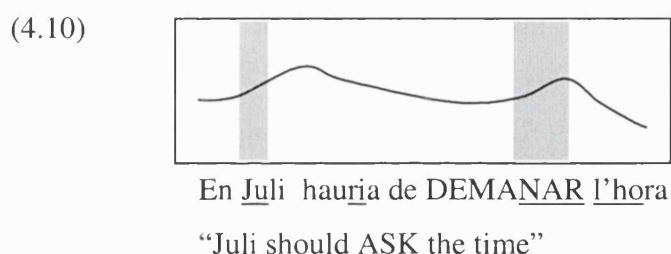


L'Elena mira de MOURE el mobiliari  
 "Elena tries to MOVE the furniture"

The accentuation of prenuclear stresses presented fluctuation among speakers. In sentences with only one potentially accentable syllable before the focal item, this tended to be accented, as illustrated in (4.7) above. In sentences with two or more potential prenuclear accents, only the first stress tended to become accented. This is



illustrated in (4.10) for an utterance with two prenuclear stresses. Even though this tendency was quite consistent in the data of all informants, inter-speaker variability was observed. For example, whereas speaker NG produced some of her utterances with as many prenuclear accents as possible, other speakers (mainly NM) produced a few sentences with no prenuclear accents. These two patterns are illustrated in (4.11) and (4.12) respectively. Differences in the number of prenuclear accents did not trigger major semantic differences, although a more marked nuance was sometimes perceived when the number of pitch accents was higher. As in broad focus sentences, prenuclear accents showed a rising contour, which started during the stressed syllable and finished at the end of the word.



Overall, narrow focus conveyed by accentual means in Central Catalan involves the location of the nuclear accent over the focal element with optional deaccenting of postfocal material. Prefocal material tends to show a pitch accent on the stressed syllable of the first lexical word of the sentence. If there are any other potentially accentable words between the first accented word and the focussed element, these tend to be unaccented. In the next two sections, the phonological interpretation of both

nuclear and prenuclear accents in Central Catalan narrow focus sentences will be investigated and compared to the representation proposed for broad focus sentences.

#### **4.4.1.1.1 The nuclear accent**

A comparison between the phonetic characteristics of the nuclear accent in broad and in early narrow focus sentences shows that, whereas in broad focus sentences the last pitch accent is realised as a progressively falling F<sub>0</sub> contour, in early narrow focus it exhibits a clear F<sub>0</sub> peak anchored within the limits of the accented syllable. The question that arises from these observations is whether the nuclear accent of narrow focus utterances in Central Catalan is the same as that of neutral sentences, as is the case of English (Ladd 1996), Dutch (Gussenhoven 1983b, 1984) or German (Féry 1992, 1993), or it involves a different accent type, as in Palermo Italian (Grice 1995a), Bari Italian (Grice and Savino 1997) or Portuguese (Frota 1998).

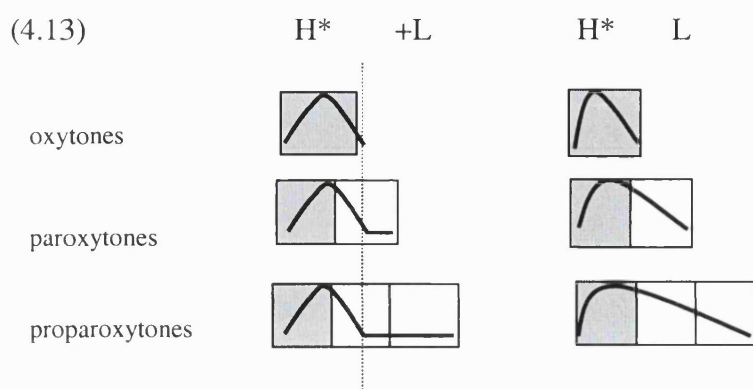
Three options were considered as possible phonological representations for the nuclear accent in narrow focus sentences, namely, H\*+L, H\*(L) and L+H\*. In all cases, an H\* is associated to the accented syllable of the focussed item but a different phonological interpretation is involved for the preceding or following F<sub>0</sub> valleys. The first option consists of a bitonal pitch accent, H\*+L. In early versions of the AM model (Pierrehumbert 1980), the L tone of the H\*+L accent was not manifested in the F<sub>0</sub> contour but only functioned as a downstep trigger of a succeeding H accent. The special status of this accent was acknowledged as problematical (see Ladd 1983, 1996) and, in later developments of the theory, this accent was eliminated from the English tone inventory (Beckman and Hirschberg 1994). However, work on the intonation of Italian and Portuguese shows that the H\*+L is required as the notation for the focal accent, since clear H and L targets are observed in the F<sub>0</sub> contour. In these languages, H\*+L accounts for a peak aligned within the accented syllable followed by an immediate fall. Given the presence of H\*+L in the tonal inventory of these Romance languages, the possibility of analysing Central Catalan focal accents as H\*+L is investigated in this study. If the data provides evidence that H\*+L is the pitch accent used in Central Catalan for narrow focus utterances, the immediate implication for intonational theory is that two different accent types are used for broad and narrow focus nuclear accents, as proposed for Portuguese, Italian and Spanish.

The second option involves analysing the nuclear accent of narrow focus as in broad focus, namely, as a sequence of an H\* pitch accent and an L word edge tone. If this is the case, broad focus and late narrow focus are expected to be ambiguous, as reported for English, Dutch or German.

Finally, the third option involves the bitonal accent L+H\*. This accent consists of a peak that rises from a low level. This accent type has been proposed by Prieto (1995) to account for focal accents in Central Catalan with a surprise and contrastive nuance. This possibility is considered because a valley was usually observed between the peak of the focussed element and the prenuclear peak (see (4.7) or (4.9) above). This valley can respond to two interpretations: either a sagging transition between the two peaks or an intended L target. D'Imperio (1999a) for Neapolitan Italian shows that in similar sequences, the valley is intended and hence it is significant for phonological structure. Again, if the L+H\* accent is observed in the data, two different pitch accents would be postulated in Central Catalan to account for broad and narrow focus sentences.

In the following pages, the three possibilities will be examined through a more detailed acoustic analysis of the data. First, options H\*+L and H\*(L) will be analysed and second H\* vs L+H\* will be examined.

In order to identify the phonological status of the nuclear accent in narrow focus sentences as H\*+L or H\*(L), the location of the L was examined in more detail, with the following assumptions. (i) If L is part of the bitonal accent, it is expected to be placed at a fixed distance with respect to the accented syllable, since trailing tones are observed to have a fixed location after the starred tone (Grice 1995a, Arvaniti et al to appear). (ii) If, on the other hand, L demarcates a boundary, its location is independent of the accented syllable, yet it should be aligned with the end of the focussed word, no matter the number of poststressed syllables. As in chapter 3, we assume that the starred tone is located at the onset of the accented syllable. The predictions of these two tonal possibilities are schematised in (4.13) below, for words with different stress patterns (oxytones, paroxytones and proparoxytones).



Since the initial data only contained focussed words with stress on the ultimate syllable (oxytones) and on the penultimate syllable (paroxytones), a few sentences where the focussed item had stress on the antepenultimate syllable (proparoxytones) were gathered in a later recording session for one speaker (DV)<sup>4</sup>. This was done since it was considered necessary to have all stress distributions to examine the location of L. Thus, 12 sentences with narrow focus on the subject and 6 with narrow focus on the verb were recorded following the same methods as reported in section 2.3.1<sup>5</sup>. See Appendix 3 (section 3.2.1) for the list of utterances. A few more sentences with oxytones were also gathered since the number of oxytones in focal position was not very high in the initial recording. See section 3.2.2 in Appendix 3 for the sentences with oxytones.

For all focal words, the following measurements were performed:

1. Distance between the onset of the accented syllable and the location of the valley (on-valley).
2. Distance between and the onset of the accented syllable and the end of the word (on-end).

As in broad focus sentences, in order to see the relationship between the two variables, correlations were performed for all the measurements. If L is part of the bitonal accent, a very low correlation between the two distances is expected, as L will be anchored at a fixed location with respect to the starred tone (or onset of the accented syllable). If, on the other hand, L marks the presence of some kind of

<sup>4</sup> This informant also recorded broad focus sentences with proparoxytones.

<sup>5</sup> Only 6 sentences with verb proparoxytones were gathered because it is difficult to find verb forms with stress on the antepenultimate syllable in Central Catalan.

boundary, L will be anchored at the end of the word irrespective of the number of poststressed syllables and hence a high correlation between the two measurements is expected.

The results are presented in Figures 4.1-4.3. Figure 4.1 includes the measurements for nuclear oxytones both in subject and in verb position for all speakers. For speaker DV, the oxytones of the second recording are added to the original number. Figure 4.2 shows the results for paroxytones. Finally, Figure 4.3 exhibits the results of proparoxytones produced by speaker DV both in subject and verb positions. Each graph plots the distance between the onset of the accented syllable and the F0 valley (on-valley) against the distance between the onset of the accented syllable and the end of the word (on-end) in seconds. Regression lines and  $R^2$  are included in each graph.

OXYTONES

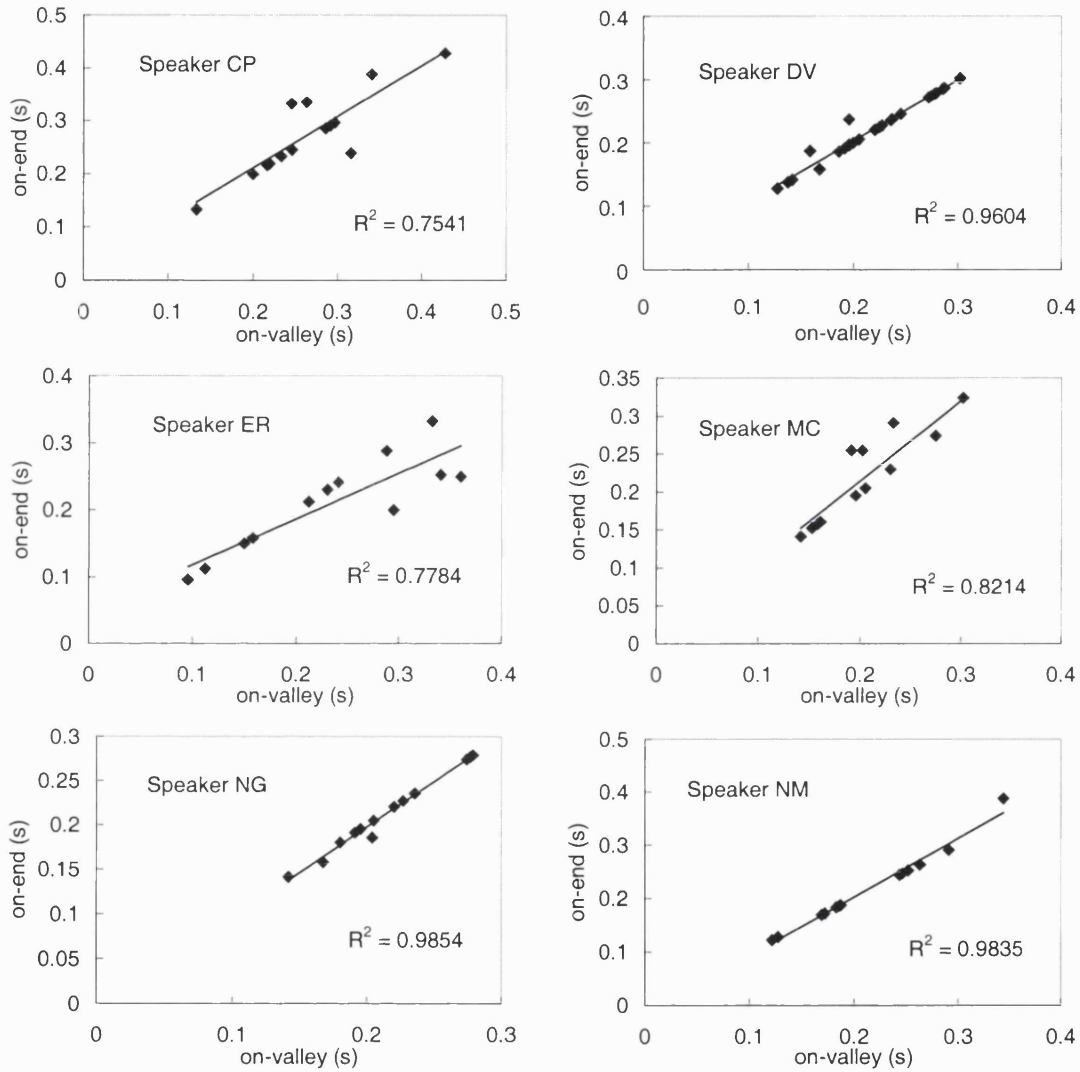


Figure 4.1. Distance between the onset of the accented syllable and the valley (on-valley) against distance between the onset of the accented syllable and the end of the word (on-end) in seconds for focal oxytones in subject and verb focal domains for all speakers.

## PAROXYTONES

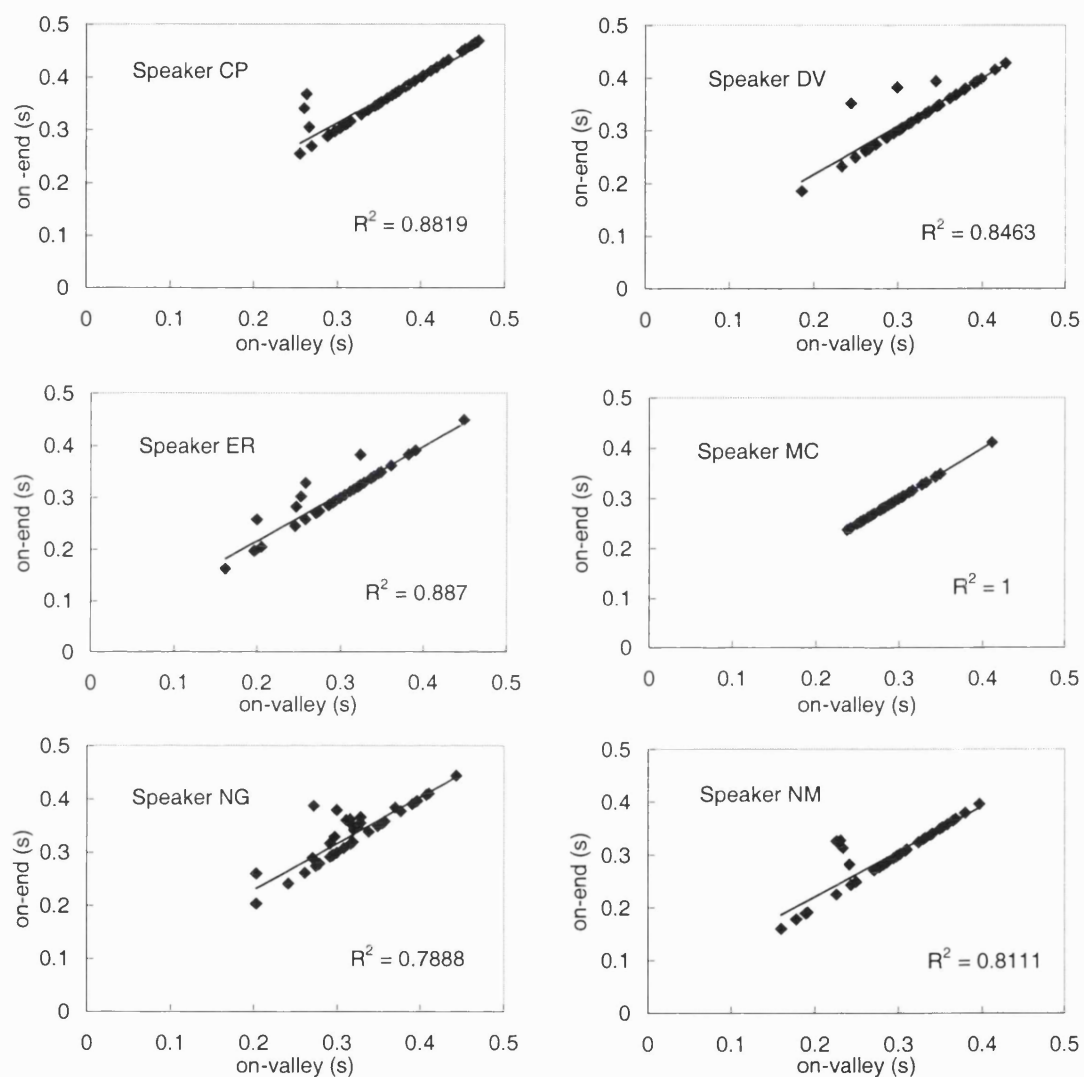


Figure 4.2. Distance between the onset of the accented syllable and the valley (on-valley) against distance between the onset of the accented syllable and the end of the word (on-end) in seconds for focal paroxytones in subject and verb focal domains for all speakers.

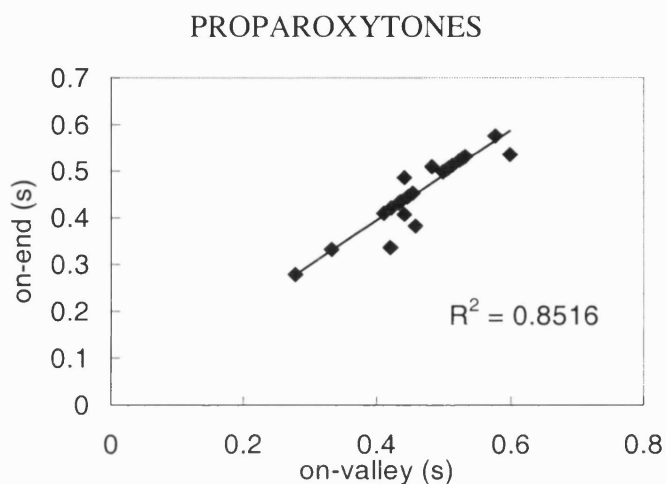
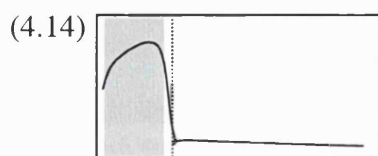


Figure 4.3. Distance between the onset of the accented syllable and the valley (on-valley) against distance between the onset of the accented syllable and the end of the word (on-end) in seconds for focal proparoxytones in subject and verb focal domains for speaker DV.

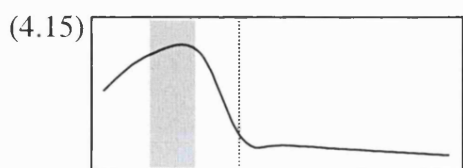
In all cases, the results show high correlations ( $R^2$  ranges from 0.75 to 1.0) between the "on-end" distance and the "on-valley" distance, indicating that the valley is consistently anchored at the end of the focussed word, no matter the number of poststressed syllables. This suggests that the L signals some kind of prosodic boundary, rather than the trailing tone of a bitonal accent. Illustrations of the alignment of L at the word offset are provided in (4.14) for a focussed oxytone, in (4.15) for a focussed paroxytone, and in (4.16) for a focussed proparoxytone in subject position (produced by speaker DV). See Appendix 5 for the original F0 traces. In all examples, the valley in the F0 contour coincides with the offset of the word. Thus, the nuclear accent of narrow focus has to be described as H\* followed by some kind of L edge tone.



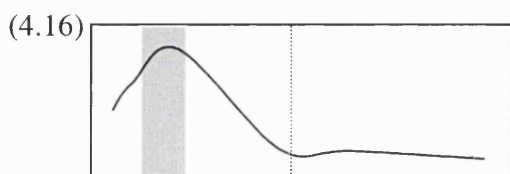
NOU miraven el mar

"NINE were looking at the sea





LA MARE menja arengades  
 "THE MOTHER eats herrings"



EL NOMADA menjava al carrer  
 "THE NOMAD ate in the street"

The clear anchoring of the L tone at the end of the focussed word in narrow focus sentences is consistent with the results observed in chapter 3, where word edge tones marked the right edge of words containing a pitch accent. The possibility that a higher level of prosodic phrasing, such as an ip, signals the end of the focussed domain will be examined in more detail in section 4.4.1.3. So far, the results presented above suggest that the nuclear accent in narrow focus is better described as an H\* pitch accent associated to the accented syllable of the focussed word followed by an L edge tone, rather than as a bitonal accent (H\*+L). The possibility that the tonal structure of the focal pitch accent is L+H\* rather than H\* is investigated below.

The idea that the nuclear accent of narrow focus in Central Catalan declaratives could be accounted for as L+H\* rather than H\* derives from the observation that the F0 traces showed an F0 minimum between the peak of the focal accent and the peak of a prenuclear accent. This F0 valley can have two possible explanations. First, the valley can be the result of a sagging interpolation between two peaks (Pierrehumbert 1980). Second, it can be the manifestation of a tonal target as observed in Neapolitan Italian focal structures (D'Imperio 1999a). For Central Catalan focal accents, Prieto (1995) proposed the L+H\* notation on the basis that an F0 valley was always observed before the peak.

In order to decide whether the valley could be explained as a phonetic phenomenon or as a phonological target, the values of L (or F0 minimum between peaks) were obtained for all sentences with two peaks. The results were divided into two groups according to the number of intervening syllables between the peaks (i.e. 1-2 syllables and 3-4 syllables). The hypothesis was that if L is the result of a sagging interpolation between two peaks, the value of L will be smaller as the number of syllables (and hence the duration) between peaks increases, due to a time-dependent F0 downtrend. Alternatively, if L is a target, a fairly constant value is expected irrespective of the number of intervening syllables<sup>6</sup>.

The mean F0 values of valleys corresponding to a different number of intervening syllables between the two peaks is presented in Figure 4.4 for all speakers. The results showed that the mean F0 value of the valley decreased as the number of segmental material between the two peaks increased. This was observed in the data of all informants.

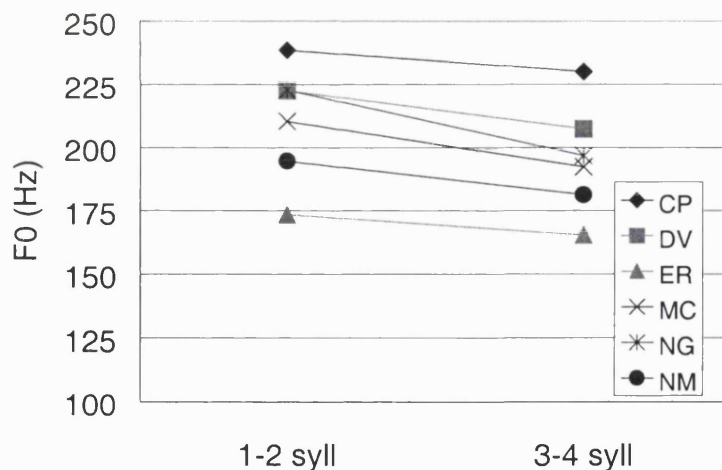


Figure 4.4. Mean F0 values of the valley between two peaks separated by a different number of intervening syllables (1-2 or 3-4 syllables) for all speakers.

T-tests were performed to check whether differences in the two samples (1-2 syllables vs 3-4 syllables) were significant or not. The results of the t-tests are presented in Table 4.1. For 5 speakers, significant differences are observed between the two variables at 1 per cent significance level. For speaker CP, significant differences are observed at 5 per cent significance level. These results indicate that the valley

<sup>6</sup> This hypothesis is based on the fact that, as will be observed in chapter 5, the heights of H\* peaks themselves are constant regardless of the number of intervening syllables.

observed between the focal accent and the prenuclear accent in Central Catalan sentences produced with narrow focus is not a phonological target, since it is affected by temporal constraints. That is, the more syllables between the peaks (i.e. the bigger the temporal distance)<sup>7</sup>, the lower the F0 values of the valley. This suggests that the valley is a phonetic induced effect rather than an intended target. These observations support the claim that the phonological representation of the focal accent in Central Catalan is H\* rather than L+H\*.

	CP	DV	ER	MC	NG	NM
p	=0.05	<0.001	<0.001	<0.01	<0.001	<0.001
t	1.7	3.9	3.5	3.0	4.2	3.6
df	43	31	43	36	15	27

Table 4.1. Results of the t-tests comparing the F0 values of valleys with a different the number of intervening syllables between peaks for all speakers.

Anticipating some of the results on F0 downtrends presented in chapter 5, it will be observed that the scaling of L in L\*H prenuclear rises is fixed irrespective of the number of intervening syllables between peaks. The stable scaling of L when it is a target (L\*H), as opposed to the variable scaling of L observed in Figure 4.4, corroborates the idea that the nuclear accent of narrow focus sentences does not consist of an L leading tone but is a monotonal pitch accent (H\*). Furthermore, observations of the focal accent when it is not preceded by any other prenuclear accents show that the peak of the focal accent does not rise from a low F0 but from a mid F0 (see examples (A20)-(A22) in Appendix 5, corresponding to sketches (4.14)-(4.16)). These observations confirm the analysis of the focal accent as H\* rather than L+H\*.

According to these results, Central Catalan seems to use the same pitch accent as the nuclear accent of both narrow and broad focus sentences. This behaviour is also reported for languages such as English, Dutch or German but differs from the results observed in other Romance languages (Italian and Portuguese). Thus, the immediate

<sup>7</sup> The fact that differences in the number of syllables involve differences in duration is confirmed in chapter 5 (section 5.3.1.2.2).

implications derived from the data is that focus in Central Catalan is expressed by the same accent type in both neutral and marked sentences.

#### **4.4.1.1.2 Prenuclear accents**

As in broad focus utterances, prenuclear accents showed a rise whose peak is consistently anchored at the end of the word. This suggests that the phonological representation of such accents is the same one as in broad focus, namely, an L\* pitch accent associated with the accented syllable and an H word edge tone associated with the end of the accented (or phonological) word. The data only included prenuclear oxytones and paroxytones but, given the consistent behaviour found in the productions of all speakers and given the results presented in chapter 3 where proparoxytones were also analysed, we concluded that the pattern of prenuclear accents in narrow focus declaratives is the same one as in neutral declaratives. As an illustration, the results of two speakers are presented in Figure 4.5 below. Not all speakers are included because the behaviour was the same as that observed in the last chapter. The data show the correlations between the onset of the accented syllable and the peak (on-peak) and the onset of the accented syllable and the end of the word (on-end) for prenuclear oxytones and paroxytones in sentences with a narrow focus reading.

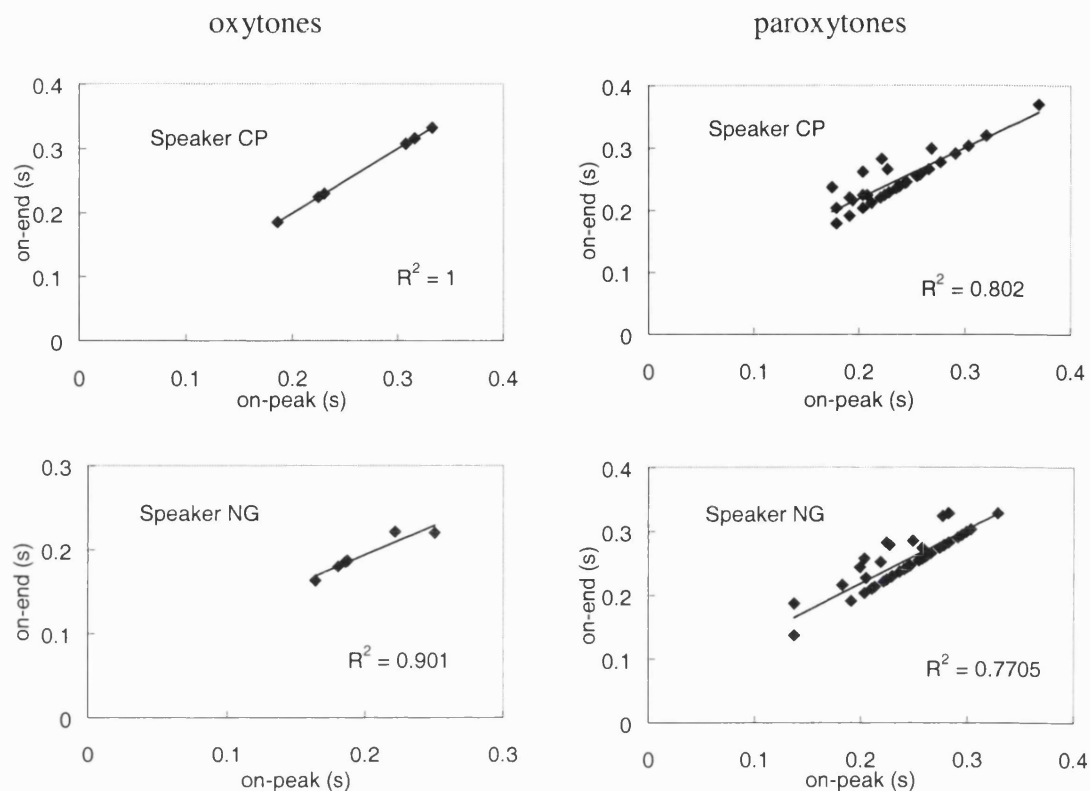


Figure 4.5. Distance between the onset of the accented syllable and the peak (on-peak) against distance between the onset of the accented syllable and the end of the word (on-end) for prenuclear oxytones and paroxytones in sentences with a narrow focus reading produced by speakers CP and NG.

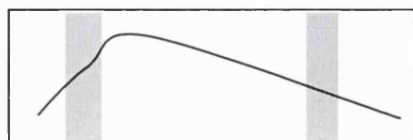
The results presented in Figure 4.5 show a high correlation between the "on-peak" distance and the "on-end" distance ( $R^2$  ranges from 0.77 to 1.0) in all cases, confirming that the location of the peak of prenuclear rises is at the offset of the word (i.e. L\*H). Thus, the intonation of narrow focus in Central Catalan seems to be very similar to that of broad focus utterances, except in the location of the nuclear accent, which is placed over the focussed word. This will be corroborated in the next section for sentences with late narrow focus.

#### 4.4.1.2 Late narrow focus

All sentences with late narrow focus triggered by an identification prompt and some of those triggered by a contrastive prompt (67% of the responses) were produced with a rising pitch in the vicinity of the first accented syllable, followed by a progressive falling contour up to the end of the utterance. Prenuclear stressed syllables other than the first one did not become accented. This is illustrated in (4.17) for an utterance with

narrow focus on the object. This pitch pattern is exactly the same as the one observed in broad focus sentences. Thus, the suggested tonal configuration involves an L\* prenuclear accent followed by an H anchored at the end of the word and an H\* (with final lowering) on the final stressed syllable followed by L edge tones. See chapter 5 for more details on final lowering.

(4.17)

La Glòria ve de VILANOVA

L\*H

H\* LL-L%

“Gloria comes from VILANOVA”

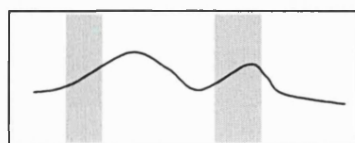
The fact that late narrow focus and broad focus present a similar F0 trace confirms the idea that the same accent type is used to express the two kinds of focus. Thus, the reading of late focus and broad focus is ambiguous in Central Catalan. As pointed out before, ambiguity of the late nuclear accent pattern between a broad and a narrow focus interpretation has been observed in English (Ladd 1996), Dutch (Gussenhoven 1983b) and German (Féry 1992, 1993). In these languages, no special tonal events are used to distinguish between neutral and marked focus and hence the *w-s* pattern can convey both broad focus and narrow focus on the object. Languages such as Bengali (Hayes and Lahiri 1991), Palermo Italian (Grice 1995a), Neapolitan Italian (D’Imperio 1997) and European Portuguese (Frota 1998), on the other hand, are non-ambiguous as far as late prominence is concerned, since different pitch accents are used to signal broad and narrow focus.

In order to confirm that Central Catalan speakers do use the same tonal events for both broad focus and late narrow focus, a small perception test including two Central Catalan native speakers was performed. The test consisted of 1) 16 utterances with narrow focus on the object obtained as answers to a contrastive prompt, 2) 16 utterances with narrow focus on the object obtained as answers to an identification prompt, and 3) 16 utterances produced with a broad focus reading. The segmental structure of the 16 sentences for each group was the same. Thus, sentences only differed in the way they were elicited (i.e. as answers to a contrast, an identification or

a narrow focus question type). Overall 48 utterances were included in the test. Sentences were produced by one informant (MC) to avoid adding speaker variability in the perception test. This informant was selected because she was rather consistent in her narrow focus productions<sup>8</sup>. Sentences of speaker MC were transferred from the Sun computer to a 60 ES DAT recorder in a randomised order (i.e. mixing the three kinds of utterances). For the perception test, the two listeners were asked to say whether the utterances they heard were appropriate as answers to three kinds of questions: 1) a broad focus question (e.g. "what happens?"), 2) an identification question, and 3) a contrastive question. Questions were written on a paper. None of the stimuli was elicited by means of the question type the listeners were judging. For example, a stimulus obtained as a response to a contrastive narrow focus question was now being judged as a possible answer to a broad focus question. The results of the test showed that in all cases the listeners found that each stimulus was appropriate as a response to the question, indicating that no differences were perceived between sentences produced with broad focus and those uttered with narrow focus on the object (contrast and identification). The results of the perception test were very consistent for the two listeners and confirmed the ambiguity in the reading of late focus in Central Catalan as a broad focus and a narrow focus interpretation.

A few object focus sentences elicited with a contrastive prompt (basically half of the productions of speaker CP) presented a smooth rise over the nuclear syllable, rather than a progressive fall. An example is provided in (4.18) below. The alignment of the peak showed exactly the same behaviour as in early focus or in single-stressed broad focus utterances. That is, the peak was aligned late within the accented syllable (see chapter 3, section 3.4.3.3 for more details).

(4.18)

La Rosa llegia LLIBRES

"Rosa read BOOKS"

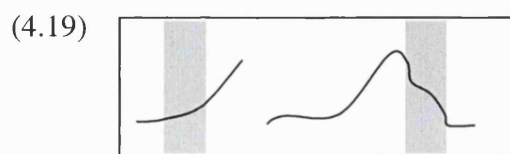
<sup>8</sup> Other speakers used a different pitch accent (H+!H\*) in contrastive narrow focus, as will be discussed later on.

These sentences were perceived (by the researcher and two other Central Catalan native speakers) as more emphatic. However, even though this emphatic accent can be used to disambiguate between a broad focus and a late narrow focus reading (i.e. focus on the object), it does not necessarily trigger disambiguation between the two focal patterns. The three listeners agreed that a sentence with a final emphatic stress can also be a more enthusiastic and committed response to a “what happens?” type of question. Thus, the appearance of an emphatic accent is not compulsory to express narrow focus in Central Catalan. In other words, emphasis is a matter of choice on the speaker's part rather than a matter of phonological necessity. In fact, most of the narrow focus sentences were produced with non-emphatic stress (see Estebas-Vilaplana and Maidment 1999b and also chapter 5 of the present study). These observations confirm the idea that there is no distinctive category to express narrow focus in Central Catalan, although there is a possibility of gradiently modifying it. Thus, the final peak in (4.18) and the progressive fall in (4.17) are gradient realisations of the same entity (H\*). This confirms some of the generalisations already provided in chapter 3 and reproduced below.

1. The nuclear accent of broad and narrow focus sentences in Central Catalan involve the same pitch accent (H\*).
2. In single-stressed broad focus structures, in early narrow focus, and in emphatic late narrow focus, the peak of the rise is aligned near the end of the accented syllable (in oxytones the peak might be aligned a little earlier to be able to produce the subsequent fall).
3. In multi-stressed broad focus structures and in non-emphatic late narrow focus, no clear peak is observed within the accented syllable. The lack of F0 prominence will be explained as a final lowering effect (see chapter 5).

Finally, 23 % of the sentences triggered by a contrastive prompt presented a different pitch trace (some responses of speakers CP and DV). As before, an initial rise was observed at the beginning of the utterance, near the first stressed syllable (i.e. L\*H). The final pitch accent, however, presented a different pitch excursion which consisted of a slight lowering in the F0 contour over the accented syllable preceded by a high pitch. This is illustrated in the sentence below which was also produced with an ip boundary between the subject and the verb.





La Glòria ve de VILANOVA

“Gloria comes from VILANOVA”

In early versions of intonational phonology (Pierrehumbert 1980, Beckman and Pierrehumbert 1986), this pitch movement was transcribed as H+L\*. It mainly involved a pitch excursion from a high level to a mid level. In the ToBI version of Pierrehumbert’s taxonomy (Beckman and Ayers 1994, Pitrelli et al 1994, Beckman and Hirschberg 1994), this pitch movement is transcribed as H+!H\*. This involves a peak followed by a downstepped, and hence slightly lowered, H target. The new notation makes it clear that the level of the accented syllable is not necessarily low in the speaker’s range but only lower than the preceding material. The early notation (H+L\*) is either no longer included in the tone inventory (i.e. English ToBI) or is kept to signal HL sequences where the L shows a clear low target. H+L\* has recently been used as the nuclear accent of neutral declaratives in many Romance languages, such as Palermo Italian (Grice 1995a), Neapolitan Italian (D’Imperio 1997a) and European Portuguese (Frota 1998). In this study, the HL movement observed in object focus triggered by contrast is analysed as H+!H\* on the basis that the F0 movement over the accented syllable is not an L target but a lowered H target. This accent was also observed in Prieto (1995), who claimed that in Central Catalan H+!H\* is mainly used with a pedagogical intention. This nuance can also be attributed to our data.

The use of the H+!H\* pitch accent is restricted to a contrastive narrow focus and it is not accepted with an identification reading. It has a nuance of insistence and correction of something that is erroneous. Although in the production data this accent only appeared in final position, it can be used in initial (subject) and in medial (verb) positions too with the same insistent contrastive meaning. The fact that it was only used in final position may be derived from the necessity of disambiguating between the broad and the late narrow focus reading. Since in late focus there is no possibility of postfocal deaccenting, a more marked accent was considered necessary to convey the intended meaning.

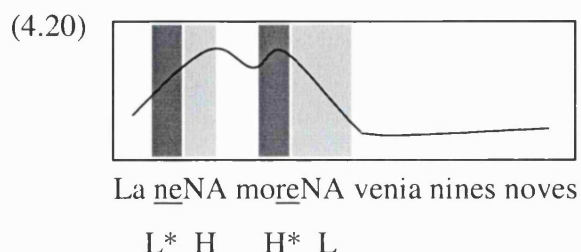
#### **4.4.1.3 Levels of phrasing**

The results presented in section 4.4.1.1 showed that the final pitch movement of narrow focus sentences involves an H\* pitch accent associated to the accented syllable of the focussed word followed by an L tone docked at the right edge of the word. This behaviour is consistent with that reported in chapter 3, where a word edge tone marks the end of the word that contains the pitch accent. However, does this L only mark a word boundary or does it signal a prosodic level above the word, such as the end of an ip boundary (L-)? Prieto (to appear b) proposed that focus in Central Catalan triggers a prosodic segmentation of the intonation phrase, right after the focussed material. This is examined in the following paragraphs.

In order to investigate whether there is an ip boundary after the focussed element in Central Catalan sentences, three pieces of evidence were taken into account: 1) preboundary lengthening, 2) the application of an assimilation process, and 3) a comparison between the intonation of sentences where narrow focus was signalled by accentual means only and that of utterances where a syntactic shift took place to attain the focus/accent association.

As reported in section 3.4.2.1, one of the cues that determines the presence of an intonation boundary is the lengthening of the final syllable of the intonation group (Cruttenden 1986). In our data, this was already observed in neutral declaratives produced with an ip boundary between the subject and the verb. In those cases, the duration of the syllable before the ip boundary was longer than the duration of comparable non-preboundary syllables. In order to clarify whether there is an ip boundary after the focussed word, the duration of the last syllable of the focussed word was obtained and compared to the duration of similar non-preboundary syllables within the same sentence. Since the data presented some variability in terms of syllable structure, only syllables with the same number of segments (basically those with a CV pattern) were examined. Furthermore, the non-preboundary syllable and the putative preboundary syllable had to have the same stress, as well as pitch distribution. This mainly involved postaccent syllables, which were unstressed but became the docking site for a word edge tone. This is illustrated in (4.20) below. The

syllables in capitals (corresponding to the lightly shaded area) stand for both the non-preboundary syllable (control syllable) and the putative preboundary syllable (target syllable). Sentences with only one focussed element were not analysed since no comparable syllables could be obtained.



The results are presented in Figure 4.6, which displays the mean duration of syllable 1 (non-preboundary) and syllable 2 (putative preboundary) for all speakers. The results show that in all cases the mean value of the putative preboundary syllable is significantly longer (see Table 4.2) than the mean duration of the non-preboundary syllable. This suggests that in Central Catalan the end of the focussed domain is signalled by the presence of a higher phrase boundary along with the word boundary. Given the fact that no major F0 breaks are observed in the F0 contour, we propose that this is an intermediate phrase boundary marked by the presence of an L- phrase accent.

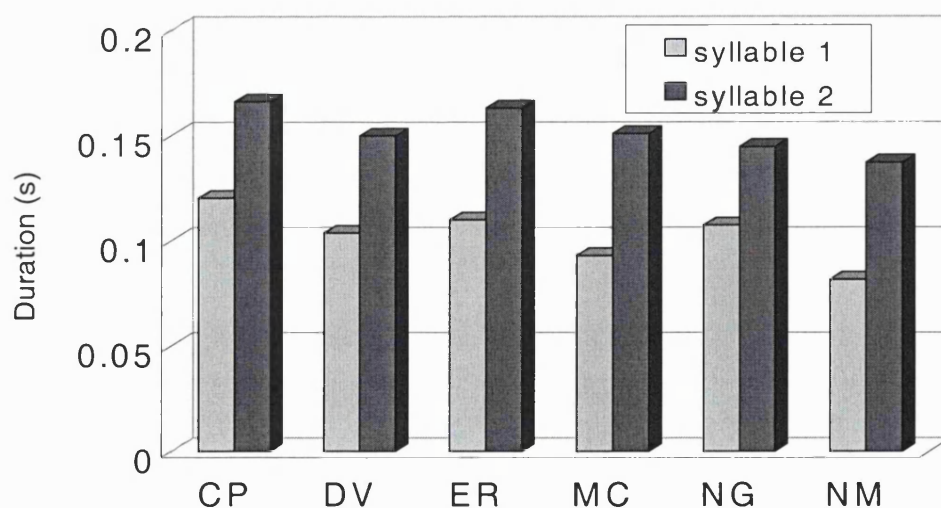


Figure 4.6. Mean duration of syllable 1 (non-preboundary) and syllable 2 (preboundary) in seconds for all speakers.

	CP	DV	ER	MC	NG	NM
p	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
t	-5.7	-7.0	-3.8	-5.2	-4.0	-4.9
df	21	28	16	17	30	9

Table 4.2. Results of the t-test comparing the mean duration of syllable 1 (non-preboundary) and syllable 2 (preboundary) for all speakers.

Even though in this study we have not used sandhi processes to define prosodic structure, in this particular case we will corroborate the presence on an ip boundary on the basis of a sandhi phenomenon. Specifically, the piece of evidence to postulate an ip boundary after the focussed domain comes from sentences with a double accented subject containing words ending in [s]. In Central Catalan, any word-final voiceless alveolar fricative becomes voiced when followed by a voiced sound (Recasens 1993). Thus, for example in the sentence *Vàries nenes lligaven els globus* the expected pronunciation is that all word final [s] (except the sentence final one) become [z] when followed by voiced material as in (4.21b). However, most of the times no assimilation of [s] was observed after the focussed word (i.e. the F0 trace was broken) indicating the presence of a boundary that blocked the assimilation process. This is illustrated in (4.21c).

- (4.21) a. *Vàries NENES lligaven els globus*  
 "Several GIRLS were tying the balloons"  
 /'barjəs 'nɛnəs li'ɣaβən əls 'ɡlɔβus/  
 b. ['barjəz 'nɛnəz li'ɣaβən əlz 'ɣlɔβus]  
 c. ['barjəz 'nɛnəs # li'ɣaβən əlz 'ɣlɔβus]

Finally, further evidence supporting the presence of a boundary after the focussed word can be obtained by comparing the intonational behaviour of sentences where the focus/accent association was attained by accentual means only to that of sentences which underwent a syntactic shift to assign focus. Syntactic focus involved a reorganisation of the constituents of the sentence in order to locate the focussed word in a prominent position or accent-bearing location. The focussed word was located at the end of the phrase, which was subsequently followed by a dislocated syntactic

phrase. This is illustrated in (4.22) below. Sentences with syntactic focus will be analysed in more detail in section 4.4.2. However, anticipating some of the results, it will be observed that the F0 traces of syntactic focus are very similar to the F0 traces of accentual focus, indicating that in both cases the focussed word seems to be at the end of a domain.

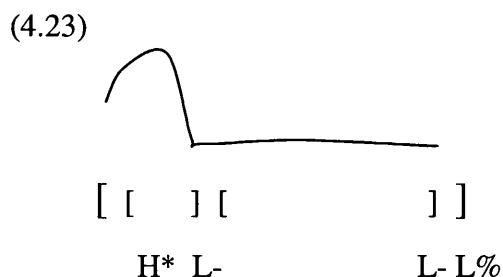
- (4.22) Hi bullia L'OU, a l'olla  
“It boiled THE EGG, inside the pot”

These results show that the valley after H\* not only marks the end of the word (L) but also the end of a higher prosodic domain, presumably an ip (L-). Thus, Central Catalan signals focus by placing the nuclear accent on the most prominent word and by creating a prosodic boundary after the focussed element. Contrary to other Romance languages, no different pitch accent seems to be used to convey narrow focus. Thus, narrow focus is expressed with the same accent type as broad focus and with the presence of an ip boundary.

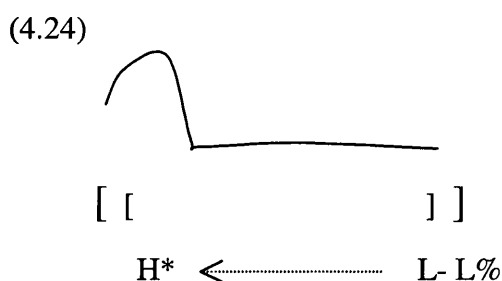
The appearance of an ip boundary after the focussed element in Central Catalan is consistent with the idea that the focal element in Romance languages tends to be located at the end of the sentence/phrase (Zubizarreta 1998, Sosa 1999). However, the results of our data show that the phrase final position cannot only be attained by a syntactic reorganisation of the elements of the sentence (as proposed by Vallduví 1990, 1991, 1994a/b and Vallduví and Zacharski 1994) but also by an accentual reorganisation (accent location and ip boundary). Thus, Central Catalan seems to be able to signal focus by the combination of both syntactic and intonational strategies or by intonational strategies only. This indicates that the idea that there should always be a syntactic reorganisation to signal focus in Central Catalan is thrown into doubt by the evidence of the data but the fact that focus is positionally determined is preserved.

The postulation of an ip boundary after the focal domain in Central Catalan creates some problems for intonational theory, since the status of postfocal material is unclear. If an IP is made up of one or more ips, it seems logical to think that after an ip boundary a new ip starts. Thus, for example in our data, the stretch of words after

the ip marked L- should constitute another ip as schematised below, most probably identified as L- followed by the L% boundary tone.



This proposal, however, is problematic for phonological structure since the second ip has no pitch accent and in principle the theory does not allow empty ips. In fact, descriptions of English focus (Beckman and Pierrehumbert 1986) have always claimed that a sentence such as *PETER brought the books* has only one ip. The nuclear accent is H\* (or L+H\*) and the L- covers the syllables between the focussed word and the boundary tone. In the early versions of the AM model (Pierrehumbert 1980), the phrase accent was not associated to a particular level of prosodic phrasing, but it was treated as a floating tone, which accounted for the intonation between the last pitch accent and the boundary tone. In this version, the appearance of L- somewhere between H\* and L% in focussed structures was justified because L- had no landing site. In most recent versions, however, the phrase accent has to be associated to an ip boundary and hence the only way to account for the contour is to propose that the L- spreads from the end of the ip leftwards (Beckman and Pierrehumbert 1986), as illustrated below.



If we analyse the Central Catalan data in the same way as the English data (that is with the L- spreading from the end of the sentence leftwards till the end of the focussed domain), the conception of the nucleus (as the last pitch accent of an ip) is

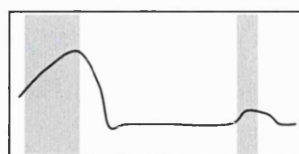
not undermined. However, there is a problem with the assignment of focus in Romance languages, since the focussed element would not be placed in phrase-final position. If, on the other hand, we postulate an ip boundary after the focussed element and hence the focal word is in final phrase position, we have problems with the phonological account of what is postfocal.

As will be observed in chapter 6, the results obtained for Central Catalan differ from English in that the L- is clearly located at the end of the focussed word and hence the stretch between the focussed item and the end of the sentence cannot belong to the same ip. Our proposal then is that there is a new ip after the focussed word. Evidence for this postfocal ip is found in the presence of postfocal accents. Postfocal accents are investigated in the next section.

#### 4.4.1.4 Postfocal accents

Most of the narrow focus utterances produced with accentual strategies obtained in the production test for all speakers showed deaccenting of the postfocal material. However, a smaller number of productions (22%) showed a pitch accent after the nuclear tone. This postfocal accent was consistently aligned with the stressed syllable of the last lexical word of the whole intonation phrase. The postfocal accent showed a severely compressed pitch range compared to that of the focal accent but presented the same tonal specification (peak described as H\*). This is illustrated in (4.25) and (4.26) below, which reproduce the examples of (4.8) and (4.9). (4.25) illustrates a postfocal accent in a subject focal sentence and (4.26) in a verb focus sentence.

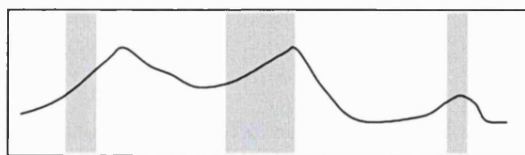
(4.25)



ELL anava a Girona

“HE went to Girona”

(4.26)



L'Elena mira de MOURE el mobiliari

"Elena tries to MOVE the furniture"

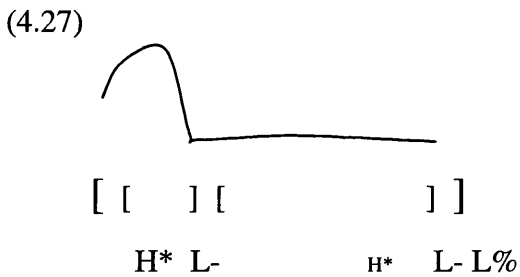
A perceptual test was designed to see whether the postfocal accent was perceived as a secondary accent and hence subordinate to the nuclear one (i.e. conserving the *s-w* pattern) or whether the postfocal accent was perceived as primary and hence having double focus. The perceptual test was performed by two Central Catalan speakers. They had to listen to 16 narrow focus sentences containing a postfocal accent. For each sentence, they had to decide how many focal points they could perceive. The test confirmed that the postfocal accent was considered as secondary since listeners perceived only the nuclear accent as the focussed element, instead of two focal points.

The presence of postfocal accents in Central Catalan justifies the claim that postfocal material is part of another ip, since within the AM model (Beckman and Pierrehumbert 1986) every ip has to have a minimum of one pitch accent. However, this raises the question of what is the status of postfocal accents: are they optional (i.e. sometimes they appear and other times they do not appear and there is deaccentuation), or is deaccenting a reduction or elision of an underlying postfocal accent? Postfocal accents have been observed in several Romance languages, such as, Italian (Grice 1995a, Grice and Savino 1997 and D'Imperio 1997b), Maltese (Vella 1995), Peninsular Spanish (García-Lecumberri 1995) and European Portuguese (Frota 1998). Frota (1998) claims that the postfocal accent is an obligatory event in Portuguese focus contours. Its apparent absence in some of the contours is due to a case of *extreme subordination* to the nuclear accent, rather than deaccenting.

Further research is needed to clarify the status of Central Catalan postfocal accents. However, from a phonological point of view, the appearance of a postfocal accent accounts for the presence of an ip after the focal material and allows us to propose that in Central Catalan, as in other Romance languages, the focal element is always placed at the end of an ip. Based on phonological grounds then, it seems plausible to



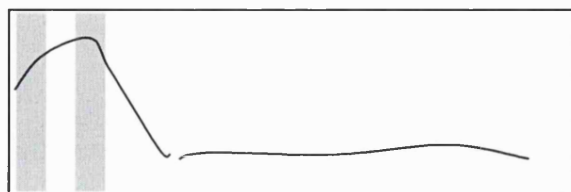
claim that there is always an underlying postfocal accent after the early focussed material in Central Catalan and that this postfocal accent tends to be reduced in most cases, although it can sporadically emerge. Thus, the notion of "postfocal deaccenting" in Central Catalan seems to be explained as a reduction of a pitch accent rather than as the lack of accentuation. With this proposal the problem of what is postfocal in Central Catalan seems to be solved since the presence of a postfocal accent confirms the presence of a postfocal ip, as schematised in (4.27). The small font for the postfocal accent indicates that this accent has a compressed pitch range which might be reduced to nothing.



#### 4.4.2 Syntactic focus

As reported in chapter 2 (section 2.3.3.2), some of the speakers in the semi-spontaneous task produced utterances where the accent/focus relationship was attained by means of a syntactic reorganisation. Two kinds of syntactic shifts were observed: 1) sentences with a right or left dislocation of the object and displacement of the focussed material to final sentence position and 2) a clefting structure. Both sentences showed a similar intonational pattern to that observed in sentences produced with accentual focus. An H\* pitch accent is associated to the stressed syllable of the focussed element followed by an L tone, marking the end of the accented word, and an L- tone marking the end of the ip. Examples of sentences with a syntactic reorganisation are provided in (4.28-4.30) below.

(4.28) *Cleft sentence*



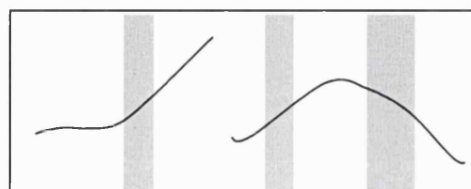
És l'ISIDRE qui du una gavardina vermella

L\*H H\*L L-

L-L%

“It is ISIDRE who wears a red waterproof”

(4.29) *Focus on sentence final position + left dislocation of the object*

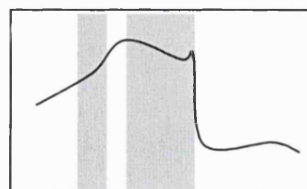


De robes negres, en duia l'AMÈLIA

L\* H H- L\* H H\* LL-L%

“Black clothes, it was AMELIA that wore them”  
(lit. black clothes, them wore AMELIA)

(4.30) *Focus on sentence final position + right dislocation of the object*



Hi bullia L'OU, a l'olla

L\*H H\* LL- L-L%

“It was THE EGG that boiled in the pot”  
(lit. in there boiled THE EGG, in the pot)

The similar intonational behaviour observed in sentences where focus is only signalled by accentual strategies and in sentences where focus is expressed by means of a syntactic reorganisation, as in the previous examples, corroborates the claim that focus in Central Catalan is marked by an ip boundary after the focal element.

## 4.5 Discussion

According to the results presented so far, the expression of narrow focus in Central Catalan declaratives has the following characteristics. First, an H\* pitch accent is associated to the stressed syllable of the focussed word. This accent is realised differently depending on the location of the focussed word. In *early* focus (focus on the subject and on the verb) the H\* shows a smooth rise over the accented syllable with the peak aligned towards the end of the syllable. In *late* focus (focus on the object) the H\* is realised as a progressive fall. In this case, the lack of a clear peak is accounted for by the sentence-level phenomenon of final lowering (see chapter 5 for more details). These results suggest that the expression of broad focus and narrow focus in Central Catalan is the same. As reported in Chapter 3, an H\* pitch accent was also used as the nuclear accent of neutral declaratives. In single-stressed utterances, the H\* was realised as a smooth peak and presented the same structure as early narrow focus. In multi-stressed sentences, the H\* was realised with final lowering, as in late focus. Sometimes H\* narrow focus nuclear accents are produced with emphasis (i.e. with an expanded F0 range). Emphatic accents, however, are not indispensable to convey narrow focus and their appearance depends on the speaker's choice. This differs from other Romance languages where emphatic accents have been incorporated into the phonological system as a distinctive category to express narrow focus.

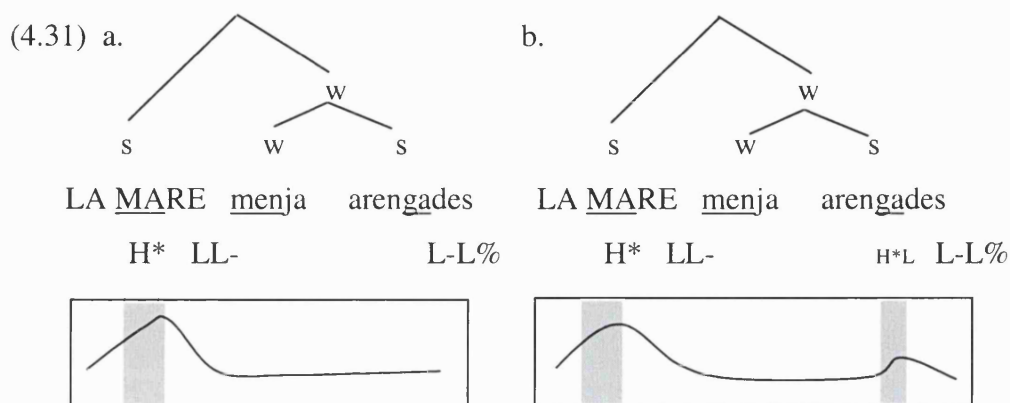
The use of the same pitch accent for the expression of both broad and narrow focus in Central Catalan disagrees with the results observed in other Romance languages, where two different pitch accents have been postulated to express broad and narrow focus. As presented in section 4.2, languages such as Palermo Italian (Grice 1995a), Bari Italian (Grice and Savino 1997) and European Portuguese (Frota 1998) use H+L\* to convey broad focus and H\*+L to express narrow focus. In Spanish, Sosa (1999) claims that L\*L% conveys a neutral intonation, whereas H\*L% and H+L\*L% a marked intonation. In this study, we propose that the same pitch accent (H\*) is used with the two focus readings and that possible differences in the realisation of this H\* are accounted for by the effects of sentence-level F0 downtrends (see chapter 5).

Second, in Central Catalan, narrow focus is also signalled by the presence of a prosodic boundary (higher than the word boundary) after the focussed word. The results suggest that this boundary corresponds to an ip boundary, which is signalled by the presence of a phrase accent (L-). The presence of a prosodic boundary after the focussed element is in line with the idea that in Romance languages focus has to be assigned at a prominent sentence position, namely, sentence/phrase final position (Zubizarreta 1998, Sosa 1999). However, the data showed that the location of the focussed element in a prominent position does not always have to be mediated by a syntactic shift, as claimed by Vallduví (1990, 1991, 1994a/b) and Vallduví and Zacharski (1994), but intonation alone can create a prominent position by introducing a prosodic boundary after the focussed item.

The location of an ip boundary after the focussed element seems to create some problems for intonational theory since the status of what is postfocal is unclear. The main problem is that if we assume that an IP is made up of one or more ips, the stretch of words after the focal ip has to constitute another ip. However, in most cases this putative ip does not have a overt pitch accent and the AM theory (Beckman and Pierrehumbert 1986) does not allow an ip without a pitch accent. Yet, the appearance of a postfocal accent, characterised as an H\* with a reduced pitch range, suggests that the idea of an ip after the focal ip is plausible. If this is the case, then, deaccenting in Central Catalan has to be explained as a reduction or elision of an underlying pitch accent rather than the lack of accentuation.

The possibility that the postfocal pitch accent was interpreted as the point of information (or major prominence) was ruled out by a perceptual test, which showed that postfocal accents were always perceived as secondary and never played the role of the nuclear accent. Thus, the prominence structure of a sentence produced with an evident postfocal accent seems to be the same as a sentence with a strongly reduced postfocal accent. This is illustrated in (4.31) for the sentence *LA MARE menja arengades* ("THE MOTHER eats herrings") produced without (or with a strongly reduced) postfocal accent (4.31a) and with an overt postfocal accent (4.31b). The fact that the two sentences have the same prominence pattern agrees with Ladd's (1980, 1996) "stress-first view", which involves that the distributions of accents within a sentence take place once the prominence relations between the elements of a sentence

have been established. The assignment of accent in relation to metrical prominence accounts for the presence of accents without focus, such as postfocal accents. Further research is needed on the status of postfocal accents in Central Catalan.



The third aspect observed in narrow focus utterances is that whenever present, prenuclear accents show a rising pitch which is characterised as L\*H. The L\* is associated to the accented syllable and the H is anchored at the end of the word. Prefocal elements tend to belong to the same ip as the focussed word, although they can also belong to another ip as in neutral sentences. Prenuclear accents in narrow focus sentences show the same pattern as in broad focus sentences.

Fourth, although the data showed that the tonal sequence H\*LL- was the most commonly used to express narrow focus in Central Catalan (both in contrast and in identification contexts), sometimes in contrastive sentences the H+!H\* pitch accent was also observed. This accent had a nuance of insistence or correction. In our data, H+!H\* was mainly used in late narrow focus to disambiguate between a broad focus and a late narrow focus reading. This accent was also reported in Prieto (1995). However, she also proposed three other tones that could be linked to a focussed word in Central Catalan: !H\* (used to express emphasis), L+H\* (used for surprise and contrast), L+H\*+H (used for contradiction). In this study, the interpretation of the data is slightly different from Prieto (1995). Here, Prieto's three phonological categories (!H\*, L+H\* and L+H\*+H) are reduced to H\*, which might be subject to realisational differences in terms of scaling and pitch range (as in emphatic accents).

Finally, the tonal notation we propose for some of the utterances presented in this chapter is recapitulated in the examples presented below. The postfocal accent is only marked when it is present in the contour (as in (4.34)). However, we assume that both the lack of an overt postfocal accent and the compressed postfocal accent are derived from the same underlying form.

(4.32)        En Joan DOMINA l'àlgebra  
                  L\* H    H\* LL-    L-L%

(4.33)        Hi bullia L'OU, a l'olla  
                  L\* H H\*LL-    L-L%

(4.34)        ELL anava a Girona  
                  H\* LL-        H\*L L-L%

#### 4.6 Conclusion

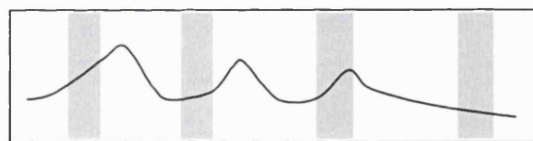
In this chapter, the phonetic and phonological properties of accentual focus have been examined. The results have shown that narrow focus in Central Catalan is expressed by means of an H\* pitch accent associated to the focussed word followed by an L word edge tone and an L- phrase accent, which demarcates the end of the focal domain. These observations confirm the idea that the assignment of focus in Romance language takes place in phrase-final position. However, they also show that intonational means alone can signal focus by creating a prosodic boundary and that there is no need for a syntactic reorganisation. The nuclear accent used with a narrow focus reading is the same as that with a broad focus reading (H\*). Differences in the realisation of H\* are accounted for by sentence-level phenomena, such as final lowering. However, an H+!H\* pitch accent is sometimes observed in contrastive focus with a meaning of correction. Finally, postfocal pitch accents are observed in the data, indicating that there is another ip after the focussed domain and that deaccenting can be explained as a reduction of an underlying accent.

## Chapter 5: F0 downtrends

### 5.1 Introduction

In the last two chapters, the F0 contours of Central Catalan utterances produced with broad and narrow focus have been described as a sequence of two distinctive levels, H and L, associated to different metrically relevant positions. Thus, for example, the F0 contour represented in (5.1) has been modelled as a series of 1) *pitch accents*, associated to accented syllables, 2) *word edge tones*, anchored at the right edge of an accented (or prosodic) word, 3) *phrase accents*, which delimited an ip boundary, and 4) *boundary tones*, which marked the end of the whole IP. Between these tonal events, the contour is phonologically unspecified but the overall F0 shape can be predicted by phonetic interpolation rules. Thus, for example the interpolation between H and L is expected to involve a falling movement, and the transition between L and H a rising movement.

(5.1)



Les nenes haurien de donar l'enhorabona

L\*H    L\*H    L\*H    H\* LL-L%

"The girls should send their congratulations"

The description of the F0 contours in terms of a string of H and L tones does not mean that these tones have to be realised always in the same way. The phonetic realisation of H and L, as the phonetic realisation of any phoneme, is subject to a variety of conditioning factors. One of these factors is the position of a particular tone within the utterance. It has been observed that F0 tends to decrease over the course of an utterance. This means that the scaling of a given tone will vary depending on its sentence position. This is illustrated in (5.1). The four H tones observed in this utterance are not scaled at the same level but each H is lower than the preceding one. Even though the tendency of an F0 decline over the course of an utterance is a well-known property of speech, the interpretation of such a downtrend varies in the

literature. Whereas sometimes the downward trend of pitch has been analysed as a global effect, conceived as a component of the overall contour, at other times it has been treated as a local, phonologically controlled mechanism, which affects accents individually and in relation to previous accents. The AM framework (Pierrehumbert 1980, Liberman and Pierrehumbert 1984) represents the most explicit case of a sequential interpretation of F0 downtrends as a string of step accents.

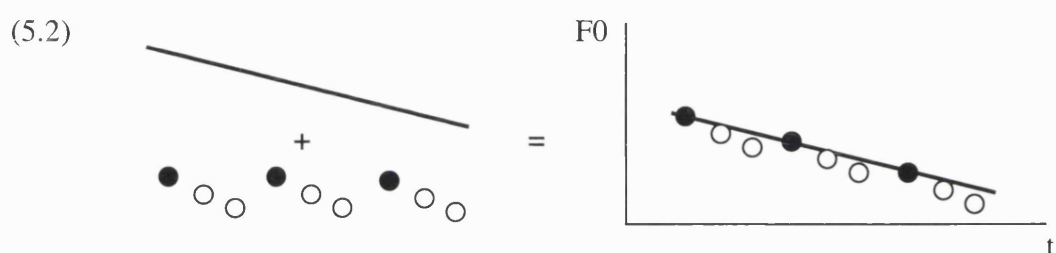
In this chapter, pitch downtrends will be analysed in Central Catalan. In particular three main aspects will be covered: 1) the lowering of peaks or H tones, 2) the lowering of valleys or L tones, and 3) the effects of focus on F0 downtrends. The organisation of the chapter will be as follows. Section 5.2 describes the different proposals for the analysis of F0 downtrends, basically, the global vs the local interpretations. Section 5.3 analyses the downward trend of peaks in Central Catalan declaratives produced both with broad and narrow focus. Some of the issues covered in this section are 1) the interpretation of F0 downtrends as a gradual declination effect or as a controlled downstep mechanism, 2) the modelling of peak downtrends within the AM framework in Central Catalan and 3) the relationship between narrow focus and the blockage of pitch downtrends. Section 5.4 investigates the nature of F0 valleys in broad focus utterances and compares it to that of narrow focus sentences. Finally, section 5.5 presents some concluding remarks for the whole chapter.

## **5.2 Global and local interpretations of F0 downtrends**

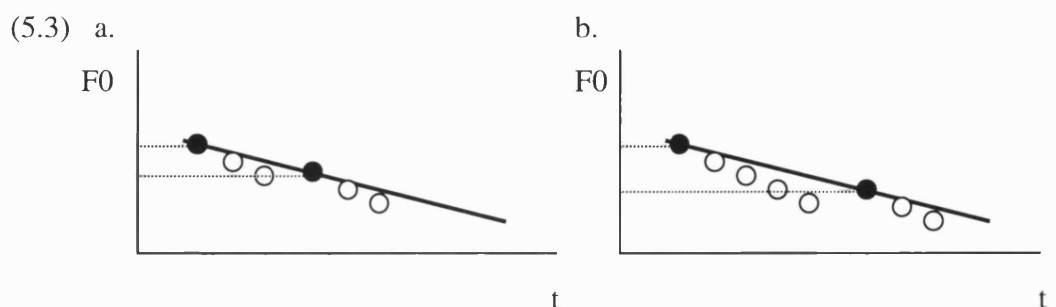
One of the most controversial issues in the study of intonation is the interpretation or modelling of declination. Declination has been described as the tendency of F0 to gradually lower over the course of an utterance. Although this pitch downtrend has been widely studied and recognised in many languages, researchers still disagree on its nature. Two major views have been followed in the description of F0 downtrends, referred to by among others Ladd (1983b) and Nolan (1995) as the *Contour Interaction* model and the *Tone Sequence* model (henceforth CI and TS respectively). Researchers within the CI tradition (Fujisaki 1983, 1988, Thorsen 1980, 1983, Bruce 1977, Liberman 1967, Cooper and Sorensen 1981, Gårding 1983, Vaissière 1983, among others) view declination as one of the components that interact to generate a given pitch contour. Within this proposal, the F0 contour of an utterance is specified



as a number of separate domains that combine to create specific pitch configurations. At least two kinds of layer are specified in all works: 1) a global component or overall line that affects the whole phrase or utterance (declination component) and 2) a succession of local F0 movements or accent units (accent layer). This is illustrated in (5.2) below (as represented in Nolan 1995). The sloping line stands for the global declining component. Filled circles represent accented syllables and open circles unaccented syllables. The final F0 contour is a combination of the global component and the accent layer. In the final contour, each accent is much lower than the preceding one due to the effect of the global component.



Within the CI model, the temporal distance between two consecutive peaks is predicted to have an effect on the F0 value of the second peak, that is, the greater the distance between two consecutive peaks, the lower the F0 level of the second peak. Thus, in this model, pitch drop can be analysed as an effect of the time interval between peaks. This is illustrated in (5.3) below, where the F0 height of a peak is lower as the time interval between peaks increases. In (5.3b) the second accent (second filled circle) is much lower than its equivalent in (5.3a), due to a greater temporal distance between accents.



The TS approach, on the other hand, analyses pitch downtrends as a linguistically controlled mechanism, which involves a deliberate use of step accents to attain



between the Hs) as in the sequence HH, the second H does not show the same amount of lowering. Pierrehumbert (1980) showed that the behaviour observed in English declaratives was similar to that observed in tone languages. She claimed that any HLH sequence, as for instance H\*+L H\*, involves a downstepped second H due to the presence of the L tone. This L is not manifested as an F0 valley but it acts as a downstep trigger of the following H tone.

Lieberman and Pierrehumbert (1984) proposed that in English downstepping contours the height of a given peak could be modelled by using a constant F0 reduction of the previous peak value. Thus, the amount of decay between peaks was calculated as a proportion of the second peak with respect to the first one, scaled above the reference line of the speaker. The reference line was described as an abstract line, which lies between the last peak of the utterance and the F0 minimum (see Liberman and Pierrehumbert 1984 for more details and also section 5.3.1.2.3 in this chapter). Additionally, Liberman and Pierrehumbert found out that the final peak of a declarative sentence undergoes a more drastic lowering in F0 than that expected by the application of the downstep rule. They proposed to account for this final lowering by means of a lowering constant, defined as a fraction of the value of the peak predicted by the downstep rule. The modelling of F0 downstepping contours by means of a downstep ratio and a final lowering constant has been applied to several languages, such as, British English (Ladd 1983b, 1984), Japanese (Beckman and Pierrehumbert 1986, Poser 1984), Mexican Spanish (Prieto et al 1996), German, (Möbius 1993, Grabe to appear) and Dutch (van den Berg et al 1992).

Some studies on intonation (Pierrehumbert and Beckman 1988, Poser 1984, Fujisaki 1983, 1988) propose that the analysis of pitch downtrends as a sequence of localised phonological events does not exclude the existence of a global declination effect and suggest that both phenomena, declination and downstep, can coexist in the same utterance. Also Grabe (to appear) suggested that what Liberman and Pierrehumbert (1984) interpreted as final lowering in American English sentences seems to be an effect of declination in British English utterances, since the greater F0 reduction observed on the utterance-final peak can be explained by a greater time interval between peaks. In the studies that claim the existence of the two mechanisms, declination is considered a residue downtrend after all other predictable lowering

mechanisms have applied. In this chapter, the nature of Central Catalan F0 downtrends will be examined.

Finally, studies on F0 downtrends have shown that the appearance of a focussed item can have an effect on downstepping sequences. Usually focus suspends the realisation of downstep. For example, Beckman and Pierrehumbert (1986) showed that in the American English utterance presented in (5.6) below, the presence of narrow focus on the word *eighty* blocks the downstepping sequence.

(5.6)

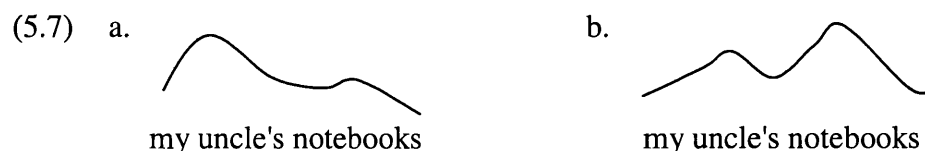


It's eleven and nine and one and EIGHTY

In several Romance languages, the relationship between broad focus and downstepped accents and narrow focus and non-downstepped accents is so consistent that some researchers have postulated different accent types for downstepped and non-downstepped accents. As pointed out in the last chapter, in European Portuguese (Frota 1995, 1998) or Palermo Italian (Grice 1995a) the "downstepped" accent observed in broad focus has been described as H+L\* and the "non-downstepped" accent of narrow focus as L+H\*. The categorical distinction between accent types that express broad focus and those that convey narrow focus avoids ambiguities between a neutral reading and a late narrow focus reading.

However, the relationship between the presence of narrow focus and the blockage of downstep is not always fulfilled. Ladd (1996) claims that for English both downstepped and non-downstepped accents can express a broad focus reading. For example, even though the F0 contours presented in (5.7) show different behaviour in terms of downstep (the second peak in (5.7a) is downstepped and in (5.7b) is not downstepped), their semantic interpretation is the same: either broad focus on the whole sentence or narrow focus on *notebooks*. According to this, patterns (5.7a) and (5.7b) are linguistically identical, the only difference being that of emphasis. In (5.7b),

*notebooks* is produced with an emphatic stress. In this chapter the relationship between focus and downstep will be examined for Central Catalan.



Overall, this chapter will consist of three main parts. First, we will investigate whether the F0 downtrend observed in Central Catalan neutral declaratives responds to a global, time-dependent declination effect or to a local linguistically-controlled downstep mechanism. This will be done by comparing the F0 maximum of consecutive peaks separated by different temporal distances (or number of intervening syllables). If declination is active, a greater amount of F0 lowering is expected as the temporal distance between Hs increases. Alternatively, if the height of peaks is stable and not affected by temporal distance, then the pitch downtrend would be the result of a controlled downstep or accent-by-accent decay. Anticipating some of the results, it will be observed that Central Catalan downtrend utterances suggest the presence of a downstepping mechanism rather than a declination effect. Given this evidence, we shall investigate to what extent the TS approach, developed within the tenets of the AM framework, can be used to predict the location of peaks in Central Catalan downstepping contours.

The second aim of this chapter is to examine the effects of narrow focus on pitch downtrends. As has already been mentioned, a focussed element tends to block the downstepping sequence. In this chapter, we will analyse to what extent the blockage of downstep in Central Catalan marked sentences is exclusively linked to the presence of narrow focus or to other paralinguistic phenomena, such as emphasis, and how the model can predict downstep in narrow focus.

Finally, the third goal of this chapter is to analyse the behaviour of the lower edge of the tonal space and in particular to confirm the phonological status of valleys proposed in chapters 3 and 4 for both neutral and marked declaratives. In sentences with a broad focus reading, prenuclear valleys have been described as L\* tonal targets on the basis of their alignment evidence. On the other hand, the F0 dipping observed

before the H\* nuclear accent of narrow focus sentences has been analysed as the result of a phonetic transition between peaks but not as an intended low target. Given this distinction, we would expect to find the following behaviour. When valleys are just transitions between H tones, the amount of dipping between two peaks is expected to increase as the temporal distance increases. When valleys are tonal targets, no correlation is expected between the temporal distance between peaks and the amount of dipping. In this case, Ls are expected to have a rather stable scaling.

### **5.3 The scaling of peaks**

In this section, the nature of peak downtrends in Central Catalan declarative utterances is analysed. The goal of this section is twofold: 1) to determine whether the F0 descent observed in the utterances is a gradual time-dependent declination effect or a linguistically-controlled downstep, and 2) to model the downstepping contours within the tenets of the AM framework (basically the TS tradition of downtrend modelling). In order to analyse the nature of downtrends in Central Catalan declaratives, we will examine the scaling of F0 values as a function of temporal distance between H targets in sentences of different lengths (three stresses and four stresses).

#### **5.3.1 Broad focus sentences**

##### **5.3.1.1 Data**

The data used for the analysis of pitch downtrends in Central Catalan declaratives involves the broad focus sentences gathered in the second recording session for 4 speakers (CP, DV, MC and NG), with the exception of utterances produced with an ip boundary between the subject and the predicate. Those sentences were excluded from the analysis of F0 downtrends because a pitch reset was observed after the first ip. Thus, the utterances analysed in this section involve all sentences of section 1.1 in Appendix 1 (produced with a neutral intonation), except those consisting of a double stressed subject, which were uttered with an ip boundary for most speakers. The data of speakers NM and ER were not analysed because they produced the ip boundary between the subject and the predicate in most cases. Thus, the number of sentences examined for each speaker is 48. This makes an overall number of 192 sentences.

The number of stresses (and hence potential accents) per sentence ranged between three and four. The number of sentences with three stresses was 25 and that of four stresses 23. The number of unstressed syllables between stressed syllables varied between one and two. An example of each kind of sentence is illustrated in (5.8).

- (5.8) a. *Three stresses*  
La Mireia remena l'olla  
"Mireia stirs de casserole"
- b. *Four stresses*  
Les noies volen rebre la reina  
"The girls want to welcome the queen"

The use of these data for the analysis of pitch downtrends in Central Catalan has advantages and disadvantages. One of the major drawbacks is that sentences have not been specifically designed for the analysis of F0 downtrends. Thus, for example, the number of unstressed syllables (or the temporal distance) between peaks has not been controlled. Moreover, the speakers' productions were rather variable in terms of the number of accents. However, the use of this corpus of data has the benefit of spontaneity since all downstepping sequences were naturally and unconsciously produced. Furthermore, as will be discussed in section 5.3.1.2.2, the variability observed in some of the speakers' productions concerning the number of accented syllables per sentence was favorable since appropriate contexts to analyse the effect of temporal distance on F0 descents were created.

### 5.3.1.2 Results

For each sentence, the following points were labelled in the F0 contours:

1. Phrase-initial F0 value
2. Highest F0 value for each pitch accent
3. Lowest F0 value between peaks

4. Phrase final F0 value<sup>1</sup>

The mean schematised F0 contours of utterances with different numbers of stresses (and accents) are presented in Figure 5.1 for all speakers. In each graph, the number of stresses and the number of accents (in brackets) are specified. Thus, for example, “4(3)” stands for a sentence with four stresses but only three accented syllables. The L and H letters indicate the mean values of F0 minimum and F0 maximum obtained in the analysis and also the tonal interpretation proposed for those values. This explains why sometimes a quite low F0 value is marked as H. The metrical structure of H and L is not specified. This implies that when we compare the scaling of two Hs or two Ls, they can be pitch accents or edge tones indistinctively. Finally, when a given F0 point is specified by means of two letters (as in L/H), the first letter corresponds to the sentence with more accents and the second letter to the sentence with less accents.

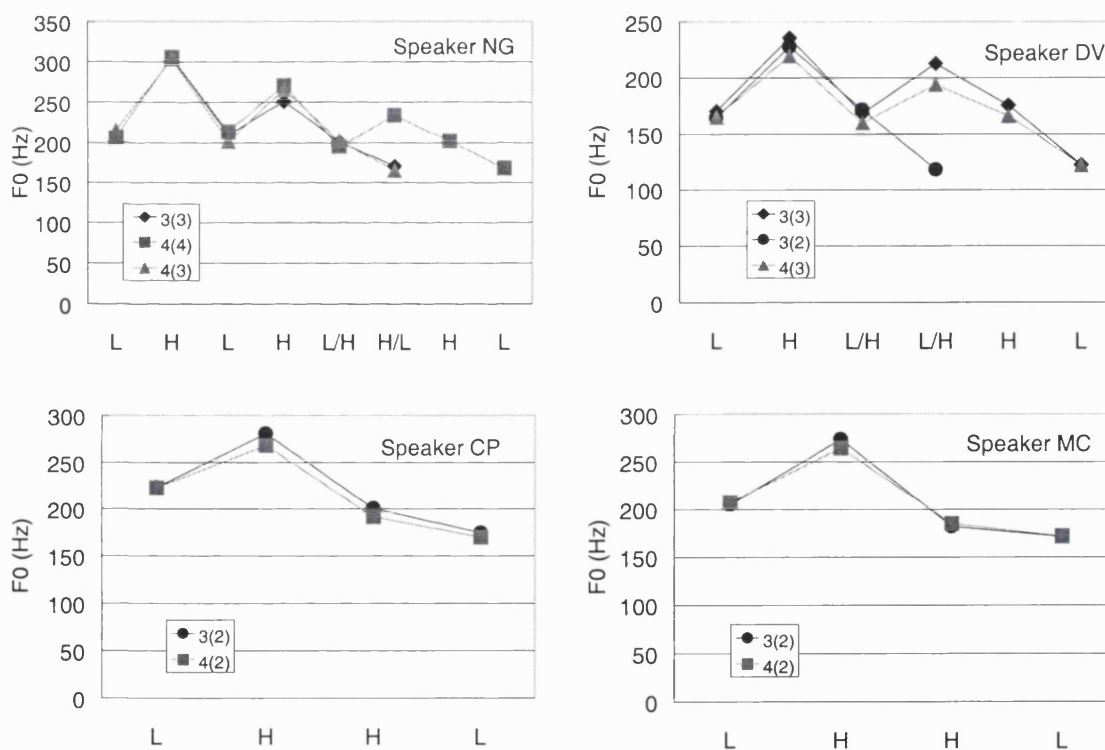


Figure 5.1. Mean schematised F0 contours for neutral declarative sentences with different number of stresses and accents (in brackets) for four speakers.

<sup>1</sup> Irregularities at the beginning and at the end of the F0 contour were ignored when identifying these points.



The results displayed in the graphs show inter-speaker variability as far as the number of accented syllables per sentence is concerned. Whereas speakers NG and DV tended to accent most of the potentially accentable syllables, speakers CP and MC only accented the first and last syllable of each utterance. This means that most of the research on Central Catalan downtrends will be performed on the data of speakers NG and DV. At first glance, the F0 descent observed in their schematised contours show a gradual F0 lowering of peaks and a more constant F0 value for valleys, except for the last L which is lower than the others. For all speakers, the last H is scaled at a very low frequency close to the F0 level to non-final L values. In the next sections, the scaling of peak F0 values in Central Catalan declaratives is examined in more detail. The behaviour of valleys is analysed later on in this chapter.

#### **5.3.1.2.1 Utterance initial and final F0 values and pitch range**

Before analysing the nature of Central Catalan F0 downtrends, it was considered necessary to see whether differences in sentence length had an effect on the utterance initial and final F0 values and on the pitch range with which sentences were produced. Whereas in some studies (Thorsen 1980, Cooper and Sorensen 1981), utterance length seems to have an effect on the F0 range of the first peak (and subsequently of all other peaks), in other studies (Lieberman and Pierrehumbert 1984, Prieto et al 1996) sentence length does not influence the degree of prominence in sentences. Since in our data sentences with a different number of stresses (and hence with different length) are compared, we first need to prove that pitch accents are actually comparable and hence that utterance length does not affect pitch range. In order to do so, utterance-initial and utterance-final F0 values were calculated as well as the scope of the first F0 rise (measured as the difference in Hz from the lowest point at the beginning of a contour to the highest F0 value of the first peak). If the F0 excursion between the initial F0 value and the first peak is the same in sentences with different length, this will show that utterances with a different number of stresses are produced with the same pitch range and hence accents are comparable.

Figure 5.2 shows the mean values of utterance-initial and utterance-final F0 values for sentences of difference length (i.e. with 3 and 4 stresses). The results show that the initial and final F0 values are independent of the utterance length for all speakers. T-

tests performed for each speaker showed no significant differences in utterance-initial F0 values for phrases of different lengths (where  $p < 0.01$ ): CP ( $p = 0.03$ ,  $t = 1.91$ ), DV ( $p = 0.23$ ,  $t = 0.72$ ), MC ( $p = 0.29$ ,  $t = -0.55$ ) and NG ( $p = 0.49$ ,  $t = -0.01$ ). Similarly, utterance-final F0 values were also rather constant and no significant differences were observed for any speaker: CP ( $p = 0.07$ ,  $t = 1.44$ ), DV ( $p = 0.18$ ,  $t = -0.9$ ), MC ( $p = 0.47$ ,  $t = -0.06$ ) and NG ( $p = 0.14$ ,  $t = 1.06$ ). These results are similar to those found by Liberman and Pierrehumbert (1984) for English and Prieto et al (1996) for Mexican Spanish: utterance-initial and utterance-final F0 values were nearly constant for a given speaker despite differences in sentence length.

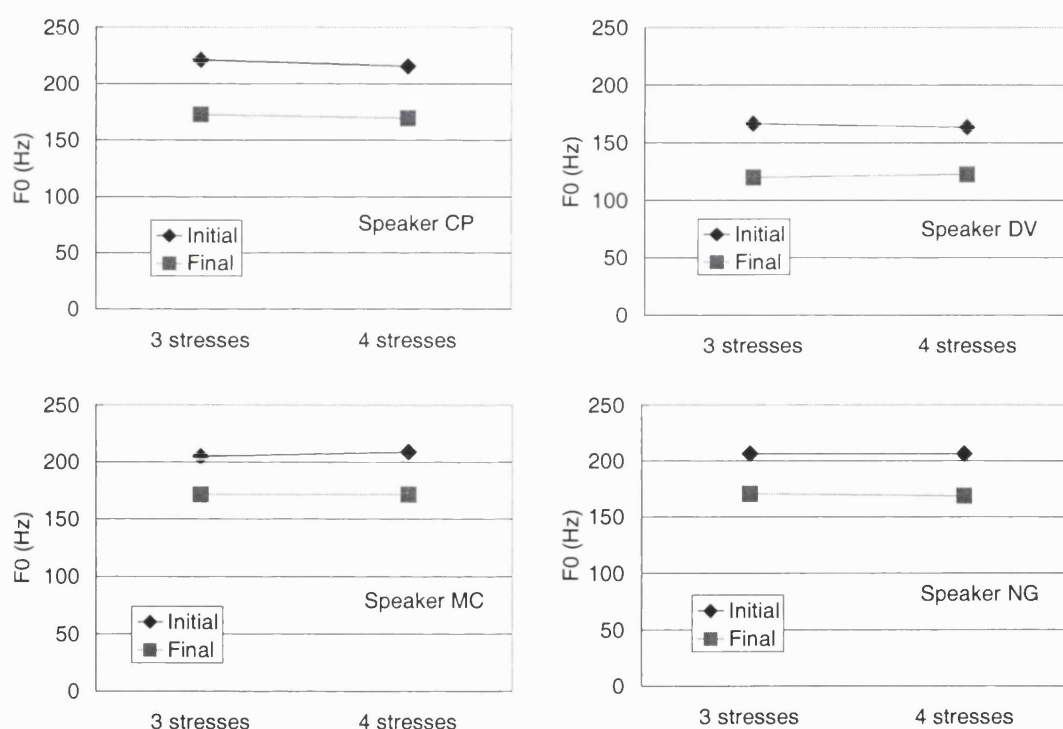


Figure 5.2. Mean values (in Hz) of utterance-initial and utterance-final F0 values of sentences of different lengths (3 to 4 stressed syllables).

Given the constant initial F0 values observed for all speakers, F0 range was calculated as the difference in Hz from the lowest F0 value at the beginning of an utterance to the highest F0 point of the first peak, as in Prieto et al (1996). The results are displayed in Table 5.1, which shows the mean F0 range in utterances with three and four stresses and the results of t-tests comparing the F0 range values for all speakers.

Stresses	Speaker CP			Speaker DV			Speaker MC			Speaker NG		
	F0 range	p	t	F0 range	p	t	F0 range	p	t	F0 range	p	t
3	80.4	0.23	0.74	65.4	0.06	1.6	88.9	0.02	2.11	99.4	0.48	-0.03
4	75.9			56.1			80.5			99.6		

Table 5.1 Mean F0 range in utterances with three and four stresses and results of t-tests comparing the F0 range values for all speakers.

The results presented in Table 5.1 show that utterances were produced with a rather constant pitch range irrespective of the number of stresses per sentence. For all speakers, t-tests comparing the F0 range values in utterances with three and four stresses show that differences in pitch range are not significant ( $p > 0.01$  for all speakers). This means that H values in sentences with different lengths are comparable, since no pitch range variation is observed in the data.

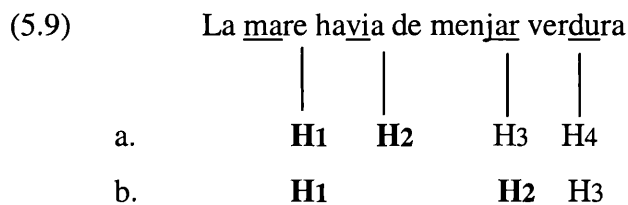
### 5.3.1.2.2 Effects of temporal distance between peaks

One of the strategies used by some researchers (Prieto et al 1996, Grabe to appear, among others) to detect whether declination is active in downtrend F0 contours is the analysis of the temporal distance (examined in terms of number of intervening syllables) between pitch accents. If declination is present, the F0 reduction is expected to be bigger as the time interval (or number of intervening syllables) between peaks increases. Otherwise, if there is no declination, the F0 reduction is expected to have a constant value irrespective of the time interval between peaks.

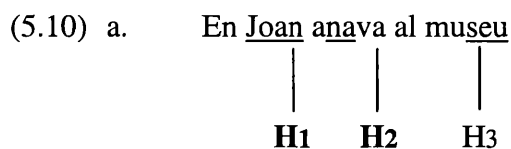
Since our first aim in this study was not to examine the characteristics of downtrend in Central Catalan, we did not control the number of unstressed syllables (or time distance) between potential peaks (or stresses). Usually the number of unstressed syllables ranges between one and two in our data and this is not enough to see the effects of time distance on the downward movement. However, since most of the speakers did not accent all the potentially accentable syllables but showed a rather heterogeneous behaviour in the number of peaks, two appropriate contexts to examine declination appeared in the results.

Context 1

The first context is observed in the productions of speakers NG and DV. It involves the analysis of H1 (first peak) and H2 (second peak) in sentences where all stresses got an accent as opposed to those sentences where not all lexical stresses became accented. For example, sentences with four-lexically stressed words, as the one illustrated in (5.9), are sometimes produced with four pitch accents (as in (5.9a)) and sometimes with three (as in (5.9b)) for the same speaker. This means that the number of intervening syllables (and hence temporal distance) between H1 and H2 is bigger depending on whether a peak is observed at the end of *havia* or not. When all pitch accents are produced, the number of intervening syllables between H1 and H2 varies from one or two. However, when a pitch accent is not produced, the number of intervening syllables increases from three to five. This case seems to provide an appropriate context to analyse declination. If declination is present, there should be a larger F0 reduction of the second peak as the temporal distance between peaks increases. If there is no declination, the amount of lowering of the second peak is expected to remain constant despite differences in the time interval between peaks.



Similarly, since the first peak has been proved to have a similar scaling irrespective of the length of the sentence, another way to look at the same aspect is to compare three-stressed/three-accented sentences with four-stressed/three-accented sentences, as illustrated in (5.10). As before, the number of syllables (and temporal distance) between H1 and H2 is bigger in (5.10b) than in (5.10a) and hence the effects of temporal distance on F0 scaling can be tested.





syllables the longer the time to produce them. In the same way, for speaker NG differences in the mean duration of 1-2 syllables in conditions (4,4) and (3,3) are non-significant ( $p=0.3$ ,  $t=0.5$ ). This proves that the analysis of declination with respect to the number of intervening syllables is reliable.

Once it had been confirmed that differences in the number of syllables involve differences in time, we compared the mean peak height of H2 preceded by 1-2 and 3-5 unaccented syllables as well as the amount of pitch drop (F0 mean difference) between H1 and H2. Table 5.3 presents the mean F0 values of H2 according to the number of intervening syllables for the two speakers. Table 5.4 exhibits the mean F0 difference between H1 and H2. Finally, Figure 5.3 shows the mean F0 values for H1 and H2 corresponding to an increase in the number of intervening syllables.

	1-2 (4,4)	3-5 (4,3)	t-test	1-2 (3,3)	3-5 (4,3)	t-test
NG	270.6	265.4	$p=0.31$ $t=-0.48$	250.3	265.4	$p=0.08$ $t=1.4$
DV	-	-	-	212.9	190.8	$p<0.001$ $t=4.6$

Table 5.3. Mean F0 values of H2 preceded by 1-2 or 3-5 unaccented syllables for the two speakers in different sentence conditions.

The results of Table 5.3 show that for speaker NG no significant differences are observed in the mean F0 value of the second peak preceded by 1-2 or 3-5 unaccented syllables. This seems to indicate that for this speaker declination is not present in her productions. Speaker DV, on the other hand, exhibits a significantly lower F0 value in H2 when the number of intervening syllables increases. We might first think that for this speaker declination is present in addition to downstep. However, if we look at Figure 5.3, we see that not only H2 but also H1 is lowered in her (4,3) sentences. This seems to indicate that the significant differences between the two H2s in Table 5.4 are not due to the differences in the number of intervening syllables, but to the fact that the whole pitch range of (4,3) sentences is a little lower than (3,3) sentences<sup>2</sup>. This is confirmed by the results displayed in Table 5.4, where the mean F0 difference between H1 and H2 with 1-2 and 3-5 intervening syllables is not significant for

<sup>2</sup> This was already observed in Table 5.1 where the F0 range between the utterance-initial F0 value and the first peak was lower in utterances with 4 stresses than in utterances with 3 stresses. This difference, however, was shown to be not significant ( $p=0.06$ ).

speaker DV. Similarly, no significant differences are observed for speaker NG. Even more, in condition (3,3) vs (4,3) speaker NG shows a higher amount of pitch drop when the number of intervening syllables is smaller than when it is bigger. This corroborates the idea that there is no declination effect in the data.

	1-2 (4,4)	3-5 (4,3)	t-test	1-2 (3,3)	3-5 (4,3)	t-test
NG	35.2	37.8	p=0.34 t=-0.4	55.3	37.8	p=0.04 t=1.78
DV	-	-	-	22.7	26.2	p=0.12 t=-1.27

Table 5.4. Mean F0 difference between H1 and H2 with 1-2 and 3-5 intervening syllables for the two speakers and in different sentence conditions.

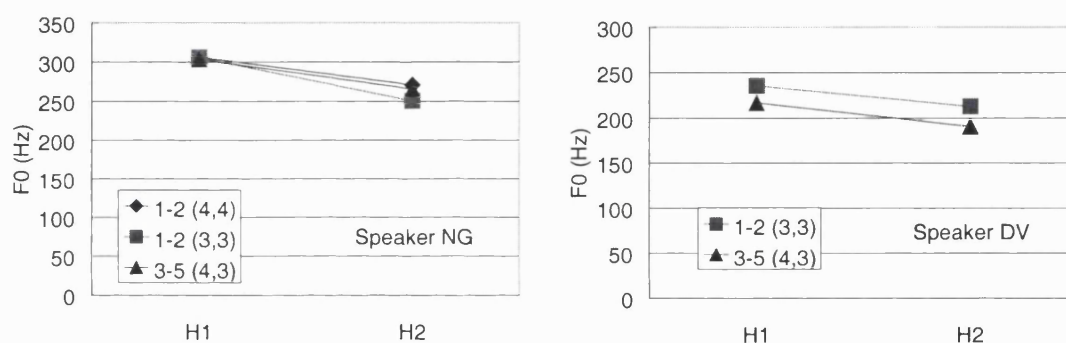


Figure 5.3. Mean F0 values for H1 and H2 corresponding to an increase in the number of intervening syllables (1-2 or 3-5) and in different sentence conditions for speakers NG and DV.

### Context 2

The second context to analyse the presence/absence of declination is observed in those speakers (CP and MC) who produced only two pitch accents in sentences with three and four stresses. In this case, the number of syllables between H1 and H2 varies between 3 and 5 in three-stressed sentences and between 6 and 9 in four-stressed sentences. As before, the temporal distance between the first and last pitch accent is expected to be longer in four-stressed than in three-stressed utterances. This is confirmed in Table 5.5, where the mean duration distance between H1 and H2 is significantly longer for the two speakers as the number of intervening syllables increases (CP:  $p < 0.001$ ,  $t = 8.9$ ; MC:  $p < 0.001$ ,  $t = 10.9$ ).

	Number of syllables (sentence condition)	
	3-5 (3,2)	6-9 (4,2)
CP	0.602	0.950
MC	0.622	1.028

Table 5.5. Mean duration (in seconds) between peaks with 3-5 intervening syllables and with 6-9 intervening syllables in sentences with three and four stresses but two pitch accents.

As before, once it had been confirmed that differences in the number of syllables correspond to differences in time, the declination effect was tested in the data of these speakers by means of the same hypothesis: if the scaling of the last pitch accent decreases as the number of intervening stresses increases, declination is present. If, on the other hand, the last pitch accent shows no differences in F0 height or in the amount of pitch drop, declination is absent.

Table 5.6 presents the mean F0 values of H2 according to the number of preceding unaccented syllables for the two speakers. Table 5.7 exhibits the mean F0 difference between H1 and H2. Finally, Figure 5.4 shows the mean F0 values for H1 and H2 corresponding to an increase in the number of intervening syllables.

	3-5 (3,2)	6-9 (4,2)	t-test
CP	200.7	191.5	p=0.001 t=3.2
MC	181.6	184.7	p=0.06 t=-1.58

Table 5.6. Mean F0 values of H2 preceded by 3-5 or 6-9 unaccented syllables for the two speakers.

	3-5 (3,2)	6-9 (4,2)	t-test
CP	80	76.6	p=0.23 t=0.74
MC	91.8	79.6	p=0.002 t=3.08

Table 5.7. Mean F0 difference between H1 and H2 with 1-2 and 3-5 intervening syllables for the two speakers and in different sentence conditions.

The results of Table 5.6 show that for speaker MC, no significant differences are observed in the mean F0 value of H2 according to the number of preceding syllables. This indicates the lack of declination in the productions of MC. For speaker CP, on



the other hand, H2 is significantly lower when the number of preceding syllables is bigger. However, this does not indicate the presence of declination since the mean F0 difference between H1 and H2 displayed in Table 5.7 for this speaker shows no significant results. Even more, both speakers exhibit a tendency which contradicts the declination hypothesis. The pitch drop is slightly bigger with a shorter time interval (3-5 syllables) than with a longer duration (6-9 syllables). This corroborates the findings reported for speakers NG and DV. Thus, no declination effect is observed in Central Catalan neutral declaratives. This suggests that downtrend contours in this language originate from the repeating occurrence of downstep accents, rather than from a global declination component.

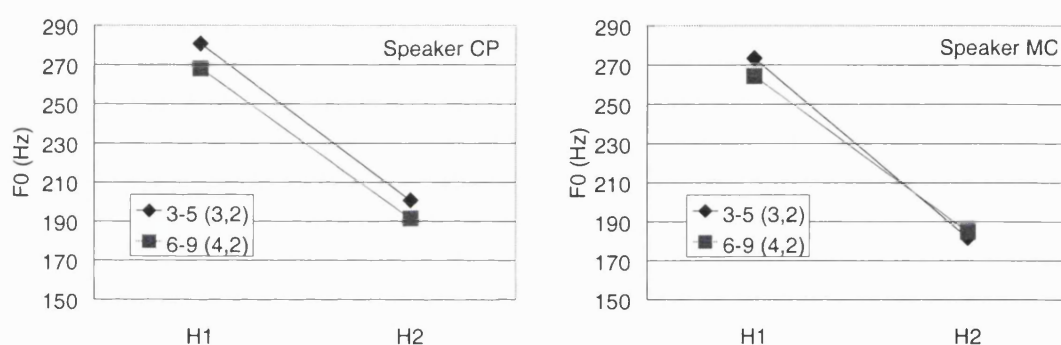


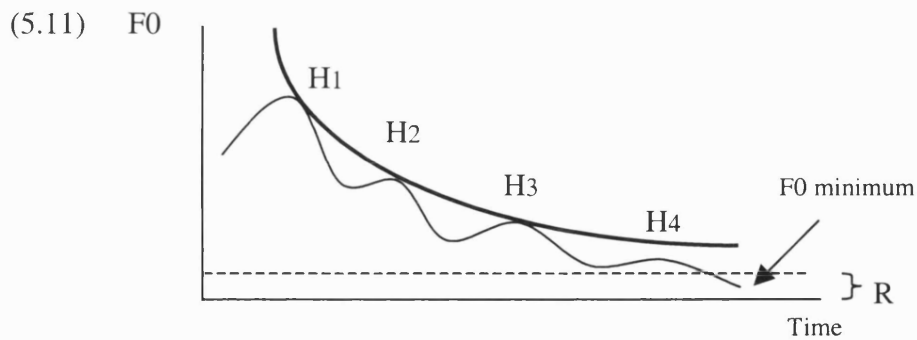
Figure 5.4. Mean F0 values for H1 and H2 corresponding to an increase in the number of intervening syllables (3-5 or 6-9) and in different sentence conditions for speakers CP and MC.

### 5.3.1.2.3 Predicting F0 values

The results observed so far show that Central Catalan downtrend patterns have to be explained by a linguistically controlled downstep rather than by a gradual declination effect. This supports the proposal of the TS view, which analyses pitch downtrends as the result of an accent-by-accent decay. Given these results, the next aim is to find out whether the model proposed by Liberman and Pierrehumbert (1984) to predict downstep in American English can be used to determine the scaling of peaks in Central Catalan downstepping sequences. Liberman and Pierrehumbert's model is examined since it has successfully predicted the scaling of downstepped accents in several languages, namely, Dutch (van den Berg et al 1992), Mexican Spanish (Prieto et al 1996), Japanese (Beckman and Pierrehumbert 1986, Poser 1984), and German, (Möbius 1993, Grabe to appear). The ability of Liberman and Pierrehumbert's model

to predict the F0 height of downstepped peaks in Central Catalan can only be tested in the utterances of speakers NG and DV, who were the ones that produced sequences of downstepped accents in neutral declaratives. As observed in Figure 5.1, speaker DV produced sentences with 3 peaks and speaker NG with 3 and 4 peaks. In this section, only the prediction of peaks in non-final utterance position will be analysed since the last peak of a sentence exhibited a much more abrupt F0 decay, which will be dealt with in section 5.3.1.2.4.

Liberman and Pierrehumbert (1984) noticed that in American English downstepping contours the height of a given peak was dependent on the F0 value of the preceding peak. That is, a correlation was observed between the height of two adjacent peaks: in general, the higher the first peak, the higher the second peak. Given this evidence, they decided to model downstep as an exponential decay to a constant nonzero asymptote, which they called the *reference line* of the speaker. The reference line (R) is a value lying between the last peak and the speaker's F0 minimum (i.e. final F0 value). Liberman and Pierrehumbert decided to introduce the notion of the reference line in their model to avoid predicting peaks that fall below the speaker's F0 minimum. The reference line and the F0 minimum of an asymptotically downstepping contour are illustrated in (5.11).



Then, the ratio of decay between adjacent peaks (or downstep constant) is calculated by using the following equation:

$$\text{Downstep ratio } (r) = (P(x+1)-R)/(P(x)-R)$$

where  $P(x)$  = peak height of a peak in a position  $x$ , and  $R$  = the reference line value.

For each speaker, the reference line was calculated as the mid value between the mean F0 of the last "peak"<sup>3</sup> and the mean F0 minimum for all sentences. For speaker NG, R=184Hz and for speaker DV, R=145Hz. Once the reference values were obtained, the downstep ratios (r) between peaks were calculated for each sentence using the equation above. The means of downstep ratios are 0.64 for speaker NG and 0.7 for speaker DV.

Having obtained these values, the F0 height of a given peak was calculated as a constant fraction (downstep ratio) of the previous one, scaled above the reference line of each speaker. This was obtained by means of the following equation:

$$P(x+1) = R + r \cdot (P_x - R)$$

where  $P(x)$  = peak height of a peak in position  $x$ ,  $r$  = downstep ratio, and  $R$  = reference line.

In Figure 5.5, the F0 values predicted by the model for a given peak are plotted against the observed F0 values for both speakers. For speaker NG, the predicted vs the observed values of peak 2 and 3 are represented. For speaker DV the values of peak 2 are shown. In each graph, the line is a reference line which indicates the values when  $x = y$  (observed = predicted).

---

<sup>3</sup> In final H\* accents, the highest F0 value was considered to be at the onset of the accented syllable, since, as pointed out in chapter 3 and also in the next section, the last H\* of neutral declaratives does not show a clear peak.

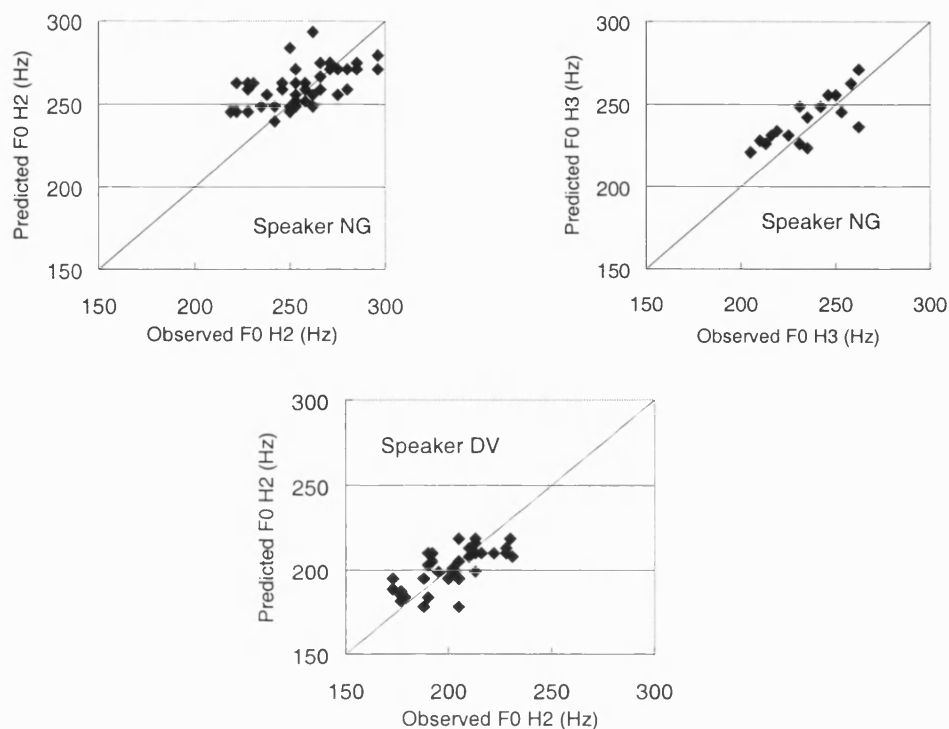


Figure 5.5. Predicted F0 values of a given peak (H2 or H3) as opposed to the observed F0 values of the same peak for speakers NG and DV. The line indicates  $x = y$ .

The results observed in the graphs of Figure 5.5 showed that the model predicts the F0 height of a given peak quite successfully since most observations are located quite close to the  $x = y$  line. This indicates that the proposal of modelling downstep as a fixed proportion of the previous peak relative to the speaker's reference line seems to work in Central Catalan. These results support the proposal of the AM model (Pierrehumbert 1980, Liberman and Pierrehumbert 1984) to analyse F0 contours as a series of H (and L) tones that are controlled at a local level.

#### 5.3.1.2.4 H with final lowering or L target?

One of the questions that arises from the data is whether the last accent in neutral declaratives is a strongly downstepped H\* (i.e. with final lowering), as has been proposed so far in this study, or an L\* target, as suggested in Prieto (1995) for Central Catalan, or in Sosa (1999) for Spanish. The main problem in classifying the last pitch accent as H\* derives from the fact that there is no evident peak in the F0 contour. Instead the last accent is realised as a progressively falling slope. In chapter 3, we showed that even though the last (nuclear) accent of a neutral declarative in Central

Catalan showed no peak, it was still classified as a strongly lowered H\* on three grounds: 1) perceptual, 2) acoustic, and 3) cross-linguistic evidence. In this section, scaling evidence provides further support for interpreting the last pitch accent as H\* rather than L\*.

If we look at the F0 values corresponding to what has been classified as the last H of a declarative sentence in Figure 5.1, we observe that the mean F0 of the last H is in fact very similar to the mean F0 of the previous L targets. However, the last H is much higher than the following L. If the last pitch accent of a declarative was L\*, then the expected transition between an L\* pitch accent and the L edge tone would be low and level, but not falling. In the data of all speakers, however, there is a falling movement from the last pitch accent to the L edge tone, indicating that the last pitch accent is a high (although strongly lowered) target. The falling transition from H to L is illustrated in Figure 5.6 for the data of all speakers.

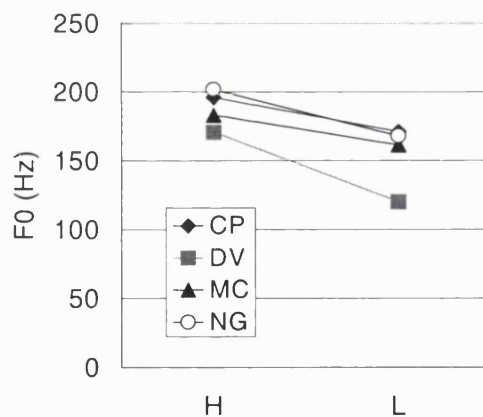


Figure 5.6. Mean F0 values of the last H (measured as the F0 height at the onset of the accented syllable) and the last L of a neutral declarative contour.

In Mexican Spanish downstepping contours, Prieto et al (1996) reported different behaviour as far as the last peak was concerned. Whereas two speakers produced a clear final peak, one speaker realised the final accent as a continuously falling slope, similar to that observed in the data of Central Catalan speakers. Prieto et al showed that the final fall behaved as the final peaks: even though there was no F0 target in the F0 contour, a model which assumed a target value better predicted the F0 value than other models.

The fact that the last peak of a contour shows a more drastic F0 descent is observed in many languages (American English, Liberman and Pierrehumbert 1984, Mexican Spanish, Prieto et al 1996). This final lowering involves a more abrupt F0 lowering of the pitch in the final peak of the utterance, which falls below the values predicted by the downstep rule. This means that the F0 value of the last peak in an utterance cannot be modelled using the downstep ratio as for the non-final peaks. Thus, Liberman and Pierrehumbert (1984) proposed that utterance final peaks should be modelled by using a lowering constant, which is defined to be a fraction of the value of the peak predicted by the downstep rule above the reference line. The F0 value of the last peak is calculated as follows:

$$P = R + l \cdot (P_{\text{(down)}} - R)$$

where P = peak height of the last peak of the utterance, P<sub>(down)</sub> = peak height of the last peak predicted by the downstep rule, R = the reference line and l = the final lowering constant. The l was obtained by dividing the distance between the reference line and the observed final peak value by the reference line and the predicted downstepped peak value as below:

$$l = (P_{\text{(obs)}} - R) / (P_{\text{(down)}} - R)$$

In languages where a clear peak is observed in the F0 contour, P<sub>(obs)</sub> equals the observed peak height. However, since in Central Catalan there is no surface F0 peak on the last pitch accent of a declarative, P<sub>(obs)</sub> corresponds to the F0 height at the onset of the accented syllable. According to this, the mean l value for the two speakers is 0.85 for speaker NG and 0.62 for speaker DV.

Once the l value was obtained for the two speakers, the predicted height of the last "peak" was calculated for all sentences and compared to the observed or real last peak value. The results are displayed in Figure 5.7, which shows predicted against observed F0 values for the last H with reference to an x = y line. The plots show a quite close association between predicted and real values since most observations are clustered fairly equally round the diagonal. This suggests that the model predicts the F0 height of the last peak quite accurately.

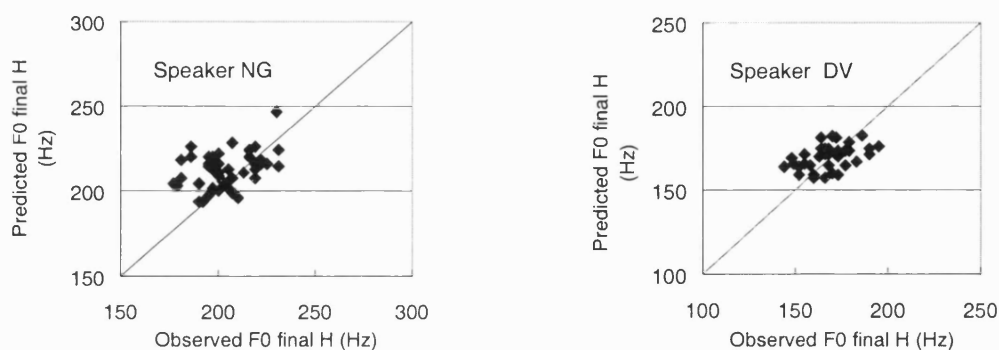


Figure 5.7. Predicted F0 values for the last H of an utterance as opposed to the real F0 values observed in the data for speakers NG and DV. The line indicates  $x = y$ .

### 5.3.1.3 Discussion

The results of the scaling of F0 peak height in Central Catalan declaratives have shown that downtrend patterns are better explained as a linguistically controlled downstep than as a global declination component. This has been postulated due to the absence of a significant effect of temporal distance between peaks on the F0 value of a given peak. The F0 height of consecutive H values has proved to be independent of the number of intervening syllables (or temporal distance) between peaks, indicating that time-dependent declination is not active. This behaviour is consistent with the models that claim that F0 downtrends are better explained as a deliberate use of step accents (Pierrehumbert 1980, Liberman and Pierrehumbert 1984, Pierrehumbert and Beckman 1988). Thus, the data support a linear interpretation of F0 contours, as claimed by the TS models, rather than a superpositional interpretation, as in the CI tradition.

Consequently, in line with Liberman and Pierrehumbert's proposal, downstepping H values in Central Catalan were predicted as a constant fraction (downstep ratio) of the previous peak, scaled above the reference line of each speaker. This model has successfully accounted for downstepping patterns in several languages (e.g. English: Pierrehumbert 1980, Liberman and Pierrehumbert 1984, Ladd 1990, 1993; Japanese: Pierrehumbert and Beckman 1988; Mexican Spanish: Prieto et al 1996; or Dutch: van den Berg et al 1992). In all these languages, descending H patterns could be explained as the repeated application of the downstep rule.

Similar to other languages, utterance-final peaks in Central Catalan underwent a much greater F0 lowering than that predicted by the downstep rule. In the data of all speakers, this final lowering was not realised as a compressed final peak but as continuously falling slope. This behaviour was also observed in Mexican Spanish (Prieto et al 1996). Despite the lack of a clear peak, final accents in Central Catalan declaratives have been analysed as H targets and have been modelled by means of a lowering constant, which applies to the output of the downstep rule. Following Prieto et al, the F0 value at the syllable onset was taken as the target "peak" value.

Overall, the results obtained in the data support the main tenets of the AM model that F0 contours can be analysed as a series of primitives controlled at a local level (as in Pierrehumbert 1980 and Liberman and Pierrehumbert 1984).

### **5.3.2 Narrow focus sentences**

The main goal of this section is twofold. First, the effects of narrow focus on the realisation of downstep will be analysed. It has been argued cross-linguistically that different focus structures have an effect on the height of different peaks (Liberman and Pierrehumbert 1984, Välimaa-Blum 1993, Ladd and Terken 1995, Rump and Collier 1996). In most cases, the presence of narrow focus on a particular item involves a higher pitch excursion, which blocks the downstepping sequence. In this section, we will examine whether the relationship between narrow focus and non-downstep also applies in Central Catalan marked sentences and in particular whether the blockage of downstep is exclusively linked to the presence of narrow focus or is associated to other phenomena, such as emphasis.

Second, we will analyse to what extent the peak height of downstepping sentences produced with narrow focus can be predicted by the same model used in downstepping broad focus sentences, that is, by means of the downstep ratio and/or the final lowering constant.



### 5.3.2.1 Data

The data used to analyse the relation between narrow focus and downstep in Central Catalan declaratives consist of sentences with narrow focus in *medial* position and in *final* position produced by the six informants (CP, DV, ER, MC, NG and NM). Sentences with narrow focus in utterance initial position were not included because they were produced with only one peak. The data with medial focus contain SVO structures with narrow focus on the verb as well as sentences with narrow focus on the second element of a double accented subject. An example of the two structures is reproduced in (5.12) below. These utterances were gathered during the first recording session as part of the whole corpus (refer to section 2.3.1.4 for a detailed account of the recording procedures and to section 1.1 in Appendix 1 for the list of sentences). The data with final focus includes SVO structures with narrow focus on the object. Those were gathered during the second recording session (see section 3.1 and Appendix 3). An utterance with final narrow focus is illustrated in (5.13). Sentences with medial and final focus were obtained by two kinds of triggering questions: a contrastive prompt and an identification prompt. The number of utterances used for the analysis of downstep and narrow focus amounts to 64 for each speaker: 48 utterances with medial narrow focus and 16 with final narrow focus.

(5.12) *Medial narrow focus*

- a. LA NENA MORENA venia nines noves  
"The girl WITH DARK HAIR was selling new puppets"
- b. La Rosa REGA els geranis  
"Rosa WATERS the geraniums"

(5.13) *Final narrow focus*

- L'Emília vol AMANIDA  
"Emilia wants SALAD"

### **5.3.2.2 Results**

As in broad focus sentences, the following points were labelled in the F0 contours of narrow focus sentences:

1. Phrase-initial F0 value
2. Highest F0 value for each pitch accent
3. Lowest F0 value between peaks
4. F0 value at the end of the focal domain
5. Phrase final F0 value

The mean schematised F0 contours of narrow focus utterances are presented in Figures 5.8 and 5.9. Figure 5.8 exhibits the schematised contours of sentences with narrow focus in medial position and Figure 5.9 those of focus in final position. In all graphs, the mean values of sentences with narrow focus triggered by contrast are plotted separately from the mean values of sentences triggered as answers to identification prompts. This was done to see whether the kind of triggering question had an effect on the realisation of downstep on the focal item. As reported in the literature, contrastive focus tends to be more salient than identification focus (Brown et al 1980, Couper-Kuhlen 1986, García-Lecumberri 1995). Mean values for postfocal accents were not plotted in the graphs, since the number of postfocal accents was very small and not all speakers used them.

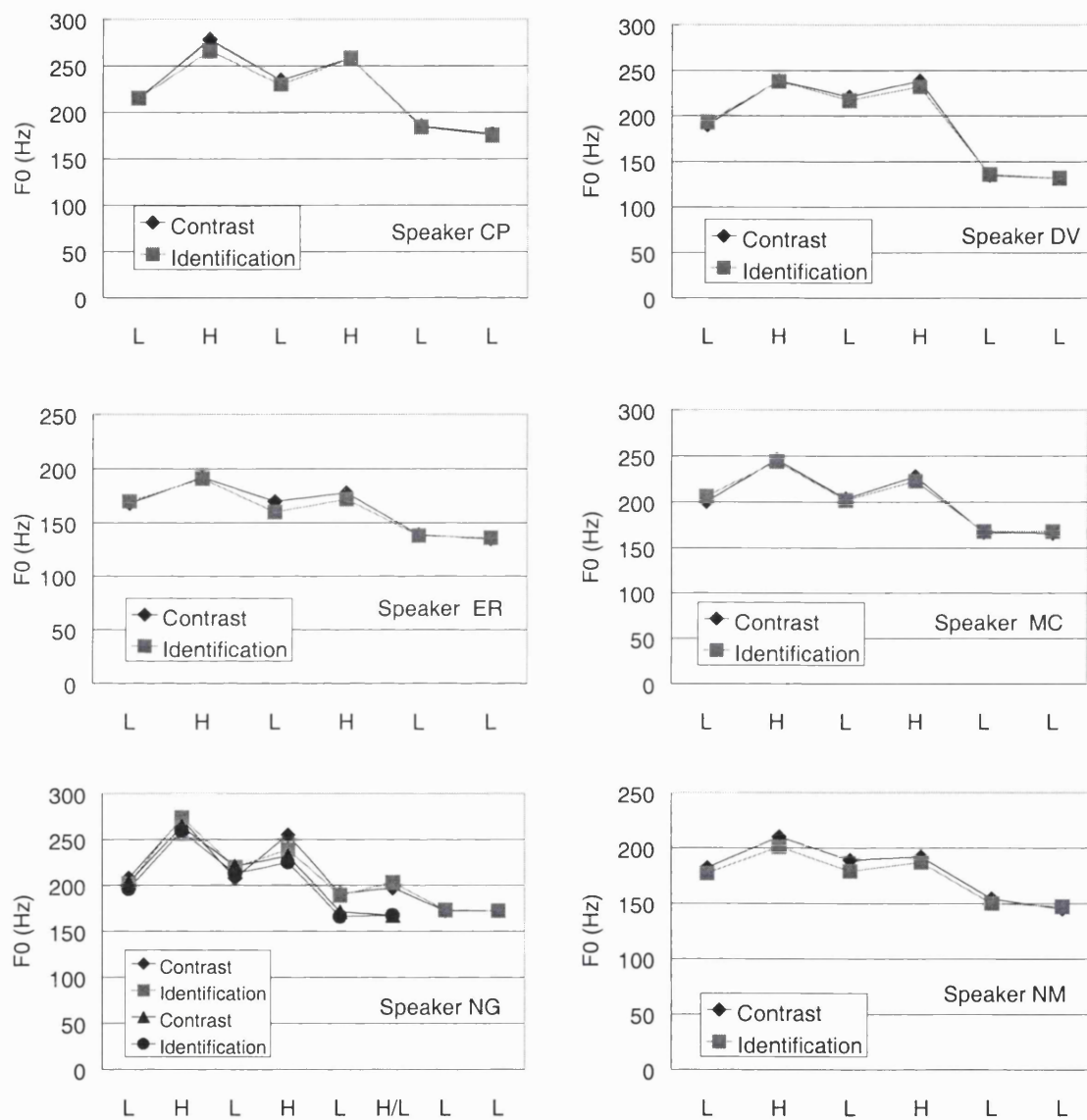


Figure 5.8. Schematised F0 contours for sentences produced with focus in medial position triggered by contrast and identification.

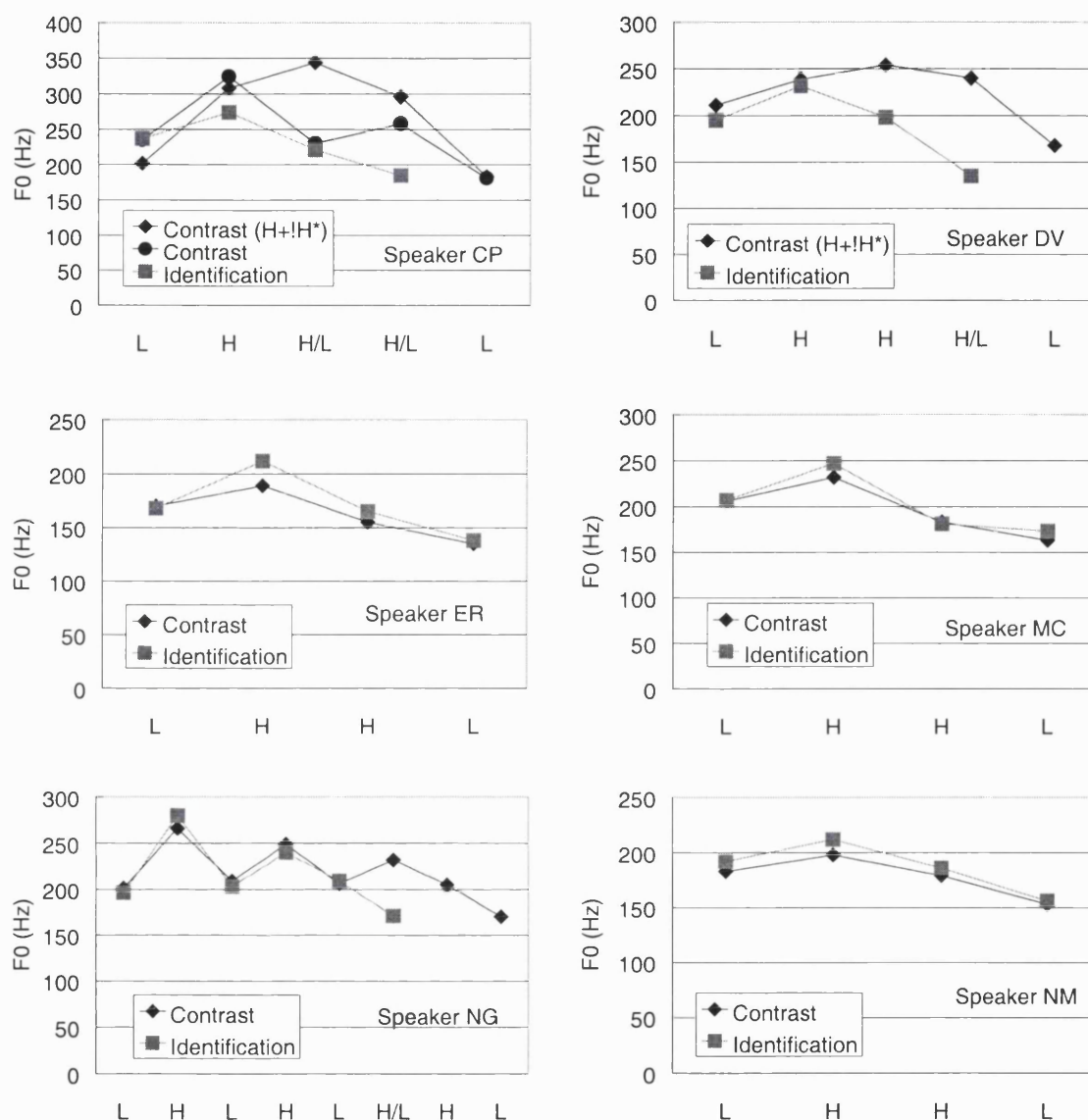


Figure 5.9. Schematised F0 contours for sentences produced with focus in final position triggered by contrast and identification.

In general terms, the results plotted in Figures 5.8 and 5.9 show no major differences in the contours of sentences triggered by a contrast and those triggered by identification, except for final focus utterances for speakers CP and DV who used a different pitch accent in contrastive reading. As pointed out in chapter 4, all contrastive final focus sentences of speaker DV and a few of speaker CP (50%) were produced with an H+!H\* nuclear accent. Also, speaker CP uttered the other half of her contrastive final focus sentences with an emphatic H\*, that is, an H\* pitch accent with no final lowering but with a clear F0 rising trajectory. Furthermore, the data present some variability in the degree of downstep of the focal accent. Whereas some

speakers (mainly speaker DV) seem to block downstep on the focal accent, others do not. These issues are examined in more detail in the following sections.

### **5.3.2.2.1 Effects of narrow focus on downstep**

This section analyses the effects of narrow focus on the realisation of downstepping accents. In particular, it investigates the claim that downstep tends to be blocked by narrow focus (Beckman and Pierrehumbert 1986, Frota 1995, 1998, among others). In order to examine the relationship focus/downstep, the data were divided into two groups: 1) focus in sentence medial vs in sentence final position, and 2) focus triggered by contrast vs focus triggered by identification. The division medial vs final position was established due to the different realisation of the H tone according to the utterance position (i.e. with a peak in medial position vs with a falling slope in final position). The subgroups focus triggered by contrast vs focus triggered by identification were arranged to detect possible effects of the kind of narrow focus on the degrees of downstep. In principle, contrastive focus is more marked than identification focus (Brown et al 1980, Couper-Kuhlen 1986, García-Lecumberri 1995) and hence the possibility of blocking downstep might be higher. Some preliminary results on these issues were reported in Estebas-Vilaplana and Maidment (1999b).

In order to analyse the effects of narrow focus on the realisation of downstep, the scaling of H2 in relation to H1 was examined for all speakers. The mean F0 values of H1 and H2 for sentences produced with narrow focus in medial position triggered both by contrast and identification prompts are displayed in Figure 5.10. T-tests comparing the F0 values for the two peaks in the two triggering conditions are shown in Table 5.8. Setting the significance level at 0.001, the results show that the majority of speakers produced a significantly lower second peak in narrow focus utterances. Only speakers DV and CP blocked the downstepping effect of H2. Speaker DV produced a non-downstepped H2 both in contrast and in identification contexts and speaker CP only in sentences with a contrastive reading. These results show that the expected relationship between the presence of narrow focus and the blockage of downstep is not fully attained in Central Catalan. Even though the blockage of downstep on a peak with narrow focus is possible, the data do not show an immediate relationship

between narrow focus and lack of downstep. This indicates that in Central Catalan there is no categorical relationship between narrow focus and non-downstep. Alternatively, downstep appears to be a gradient phenomenon, whose realisation seems to be linked to the speaker's decision to emphasise a particular item or not. Thus, narrow focus only affects downstep if it is produced with an emphatic reading. These observations are confirmed with the results of narrow focus in final position.

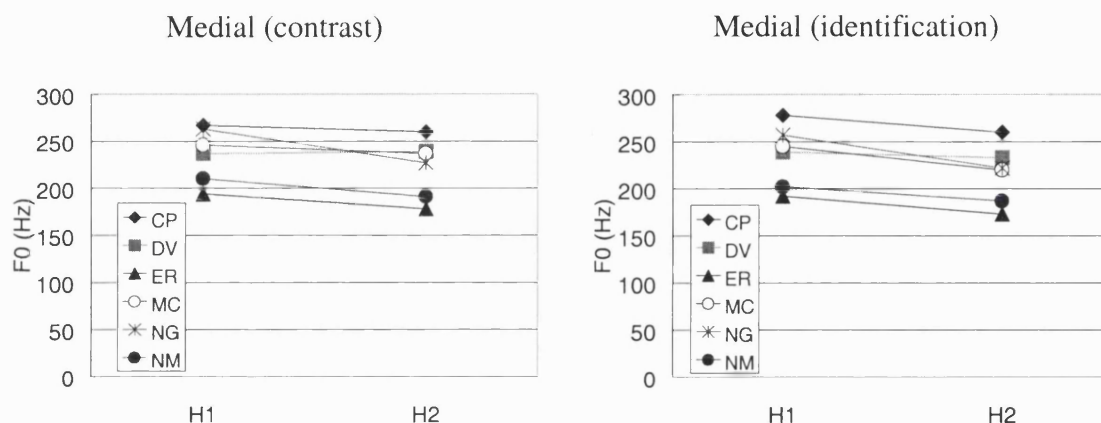


Figure 5.10. Mean H1 and H2 for medial narrow focus declaratives triggered by a contrastive and an identification prompt.

	CP	DV	ER	MC	NG	NM
Contrast	p=0.01 t=2.50	p=0.21 t=-0.8	p<0.001 t=4.26	p=0.001 t=3.55	p<0.001 t=7.27	p<0.001 t=7.7
Identification	p<0.001 t=3.73	p=0.07 t=1.49	p<0.001 t=7.39	p<0.001 t=4.9	p<0.001 t=8.11	p<0.001 t=6.59

Table 5.8. Results of the t-test comparing the height of H1 and H2 in medial position for narrow focus triggered by contrast and by identification.

Figure 5.11 and Table 5.9 present the results on the scaling of peaks in utterances with final narrow focus. As before, Figure 5.11 plots the mean F0 value of H1 and H2 in sentences with final narrow focus produced with a contrastive and an identification reading. The displays show that all speakers uttered a downstepped H2 in utterances triggered by an identification prompt. This is corroborated by the statistical results presented in Table 5.9 where significant differences ( $p<0.001$ ) are observed between H1 and H2 in identification productions. Sentences elicited with a contrastive question showed more variation. Speakers DV and CP blocked the downstepping effect of H2, whenever it corresponds to the leading tone of the bitonal accent H+!H\*. As observed

in Table 5.9, speaker DV shows no significant differences ( $p=0.02$ ) between the scaling of H1 and H2, indicating that the second peak is not lowered with respect to the first one. Speaker CP does show significant differences in the scaling of the two peaks ( $p=0.001$ ). However, as displayed in Figure 5.11, this is due to the fact that the second peak is scaled much higher than the first one in her productions. When speaker CP produced an emphatic H\*, significant differences are observed between H1 and H2, indicating that downstep is not suspended despite uttering the final H\* of the sentence with no final lowering. All other speakers also produced a downstepped H2 in their contrastive utterances.

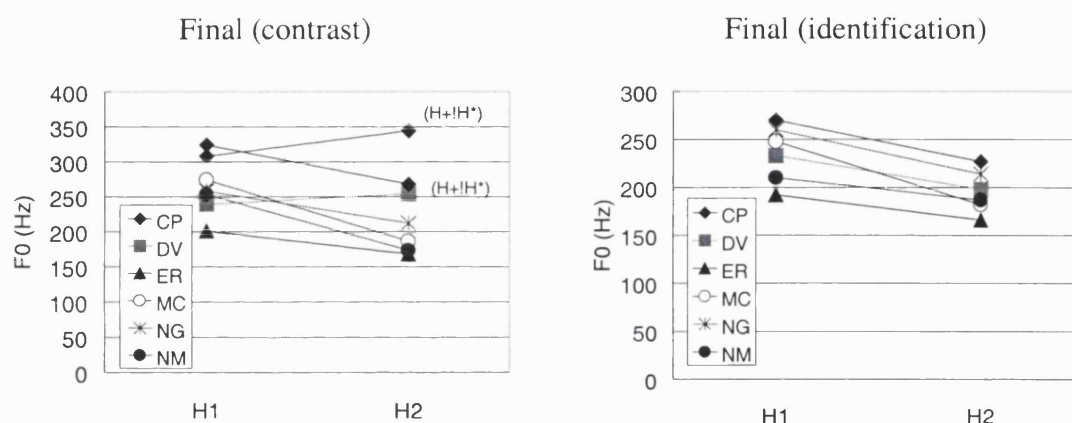


Figure 5.11. Mean H1 and H2 for final narrow focus declaratives triggered by a contrastive and an identification prompt.

	CP	DV	ER	MC	NG	NM
Contrast	$p<0.001$		$p<0.001$	$p<0.001$	$p<0.001$	$p<0.001$
H*	$t=10.8$		$t=6.61$	$t=16.7$	$t=5.15$	$t=5.3$
H+!H*	$p=0.001$	$p=0.02$				
	$t=-4.76$	$t=-2.8$				
Identification	$p<0.001$	$p=0.001$	$p<0.001$	$p<0.001$	$p=0.001$	$p<0.001$
	$t=7.76$	$t=4.3$	$t=6.49$	$t=25.9$	$t=4.3$	$t=4.89$

Table 5.9. Results of the t-test comparing the height of H1 and H2 in final position for narrow focus triggered by contrast and by identification.

These results confirm the idea that downstep is not necessarily blocked by narrow focus in Central Catalan. The blockage of downstep seems to depend on the speaker's decision to emphasise a particular item or not. Furthermore, as expected, the data

show that narrow focus derived from a contrastive question is more likely to be emphasised (and hence produced with a non-downstepped peak) than narrow focus triggered by identification. This behaviour, however, is not categorical and it is only observed in the data of two speakers. This corroborates the idea that there is no absolute association in Central Catalan between non-downstepped accents and narrow focus.

#### **5.3.2.2.2 Predicting F0 values**

Given the fact that downstep is not always blocked in Central Catalan narrow focus sentences, in this section we will try to investigate to what extent the same model proposed for the prediction of downstep in broad focus utterances could be used in narrow focus sentences. Two hypotheses will be examined:

- 1) peaks before the focussed element can be predicted in the same way as the prenuclear peaks in neutral sentences, that is, prefocal peaks are calculated as a constant fraction of the previous one;
- 2) peaks on the focal item are predicted in the same way as in broad focus utterances, except when focus is produced emphatically. According to this, focal peaks in medial sentence position will be predicted by the downstep rule and focal peaks in final position will be predicted by the final lowering rule.

#### *Prenuclear peaks*

The hypothesis that prenuclear peaks in narrow focus sentences can be predicted in the same way as prenuclear peaks in broad focus sentences could only be analysed in the data of speaker NG, since she was the only informant that produced more than one prenuclear peak before the focal element (see Figures 5.8 and 5.9). In her data, prenuclear peaks were observed both in sentences with medial focus and in sentences with final focus. An example of the two kinds of utterances is provided below. The Hs in bold are the prenuclear peaks that are expected to be predicted by the model.



- (5.14) a. Les nenes volen REBRE la reina  
           H1    **H2** H3  
 "The girls want to WELCOME the queen"

- b.    L'home venia llimones MADURES  
           H1       **H2**    **H3**    H4  
 "The man sold RIPE lemons"

As in broad focus sentences, the F0 height of a given peak was calculated as a constant fraction (downstep ratio) of the previous peak, scaled above the reference line of the speaker. The downstep ratio obtained in broad focus sentences for speaker NG (i.e. 0.64) was used to predict the prenuclear peaks in her narrow focus productions since we hypothesised similar downstepping patterns in the prenuclear peaks of the two focal structures. The results are displayed in Figures 5.12 and 5.13. Figure 5.12 shows the observed against the predicted prenuclear peak values of utterances with narrow focus on the verb. Figure 5.13 shows the observed against the predicted prenuclear peak values in sentences with narrow focus on the object. The line indicates the values when  $x = y$  (observed = predicted).

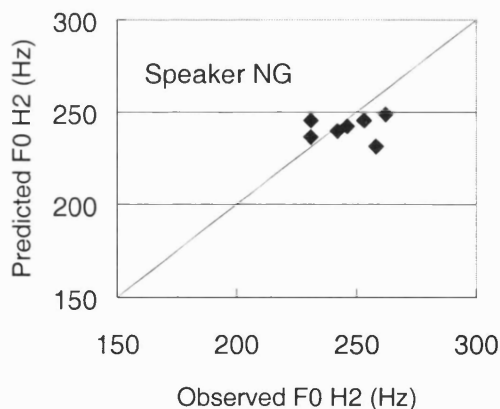


Figure 5.12. Observed against predicted prenuclear H2 values in verb narrow focus sentences for speaker NG. The line indicates  $x = y$ .

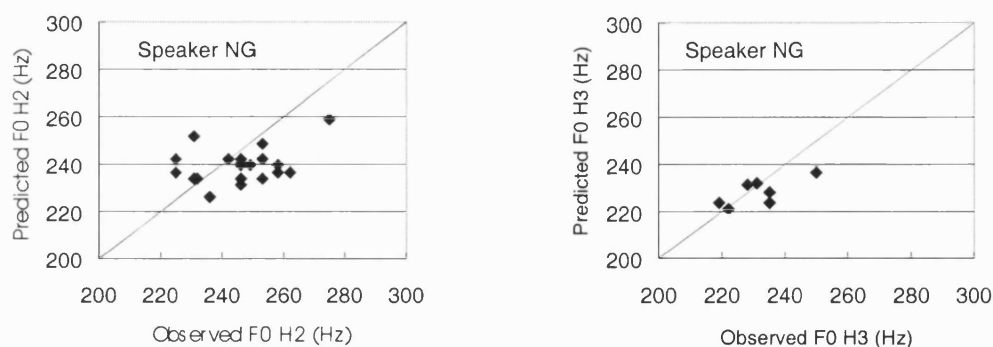


Figure 5.13. Observed against predicted prenuclear peak values for H2 and H3 in sentences with narrow focus on the object for speaker NG. The line indicates  $x = y$ .

Even though the number of observations is not very large in any of the displays, the results show that the F0 values of downstepping prenuclear peaks in narrow focus sentences are quite successfully predicted by the downstep ratio. In all cases, predicted and observed values are quite close to the  $x = y$  line. This confirms the hypothesis that prenuclear H values behave very similarly in broad and in narrow focus sentences, as far as the scaling of F0 values is concerned. Hence, a model that accounts for prenuclear peaks in broad focus can also be used to predict prenuclear H values in narrow focus.

### *Focal Accents*

Given the results observed in this study that narrow focus does not necessarily entail the blockage of downstep in Central Catalan, it seems plausible to hypothesise that, whenever the focussed item is not produced with emphasis, the F0 height of the focussed word can be predicted by means of the downstep rule in medial focus or by means of the final lowering rule in final focus. This is examined in the data of speakers NG and DV, using the same reference line, downstep ratio and final lowering constant as calculated in sections 5.3.1.2.3 and 5.3.1.2.4. Since speaker DV produced most narrow focus sentences with emphasis on the focussed word, it is expected to find a mismatch between the predicted values and the observed values in her productions. On the other hand, in NG's sentences, the peak height of the focal accent is expected to be successfully predicted by the model, since she did not emphasise the focal accents. Figure 5.14 plots the predicted against the observed F0 H values on the focussed item for speakers NG and DV in sentences with focus in medial position. As expected, the results show that for speaker NG the model can

predict the scaling of the focal peak quite successfully since the focal accent was uttered without emphasis and hence downstep was not blocked. The results of speaker DV, on the other hand, do not concentrate around the  $x = y$  reference line, indicating that the downstep rule cannot predict the peak height of emphatic accents. In this case, observed peaks were much higher than predicted peaks.

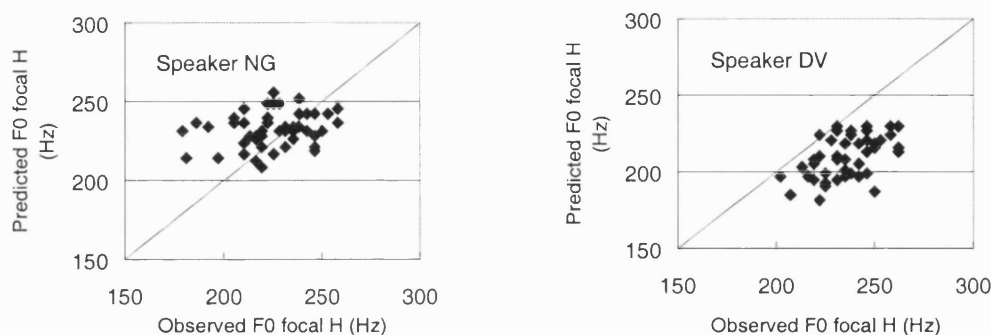


Figure 5.14. Observed against predicted values for H in sentences with narrow focus in medial position produced with no emphasis (speaker NG) and with emphasis (speaker DV). The line indicates  $x = y$ .

Similarly, Figure 5.15 plots the predicted against the observed F0 H values on the focussed item for speakers NG and DV in sentences with focus in final position. In this case, the predicted values were calculated with the final lowering ratio rather than the downstep ratio. The results show some correspondences between predicted and observed values, although not all observations are accurately predicted. This might be due to possible emphatic productions of some of the accents. More data would be needed to be able to draw clearer conclusions.

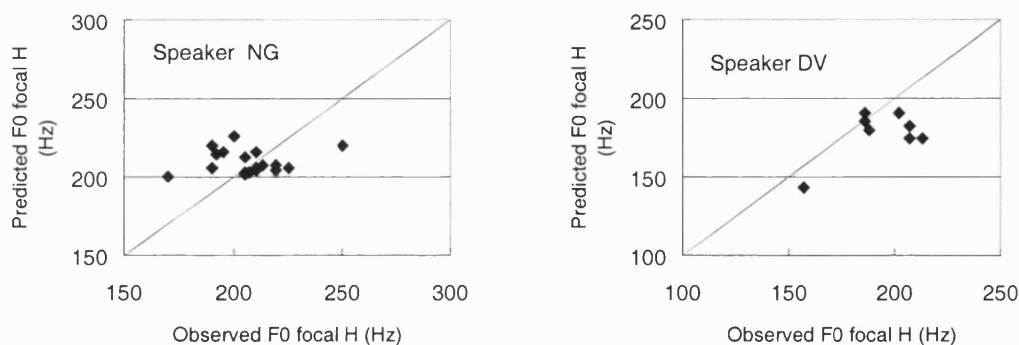


Figure 5.15. Observed against predicted values for H in sentences with narrow focus on the object for speakers NG (both contrast and identification) and DV (only identification sentences).

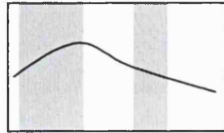
### 5.3.2.3 Discussion

In the previous sections, the effects of narrow focus on downstepping accents have been analysed. The data showed that, as opposed to the behaviour observed in some other Romance languages (e.g. Portuguese: Frota 1995, 1998; or Palermo Italian: Grice 1995a), there is no categorical relationship between the presence of narrow focus and the presence of a non-downstepping accent. In Central Catalan, the blockage of downstep is not immediately triggered by narrow focus but by the speaker's decision of adding an emphatic value to a particular item. Thus, the suspension of downstep does not have a phonological value as in Portuguese or Italian but is a gradient phenomenon, which can present a continuum of forms depending on the degree of emphasis. That is, the more emphatic an accent is, the less downstepped it will be. This behaviour confirms the idea that the same accent type is used in Central Catalan to convey broad and narrow focus.

Given the fact that there is no direct connection between narrow focus and the blockage of downstep, the scaling of peaks in narrow focus downstepping utterances was predicted in the same way as in broad focus utterances, that is, by means of the downstep rule or the final lowering ratio, depending on the location of the focussed item within the sentence. As expected, the downstep ratio or the final lowering constant could not be used to predict focussed items produced with emphasis.

The results on the scaling of focal peaks observed in Central Catalan throw up the issue of *final ambiguity*, that is, the ambiguous reading of sentences produced with final narrow focus and sentences produced with broad focus. The data have shown that speakers produce final narrow focus and broad focus in a very similar way both phonologically (with a final H\* pitch accent) and phonetically (the H\* is realised as a progressively falling slope). These observations agree with the perception test presented in chapter 4, where the same pattern could be used to respond to different focus questions. Thus, a sentence such as *vaig al cinema*, as illustrated in (5.15) below, is ambiguous since it can be the response of a broad focus question (as in 5.15a), an identification narrow focus question (as in 5.15b) and a contrastive narrow focus question (as in 5.15c).

(5.15)

Vaig al cinema

"I'm going to the cinema"

- a. Què fas? "What are you doing"?
- b. On vas? "Where are you going?"
- c. Vas al teatre? "Are you going to the theatre?"

This agrees with Ladd (1996) who pointed out that narrow focus can be conveyed by pronunciations that are not phonetically distinct from broad focus readings and that sometimes in order to disambiguate between the two readings emphatic stress is used. In the Central Catalan data, late emphatic stress has only been observed in half of the contrastive utterances of one speaker. The fact that emphatic stress has been used so little corroborates the idea that there is no need of a different accent type to convey narrow focus in Central Catalan. In other words, emphatic (or non-downstepped) accents are optional and, contrary to other languages, they have not been incorporated into the phonological system of Central Catalan as narrow focus markers. Finally, two speakers used a different pitch accent, namely, H+!H\*, in final contrastive focus. This accent has an assertive and pedagogical nuance and conveys correction or rectification.

#### 5.4 The scaling of valleys in broad focus and in narrow focus sentences

So far, we have concentrated on the analysis of the upper edge of the tonal space in Central Catalan, but little has been said on the behaviour of the L values or F0 minima between peaks. In this section, the scaling of prenuclear F0 valleys observed in the F0 contours of Central Catalan neutral declaratives will be examined. L values in narrow focus sentences are not analysed because they were already investigated in chapter 4 (section 4.4.1.1.1). In particular, we expect to confirm differences in the scaling of L when it is aimed as a phonological target or when it simply acts as a phonetic transition between Hs. When L is a target, valleys are expected to show a rather stable scaling. When L is a phonetic transition, its scaling is expected to vary according to the temporal distance between peaks.

In English, for instance, whenever an F0 valley is not a phonological target but a transition between two H tones, the amount of dipping tends to increase as the temporal distance between peaks increases (Pierrehumbert 1980). This was also observed in chapter 4 for Central Catalan: the amount of sagging between the word edge tone H and the pitch accent H\* in narrow focus sentences increased as the number of intervening syllables increased. This allowed us to classify the nuclear accent of a narrow focus sentence as H\*, rather than L+H\*. However, in broad focus utterances, prenuclear Ls have been described as phonological targets (L\*). In these cases, then, it is expected that L\* is scaled at a relatively constant value, and that the temporal distance between peaks has little effect on the location of the valley. The constant scaling and alignment of L was observed in Greek prenuclear accents, described as L\*+H in Arvaniti and Ladd (1995). In Arvaniti and Ladd, it was observed that there was no correlation between the amount of F0 dipping and the temporal distance (or number of intervening syllables) between peaks. In fact, valleys were located at a rather fixed F0 height, providing evidence for an intended tonal L target. In the next sections, the scaling of prenuclear Ls in Central Catalan broad focus declaratives is investigated. The interpretation of L values as tonal targets (i.e. L\*) is expected to be confirmed by the scaling evidence.

#### **5.4.1 Data**

In Central Catalan neutral declaratives, the scaling of valleys could only be analysed in the productions of speakers NG and DV since, as reported in section 5.3.1.2.2, they were the only ones that uttered more than one prenuclear accent. Thus, the temporal effects on the scaling of valleys will be examined in the two contexts used for the analysis of peaks. These contexts involve the analysis of valleys between H1 and H2 with a varying number of intervening syllables (1-2 or 3-5) between the peaks. Context 1 compares the scaling of L between H1 and H2 in four-stressed/four-accented sentences (4,4) with that of four-stressed/three-accented sentences (4,3). Context 2 compares the valleys in three-stressed/three-accented sentences (3,3) with those of four-stressed/three-accented sentences (4,3). For an example of these utterances, see (5.9) and (5.10). Since valleys in prenuclear position have been

described as L\* targets in this study, their scaling is expected to be quite constant despite temporal differences.

#### 5.4.2 Results

The results of both speakers are presented in Figure 5.16 and in Table 5.10. Figure 5.16 displays the mean F0 values of the valley between two peaks separated by a different number of intervening syllables. The results of t-tests comparing the two samples are shown in Table 5.10.

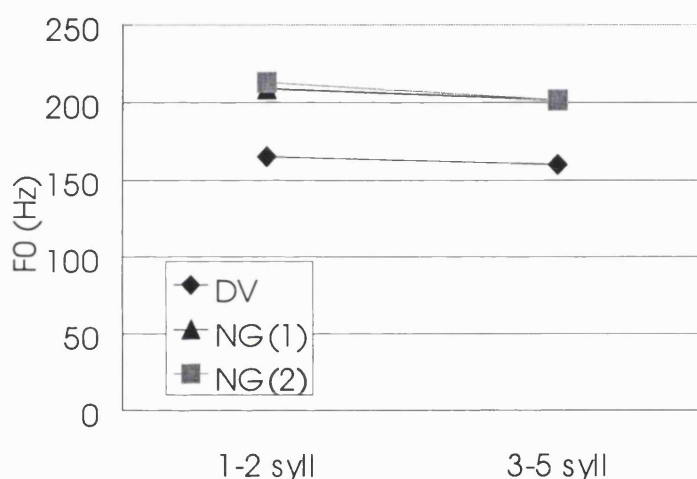


Figure 5.16. Mean F0 values of the valley between two peaks separated by a different number of intervening syllables (1-2 and 3-5) for speakers DV and NG in broad focus sentences.

	1-2 (4,4)	3-5 (4,3)	t-test	1-2 (3,3)	3-5 (4,3)	t-test
NG	213.2	201.6	p=0.06 t=1.88	209.1	201.6	p=0.03 t=1.96
DV	-	-	-	168.2	160	p=0.027 t=2.06

Table 5.10. Mean F0 value of L with 1-2 and 3-5 intervening syllables between H tones for the two speakers and in different sentence conditions and results of the t-tests.

The results of Figure 5.16 show that the scaling of L is rather fixed irrespective of the number of intervening syllables. The statistical results presented in Table 5.10 show that at 1 per cent significance level ( $p < 0.01$ ), no significant differences are observed in the L values of sentences with a different number of intervening syllables ( $p > 0.01$  in all cases). The stability in the scaling of L confirms the interpretation of prenuclear valleys in broad focus productions as L\* phonological targets. These results contrast

with those observed in chapter 4 (Figure 4.4) for the valley between the last prenuclear rise (L\*H) and the nuclear accent (H\*) of narrow focus utterances. In this case, the F0 dipping between peaks increased as the amount of segmental material increased, showing that the valley was a phonetic transition rather than an L target.

### **5.4.3 Discussion**

The results on the scaling of L values observed in broad focus utterances confirm the analysis of prenuclear valleys as L\* targets since L points are scaled at a rather fixed height, irrespective of the number of intervening syllables. These observations agree with the results reported in Arvaniti and Ladd (1995) and Arvaniti et al (1998, to appear) for Greek prenuclear accents. In Greek, the height of prenuclear F0 valleys was not affected by the number of intervening syllables between accents, suggesting a low phonological target. The results of Central Catalan differ from those of Mexican Spanish (Prieto 1998) where L values were more variable and affected by temporal distance between peaks. Thus, according to the scaling evidence, a different phonological interpretation of prenuclear rises has to be postulated in these languages, namely, H\* in Mexican Spanish, L\*+H in Greek and L\*H in Central Catalan.

The results of prenuclear valleys in neutral declaratives differ from those of narrow focus sentences, where the amount of sagging between the nuclear H\* and the preceding H word edge tone increased as the number of intervening syllables increased. This suggested that the final or nuclear accent of narrow focus was not L+H\* but H\*. This time-dependent effect of L values in narrow focus utterances might indicate the presence of an overall declination effect. If this was the case, declination would be understood as a residue downtrend after all other lowering mechanisms (downstep, final lowering) have applied, as suggested in Pierrehumbert and Beckman (1988). Further research is needed on this topic.

## **5.5 Conclusion**

In this chapter, the scaling of peak and valley F0 values in Central Catalan has been analysed for utterances produced with broad focus and narrow focus. In particular, three main issues have been covered: 1) the interpretation of H downtrends as a global



declination effect or as a local downstep mechanism, 2) the effects of narrow focus on the production of downstepping contours and 3) the differences and similarities of L values in broad and narrow focus.

As far as descending H patterns are concerned, the results showed no effect of temporal distance (or number of intervening syllables) between peaks on the scaling of H values. This indicates that peak downtrends in Central Catalan are better explained as a linguistically controlled accent-by-accent downstep than as a global time-dependent declination. As in English (Lieberman and Pierrehumbert 1984), Mexican Spanish (Prieto et al 1996) and Dutch (van den Berg et al 1992), peak height in Central Catalan can be accurately predicted as a constant proportion of the height of the previous peak. Similarly to these languages, utterance final peaks (or progressively falling slopes) undergo a greater amount of lowering than that expected by the downstep rule. Hence, the final Hs are better explained by means of a particular ratio of decay (lowering constant), which is higher than the downstep ratio.

The results concerning the relationship between narrow focus and downstep reveal that narrow focus does not necessarily block downstep on the focal accent. In some Romance languages, the presence of focus involves a much higher F0 peak on the focal accent, which has been classified as a different accent type from that used in neutral sentences ( $H^*+L$  for narrow focus and  $H+L^*$  for broad focus). However, evidence from Central Catalan shows that the same accent type is used in the two kinds of sentences (namely  $H^*$ ). In broad focus, this accent is realised with final lowering since it is the last accent of the utterance. In narrow focus,  $H^*$  can be realised with downstep or final lowering depending on the sentence position (medial or final respectively). The possibility of blocking downstep (or final lowering) on the focal accent responds to a decision of the speaker (e.g. emphasis), rather than being an intrinsic property of the accent.

Finally, the results on the scaling of L values seem to corroborate the phonological analysis of valleys provided in chapter 3 for broad focus sentences. In these utterances, prenuclear valleys tend not to be affected by temporal distance (or the number of unaccented syllables between peaks) but show a rather fixed F0 height. This seems to confirm the analysis of prenuclear valleys as  $L^*$  targets rather than

simply transitions between peaks. These results differ from those on the scaling of the L between the last prenuclear rise and the nuclear accent in narrow focus sentences, where an increase in temporal distance between peaks involved a higher amount of dipping.

## ***Chapter 6: Comparison to English***

### **6.1 Introduction**

In the previous chapters, we have observed that the use of accentual strategies to convey narrow focus is possible in Central Catalan and we have described the phonetic and phonological properties of sentences produced both with a broad focus and a narrow focus reading. The aim of this chapter is to compare the phonetic and phonological characteristics of accentual focus in Central Catalan to those of a language that predominantly uses accentual devices to express focus, such as English. The choice of English is due to the fact that it is a language which benefits from several descriptions on focal and intonational matters and, even though in this study some new data will be gathered, previous works will also be taken into account for the cross-linguistic analysis.

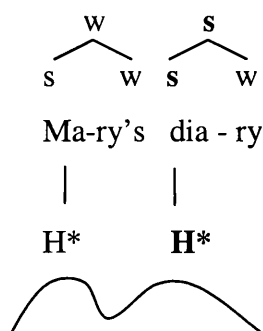
The organisation of this chapter is as follows. Section 6.2 recalls the main characteristics of English and Central Catalan accentual focus. Section 6.3 describes the English data analysed in this chapter. Section 6.4 presents the results in two parts. First, a phonetic and phonological analysis of broad and narrow focus English F0 contours is carried out, with particular attention to the questions of tonal association and tonal alignment. Second, the main properties of F0 downtrends in English broad and narrow focus sentences are examined. Section 6.5 discusses the results and compares the English data with the Central Catalan data. Finally, section 6.6 includes a few concluding remarks.

### **6.2 Accentual focus in English and in Central Catalan**

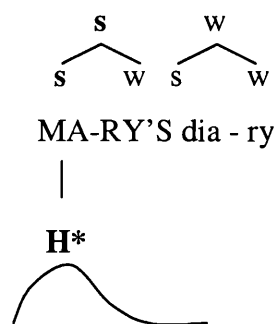
As reported in previous chapters, the expression of focus by accentual means can be manifested in three different ways: 1) the deletion of pitch accents or deaccentuation, 2) a reorganisation in the levels of phrasing, and 3) the selection of a particular pitch accent type. Studies on the signalling of information structure in English, and in most Germanic languages, show that the most common device to convey focus in these languages is the accentuation of the focal element and the deletion of the pitch accents

(or deaccentuation) of the postfocal items (Ladd 1980, 1996, Gussenhoven 1983, Féry 1992, 1993). According to Ladd (1996), deaccentuation is the result of a reversal in the abstract prominence relation between the constituents of a phrase or sentence. This is illustrated in (6.1), which reproduces the structures presented in chapter 1 (example (1.6)). (6.1a) shows a phrase with a neutral or broad focus reading. This structure is characterised by a “normal” stress pattern (i.e. *w(eak)-s(trong)*), used when there is no reason to highlight any element in particular and hence the focus is on the phrase as a whole. The stress pattern of the phrase reflects the organisation of the syllables into a hierarchical metrical structure. This structure specifies abstract relations of prominence or strength between syllables and between higher constituents, such as words or phrases. Given the prominence pattern of a phrase, the match between focus and accent is accomplished by assigning the nuclear accent on the syllable that bears the sentence stress (i.e. the syllable dominated by the major number of *s* nodes). Thus, in (6.1a) the nuclear accent falls on the first syllable of *diary*. On the other hand, (6.1b) conveys a narrow focus reading. Now the focussed element is *MARY* since the prominence relation between the two elements of the phrase has been altered (i.e. *s-w*). The modification of the prominence pattern between constituents also involves the location of the nuclear accent at the most prominent element and the subsequent deaccenting of the following, non-prominent constituent(s). Whereas the *s-w* pattern always conveys a narrow focus reading, the *w-s* pattern presents ambiguity between focus on the whole sentence and late narrow focus.

(6.1) a. *Neutral sentence*



b. *Narrow focus sentence*



So far, the results observed for Central Catalan have shown that focus involves two of the aforementioned effects: 1) differences in phrasing (an intermediate phrase boundary is observed after the focussed element), and 2) deaccentuation of non-

focussed constituents. Whereas the introduction of a phrase accent has shown to be obligatory in Central Catalan, deaccentuation seems to be optional. Postfocal accents may or may not be present and whenever present they tend to be reduced in pitch range. This means that the lack of postfocal accents in Central Catalan is not categorical but gradient. Furthermore, unlike other Romance languages, in Central Catalan there are no differences in the pitch accent type but the same pitch accent is used for broad focus and narrow focus. Sometimes the narrow focus accent can be produced with more prominence (and hence with a higher pitch range) than the broad focus one. However, this is the result of a phonetic or emphatic boost rather than a categorical distinction. An example from Central Catalan is reproduced in (6.2).

(6.2) a. *Neutral sentence*

[[La mare menja arengades]]

L\*H                      H\* LL-L%

“The mother eats herrings”

b. *Narrow focus sentence*

[[La MARE] [menja arengades]]

H\*    LL-                      LL-L%

“The MOTHER eats herrings”

According to this, Central Catalan coincides with English in that no different accent type is used to convey narrow focus. In the two languages, the boost observed in narrow focus seems to be phonetic rather than phonological. However, Central Catalan and English appear to differ in two aspects: 1) the levels of phrasing and 2) the nature of deaccenting. Whereas Central Catalan adds a phrase boundary after the focal constituent, English modifies the location of nuclear accent (as a result of an alteration in the prominence pattern of the constituents) but no phrase boundary seems to be observed after the focal element. Contrary to Central Catalan, in English, deaccenting is obligatory and categorical.

In this chapter, the phonetic and phonological means to express broad and narrow focus in English will be compared to those of Central Catalan. Particular attention will be paid to the aspects of phrasing, deaccentuation and accent type.

## 6.3 Experimental design

Even though English is a language which has been thoroughly investigated with regard to the intonational characteristics of broad and narrow focus structures, a few English sentences were recorded in this study. This was done for two reasons. First, in order to perform a cross-linguistic analysis, it was considered necessary to gather similar kinds of structures in the two languages, obtained by similar means (i.e. reading style, responses to contrastive and identification prompts, speech and Lx signals). Second, the collection of some English sentences allowed us to analyse the data in the same way as we did for Central Catalan concerning aspects such as tonal alignment and association or F0 downtrends. In the next sections, the experimental procedures used for the gathering of the English utterances are described.

### 6.3.1 Materials

The English data consisted of S(subject) V(erb) O(object) structures produced both with a broad focus and a narrow focus intonation. As for Central Catalan, the same sentences were used for the two kinds of focal readings and hence the only source of variation was expected to be intonational. In narrow focus sentences, three kinds of focal domains were examined: 1) focus on the *subject*, 2) focus on the *verb*, and 3) focus on the *object*. In all domains, narrow focus was elicited by means of two questions: a *contrastive* prompt and an *identification* prompt. An example of the three focal domains is presented in (6.3). (6.3a) and (6.3b) are responses to a contrastive prompt. (6.3c) is an answer to an identification question.

- (6.3) a.     *Subject focus (contrastive prompt)*  
          Did Helen make a marvellous dinner?  
          MOLLY made a marvellous dinner
- b.     *Verb focus (contrastive prompt)*  
          Does the rain benefit the harvest?  
          The rain DAMAGES the harvest

c. *Object focus (identification prompt)*

What's the nanny ironing?

The nanny's ironing THE LINEN

For each focal domain (S-V-O), 24 sentences were designed (12 triggered by contrast and 12 by identification). Additionally, each domain included words with three possible stress distributions: oxytones (words with lexical stress on the last syllable), paroxytones (words with lexical stress on the penultimate syllable), and proparoxytones (words with lexical stress on the antepenultimate syllable). As for Central Catalan, words with different stress patterns were included to examine the effects of tonal alignment. Overall, 72 sentences were recorded for narrow focus and 72 for broad focus. The number of sentences within each domain and each subcategory are summarised in Table 6.1. The overall list of sentences is presented in Appendix 4 (sections 4.2 and 4.3).

Focus Trigger	Narrow focus									Broad focus
	Subject			Verb			Object			
	Oxy.	Par.	Pro.	Oxy.	Par.	Pro.	Oxy.	Par.	Pro.	
Cont.	4	4	4	4	4	4	4	4	4	72
Ident.	4	4	4	4	4	4	4	4	4	

Table 6.1. Number of English sentences recorded for broad focus and narrow focus. In narrow focus, sentences are divided according to the focal domains (subject, verb, object), the triggering question (contrast, identification) and the stress pattern of the word (oxytone, paroxytone, proparoxytone).

The number of stresses per sentence varied between 3 and 5. Longer sentences were not included to avoid the division of utterances in more than one tone unit. A few structures consisting of only one stressed word were also included in the materials to compare the intonation of single-stressed utterances with that of multi-stressed utterances. The number of single words was 30, 10 for each stress category (oxytones, paroxytones and proparoxytones). The list of structures with only one lexical word is displayed in Appendix 4 (section 4.1).

Sentences were designed with the maximum number of voiced segments possible to avoid interrupted F0 contours.

### **6.3.2 Informant**

Sentences were produced by a 30 year-old female speaker of a Southern (London) variety of British English. Although the speaker was born in the county of Merseyside, she moved into London when she was very young and her accent is Standard Southern English. At the time of the recording, the speaker was following postgraduate studies in linguistics. Her knowledge of linguistics did not interfere with her productions, which were totally naive. In the following sections, the speaker will be identified as KF.

### **6.3.3 Data collection procedures**

As in Central Catalan, sentences were gathered by means of a reading activity. For broad focus sentences and for single words, the speaker was given a list of structures and was asked to read them in a neutral way, as if they were answers to the question "what happens?". For narrow focus, the speaker was given the same list of structures (except for single words) but had to produce them as responses to questions posed by the researcher. The kinds of questions were similar to those used for Central Catalan, that is, questions that triggered a contrastive response and an identification response. Questions were not pre-recorded but were asked live. Contrary to Central Catalan, the order of the responses was the same one as that of the questions. Also no distractors were included in the data. Neither the use of distractors nor the use of a randomised order of responses was considered necessary in this case since here we were not testing the possibility of using accentual focus but we simply wanted to record utterances with a narrow focus intonation.

### **6.3.4 Instrumentation and recording**

The data were recorded in the anechoic room of the Department of Phonetics and Linguistics of University College London. For each sentence, speech and laryngeal (Lx) signals were obtained. Speech was recorded to a B&K sound level meter of the type 2231, which was fitted with a 4165 microphone. The microphone was placed at about twenty centimetres from the speaker. The laryngeal signal was obtained by



means of a laryngograph processor. Both speech and Lx signals were recorded on a Sony 1000 ES DAT recorder.

### 6.3.5 Acoustic analysis

As for Central Catalan sentences, the acoustic analysis of the data was done by means of the Speech Filing System (SFS) program. Speech and laryngeal signals were played on a Sony 1000 ES DAT recorder and transferred into a Sun Sparc-10 computer in which SFS ran. Acquisition of the signals was done at 16 KHz sampling rate, following the routines of the program. An anti-aliasing filter set at 6.4 KHz cut-off point was used to prevent signals half above the sampling rate from being acquired.

F0 traces were obtained from the laryngeal signal by means of the VTX and FX programs. VTX converted the laryngeal waveform (Lx) into excitation period measurements (Tx). FX converted excitation period measurements into F0 traces. Within this program the option -s (smoothing) was used in order to get a pitch contour with less abrupt changes. This option uses a 5-point median filter.

## 6.4 Results

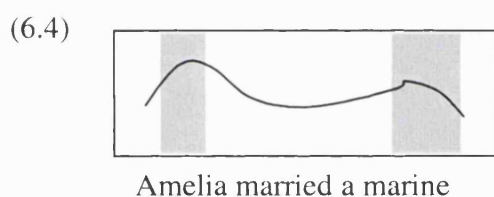
### 6.4.1 Phonetic and phonological properties

#### 6.4.1.1 Broad focus

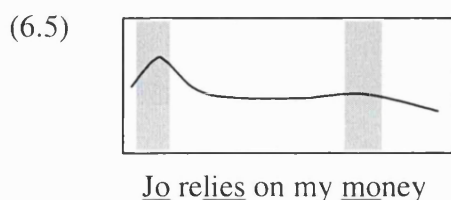
One of the main problems with the analysis of the English data was deciding where to segment syllables. Some theories (Wells 1990, 2000) propose to divide syllables according to the principle of syllable strength (i.e. consonants are syllabified with whichever of the adjacent vowels is more strongly stressed or if they are equally stressed with the leftward one). Hence, a word such as *Melanie* would be divided as *Mel-an-ie*. Other researchers (e.g. Fudge 1984) advocate a CV pattern (i.e. an intervocalic consonant is syllabified with the following vowel). Thus, *Melanie* would be syllabified as *Me-la-nie*. Finally, recent work on intonation (Coleman 1993, Local 1995) supports the ambisyllabicity proposal, according to which the *l* and the *n* belong to two syllables at the same time. In the present study, syllables have been segmented according to the CV account. This has been done for two reasons. First, the CV

perspective has been (implicitly) used in most studies within the AM tradition (Silverman and Pierrehumbert 1990, Grabe 1998a among others). Second, the CV pattern is similar to the analysis performed for Central Catalan and this makes the comparison between the two languages more compatible.

The F0 traces of English declaratives produced with a broad focus reading showed an F0 rise in the vicinity of the stressed syllables, followed by a fall over the poststressed syllables. Not all stressed syllables became accented. As in Central Catalan, the first and the last stressed syllables were always accented and medial stressed syllables were more likely not to get an accent. This implies that utterances with three stresses tended to get two pitch accents and those with four stresses used to have three pitch accents (the second stress did not become accented). Sentences were uttered with no evident prosodic breaks. An example is provided in (6.4), which sketches the F0 trace for the sentence *Amelia married a marine*. The real contour is shown in Appendix 5.

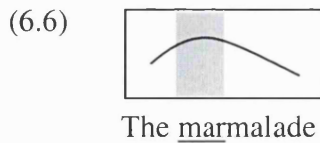


Whereas non-final or prenuclear accents were always characterised by a rising pitch around the stressed syllable followed by a fall, nuclear accents did not always show the rising trajectory. Instead, the F0 remained relatively low and level from the end of the fall of the preceding accent up to the nuclear one. Postnuclear syllables showed a fall from this low and level pitch. This is pictured in (6.5) for the sentence *Jo relies on my money*. In this utterance, no rise is observed over the last accented syllable.



Finally, utterances consisting of only one stress were produced with an F0 peak over the accented syllable. As before, the fall occurred on the poststressed syllable/s. This

is illustrated in (6.6). If there were no poststressed syllables, the whole rise-fall movement was compressed on the accented syllable.



So far, one of the basic differences observed in the data between English and Central Catalan is that whereas in Central Catalan prenuclear and nuclear accents showed very different pitch patterns and hence they were analysed as different phonological categories, in English, prenuclear accents present similar characteristics to nuclear accents. In the following sections, the phonetic and phonological properties of English accents in neutral declaratives are analysed in more detail. Even though prenuclear and nuclear accents will have the same phonological interpretation, they will be analysed in separate sections, as we did for Central Catalan.

#### 6.4.1.1.1 Prenuclear accents

English prenuclear accents are realised with an F0 peak, which tends to be aligned near the end of the accented syllable and which is followed by a fall. As opposed to Central Catalan prenuclear accents, the location of the peak tends not to exceed (or exceeds by very little) the right-hand boundary of the accented syllable. Between peaks the F0 contour descends into a dip. The phonological interpretation of this rise-fall F0 movement clearly involves an H\* tone associated to the accented syllable, since, as will be observed later in this section, the F0 peak is unequivocally positioned within (or very close to) the limits of the accented syllable. However, before analysing the properties of the H\*, further research is needed on the status of the surrounding valleys. Basically, is the F0 dip a mere phonetic transition between peaks or is it specified as a tonal target? And if so, is the L a leading or a trailing tone of the starred H (i.e. L+H\* or H\*+L)?

In order to decide whether the F0 dip observed between peaks is basically due to a phonetic transition or is actually an intended tonal target, the scaling and the alignment of the valley in prenuclear accents were analysed in more detail. The

parameter used to investigate the status of the F0 dip was the number of unaccented syllables (or temporal distance) between accents. This parameter has been used cross-linguistically (English: Pierrehumbert 1980; Greek: Arvaniti and Ladd 1995, Arvaniti et al 1998; or Neapolitan Italian: D'Imperio 1999, among others) to clarify the tonal status of particular F0 turning points. The rationale proposed in these studies is that if the number of unaccented syllables between accents affects the alignment and the scaling of an F0 valley, then the dip is the result of a phonetic transition (i.e. the higher the number of syllables, the later and the lower the dip). If on the contrary, the F0 valley is fixed in time and F0 irrespective of the number of unaccented syllables, then the dip has to be interpreted as an L tonal target.

According to these statements, the following predictions can be made. If prenuclear accents correspond to H\*, 1) the F0 dip between peaks should fall gradually until just before the second H\* at which point the F0 rises abruptly, and 2) the F0 dip should become deeper as the temporal distance between peaks (or the number of intervening syllables) increases. If prenuclear accents are H\*+L, the valley is expected to have a rather constant F0 value irrespective of the number of intervening syllables between peaks and the L is also expected to be fixed in time with respect to the *preceding* H\*. If prenuclear accents are L+H\*, the L has to have a constant F0 value and has to be aligned at a fixed position with respect to the *following* H\*.

Before analysing the scaling and alignment of L values, it was considered necessary to confirm the fact that an increase in the number of syllables does correspond to an increase in temporal distance. Thus, productions were divided into three sets of data according to the number of intervening syllables between accents: i) 1-2 syllables (29 observations), ii) 3-4 syllables (39 observations), and iii) 5-8 syllables (18 observations). For each set, the duration of the segmental material between accents was measured. The mean duration of the three sets of unaccented syllables are displayed in Table 6.2. As expected, an increase in the number of unaccented syllables or segmental material between accents corresponds to an increase in the temporal distance.

	Number of unaccented syllables		
	1-2	3-4	5-8
Mean duration (s)	0.482	0.766	1.026

Table 6.2. Mean duration of various numbers of unaccented syllables between accents.

Given these results, the scaling of the F0 dip was analysed in the first place. F0 values were obtained for all valleys (or lowest F0 point between peaks) for the three sets of data. Figure 6.1 plots the mean values and the standard deviations for the scaling of L as a function of the number of unaccented syllables between accents. The results suggest a rather fixed scaling of the L value, since the F0 dip does not increase as the number of intervening syllables increases. The data were analysed statistically using an analysis of variance in which the independent variable was the number of unaccented syllables. Setting the  $p$  value at 0.01, the statistical analysis confirms that the number of intervening syllables between accents does not significantly affect the scaling of L ( $p=0.65$ ,  $F=1.83$ ).

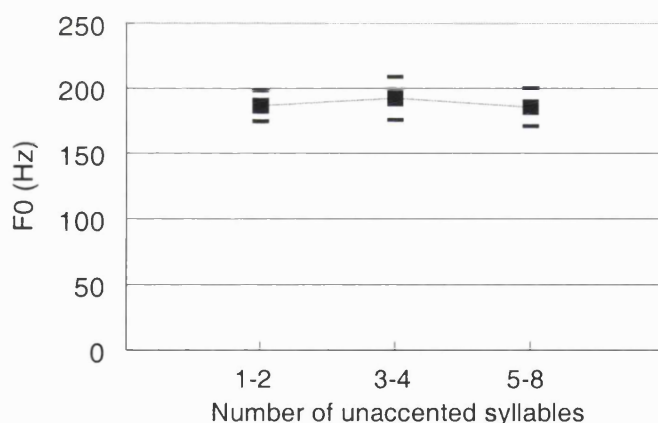
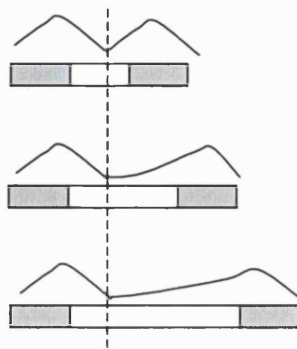


Figure 6.1. Mean values and standard deviations for the scaling of L as a function of the number of unaccented syllables between accents.

These findings indicate that the L is not the result of a phonetic transition but seems to be specified as a phonological target. Thus, the possibility of analysing prenuclear accents as monotonal accents ( $H^*$ ) is ruled out by the scaling evidence. Next, the status of the L as a leading tone ( $L+H^*$ ) or as a trailing tone ( $H^*+L$ ) is examined according to the alignment evidence. As stated before, 1) if L is part of  $H^*+L$ , its location has to be fixed with respect to the preceding  $H^*$ , and 2) if L is part of  $L+H^*$ , its location has to be fixed with respect to the following  $H^*$ . This is schematised in (6.7) below, where the shaded areas correspond to the accented syllables ( $H^*$ ) and the white areas show an increase in the number of unaccented syllables between accents.

(6.7) H\*+L H\*



H\* L+H\*

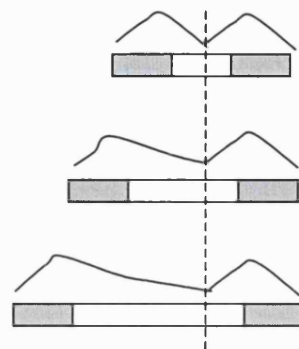


Figure 6.2 plots the mean values for the following measurements: 1) duration between the onset of the first H\* and L (on1-valley), and 2) duration between L and the onset of the second H\* (valley-on2) as a function of the number of unaccented syllables between accents. The results show that the L is located at a fixed position from the onset of the first accent and varies from the onset of the second accent. In other words, whereas the mean duration of distance “on1-valley” is very stable, the mean duration of the distance “valley-on2” alters according to the amount of segmental material. In order to confirm that the alignment of the L was fixed in time with respect to the previous H\* and that the number of unaccented syllables had no effect on its location, an analysis of variance was performed for the “on1-valley” sample. The statistical results confirmed that the number of unaccented syllables does not have an effect on the alignment of the F0 valley ( $p=0.079$ ,  $F=2.6$ ). Thus, the location of the L is stable in time with respect to the preceding H\* and hence, the phonological entity that better describes prenuclear accents in the data is H\*+L.

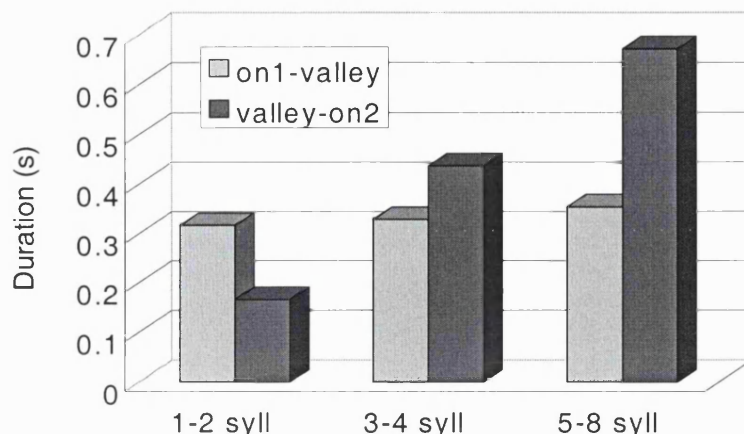


Figure 6.2. Mean values for the duration between the onset of the first accent and L (on1-valley) and for the duration between L and the onset of the second accent (valley-on2) as a function of the number of unaccented syllables between accents.

The results on the scaling and alignment of the F0 valley in English prenuclear accents of neutral declaratives suggest that the L is phonologically specified as the trailing tone of a bitonal accent (H\*+L). This analysis is at odds with the initial versions of the AM model. In Pierrehumbert (1980) or Beckman and Pierrehumbert (1986), H\*+L could not be applied to these kinds of F0 movements since this accent type was only used in downstepping sequences. Thus, the L was not a real low target but a downstep trigger of a following H accent. However, if downstep is interpreted as a property of the lowered tone itself (Ladd 1980, 1996) rather than the result of bitonality, then H\*+L can legitimately be used to describe a fall, as that observed in the data. The use of H\*+L rather than H\* has already been proposed in studies that present a left-headed account of pitch accents, such as Gussenhoven (1984) or Grabe (1998a).

Whereas the scaling and timing of the L have been observed to be rather constant, the alignment of the F0 peak shows more variation. Analyses of English intonation have pointed out several factors that influence the location of the peak. Two of these factors are the duration of the syllable and the presence of an upcoming prosodic boundary. As reported in Silverman and Pierrehumbert (1990) and House and Wichmann (1996), studies have shown that peak location can be expressed as a function of syllable (or rhyme) duration (i.e. the longer the syllable, the later the peak) and that a prosodic boundary has retracting effects in the location of the peak.

These tendencies are confirmed in the English data analysed in this study. In order to examine the effects of syllable duration and prosodic boundaries in the location of the peak, the same measurements obtained for Central Catalan prenuclear accents were collected for the English data.

1. Distance between the onset of the accented syllable and the location of the peak (on-peak).
2. Distance between the onset and the offset of the accented syllable, i.e. duration of the accented syllable (on-of)<sup>1</sup>.

These measurements were obtained for the three kinds of words (oxytones, paroxytones and proparoxytones). The results are presented in Figure 6.3. Each graph plots the distance between the onset of the accented syllable and the location of the peak (on-peak) against the duration of the syllable (on-of) in seconds, for all prenuclear accents of words with different stress distributions. The regression lines summarise the correlation between the two variables and the coefficients of correlation ( $R^2$ ) are displayed in each graph.

---

<sup>1</sup> Silverman and Pierrehumbert (1990) measured peak delay in relation to the duration of the sonorant rhyme rather than the duration of the whole syllable. In this study, however, the duration of the syllable is examined as we did for Central Catalan.



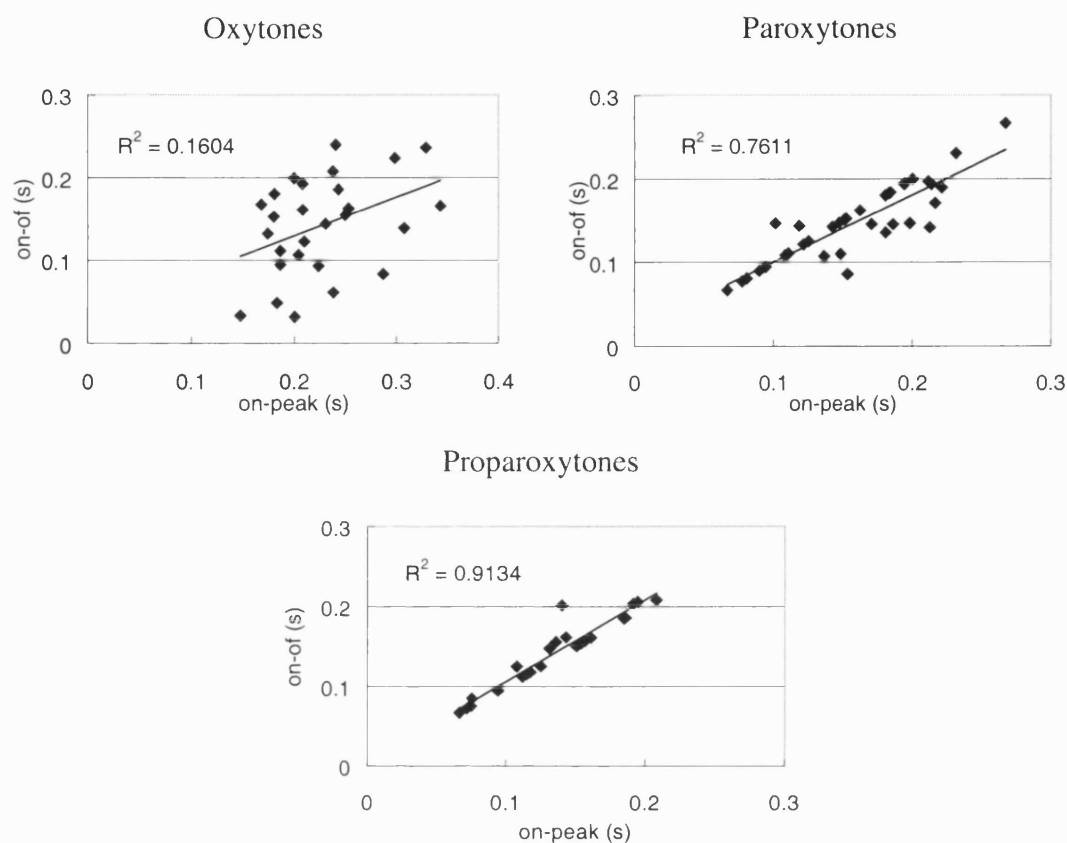


Figure 6.3. Distance between the onset of the accented syllable and the peak (on-peak) against distance between the onset and the offset of the accented syllable (on-of) in seconds for prenuclear accents in different kinds of words (oxytones, paroxytones and proparoxytones).

The plots in Figure 6.3 show that for paroxytones and for proparoxytones there is a close relationship between on-peak and on-of, indicating that the location of the peak can be expressed as a function of syllable duration. The coefficients of correlation between the two variables are very high ( $R^2=0.76$  and  $0.91$ ) in these kinds of words. For oxytones, on the contrary, the  $R^2$  value is rather low ( $R^2=0.16$ ), indicating that there is no correlation between syllable duration and location of the peak. However, this can be explained as the result of word boundary effects. It has been observed cross-linguistically (Silverman and Pierrehumbert 1990 for English, Kohler 1983 for German, Prieto et al 1995 for Mexican Spanish) that word boundaries have retracting effects on F0 movements. Since in oxytones the end of the word coincides with the end of the accented syllable, the peak is influenced by the presence of the immediate word boundary. This is observed in Table 6.3, which displays the mean values of F0 peak position (expressed as % of the duration of the accented syllable) for all kinds of

words, the mean accented syllable duration (on-of) and the mean distance between the onset of the accented syllable and the peak (on-p) in seconds. The presence of a boundary pushes the peak to an early position. Whereas in oxytones the peak is located near the middle of the accented syllable (62.3%) due to the presence of the word boundary, in paroxytones and in proparoxytones the peak is anchored towards the end of the accented syllable. In proparoxytones, the peak even exceeds the right boundary of the accented syllable in 5.7%.

	Oxytones	Paroxytones	Proparoxytones
F0 peak position (%)	62.3	93.1	105.7
on-of (s)	0.226	0.159	0.133
on-p (s)	0.141	0.148	0.139

Table 6.3. Mean values of F0 peak position (expressed as % of the duration of the accented syllable), the accented syllable duration (on-of) and the distance between the onset of the accented syllable and the peak (on-p) in seconds for prenuclear accents.

Despite variability in peak location derived from the influence of right-hand word boundaries, the English data show a relationship between syllable duration and peak placement. These observations corroborate the findings reported in many other studies (Pierrehumbert and Steele 1989, Silverman and Pierrehumbert 1990). These results, however, are very different from the results observed in Central Catalan prenuclear accents. The results presented in chapter 3 for Catalan (Figure 3.4) show that there was no relationship between the two variables. The findings of English corroborate the interpretation provided for the Central Catalan data. In the two sets of data prenuclear accents of neutral declaratives differ both in the F0 shape and in the phonological interpretation. In Central Catalan, prenuclear accents show a rising trajectory interpreted as a combination of an L\* pitch accent and an H word edge tone. In English, on the other hand, prenuclear accents show a rise-fall movement which is specified as H\*+L. The possibility of using a rising category (L\*+H) in English prenuclear declarative accents is not abandoned. Grabe (1998a) shows that both L\*+H and H\*+L can be attached to nuclear and prenuclear positions in English. However, L\*+H has not been observed in the productions of our speaker.

Finally, recent research on peak timing in British English (House et al 1999a, b, Dankovičová and House 2000) has shown that relating the peak location to the

duration of the foot (rather than the duration of the syllable) reduces variability in peak alignment. These observations are confirmed in our data. The statistical analysis of the data shows that the number of syllables per foot has some effect on the location of the peak (see Table 6.4). However, an analysis of variance shows that differences in the four samples are not significant ( $p=0.10$ ,  $F=2.12$ ). Thus, the peak seems to be located at a rather stable proportion of the foot<sup>2</sup>.

	1 syllable foot	2 syllable foot	3 syllable foot	4 syllable foot
F0 peak position (%)	52.4	45.9	36.6	34.2
Foot duration (s)	0.237	0.32	0.405	0.569
on-p (s)	0.124	0.147	0.148	0.195

Table 6.4. Mean values of F0 peak position (expressed as % of the foot duration), foot duration, and distance between the onset of the accented syllable and the peak (on-p) in seconds.

In Central Catalan, the timing of the peak in prenuclear accents as a function of the foot duration was examined for one speaker (see Table 3.2 in chapter 3). The results showed that there was no relationship between foot duration and peak location, since in feet with similar duration but different location of word boundaries, the peak was always anchored at or near the end of the word. The results obtained for English confirm the analysis proposed for Central Catalan where prenuclear peaks were interpreted as H word edge tones.

#### **6.4.1.1.2 Nuclear accents**

The phonetic properties of the nuclear accent in the English data were very similar to those of prenuclear accents, suggesting that the same phonological category ( $H^*+L$ ) might be used for both non-final and final accents. The interpretation of nuclear accents as  $H^*+L$  is based on two aspects: 1) an auditory judgement and 2) an acoustic analysis of the data.

---

<sup>2</sup> In line with House et al (1999a, b), we have assumed that a foot is made up of a stressed syllable optionally followed by less salient syllables (Abercrombie 1965). Other studies on foot structure have limited the number of syllables after the stressed one. Whereas for Selkirk (1980) feet have an upper number of three syllables, Hayes (1982) argues that feet should be limited to two syllables.

Auditorily, the last pitch accent was perceived as similar to the trajectory observed in prenuclear accents when the last part of the utterance was removed<sup>3</sup>. This indicates that the rise-fall pitch movement of nuclear and prenuclear accents should be described by means of the same tonal category.

Further evidence to support the use of the same phonological entity to describe prenuclear and nuclear accents is found in the acoustic analysis of the productions. As in Central Catalan, the phonetic properties of the final pitch accent in English neutral declaratives differs depending on the number of accents in the utterance. If the utterance consists of one stressed item, a clear rise is observed within the accented syllable followed by a fall. In this case, the peak shows late alignment within the accented syllable although the exact peak location is subject to boundary effects. That is, the nearer the word boundary (or phrase boundary) the earlier the peak. These observations are confirmed by the results presented in Table 6.5, which displays the mean values of F0 peak position (expressed as % of the duration of the accented syllable), the mean accented syllable duration and the mean distance between the onset of the accented syllable and the peak in words produced in isolation. The results show that in oxytones the location of the peak is earlier than in paroxytones and proparoxytones due to the presence of the word (and phrase) boundaries. In oxytones, the peak is located at 33.7% within the accented syllable, showing a rather early peak. In the other kinds of words, the peak is located at about 73% of the syllable duration.

	Oxytones	Paroxytones	Proparoxytones
F0 peak position (%)	33.7	73.5	72.8
on-of (s)	0.371	0.202	0.176
on-p (s)	0.125	0.147	0.128

Table 6.5. Mean values of F0 peak position (expressed as % of the duration of the accented syllable), the mean accented syllable duration (on-of) and the mean distance between the onset of the accented syllable and the peak (on-p) in seconds in isolated words.

These results are similar to those observed in Central Catalan (Table 3.2) for words in isolation. In Central Catalan, the peak in oxytones was located at 50% of the syllable duration and in paroxytones and proparoxytones at about 88%. The similar behaviour

---

<sup>3</sup> This comment is based on the researcher's own perception and on an informal perception test carried out by a native English speaker. Experimental verification of the claim is needed.

observed in the two languages validates the interpretation of the Central Catalan nuclear accents as H\*L.

In English multi-stressed sentences, the phonetic properties of the nuclear accent are different from those of single-stressed sentences. The final accent of multi-stressed utterances presents two realisations: either a low (downstepped) peak aligned very early within the accented syllable or no peak at all. The percentage of sentences produced with no peak is 54.3% and that of structures with a final peak is 45.7%. These two patterns can be interpreted as phonetic realisations of the same pitch accent, since no relevant differences are perceived between sentences produced with the two trajectories. The lack of peak then is understood as the extreme end of a downstepping continuum.

The alignment of the peak in those sentences which present an F0 maximum is shown in Table 6.6, which displays the mean values of F0 peak position (expressed as % of the duration of the accented syllable), the mean accented syllable duration and the mean distance between the onset of the accented syllable and the peak. As expected, the results show that in oxytones the peak is earlier within the accented syllable (at 28.8%) than in paroxytones (43.2%) and proparoxytones (42.4%).

	Oxytones	Paroxytones	Proparoxytones
F0 peak position (%)	28.8	43.2	42.4
on-of (s)	0.267	0.194	0.186
on-p (s)	0.077	0.084	0.079

Table 6.6. Mean values of F0 peak position (expressed as % of the duration of the accented syllable), the mean accented syllable duration (on-of) and the mean distance between the onset of the accented syllable and the peak (on-p) in seconds for the nuclear accent of multi-stressed sentences.

A comparison of these percentages to those obtained for words in isolation in English (Table 6.5 above) shows that the peak is aligned much earlier in nuclear accents of multi-stressed sentences than in nuclear accents of isolated items. These differences might be accounted for by the presence of final lowering on the last peak of multi-stressed utterances. It has been observed (Pierrehumbert and Beckman 1988) that the last H in declarative sentences undergoes a substantial F0 reduction. This implies a shorter F0 excursion and most probably an early peak location.

The intonation at the end of English sentences is interpreted by means of an L- phrase accent L- and an L boundary tone. Studies that analyse English intonation by means of H\*+L and L\*+H only (e.g. Grabe 1998a) consider the presence of phrase accents unnecessary and only specify the tonal characteristics at the end of an IP. Since this study does not intend to analyse the levels of phrasing in English, we will keep the convention of having two edge tones at the end of a contour, indicating two levels of prosodic phrasing.

Finally, the phonological interpretation of the sentences in (6.4)-(6.6) above would be as presented in (6.8)-(6.10).

(6.8) Amelia married a marine  
           H\*+L                      H\*+L L-L%

(6.9) Mary learns modern languages  
           H\*+L                      H\*+L L-L%

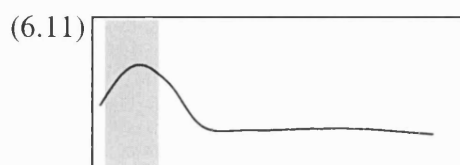
(6.10) marmalade  
           H\*+L L-L%

Cross-linguistically, English and Central Catalan nuclear accents are interpreted in a similar way, namely, as a combination of two tones: H\*(+)L. Whereas in English the L target is interpreted as the trailing tone of a bitonal accent (H\*+L), in Central Catalan the L is an edge tone marking the end of a word (H\*L). In both languages, the H tone is downstepped (or strongly lowered) with respect to the preceding H target.

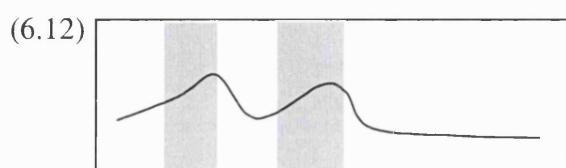
#### **6.4.1.2 Narrow focus**

As expected, narrow focus structures were produced with an alteration of the “normal” stress prominence pattern ( $w-s > s-w$ ), with the subsequent attachment of the nuclear accent to the syllable that bears the main sentence stress (i.e. the syllable dominated by the maximum number of *s* nodes) and with deaccenting of old, postfocal material. Sometimes prefocal elements were also deaccented, although deaccenting of prenuclear items was optional. Examples of utterances produced with

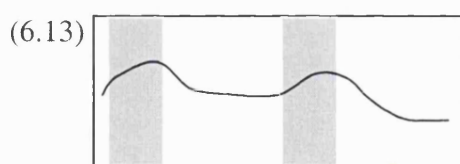
narrow focus are presented in (6.11)-(6.13) below. (6.11) illustrates an utterance with narrow focus on the subject; (6.12) on the verb and (6.13) on the object. (6.14) presents an example with narrow focus on the object and deaccenting of prefocal material.



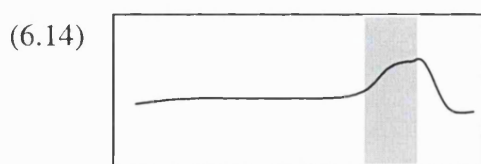
MILLIE lives in Ireland



The neighbour MARRIED the model



John married MELANIE



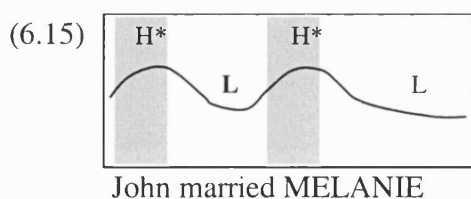
Gary borrowed the MONEY

The phonological interpretation of the nuclear accent in narrow focus structures varies in the literature. In Ladd (1980, 1996) and Gussenhoven (1983b, 1984, 1994), the nuclear accent of narrow focus sentences has been classified in the same way as the nuclear accent of a broad focus utterance. This involves ambiguity between a late narrow focus and a broad focus interpretation. Furthermore, possible differences in the pitch range between narrow focus and broad focus accents (i.e. in narrow focus, the pitch range tends to be wider than in broad focus) are explained as the result of emphasis but not as a distinct linguistic choice. Other times, however, two different accents have been used for the two focal readings (Venditti et al 1996): i) H\* for

broad focus and ii) L+H\* for narrow focus. Basically, the main difference between H\* and L+H\* is that in the bitonal accent the peak rises from a much lower level than in the monotonal one. According to this interpretation, differences in the pitch excursion of the two accents are incorporated in the phonological system and hence the two focal readings are linked to two different accent types. This involves no ambiguity between late narrow focus and broad focus.

The results obtained in this study support the claim of a unique accent type for broad and narrow focus nuclear accents and hence differences in the pitch range of the accent are interpreted as the result of a phonetic boost rather than a different tonal category. However, the accent type that we propose as the nuclear tone of narrow focus utterances is neither H\* nor L+H\*, but H\*+L as in broad focus sentences. The choice of H\*+L as the nuclear accent of narrow focus productions derives from three pieces of evidence: 1) the scaling and the alignment of the valley before the nuclear peak, 2) the alignment of the L after the nuclear peak, and 3) the alignment properties of the nuclear peak itself. These three aspects are analysed below.

The first parameter that we will use to identify the status of the L before the nuclear peak involves an analysis of its scaling and alignment properties. As in broad focus sentences, the context used to examine the F0 and timing characteristics of the valley is between peaks, as in the sequence H\*LH\*. This is illustrated (6.15) for the utterance *John married MELANIE* with narrow focus on the object.



As argued in section 6.4.1.1.1, the L in this kind of context can be the result of three aspects: 1) a phonetic transition between peaks, 2) a tonal target specified as a leading tone of the second H, as in H\* L+H\* or 3) a tonal target specified as a trailing tone of the first H, as in H\*+L H\*. Similar to broad focus sentences, the first option is thrown into doubt by the fact that an increase in the amount of segmental material (or in the number of unaccented syllables) between accents did not affect the scaling of the L.



This is shown in Figure 6.4, which plots the mean values and standard deviations for the scaling of the valley as a function of the number of unaccented syllables between peaks. The results show that the scaling of the L is independent of the number of intervening syllables. The results of an ANOVA confirmed that there are no significant differences in the scaling of L in the three sets of unaccented syllables ( $p=0.69$ ,  $F=0.36$ ). Thus, the F0 valley between H tones is not the result of a phonetic transition but seems to be a linguistic choice.

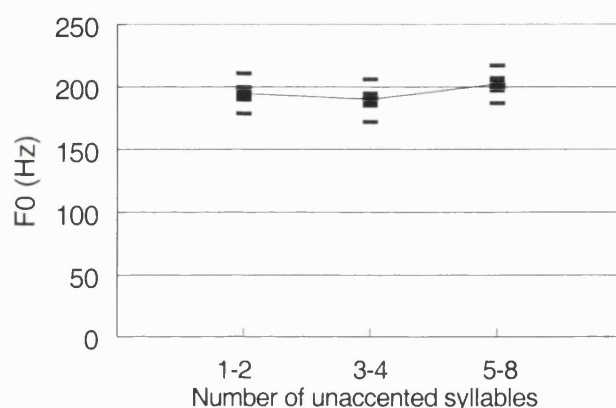


Figure 6.4. Mean values and standard deviations for the scaling of L as a function of the number of unaccented syllables between accents in sentences with narrow focus on the second element.

The next step then is to decide whether the L is a leading tone of the prenuclear accent ( $H^*+L$ ) or the trailing tone of the nuclear accent ( $L+H^*$ ), as has been proposed by some researchers. Thus, as in broad focus utterances, we compared the alignment of L with respect to the onset of the first  $H^*$  and to the onset of the second  $H^*$ , with the following expectations: 1) if L is part of  $H^*+L$ , its location has to be fixed in time with respect to the first  $H^*$ , and 2) if L is part of  $L+H^*$ , its location has to be fixed in time with respect to the second  $H^*$ .

Figure 6.5 plots the mean values for 1) the duration between the onset of the first accent and L (on1-valley) and 2) the duration between L and the onset of the second accent (valley-on2), as a function of the number of unaccented syllables between accents. As in Figure 6.2, the results show that the L is fixed in time with respect to the onset of the first accent and varies from the onset of the second accent. An ANOVA performed on the “on1-valley” data showed that differences in the number of unaccented syllables between peaks have no significant effect on the alignment of

L ( $p=0.85$ ,  $F=0.162$ ). Thus, the L is attached to the first H\* rather than to the second H\*. These findings have two implications. First, they confirm the analysis of prenuclear accents as H\*+L in both broad and narrow focus sentences. Second, they throw into doubt the phonological interpretation of the nuclear accent in narrow focus as L+H\*.

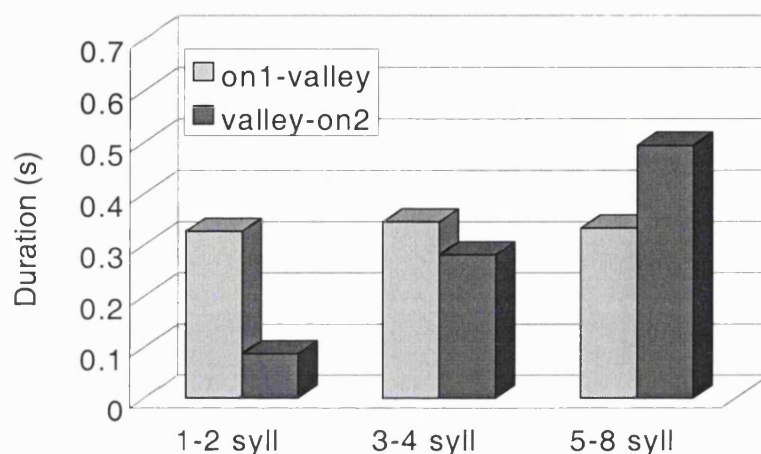


Figure 6.5. Mean values for the duration between the onset of the first accent and L (on1-valley) and for the duration between L and the onset of the second accent (valley-on2) as a function of the number of unaccented syllables between accents in narrow focus.

So far, the results on the scaling and alignment of L have shown that the valley in H\*LH\* sequences is part of the first accent, and hence the sequence is re-examined as H\*+L H\*. However, the nuclear accent can still be specified as H\*+L or an H\* followed by L edge tones. In order to determine the tonal characteristics of the H\*L sequence, the location of the valley after the nuclear accent was analysed in more detail. As in Central Catalan, the following measurements were performed for all nuclear words:

1. Distance between the onset of the accented syllable and the location of the valley (on-valley).
2. Distance between and the onset of the accented syllable and the end of the word (on-end) for oxytones, paroxytones and proparoxytones.

As for the Central Catalan data, the hypotheses are the following: if the L is part of the bitonal accent a very low correlation between the two distances is expected, as the

L would be anchored at a fixed location with respect to the starred tone. If, as in Central Catalan, the L marks the presence of some kind of boundary, it will be anchored at the end of the word irrespective of the number of poststressed syllables and hence a high correlation between the two measurements is expected.

The results are displayed in Figure 6.6. Each graph plots the distance between the onset of the accented syllable and the location of the valley against the distance between the onset of the accented syllable and the end of the word for oxytones, paroxytones and proparoxytones. The coefficients of correlation between the two variables and the regression lines are displayed in each graph. The results show a very low correlation between the two variables for all kinds of words ( $R^2$  varies from 0.02 to 0.25), suggesting that the L is not anchored at the end of the word. This behaviour is different from the Central Catalan data where the L was clearly located at the offset of the word and marked the right boundary of an ip.

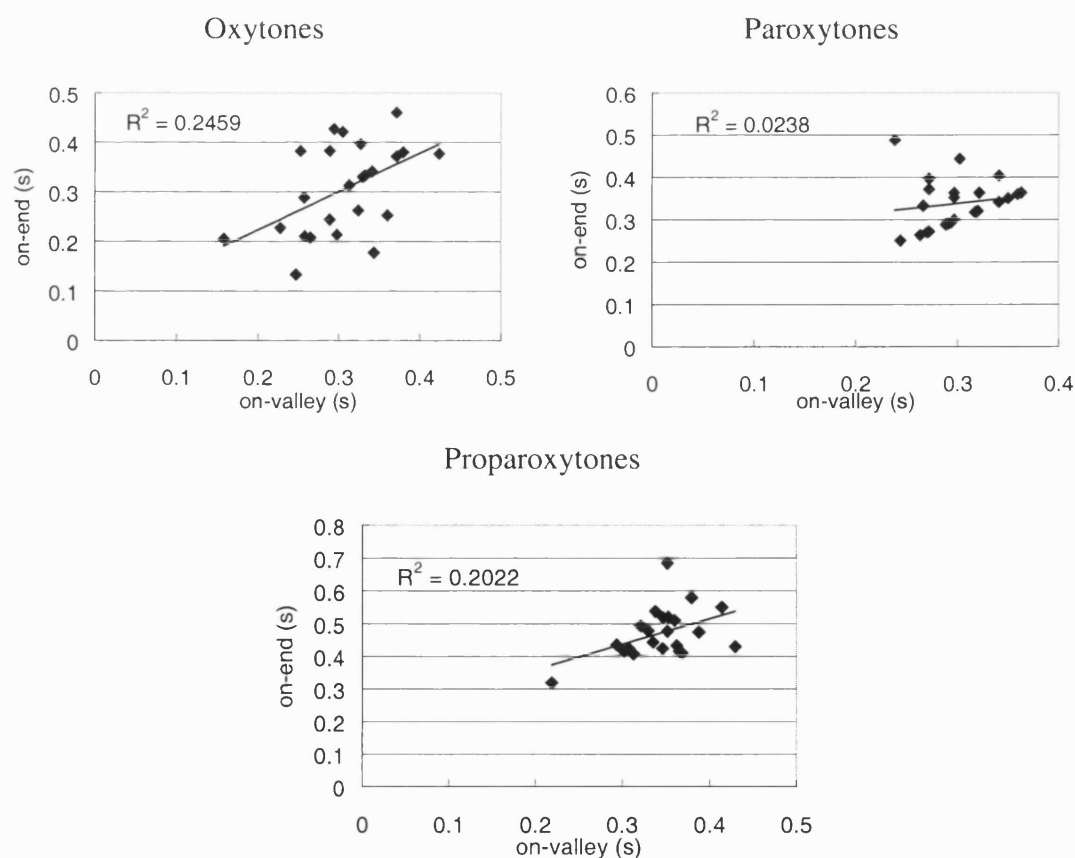


Figure 6.6. Distance between the onset of the accented syllable and the valley (on-valley) against distance between the onset of the accented syllable and the end of the word (on-end) in seconds for nuclear accents in different kinds of words (oxytones, paroxytones and proparoxytones).

The results observed so far indicate that the nuclear accent of English narrow focus sentences cannot be interpreted as H\*L but seems to respond to a bitonal accent (H\*+L). In order to confirm this possibility, we performed an analysis of variance on the “on-valley” measurements for the three groups of words (oxytones, paroxytones and proparoxytones). The expectations were that if L is the trailing tone of the bitonal accent, no significant differences should be observed on the “on-valley” for the three types of words. The results of the ANOVA show no significant differences in the location of L in the three word types ( $p=0.03$ ,  $F=6.08$ ). The mean and the SD for the three word types are as follows: oxytones (mean=0.31, SD=0.002), paroxytones (mean=0.30, SD=0.001) and proparoxytones (mean =0.33, SD=0.001). These results indicate that the L has a fixed location in time with respect to the starred tone (or onset of the accented syllable) and confirm the interpretation of the nuclear accent in English narrow focus sentences as H\*+L.

The results observed in the English data led us to review some of the former phonological interpretations of narrow focus utterances. Within the AM model, the traditional characterisation of the valley (and subsequent low plateau) after the nuclear accent in narrow focus structures was an L- phrase accent (Pierrehumbert 1980, Beckman and Pierrehumbert 1986). This low phrasal tone covered the syllables between the focussed word and the phrase boundary. In early versions of the model (Pierrehumbert 1980), phrase accents were treated as floating tones which accounted for the pitch between the last pitch accent and the boundary tone. However, in later proposals, phrase accents became markers of an intermediate level of prosodic phrasing (Beckman and Pierrehumbert 1986, Pierrehumbert and Beckman 1988) and subsequently the L- had to be associated at the end of the phrase. This new status of the phrase accent is not consistent with the former account of postfocal material and the idea of a floating tone. Now the L- is not floating any more but it is associated to a fixed location. This meant that in order to account for the low plateau between the nuclear H\* and the end of the phrase, a tone spreading rule had to be formulated, which claimed a leftward spreading of the L- phrase accent up to the nuclear accent. However, if we analyse the nuclear accent of narrow focus productions as H\*+L, the discrepancy derived from the two versions of the theory seems to be solved. On the one hand, H\*+L accounts for the immediate valley after the focal peak. On the other, the L- phrase accent is associated to the ip domain and there is no need to postulate a

spreading of the tone. Between H\*+L and L the F0 is low and level as predicted by interpolation rules.

Finally, the third piece of evidence that corroborates the analysis of the focal accent as H\*+L is the similar characteristics as far as the timing of the peak is concerned to the neutral nuclear accent. This is observed in Tables 6.7-6.9, which show the mean values of F0 peak position (expressed as % of the duration of the accented syllable), the mean accented syllable duration and the mean distance between the onset of the accented syllable and the peak for all kinds of words (oxytones, paroxytones and proparoxytones) in different focal positions (subject, verb and object).

SUBJECT	Oxytones	Paroxytones	Proparoxytones
F0 peak position (%)	61.9	81.2	89.3
on-of (s)	0.21	0.164	0.159
on-p (s)	0.13	0.13	0.13

Table 6.7. Mean values of F0 peak position (expressed as % of the duration of the accented syllable), the accented syllable duration (on-of) and the distance between the onset of the accented syllable and the peak (on-p) in seconds for the nuclear accent of a subject focal domain.

VERB	Oxytones	Paroxytones	Proparoxytones
F0 peak position (%)	68.3	84	93.5
on-of (s)	0.223	0.182	0.139
on-p (s)	0.15	0.15	0.13

Table 6.8. Mean values of F0 peak position (expressed as % of the duration of the accented syllable), the accented syllable duration (on-of) and the distance between the onset of the accented syllable and the peak (on-p) in seconds for the nuclear accent of a verb focal domain.

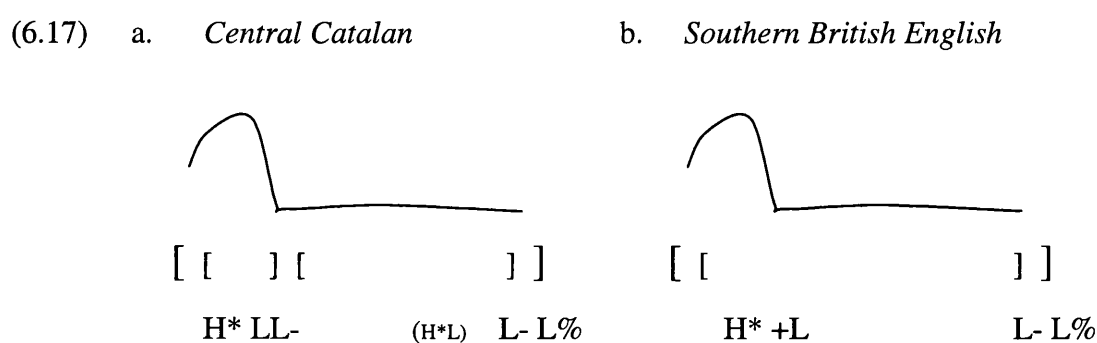
OBJECT	Oxytones	Paroxytones	Proparoxytones
F0 peak position (%)	33.2	71.4	68.7
on-of (s)	0.37	0.215	0.16
on-p (s)	0.12	0.15	0.11

Table 6.9. Mean values of F0 peak position (expressed as % of the duration of the accented syllable), the accented syllable duration (on-of) and the distance between the onset of the accented syllable and the peak (on-p) in seconds for the nuclear accent of an object focal domain.

The results show that the timing of the F0 peak in items produced with narrow focus is consistent with the findings reported so far for broad focus sentences. The peak is

aligned late within the syllable in paroxytones and proparoxytones and early in oxytones due to the presence of the word boundary. In items with narrow focus on the object, the location of the peak is earlier in all kinds of words due to the upcoming phrase boundary. The results of narrow focus on the object are very similar to those of single-stressed structures (Table 6.5) and differ from the results on the nuclear accent of multi-stressed sentences (Table 6.6). This is probably due to the F0 boost observed in the nuclear accent of some narrow focus productions. The similar behaviour in terms of peak alignment between single-stressed structures and late narrow focus confirms the claim that the early peak on broad focus final accents is due to the small F0 excursion caused by final lowering.

The English results corroborate the idea that narrow focus utterances are produced with an alteration of the abstract prominence pattern of the sentence involving the location of the nuclear accent to the strongest element of the tree and deaccenting of old, postfocal material. As in Central Catalan, the same accent type seems to be used for both neutral and marked structures. However, whereas Central Catalan introduces an ip boundary after the focal element with optional deaccenting of postfocal material, English does not signal focus by means of phrasing and hence deaccentuation is compulsory in this language. A sketched contour of the different ways of signalling focus in the two languages is presented in (6.17) below.



### 6.4.2 Downtrends

As for Central Catalan, the behaviour of F0 downtrends was examined in English broad focus and narrow focus sentences. With this analysis we expected to confirm the results obtained in other studies on English downtrends and to corroborate the

validity of the Central Catalan data by analysing the English utterances in a similar way. Since the behaviour of L values has already been examined in sections 6.4.1.1.1 and 6.4.1.2, in this section we will basically concentrate on the analysis of the higher part of the tonal space or H values.

As stated in chapter 5, studies on English downtrends (Lieberman and Pierrehumbert 1984, Beckman and Pierrehumbert 1986, Pierrehumbert and Beckman 1988) have reported three sources of F0 lowering in American English sentences:

- 1) declination (global phenomenon which operates in time from the beginning of the utterance, irrespective of the tonal specification);
- 2) downstep (local, stepwise lowering of pitch at specific accents);
- 3) final lowering (compression of the final accent of a declarative whose F0 is lower than predicted by the scaling of the preceding accent).

According to these studies, broad focus English declaratives are basically described as a series of local downstepping accents, which decay by a fixed proportion. The last accent of the declarative is expected to undergo a greater amount of lowering than the predicted by the downstep ratio. Finally, declination, if present, is a residue downtrend that appears after all other lowering mechanisms have applied. In narrow focus sentences, the presence of a focussed element suspends the downstepping sequence.

In line with these studies, H downtrends in our data are expected to respond to a series of localised phonological events. In narrow focus sentences, we anticipate the blockage of downstep on the focal constituent. However, it is expected to find differences in the degree of suspension concerning the kind of narrow focus: contrastive vs identification. As reported in the literature, contrastive focus tends to be more salient than identification focus (Brown et al 1980, Couper-Kuhlen 1986, García-Lecumberri 1995) and hence downstep is more likely to be blocked.

In order to analyse the nature of F0 downtrends in the English productions, the same points as for Central Catalan sentences were labelled in the F0 contours of both broad focus and narrow focus utterances:

1. Phrase-initial F0 value
2. Highest F0 value for each pitch accent
3. Lowest F0 value between peaks
4. Phrase final F0 value

### 6.4.2.1 Broad focus

Unlike Central Catalan utterances, English sentences were more homogeneous as far as the number of stresses per utterance are concerned. The majority of sentences consisted of 3 stresses (88% of sentences) and only a few of them had 4 stresses (12% of sentences). As pointed out in section 6.4.1.1, not all stresses became accented. Sentences with 3 stresses tended to get two pitch accents (on the first stressed syllable and on the last stressed syllable). Sentences with 4 stresses had 3 accents (usually the second stressed syllable did not become accented). The schematised mean F0 contours of utterances with a different number of stresses (and accents) are presented in Figure 6.7. The L and H letters indicate the mean values of F0 minimum and F0 maximum obtained in the analysis respectively. The number of stresses and accents (in brackets) are specified on the graphs.

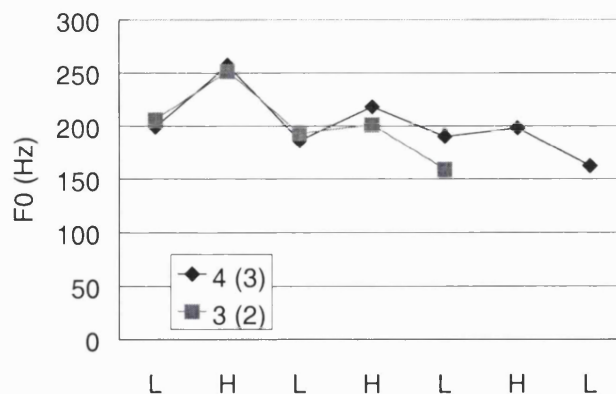


Figure 6.7. Schematised mean F0 contours for neutral declarative sentences with different number of stresses and accents (in brackets).

The results in Figure 6.7 show that neutral declaratives are produced with a sequence of downstep accents. As may be seen on the graph, the number of stresses per



sentence (and subsequently the sentence length) does not seem to affect the degree of prominence or pitch range with which sentences were produced. T-tests showed no significant differences in utterance-initial F0 values for phrases of different lengths at 1 per cent significance level ( $p=0.07$ ,  $t=1.66$ ). Similarly, utterance-final F0 values were also stable and no significant differences were observed with respect to sentence length ( $p=0.16$ ,  $t=1.66$ ). In addition, the F0 range (difference in Hz from the lowest F0 value at the beginning of an utterance to the highest F0 point of the first peak) was rather constant in sentences with different number of stresses. This is observed in Table 6.10. These results agree with those of Liberman and Pierrehumbert (1984) for American English and those of Central Catalan presented in chapter 5. This indicates that there is no preplanning of downstep at the sentence level and corroborates the local nature of downstep.

Stresses	F0 range	p	t
3	46.9	0.09	1.66
4	55		

Table 6.10. Mean F0 range in utterances with three and four stresses and results of t-tests comparing the F0 range values.

In order to see whether a declination component was operating along with downstep in the English sentences, the values of H1 and H2 were compared in sentences with two accents but with a different number of intervening syllables between them. Sentences were divided into three groups according to the number of unaccented syllables between accents: i) 1-2 syllables, ii) 3-4 syllables and iii) 5-8 syllables. Given the results reported in other studies on English downtrends, it was expected to find little effect from the number of unaccented syllables between accents (or temporal distance) on the F0 value of H2. The results are plotted in Figure 6.8, which shows the mean F0 values for H1 and H2 with different number of intervening syllables and in Table 6.11, which displays the mean F0 difference between H1 and H2 with 1-2, 3-4 and 5-8 unaccented syllables. The results show that the scaling of H2 is very little affected by the number of unaccented syllables. Even though the values in Table 6.11 increase as the number of syllables increases, the statistical analysis (ANOVA) shows that these differences are not significant ( $p=0.13$  and  $F=2.11$ ). Thus, the data confirm the analysis of English F0 downtrends as a series of step accents.

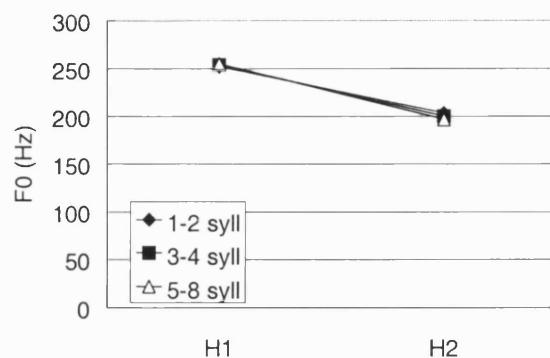


Figure 6.8. Mean F0 values for H1 and H2 corresponding to an increase in the number of intervening syllables (1-2, 3-4 or 5-8).

1-2 syllables	3-4 syllables	5-8 syllables
48.5	54.1	58.2

Table 6.11. Mean F0 difference between H1 and H2 with 1-2 and 3-5 intervening syllables for the two speakers and in different sentence conditions.

One final issue derived from the data is the question whether the last pitch accent undergoes final lowering or not. Liberman and Pierrehumbert (1984) carried out several experiments, which showed that the final accent of a declarative is lower in F0 range than predicted by the location of the immediately preceding accent. However, recent work on British English (Grabe 1998a, Grabe to appear) has shown that, unlike American English, Southern Standard British English does not have final lowering. The results of her data demonstrate that the last pitch accent was affected by the amount of segmental material between accents (steps between successive accents increased as more material was inserted) suggesting the presence of declination. Obviously, to be able to support one of these two views more data and more speakers would be needed. However, from the observations gathered so far, final lowering does seem to operate in the data. This is illustrated in Figure 6.9, which plots 1) the mean peak values of H1 and H2 in sentences with three stresses and two pitch accents (produced with a different number of intervening syllables between accents), and 2) the mean peak values of H1, H2 and H3 in sentences with four stresses and three pitch accents. The results show that H2 (or final accent) in (3,2) sentences is much lower than the second H2 (non-final accent) in (4,3) utterances, indicating the presence of final lowering on the last accent in (3,2) utterances.

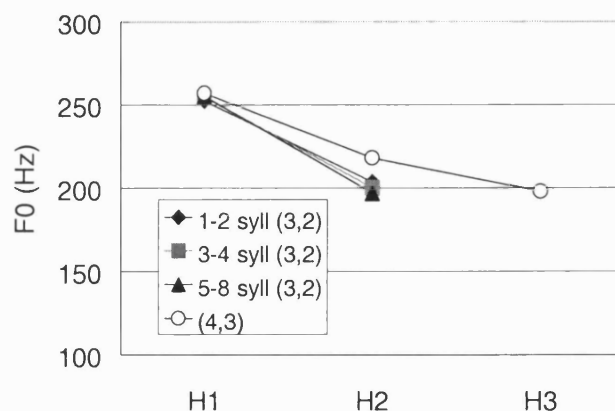


Figure 6.9. Mean F0 values for H1 and H2 in sentences with 3 stresses but 2 accents (produced with a different number of unaccented syllables between accents) and mean F0 values for H1, H2 and H3 in sentences with 4 stresses but 3 accents.

Overall, the results of H F0 downtrends in the English data confirm the findings of Liberman and Pierrehumbert (1984), Beckman and Pierrehumbert (1986), and Pierrehumbert and Beckman (1988) for American English. Broad focus English declaratives are produced with a series of local downstepping accents. The last accent of the declarative shows a greater amount of lowering than the predicted by the downstep rule. This differs from the results of Grabe (1998a, to appear) who shows that in British English final lowering is absent. The results observed so far for the English data are similar to the Central Catalan findings. In the two languages, F0 downtrends emerge from a sequence of step accents. However, whereas in English the step accent is always a pitch accent (!H\*+L) in Central Catalan it can be both a word edge tone (!H) and a pitch accent (!H\*).

#### 6.4.2.2 Narrow focus

Since English broad focus sentences were produced with a series of downstepped pitch accents, in narrow focus it is expected that downstep is blocked, as reported in Beckman and Pierrehumbert (1986). Furthermore, the suspension of downstep is expected to be more common in contrastive responses rather than in identification contexts since contrastive focus tends to be more salient than identification focus (Brown et al 1980, Couper-Kuhlen 1986, García-Lecumberri 1995).

The schematised mean F0 contours of narrow focus utterances are presented in Figure 6.10. The graph on the left-hand side shows the contours for sentences with narrow focus in medial position (focus on the verb). The graph on the right-hand side shows sentences with in final narrow focus or focus on the object. In both graphs, the mean values of sentences with narrow focus triggered by contrast are plotted separately from the mean values of sentences triggered as answers to identification prompts.

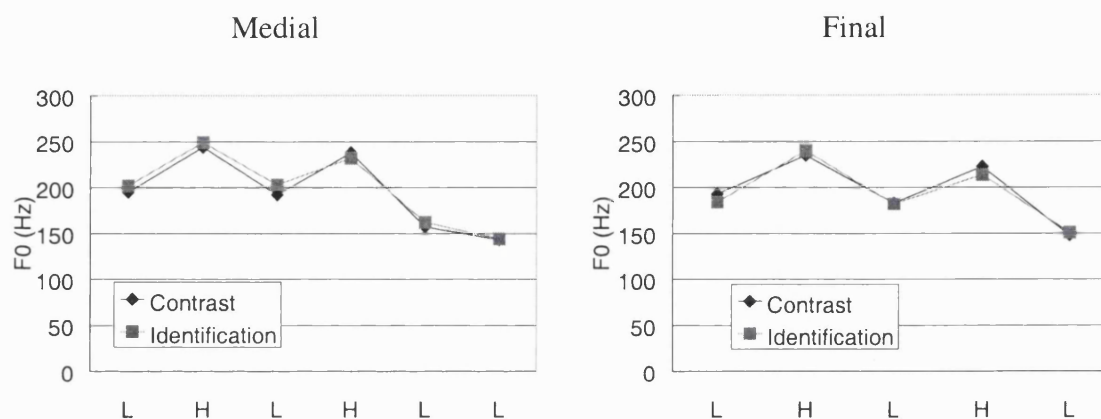


Figure 6.10. Schematised F0 contours for sentences produced with focus in medial position and in final position triggered by contrast and by identification

In Figure 6.10, no apparent differences were observed in the F0 contours of sentences triggered by contrast and those triggered by identification. However, a more thorough analysis of the data showed that the kind of triggering question did have an effect on the realisation of downstep on the focal item. As reported in the literature, the suspension of downstep is expected to be more common in contrast than in identification contexts. This is observed in Figure 6.11, which plots the mean F0 value of H1 and H2 for the two kinds of triggers and in Table 6.12 which shows the results of a t-test comparing the height of H1 and H2 in the two contexts and in different focal positions.

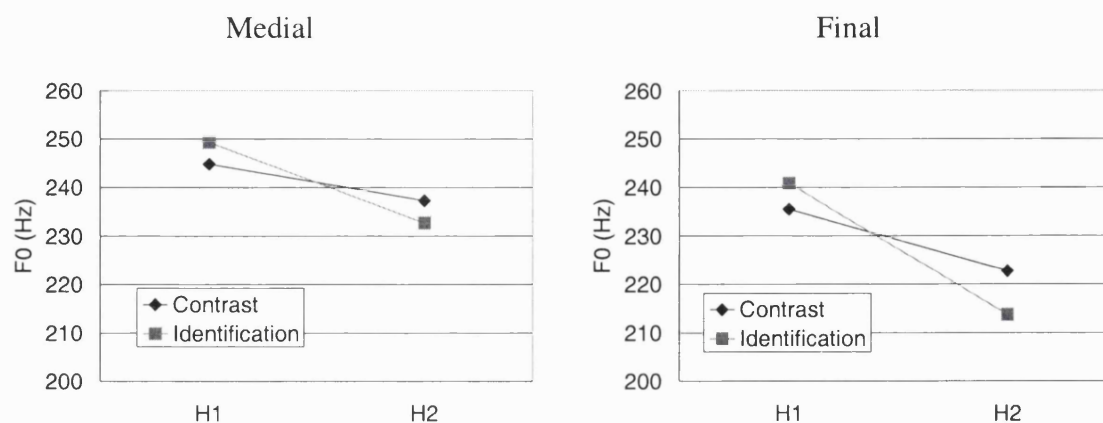


Figure 6.11. Mean H1 and H2 for medial and final narrow focus declaratives triggered by a contrastive and an identification prompt.

	Medial	Final
Contrast	p=0.17 t=1.73	p=0.069 t=1.73
Identification	p=0.002 t=1.72	p=0.002 t=1.75

Table 6.12. Results of the t-test comparing the height of H1 and H2 in medial and in final position for narrow focus triggered by contrast and by identification.

The results show that downstep is blocked when narrow focus is triggered by contrast but not when it is triggered by an identification prompt. As observed in Table 6.12, the height of H1 and H2 in contrastive sentences is not significantly different (at 1 per cent significance level), indicating that the H2 is not much lowered with respect to H1. In identification sentences, on the other hand, the height of H1 and H2 is significantly different. H2 is produced at a much lower frequency compared to H1 and hence it is downstepped. These results are quite similar to the Central Catalan results. Basically, they throw into doubt the idea that narrow focus always blocks downstep but support the claim that different kinds of narrow focus might be more likely to suspend the downstepping sequence. As expected, downstep is more inclined to be blocked in contrastive contexts than in identification contexts.

## 6.5 Discussion

In this chapter, the phonetic and phonological properties of broad focus and narrow focus sentences have been analysed for one speaker of Southern British English. The main motivation to include some English data in the final part of this study was

basically to detect similarities and differences between the way Central Catalan and English use accentual strategies to signal focus. Studies such as Vallduví (1990, 1991) and Vallduví and Zacharski (1994) have put English and Catalan as examples of languages that convey focus by opposite means. Whereas English achieves the focus/accent association through accentual means, Central Catalan is supposed to alter the word order of the sentence to locate the focussed element in an accent-bearing position. This study, however, has shown that a word order shift is not always necessary in Central Catalan to signal focus and that intonational means alone can also be used as a focussing device. According to these findings, the obvious question then was whether the intonational strategies used in Central Catalan were the same ones as those of a language such as English, where accentual focus is predominantly used.

Even though the issue of focussing in English has been widely studied, a few sentences were recorded and analysed following the same parameters as for the Central Catalan productions. This allowed us both to compare the two corpora and to validate the analysis and the findings of Central Catalan. Furthermore, the gathering of some English utterances not only enabled us to carry out a cross-linguistic study of the focussing strategies but also provided new details on the intonational properties of English. In this section, both the findings on intonational phonology and on accentual focus will be discussed and compared to the Central Catalan results.

As far as the phonological analysis of the data is concerned, the English F0 contours of broad and narrow focus productions were analysed within the tenets of the AM approach. According to Pierrehumbert and Beckman (1988), Beckman and Hirschberg (1994) and Ladd (1996) among others, the F0 contours were described as a series of pitch accents and edge tones. Two levels of prosodic phrasing were proposed for the analysis of the English data: an IP (signalled by the presence of a boundary tone) and an ip (marked with a phrase accent)<sup>4</sup>. As opposed to Central Catalan, no smaller level of phrasing, such as the word domain, appeared to be relevant for the prosodic structure of English.

---

<sup>4</sup> Even though some studies (Grabe 1998a) consider the ip level unnecessary, the possibility of eliminating this domain was not examined in this study and hence the ip was kept, as suggested by several researchers within the AM model.

The intonation of English neutral declaratives was analysed as a series of pitch accents of the type H\*+L. The right-hand edge of the contour was marked by means of an L- phrase accent and an L% boundary tone. The choice H\*+L rather than H\* or L+H\* to describe the F0 movements observed in the English data was essentially based on two criteria: 1) the F0 rise occurring within the limits of the accented syllable, and 2) the stability in the alignment and in the scaling of the L after the peak.

As far as the F0 rise is concerned, the peak tended to be anchored within the accented syllable and, as observed in Pierrehumbert and Steele (1989) and Silverman and Pierrehumbert (1990), its exact alignment was affected by the duration of the accented syllable, as well as by the presence of an upcoming word or phrase boundary. The peak was more displaced as the accented syllable was longer and was retracted when followed by a word or phrase boundary. Also, the analysis of the peak with respect to the foot duration was taken into consideration and, as reported in House et al (1999a/b) and Dankovičová and House (2000), analysing the peak as a function of the duration of the foot reduced the amount of variability in its timing. These findings not only confirmed previous analyses of English intonation but also corroborated the interpretation of the Central Catalan rising prenuclear accents as L\*H rather than H\* (with peak delay). In Central Catalan, the alignment of the peak was not affected by syllable duration and could not be interpreted as a function of the foot, since the peak was anchored at the end of the word.

The analysis of the timing and the scaling of the F0 valley between peaks provided further evidence for interpreting the rise-fall movements of the English data as H\*+L. The L was constant in F0 irrespective of the number of intervening syllables between accents and was fixed in time with respect to the preceding peak. The H\*+L notation is problematic with the initial taxonomy of the model (Pierrehumbert 1980, Beckman and Pierrehumbert 1986), where the L had no phonetic target value but only functioned as a downstep trigger of a succeeding H. However, with the treatment of downstep as a property of the lowered tone itself (Ladd 1980, 1996; Beckman and Hirschberg 1994; Grice 1995a; Grabe 1998a), the H\*+L accent was relieved of its downstepping duty and could be used as a phonetically transparent notation. Furthermore, the analysis of the pitch movements observed in the data as H\*+L

entities agrees with those analyses that describe English F0 contours by means of left-headed accents (Gussenhoven 1984 or Grabe 1998a). In this approach, only two pitch accents (H\*+L and L\*+H) are taken as basic tones and all other patterns are derived from them by means of a set of phonological adjustment rules. In our data, the L\*+H entity was not observed. However, this is due to the limited number of sentences (only produced by one speaker) and to their nature (only declaratives). Even though L\*+H was not found in the productions, the behaviour of H\*+L does seem to support a left-headed analysis of intonation.

A left-headed approach also seems to apply to Central Catalan, although the phonological primitives differ slightly from those of English. In Central Catalan, the F0 trajectories are also described by means of two tones. However, whereas in English the two entities seem to be part of a bitonal pitch accent, in Catalan the first category is a pitch accent and the second one a word edge tone, anchored at the end of the accented word. Thus, English intonation is described by means of a bitonal notation (H\*+L and L\*+H), and Catalan by means of two separate tones (H\*L and L\*H). In the Central Catalan data, the sequence H+!H\*L was also observed involving a bitonal pitch accent.

Finally, the results on F0 downtrends in neutral declaratives were quite similar in the two languages. Peak downward tendencies were analysed as a series of downstepping accents. The final accent of neutral declaratives underwent final lowering in the two languages. Whereas in English, different degrees of final lowering could be observed in the productions (from a low peak to the absence of a peak), in Central Catalan final lowering always took the form of a progressively falling contour. These cross-linguistic differences can be accounted for as different ways of implementing a similar phonological structure (H\*(+L).

With respect to the use of accentual strategies to signal focus, English and Central Catalan showed both similarities and differences. They agreed in using the same accent type as the nuclear accent of broad and narrow focus sentences. This differs from the results on languages such as European Portuguese (Frota 1998), Bengali (Hayes and Lahiri 1991), Palermo Italian (Grice 1995a), and Neapolitan Italian (D'Imperio 1997), where two different accents are used for the two focal readings. In



English and in Central Catalan, possible differences in the pitch range of broad and narrow focus nuclear accents do not trigger a phonological contrast but seem to be phonetic realisations of the same tonal entity. This means that the blockage of downstep in narrow focus sentences is optional in the two languages, although differences might be observed with respect to the kinds of focal readings: contrast or identification. As reported in the literature (Brown et al 1980, Couper-Kuhlen 1986, García-Lecumberri 1995), contrastive focus is more salient than identification focus and hence downstep is more likely to be blocked in contrastive contexts than in identification contexts. This is confirmed in the data of the English speaker as well as in the data of two Central Catalan informants.

Even though English and Catalan agree in using the same accent type for the nuclear accent of neutral and marked utterances, they differ in other focussing strategies, namely, in the levels of phrasing and in the nature of deaccenting. Whereas in Central Catalan a prosodic boundary is introduced after the focal element, no boundary is observed in English. This means that in English the deaccentuation of postfocal material is compulsory since the only way to signal narrow focus is by the lack of pitch accents. Furthermore, if an accent were located on a postfocal element in English, this would immediately involve two focal points or double focus (Ladd 1996). In Central Catalan, on the other hand, deaccenting is optional and the presence of postfocal accents does not necessarily mean two focal elements.

These results show that even though both languages can signal focus by intonational means, the kinds of accentual strategies do not coincide in the two languages. English basically uses deaccenting of postfocal material and Central Catalan phrasing after the focal element. This explains why in Central Catalan deaccenting is optional. Also the introduction of a phrase boundary after the focal constituent in this language agrees with the idea that in Romance languages focus has to be assigned in an accent-bearing location, such as the final position of a phrase.

## **6.6 Conclusion**

In this chapter, the phonetic and phonological properties of English broad and narrow focus structures have been analysed and compared to Central Catalan. A few English

sentences have been designed and gathered according to the same parameters as for the Central Catalan data. The cross-linguistic analysis was basically focussed on two aspects: 1) intonational phonology and 2) the ways of expressing accentual focus in the two languages. According to the AM model, the F0 contours of English and Catalan have been described as a series of pitch accents and edge tones. Cross-linguistic differences have been observed in the levels of phrasing. In Central Catalan a word domain has shown to be relevant for prosodic structure but not in English. As far as focussing devices are concerned, both languages use the same pitch accent for a broad and a narrow focus reading. Differences are detected in the levels of phrasing and in the status of deaccenting. In Central Catalan, an ip boundary is introduced after the focal element. This means that postfocal material is part of another ip. Deaccenting of postfocal elements then is optional and it can adopt a variety of forms (from a reduced pitch accent to its disappearance). English focus is signalled by means of pitch accent deletion on non-focal material. In English, no phrase boundary is introduced after the focal element and hence deaccenting is compulsory since postfocal elements are part of the same ip as the focal item.

## Chapter 7: Summary and conclusions

### 7.1 Introduction

In this study, the expression of focus by intonational means in Central Catalan declarative sentences has been investigated in three areas: 1) the use of accentual focus in Central Catalan, 2) its phonetic and phonological realisation, and 3) cross-linguistic differences in the expression of intonational focus.

Studies on the signalling of focus cross-linguistically agree that there is a close relationship between focus and accent in that every time an element is focussed, it is also accented. Languages, however, differ in the ways this focus/accent association is attained. Germanic languages, for example, tend to modify the accentual structure of the sentence while the syntactic structure remains unaffected. In these languages, the nuclear accent shifts to the focal item and non-focal elements are deaccented. Romance languages, on the other hand, seem to prefer a syntactic shift to achieve the focus/accent relation. The prosodic structure of the sentence remains unaltered and it is the focussed constituent that moves into a prominent position or accent-bearing location, usually at the end of the sentence. Most studies on the expression of focus in Catalan (Vallduví 1990, 1991, 1994a/b and Vallduví and Zacharski 1994) have claimed that in this language the focus/accent association is always attained through a syntactic shift and hence Catalan has been classified a [-plastic] or a [+syntactic] language. In this study, however, the strict classification of Catalan as a [-plastic] language has been questioned by examining some Catalan data. The results have shown that the use of accentual strategies to convey focus in Catalan is possible in a variety of contexts. This challenges the two-way categorisation of languages as [+/- plastic].

The first aim of this study was to analyse the *extent* of usage of accentual focus in Central Catalan and the *acceptability* of accentual strategies as opposed to syntactic strategies to convey focus. These two aspects were examined through different tests (production, perception and acceptability tests), in subject and verb focal domains and for different kinds of narrow focus (identification and contrast). The results showed

that Central Catalan speakers can produce and recognise focus conveyed by intonational means and that accentual strategies were accepted as natural in this language. These findings suggested that the choice of strategies to convey focus is rather flexible in Central Catalan and hence the categorical classification of languages as [+/-plastic] or [+/- syntactic] might need some revision.

The second aim of this study was to describe the phonetic and phonological properties of accentual focus in Central Catalan. In order to do that, both utterances produced with broad focus (or neutral reading) and utterances produced with narrow focus were examined. The analysis of the data was performed within the Autosegmental Metrical approach. The results presented evidence for three levels of prosodic structure in Central Catalan: the Intonation Phrase (IP), the intermediate phrase (ip), and the accented word domain. Thus, F<sub>0</sub> contours were described as a sequence of pitch accents, associated to the accented syllable of some lexically stressed words, followed by a word edge tone, anchored at the end of each accented word. A boundary tone marked the end of the whole IP and phrase accents signalled the right edge of each ip. Neutral declaratives were then analysed as a series of rises in prenuclear position (L\*H) and a final fall in the nuclear position (H\*L). Unlike other Romance languages, differences between broad and narrow focus sentences did not involve the selection of a different accent type. The nuclear accent of both readings was the same (H\*) and pitch range differences were basically phonetic and not incorporated into the phonological system of the language. Primarily, accentual focus in Central Catalan was expressed by means of accentuation and phrasing. Narrow focus utterances involved the location of the nuclear accent on the focal element and the creation of an ip boundary after the focal constituent. Deaccenting of pre and postfocal material was optional. Sometimes, compressed postnuclear accents appeared after the focal element. These were analysed as part of another ip. Finally, F<sub>0</sub> downtrends were interpreted as a series of linguistically controlled downstepping accents. The last accent of a neutral declarative sentence also underwent final lowering. The blockage of downstep in focal constituents was optional.

Finally, the third aim of this study was to compare the phonetic and phonological properties of Central Catalan broad and narrow focus structures to those of a language that predominantly uses accentual strategies to express focus, such as English. The

purpose of the comparison was to ascertain whether accentual strategies are the same in the two kinds of languages. The results suggested that the accentual strategies in the two languages do not operate in exactly the same way. The two languages are similar in that the same pitch accent is used as the nuclear tone of broad and narrow focus. However, differences in the levels of phrasing and in the nature of deaccenting were observed. Whereas Central Catalan introduces an ip boundary after the focal element, English does not. Cross-linguistic differences in phrasing might explain why in English deaccentuation of postfocal material is always present, whereas in Central Catalan is optional. These results confirm the claim that in Central Catalan the nuclear accent has to occupy an accent-bearing position, such as the last one of the phrase. However, evidence from the data indicated that there is no need of a word order shift to attain this position.

In the following section, a more detailed summary of each of the preceding chapters is provided. Then, section 7.3 discusses some of the theoretical implications of the findings and in section 7.4 some areas of future research are proposed.

## **7.2 Summary**

Chapter 1 presented a literature review on three main areas of research. First, a brief summary on the ways languages use to signal focus was provided. Particular attention was devoted to the expression of focus by phonological means. Second, the main tenets of the AM model were introduced and briefly compared to previous frameworks of intonational analysis. The AM model was chosen to describe the data since it allows for both a phonological and a phonetic description of the F0 contours by distinguishing between tonal association and alignment and by treating scaling as a separate phenomenon from tonal specification. Previous descriptions of the intonation of neutral and marked declaratives in Central Catalan were presented, with special reference to the work of Bonet (1984) and Prieto (1995).

In chapter 2, the extent of usage and acceptability of accentual focus in Central Catalan was examined by means of a production test, a perception test and an acceptability test. The first two tests were performed to investigate whether Central Catalan speakers could produce and recognise focus conveyed by accentual means,

that is, by an alteration of the prominence pattern of the sentence but no modification of the word order. The acceptability test was carried out to analyse how natural accentual focus is felt to be in this language and to examine differences in the acceptability of accentual strategies as opposed to syntactic strategies as means to signal narrow focus. Several variables were examined, such as subject and verb focal domains and focus triggered by contrast and by identification. The results of the production test and the perception test showed that accentual focus is easily produced and perceived by Central Catalan speakers in all focal domains and for all kinds of questions. The acceptability test demonstrated that syntactic strategies and intonational strategies are equally acceptable in this language. These findings challenge the claim that some languages can modify their prominence pattern (plastic) while others cannot (non-plastic) and must modify their word order instead.

In chapter 3, the phonetic and phonological properties of Central Catalan declaratives produced with a broad focus intonation were examined within the AM framework of intonational analysis. Three levels of phrasing were identified in the data: a major domain, such as the Intonation Phrase, a minor domain or intermediate phrase, and an accented word domain. Evidence for a word domain was derived from the consistent alignment of an edge tone at the right-hand side of accented lexical words. Words were divided into three types according to their stress pattern (oxytones, paroxytones and proparoxytones). In all cases, the word edge tone was anchored at the end of the word, irrespective of the number of poststressed syllables. Evidence for an intermediate phrase was found in sentences with a long subject constituent, which were produced with a more abrupt F0 rising after the subject than that observed at the word level. Prenuclear accents showed a rising movement whose peak was consistently aligned with the end of the word. This was analysed as an L\* pitch accent associated to the accented syllable and an H tone associated to the end of the word. The nuclear accent involved a fall characterised as H\*L. The end of the whole IP was marked as L-L% and whenever there was an ip boundary after the subject, this was indicated as H-.

Chapter 4 investigated the phonetic and phonological properties of accentual focus in Central Catalan. The results showed that narrow focus involves accentuation of the focal material and the introduction of an ip boundary after the focal constituent. The

F0 trajectory on the nuclear accent involved a falling movement analysed in the same way as the final tone sequence of neutral sentences (H\*L). Thus, unlike other Romance languages, Central Catalan does not select a different accent type to convey narrow focus. The end of the focal domain was marked by an L- phrase accent. These results suggest that the classification of languages into [+/-plastic] is only partially justified. The presence of an ip boundary after the focal element confirmed the idea that the focus/accent association in Central Catalan takes place in phrase-final position. Thus, it is true that the nuclear accent has to occupy an accent-bearing position within the phrase. However, what seems to be debatable is that this position is only attained by a modification of the word order. The data showed that the accent-bearing location at the end of the sentence can be achieved by introducing an intermediate phrase boundary after the focal element with no need of a syntactic reorganisation. Postfocal data tended to be deaccented, although sometimes a highly reduced pitch accent was observed on the last lexical word of the IP. The appearance of a postnuclear accent was taken as evidence for the presence of another ip domain after the focal element. In this case, the postnuclear accent is the metrically strongest element of the postfocal ip and it is reduced since it is *not* the nuclear accent (or metrically strongest element) of the whole IP. Thus, in Central Catalan, deaccenting can be explained as a reduction of an underlying accent. This implies that deaccentuation in Central Catalan is a gradient rather than a categorical phenomenon. Finally, the results showed that sometimes in contrastive focus in final position a different pitch accent could be used, namely, H+!H\*. Even though this accent type was only observed in final position, its usage is common in Central Catalan (as reported in Prieto 1995) and it conveys correction or rectification. As expected, this accent type was never used in identification focus, which is merely informative.

In chapter 5, the scaling of peak and valley F0 values in Central Catalan was analysed for utterances produced with broad focus and with narrow focus. As far as descending H patterns are concerned, the results showed no time effects on the scaling of consecutive peaks, suggesting an accent-by-accent downstepping sequence rather than a global declination. Utterance final peaks were realised as progressively falling slopes in neutral declaratives and underwent a greater amount of lowering than that expected by the downstep rule. This was accounted for as final lowering. The results concerning the relationship between narrow focus and downstep revealed that narrow

focus does not necessarily block downstep on the focal accent. In some Romance languages, the presence of focus involves a much higher F<sub>0</sub> peak on the focal accent. This increase in pitch range has been incorporated into the phonological system of the language, and hence a different accent type is used for neutral and marked sentences. Evidence from Central Catalan shows that the same nuclear tone is used in the two kinds of structures (H\*) and any differences in the pitch range are a matter of choice on the speaker's part rather than a matter of phonological necessity. This explains why the blockage of downstep is optional in this language. Finally, the scaling of L F<sub>0</sub> values corroborated the phonological analysis of valleys. Whenever a valley was analysed as L\*, its scaling tended not to be affected by temporal distance between accents but showed a rather fixed F<sub>0</sub> height. Whenever valleys were mere transitions between some kind of H tones, an increase in temporal distance between peaks affected the F<sub>0</sub> of the valley: the longer the distance, the higher the amount of dipping.

Finally, chapter 6 presented a cross-linguistic analysis on the ways of conveying focus in Southern British English and in Central Catalan and on the phonetic and phonological properties of broad and narrow focus structures in the two languages. As far as focussing devices are concerned, both languages use the same accent type for neutral and marked readings. However, they present differences in the levels of phrasing and in the status of deaccenting. In Central Catalan, an ip boundary is introduced after the focal element, making postfocal material part of another ip. This explains the possible appearance of postfocal accents in this language and accounts for deaccenting as a gradual phenomenon, which varies in form (from reduction to absence of pitch accent). In English, on the other hand, focus is signalled by means of accenting the focal element and deleting postfocal pitch accents. In English, no phrase boundary is introduced after the focal element. This makes deaccenting compulsory and categorical. With respect to intonational phonology, cross-linguistic differences are observed in the levels of phrasing and in the nature of pitch accents. Central Catalan provides evidence for an intonationally defined word domain. In this language, tonal features are associated at the end of accented words. In English, on the other hand, no word level of prosodic structure seems to be identified on intonational grounds. The presence of word edge tones in Central Catalan implies a different analysis of rising and falling F<sub>0</sub> movements in the two languages. In English, rises



and falls are described by means of bitonal pitch accents (L\*+H and H\*+L) and in Central Catalan, by a combination of a pitch accent and a word edge tone (L\*H and H\*L). Despite differences in the nature of the tonal sequence, both languages seem to provide evidence for a left-headed approach of intonation. Finally, the two languages agree in interpreting F0 downtrends as a sequence of downstepping accents. Neither in English nor in Central Catalan, narrow focus necessarily suspends the downstepping sequence. However, in contrastive narrow focus the blockage of downstep is more usual than in identification narrow focus in both languages.

### **7.3 Implications**

The results of this study have some implications on three areas of research: 1) on the analysis of focus and on the idea of a plasticity parameter, 2) on the tenets of the AM theoretical framework, and 3) on the description of Central Catalan intonation.

One of the theoretical implications of this thesis is that it forces to revise the notion of [+/-plasticity]. This notion involves two separate claims: 1) that there are languages that can modify their prominence pattern to convey focus (plastic) and that others cannot (non-plastic) and must modify their word order instead, and 2) that in non-plastic languages there must be an accent-bearing location, such as the end of the sentence, which is the landing site for the focal element. The data presented in this thesis confirm that in Central Catalan focal elements must be placed at the end of a prosodic phrase (ip). However, this does not necessarily need to be achieved by a syntactic shift. A prosodic shift, such as the introduction of an ip boundary, can locate the focussed item on an accent-bearing position and hence no word order alteration is needed. Thus, according to the Central Catalan data analysed in this study, [-plastic] has to be attributed to languages which have a fixed sentence position to achieve the focus/accent association. However, the strategies used to locate the focal item at phrase final position need not to be syntactic. Intonation alone can create the necessary boundaries to make a focal element final.

The phonological analysis proposed in this study for Central Catalan broad and narrow focus structures has implications for phonological theory. The main innovation is that a word level of intonational analysis can be justified on intonational

grounds, such as the presence of word edge tones in accented word domains. Evidence for this level has derived from the systematic anchoring of edge tones at the end of words with a different number of poststressed syllables. To our knowledge, this is the first time a word domain has been proposed for Central Catalan prosodic structure on the basis of intonational evidence. The word domain seems to be related to the notion of phonological or prosodic word in Prosodic Phonology (Selkirk 1984, Nespor and Vogel 1986), also suggested in Oliva (1992) for Catalan on the basis of junctural and rhythmic phenomena. These results suggest that maybe phonological rules and intonational features are not so far apart to define prosodic hierarchy. An integrated-based approach to identify the levels of prosodic structure has already been followed by Hayes and Lahiri (1991) and Frota (1998). More research is needed on this topic. The appearance of word edge tones has also implications for speech synthesis. The word domain should be taken into account in synthesising Central Catalan speech so as to obtain a more natural intonation.

Finally, the analysis proposed in this study has some advantages over previous descriptions of Central Catalan intonation (e.g. Prieto 1995 and Bonet 1984). Concerning the L\*H description of prenuclear accents, our analysis reconciles differences between Bonet's perceptual analysis and Prieto's acoustic observation. The L\* tone agrees with Bonet's perception of a low pitch over the accented syllable. The H word edge tone accounts for Prieto's proposal that the peak of the rise is placed after the accented syllable. Contrary to Prieto, our analysis uses the same tonal sequence for the nuclear accent of single and multi-stressed utterances (H\*L). The fact that in multi-stressed structures the H\* is produced (and perceived) with a rather low F0 (pitch) is explained as the result of final lowering. This has the benefit of using the same tonal category for the last accent of sentences which do not differ in the intonational meaning (and hence are not contrastive) but only differ in the number of stresses.

#### **7.4 Further research**

The results obtained in this study suggest a number of areas for further investigation. One of the major findings derived from this study is the idea that Central Catalan has a word domain of prosodic structure defined on intonational grounds. It would be

interesting to examine how intonationally defined domains (the accented word, the ip and the IP) correspond to prosodic constituents identified on other grounds, such as the application of sandhi processes or rhythmic phenomena. Furthermore, the behaviour of word edge tones should be analysed when the postsressed syllable/s is/are not part of the same content word but belong/s to a right-attached function word (such as a clitic). If the word edge tone is anchored at the end of the clitic, this will confirm the idea that F0 can be used as evidence for phonological word formation. Finally, the claim that Central Catalan has tonal features at the end of accented words has been based on an acoustic analysis of the F0 contours. It would be interesting to confirm the acoustic observations with a perceptual analysis of the data and see whether Central Catalan listeners find that there is relevant intonational information at the end of words.

Second, the production data examined in the present study was restricted to a reading style (or semi-spontaneous). It would be interesting to investigate how the findings observed in this study can be extended to other speaking styles, in particular, to spontaneous speech. This applies both to the results on intonational structure and to those on focus. As far as intonation is concerned, differences are expected to appear in the realisation of H\*L and L\*H. It is possible that in more colloquial styles (and also in faster speaking rates) these tonal sequences undergo some kind of adjustment, such as truncation or compression. As far as the expression of focus is concerned, it is necessary to examine the extent of usage and acceptability of accentual strategies in spontaneous speech.

Finally, this work has examined a standard variety of Catalan (and English) but analyses of other varieties are needed, since Central Catalan and Southern British English are only examples of a wide range of dialects. Furthermore, the data used for the acoustic analysis was only produced by female speakers between 25 and 30 years old. Further analyses covering male informants and speakers of different ages should be carried out.

Overall, this study has contributed to the expression of accentual focus cross-linguistically as well as on the intonational structure of Central Catalan. A variety of topics are left open for further research.

## References

- Abercrombie, D. 1965. *Studies in Phonetics and Linguistics*. London: Oxford University Press.
- Arvaniti, A. 1992. Secondary stress: evidence from Modern Greek. In G. J. Docherty and D. R. Ladd (eds.). *Papers in Laboratory Phonology II*. Cambridge: Cambridge University Press. 398-423.
1994. Acoustic features of Greek rhythmic structure. *Journal of Phonetics* 22: 239-268.
- Arvaniti, A. and M. Baltazani. 1999. Greek ToBI. Paper presented at the 14<sup>th</sup> *International Conference of Phonetic Sciences Satellite Meeting*. Workshop on "Intonation: Models and ToBI labeling".
- Arvaniti, A. and D. R. Ladd. 1995. Tonal alignment and the representation of accentual targets. *Proceedings of the 13<sup>th</sup> International Conference of Phonetic Sciences*. Stockholm. 220-223.
- Arvaniti, A., D. R. Ladd and I. Mennen. 1998. Stability of tonal alignment: the case of Greek prenuclear accents. *Journal of Phonetics* 26: 3-25.
- To appear. What is a starred tone? Evidence from Greek. *Papers in Laboratory Phonology V*. Cambridge: Cambridge University Press.
- Ashby, M. 1978. A study of two English nuclear tones. *Language and Speech* 21: 326-336.
- Badia i Margarit, A. 1994. *Gramàtica de la llengua catalana*. Barcelona: Edicions Proa.
- Baltazani, M. and S. Jun. 1999. Focus and topic intonation in Greek. *Proceedings of the 14<sup>th</sup> International Congress of Phonetic Sciences*. San Francisco. 1305-1308.
- Barnils, P. 1916. De l'entonació en els nostres dialectes. *Butlletí de dialectologia catalana*. 11-14.
- Beckman, M. E. 1986. *Stress and non-stress accent*. Dordrecht: Foris.
- Beckman, M. E. and G. L. Ayers. 1994. Guidelines for ToBI labelling, v. 2.0. Ms. and accompanying speech materials. Ohio State University.
- Beckman, M. E. and J. Hirschberg. 1994. The ToBI annotation conventions. Ms. and accompanying speech materials. Ohio State University.

- Beckman, M. E. and J. Pierrehumbert. 1986. Intonational structure in Japanese and English. *Phonology Yearbook* 3: 255-310.
- Benzmüller, R. and M. Grice. 1999. German ToBI. Paper presented at the 14<sup>th</sup> *International Conference of Phonetic Sciences Satellite Meeting*. Workshop on "Intonation: Models and ToBI labeling".
- Berman, A. and M. Szamosi. 1972. Observations on sentential stress. *Language* 48: 304-325.
- Bonet, E. 1984. *Aproximació a l'entonació del català*. Tesi de llicenciatura. Universitat Autònoma de Barcelona.
1986. L'entonació de les formes interrogatives in barceloní. *Els Marges* 33: 103-117.
- Bolinger, D. 1951. Intonation: levels versus configurations. *Word* 7: 199-210.
1958. A theory of pitch accent in English. *Word* 14: 109-149.
1972. Accent is predictable (if you're a mind-reader). *Language* 48: 633-644.
- Bresnan, J. 1971. Sentence stress and syntactic transformations. *Language* 47: 257-280.
- Brown, G., K. L. Currie and J. Kenworthy. 1980. *Questions of intonation*. London: Croom and Helm.
- Bruce, G. 1977. *Swedish word accents in sentence perspective*. Lund: Gleerup.
- Cabrera-Abreu, M. 1996. *A phonological model of intonation without low tone*. PhD dissertation. University College London.
- Carr, P. 1993. *Phonology*. London: Macmillan.
- Cerdà, R. 1970. L'estructura vocàlica del català comú modern. *Estudis Romànics* 12: 179-213.
1972. *El timbre vocàlico en catalán*. Madrid: Consejo Superior de Investigaciones Científicas.
- Chafe, W. L. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. L. Li (ed.). *Subject and topic*. New York: Associated Press. 25-55.
- Chomsky, N. 1971. Deep structure, surface structure, and semantic interpretation. In D. Steinberg and L. Jakobovits (eds.). *Semantics*. Cambridge: Cambridge University Press. 1-40.
- Chomsky, N. and M. Halle. 1968. *The sound pattern of English*. New York: Harper and Row.

- Clements, G. 1983. The hierarchical representation of tone features. In I. Dihoff (ed.). *Current approaches to African linguistics, Vol 1*. Dordrecht: Foris.
- Coleman, J. 1993. Polysyllabic words in the YorkTalk synthesis system. In P. Keating (ed.). *Papers in Laboratory Phonology III*. Cambridge: Cambridge University Press. 293-324.
- Contreras, H. 1976. *A theory of word order with special reference to Spanish*. Amsterdam: North Holland.
- Cooper, W. F. and J. M. Sorensen. 1981. *Fundamental frequency in sentence production*. Heidelberg: Springer.
- Cooper, W. J., S. R. Eady and P. R. Mueller. 1985. Acoustical aspects of contrastive stress in question-answer contexts. *Journal of the Acoustical Society of America* 77: 2142-2155.
- Couper-Kuhlen, E. 1986. *An introduction to English prosody*. London: Edward Arnold.
- Cruttenden, A. 1986. *Intonation*. Cambridge: Cambridge University Press.
- Crystal, D. 1969. *Prosodic systems and intonation in English*. Cambridge: Cambridge University Press.
- Dahl, Ö. 1974. Topic-comment structure revisited. In Ö. Dahl (ed.). *Topic and comment, contextual boundedness and focus*. Hamburg: Helmut Buske. 1-24.
- Dankovičová, J. and J. House. 2000. Modelling intonation for synthesis: domains of accent alignment. Paper presented in the *British Association of Academic Phoneticians Colloquium*. University of Glasgow.
- D'Imperio, M. 1997a. Breadth of focus, modality and prominence perception in Neopolitan Italian. *OSU Working Papers in Linguistics* 50: 19-39.
- 1997b. Narrow focus and focal accent in the Neopolitan variety of Italian. In A. Botininis et al (eds.). *Intonation: Theory, Models and Applications. Proceedings of an ESCA Workshop*. Athens. 87-90.
- 1999a. Tonal structure and pitch targets in Italian focus constituents. *Proceedings of the 14<sup>th</sup> International Conference of Phonetic Sciences*. San Francisco. 1757-1769.
- 1999b. Neopolitan Italian ToBI. Paper presented at the *14<sup>th</sup> International Conference of Phonetic Sciences Satellite Meeting*. Workshop on "Intonation: Models and ToBI labeling".

- Durand, J. 1990. *Generative and non-linear phonology*. Singapore: Longman.
- Eady, S. R. and W. J. Cooper. 1986. Speech intonation and focus location in matched statements and questions. *Journal of the Acoustical Society of America* 80 (2): 402-415.
- Erickson, D. and I. Lehiste. 1995. Contrastive emphasis in elicited dialogue: durational compensation. *Proceedings of the 13<sup>th</sup> International Conference of Phonetic Sciences*. Stockholm: 352-355.
- Estebas-Vilaplana, E. 1996. *The intonation of enumerations in Catalan*. MA thesis. University College London.
1997. The intonation of enumerations: a cross-linguistic study. *Actas del XXI Congreso de AEDEAN*. Universidad de Sevilla. 645-651.
- Estebas-Vilaplana, E. and J. Maidment. 1999a. The intermediate phrase in Central Catalan declaratives: a case for questioning the representation of downstep. *Speech, Hearing and Language: work in progress*. University College London. Vol. 11: 90-116.
- 1999b. Central Catalan declaratives: the relation between focus and downstep. *Proceedings of the 14<sup>th</sup> International Conference of Phonetic Sciences*. San Francisco. 1861-1864.
- Face, T. L. To appear. A Phonological analysis of rising pitch in Castilian Spanish. *Hispanic Linguistics* 11.
- Féry, C. 1992. Focus, topic and intonation in German. *Arbeitspapiere des Sonderforschungsbereichs* 340. Bericht 20.
1993. *German intonational patterns*. Tübingen: Niemeyer.
- Firbas, J. 1964. On defining the theme in functional sentence perspective. *Travaux linguistiques de Prague* 1: 267-280.
- Frota, S. 1993. On the prosody of focus in European Portuguese. *Proceedings of the Workshop on Phonology*. Lisboa: APL. 45-66.
1995. On the prosody of intonation of focus in European Portuguese. Ms., University of Lisbon.
1997. On the prosody and intonation of focus in European Portuguese. In F. Martínez-Gil and A. Morales-Front (eds.). *Issues in the Phonology and Morphology of the Major Iberian Languages*. Washington, D.C.: Georgetown University Press. 359-392.

1998. *Prosody and focus in European Portuguese*. PhD dissertation. University of Lisbon.
- Fry, D. B. 1958. Experiments in the perception of stress. *Language and Speech* 27: 765-768.
- Fudge, E. 1984. *English word-stress*. London: George Allen and Unwin.
- Fujisaki, H. 1983. Dynamic characteristics of voice fundamental frequency of speech and singing. In P. F. MacNeilage (ed.). *The production of speech*. New York and Berlin: Springer-Verlag. 39-55.
1988. A note on the physiological and physical basis for the phrase and accent components in the voice fundamental frequency contour. In O. Fujimura (ed.). *Vocal physiology: voice production, mechanisms and functions*. New York: Raven. 347-355.
- García-Lecumberri, M. L. 1995. *Intonational signalling of information structure in English and in Spanish: a comparative study*. PhD dissertation. University College London.
- García-Lecumberri, M. L., M. Cabrera-Abreu and J. A. Maidment. 1997. The realisation of focal accent in English and Spanish: similarities and differences. *Intonation Theory, Models and Applications. Proceedings of ESCA Workshop*. Athens. 141-144.
- Gårding, E. 1983. A generative model of intonation. In A. Cutler and D. R. Ladd (eds.). *Prosody: Models and measurements*. Berlin: Springer. 11-26.
- Gay, T. 1978. Physiological and acoustic correlates of perceived stress. *Language and Speech* 21: 347-353.
- Gimson, A. C. 1980. *An introduction to the pronunciation of English*. London: Edward Arnold.
- Goldsmith, J. 1976. *Autosegmental Phonology*. New York: Garland Press.
- Grabe, E. 1998a. *Comparative Intonational Phonology: English and German*. Wageningen: Ponsen and Looijen bv.
- 1998b. Pitch accent realisation in English and German. *Journal of Phonetics* 26: 129-143.
- To appear. Fundamental frequency downtrends in English and German. *Language and Speech*.
- Grabe, E., F. Nolan, and K. Farrar. To appear. Pitch accent realisation in four varieties of British English. *Journal of Phonetics*.



- Grice, M. 1995a. *The intonation of interrogation in Palermo Italian: implications for intonation theory*. Tübingen: Niemeyer.
- 1995b. Leading tones and downstep in English. *Phonology* 12: 183-234.
- Grice, M. and R. Benz Müller. 1995. Transcription of German intonation using ToBI-Tones: the Saarbrücken System. *PHONUS* (Research Report, Institute of Phonetics, University of the Saarland) 1: 33-51.
- Grice, M. and M. Savino. 1997. Can pitch accent type convey information status in yes-no questions? *Proceedings ACL Workshop on Concept to Speech Generation Systems*. Madrid. 29-38.
- Grønnum, N. *See also* Thorsen.
1991. Prosodic parameters in a variety of regional Danish standard languages, with a view towards Swedish and German. *Phonetica* 46: 188-214.
- Gundel, J. K. 1988. Universals of topic-comment structure. In M. Hammond, E. Moravcsik and J. R. Wirth (ed.). *Studies in syntactic typology*. Amsterdam/Philadelphia: John Benjamins. 209-239.
- Gussenhoven, C. 1983a. Focus, mode and nucleus. *Journal of Linguistics* 19: 377-417.
- 1983b. Testing the reality of focus domains. *Language and Speech* 26: 61-80.
1984. *On the grammar and semantics of sentence accents*. Dordrecht: Foris.
1994. Focus and sentence accents in English. In P. Bosch and R. van Sandt (eds.). *Focus and normal language processing. Working Papers of the Institute for Logic and Linguistics*. Heidelberg. 83-92.
2000. Structural effects of pitch accent placement. Paper presented in the *Workshop of Grammar and Intonation*. LAGB. University College London.
- Gutiérrez, F. 1983. Aspectos lingüísticos de la segmentación del tono en inglés, castellano y catalán. *Actas del primer Congreso de Lingüística Aplicada*. Universidad de Murcia. 179-192.
1984. Comparación de algunos aspectos de la entonación del catalán y del inglés. *Folia Phonetica* 1: 149-158.
- Hajičová, E. 1984. Topic and focus. In P. Sgall (ed.). *Contributions to functional syntax. Semantics and language comprehension*. Amsterdam/Philadelphia: John Benjamins.

- Halliday, M. A. K. 1967. *Intonation and grammar in British English*. The Hague: Mouton.
1970. *A course in spoken English: intonation*. Oxford: Oxford University Press.
- Hayes, B. 1982. Extrametricality and English stress. *Linguistic Inquiry* 13: 227-276.
1989. The prosodic hierarchy in meter. In P. Kiparsky and G. Youmans (eds.). *Phonetics and Phonology: rhythm and meter*. New York: American Press. 201-260.
- Hayes, B. and A. Lahiri. 1991. Bengali Intonational Phonology. *Natural Language and Linguistic Theory* 9: 47-96.
- Hirschberg, J. 1999. Communication and prosody: functional aspects of prosody. *Proceedings of ESCA International Workshop on Dialogue and Prosody*. Veldhoven. The Netherlands. 7-15.
- House, J. 1990. Intonation structure and pragmatic interpretation. In S. Ramsaran (ed.). *Studies in the pronunciation of English: a commemorative volume in honour of A. C. Gimson*. London: Routledge. 38-57.
- House, J. and A. Wichmann. 1996. Investigating peak timing in naturally-occurring speech: from segmental constraints to discourse structure. *Speech, Hearing and Language: work in progress* 9. University College London. 99-117.
- House, J., J. Dankovičová and M. Huckvale. 1999a. Intonation modelling in ProSynth. *Speech, Hearing and Language: work in progress* 11. University College London. 51-61.
- 1999b. Intonation modelling in ProSynth: an integrated prosodic approach to speech synthesis. *Proceedings of the 14<sup>th</sup> International Conference of Phonetic Sciences*. San Francisco. 2343-2346.
- Hualde, J. I. 2000. Intonation in Spanish and the other Ibero-Romance languages: overview and status quaestionis. Paper presented at the *Linguistic Symposium of Romance Languages* 30. Gainesville. Florida.
- Jackendoff, R. 1972. *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.
- Jun, S. 1991. An experimental phonetic study of focus of current Korean. Eonohak Yonku [Research in Linguistics], Seoul National University. Korea.

1996. *The phonetics and phonology of Korean prosody*. New York and London: Garland Publishing.
- Jun, S. and C. Fougeron. To appear. A phonological model of French intonation. In A. Botinis (ed.). *Intonation research and applications*. Cambridge: Cambridge University Press.
- Jun, S. and H. Lee. 1998. Phonetic and phonological markers of contrastive focus in Korean. *Proceedings of the 5<sup>th</sup> International Conference in Spoken Language Processing 4*. Sydney. 1295-1298.
- Kanerva, J. 1990. Focussing and phonological phrases in Chichewa. In S. Inkelas and D. Zec (eds.). *The phonology-syntax connection*. Chicago: University of Chicago Press. 145-161.
- King, P. H. 1995. *Configuring topic and focus in Russian*. Stanford: CSLI Publications.
- Kingdon, R. 1958. *The groundwork of English intonation*. London: Longmans, Green and Co.
- Kiss, K. 1995. Introduction. In K. Kiss (ed.). *Discourse configurational languages*. New York/Oxford: Oxford University Press. 3-27.
- Ladd, D. R. 1980. *The structure of intonational meaning. Evidence from English*. Bloomington: Indiana University Press.
- 1983a. Phonological features of intonational peaks. *Language* 59: 721-759.
- 1983b. Peak features and overall slope. In A. Cutler and D. R. Ladd (eds.). *Prosody: Models and measurements*. Berlin: Springer. 39-52.
1984. Declination: a review and some hypotheses. *Phonology Yearbook* 1: 53-74.
1986. Intonational phrasing: the case for recursive prosodic structure. *Phonology Yearbook* 3: 311-340.
1990. Metrical representation of pitch register. In J. Kingston and M. E. Beckman (eds.). *Papers in Laboratory Phonology I*. Cambridge: CUP. 35-57.
1992. An introduction to intonational phonology. In G. J. Docherty and D. R. Ladd (eds.). *Papers in Laboratory Phonology II*. Cambridge: Cambridge University Press. 321-334.

1993. In defense of a metrical theory of intonational downstep. In H. van der Hulst and K. Snider (eds.). *The phonology of tone: the representation of tonal register*. Berlin: Mouton de Gruyter.
1996. *Intonational Phonology*. Cambridge: Cambridge University Press.
- Ladd, D. R. and J. M. B. Terken. 1995. Modelling intra- and inter-speaker pitch variation. *Proceedings of the 13<sup>th</sup> International Conference of Phonetic Sciences*. Stockholm. 386-389.
- Lakoff, G. 1972. The global nature of the Nuclear Stress Rule. *Language* 48: 285-303.
- Lambrecht, K. 1994. Information structure and sentence form. Cambridge: Cambridge University Press.
- Leben, W. 1976. The tones in English intonation. *Linguistic Analysis* 2: 69-107.
- Leusen, N. van and L. Kálmán. 1993. The interpretation of free focus. Ms. University of Amsterdam.
- Liberman, M. 1975. The intonational system of English. PhD thesis. MIT.
- Liberman, M. and J. Pierrehumbert. 1984. Intonational invariance under changes in pitch range and length. In M. Aronoff and R. Oerhle (eds.). *Language sound structure*. Cambridge, MA: MIT Press. 157-233.
- Liberman, M. and A. Prince. 1977. On stress and linguistic rhythm. *Linguistic Inquiry* 8: 249-336.
- Lieberman, P. 1967. *Intonation, perception, and language*. Cambridge, MA: MIT Press.
- Local, J. 1995. Syllabification and rhythm in non-segmental phonology. In J. Windsor Lewis (ed.). *Studies in general and English Phonetics. Essays in honour of Professor J. D. O'Connor*. London: Routledge. 350-366.
- Maidment, J. A. 1990. Focus and tone in English intonation. In S. Ramsaran (ed.). *Studies in the pronunciation of English: a commemorative volume in honour of A. C. Gimson*. London: Routledge. 19-26.
- Mascaró, J. 1978. *Catalan phonology and phonological cycle*. Bloomington, Indiana: Indiana University Linguistics Club.
- Mascaró, I. 1986. Introducció a l'entonació dialectal catalana. *Randa* 21: 5-38.
1987. Ciutadella-Maó. Greu vs agut en dos parlants menorquins. Platejament de la qüestió. *Randa* 22: 197-211.
- Meyers, L. 1976. Aspects of Hausa tone. *Working Papers in Phonetics* 32. University of California, Los Angeles.

- Möbius, B. 1993. *Ein quantitatives Modell der deutschen Intonation. Analyse und Synthese von Grundfrequenzverläufe*. Niemeyer: Tübingen.
- Nespor, M. and I. Vogel. 1986. *Prosodic phonology*. Dordrecht: Foris.
- Nibert, H. 2000. Phonetic and phonological evidence for intermediate phrasing in Spanish intonation. PhD dissertation. University of Illinois at Urbana-Champaign.
- Nolan, F. 1995. The effect of emphasis on declination in English intonation. In J. Windsor Lewis (ed.). *Studies in general and English Phonetics. Essays in honour of Professor J. D. O'Connor*. London: Routledge. 241-254.
1999. The devil is in the detail. *Proceedings of the 14<sup>th</sup> International Conference of Phonetic Sciences*. San Francisco. 1-8.
- Nolan, F. and K. Farrar. 1999. Timing of F0 peaks and peak lag. *Proceedings of the 14<sup>th</sup> International Conference of Phonetic Sciences*. San Francisco. 961-964.
- O'Connor, J. D. and G. F. Arnold. 1973. *Intonation of colloquial English*. London: Longman.
- Oh, M. 1999. Korean prosodic structure and focus. *Proceedings of the 14<sup>th</sup> International Conference of Phonetic Sciences*. San Francisco. 1517-1520.
- Oliva, S. 1992. *La mètrica i el ritme de la prosa*. Barcelona: Quaderns Crema.
- Palmer, H. 1922. *English intonation, with syntactic exercises*. Cambridge: Heffer.
- Pierrehumbert, J. 1980. *The Phonology and Phonetics of English Intonation*. PhD dissertation. Bloomington: Indiana University Press.
- Pierrehumbert, J. and M. E. Beckman. 1988. *Japanese Tone Structure*. Cambridge, MA: MIT Press.
- Pierrehumbert, J. and S. Steele. 1989. Categories of tonal alignment in English. *Phonetica* 46: 181-196.
- Pike, K. L. 1945. *The intonation of American English*. Ann Arbor: University of Michigan Press.
- Pitrelli, J. F., M. Beckman and J. Hirschberg. 1994. Evaluation of prosodic transcription labeling reliability in the ToBI framework. *Proceedings of the International Conference on Spoken Language Processing 1*: 123-126. Yokohama, Japan.
- Poser, W. J. 1984. The phonetics and phonology of tone and intonation in Japanese. PhD dissertation. MIT.

- Post, B. 2000. *Tonal and phrasal structures in French intonation*. The Hague: Holland Academic Graphics.
- Prieto, P. 1995. Aproximació als contorns tonals del català central. *Caplletra* 19: 161-186.
1997. Prosodic manifestation of syntactic structure in Catalan. In F. Martínez-Gil and A. Morales-Front (eds.). *Issues in the Phonology and Morphology of the Major Iberian Languages*. Washington, D.C.: Georgetown University Press. 179-199.
1998. The scaling of the L values in Spanish downstepping contours. *Journal of Phonetics* 26: 261-282.
- To appear a. Entonació dialectal del català: el cas de les oracions interrogatives absolutes. *Proceedings of the 9<sup>th</sup> Colloquium of the North American Catalan Society*. Barcelona: Publicacions de l'Abadia de Montserrat.
- To appear b. Entonació. In J. Solà (ed.). *Gramàtica del català contemporani*. Barcelona: Edicions 63.
- To appear c. *Manual d'entonació*. Barcelona: Editorial Ariel.
- To appear d. Reorganització gestual i estructura tonal en català. *Working Papers in Linguistics*. Barcelona: Departament de Filologia Catalana. Universitat Autònoma de Barcelona.
- To appear e. Review of J. M. Sosa's book: *La entonación del español*. *Hispanic Linguistics*.
- Prieto, P., J. van Santen and J. Hirschberg. 1994. Patterns of F0 peak placement in Mexican Spanish. *Proceedings of the ESCA/IEEE Workshop on Speech Synthesis*. New Paltz, New York. 33-37.
1995. Tonal alignment patterns in Spanish. *Journal of Phonetics* 23: 429-451.
- Prieto, P., C. Shih, and H. Nibert. 1996. Pitch downtrend in Spanish. *Journal of Phonetics* 24: 445-473.
- Recasens, D. 1977. Aproximació a les cadències tonals del català. *Anuario de Filología* 3: 509-516.
1993. *Fonètica i fonologia*. Barcelona: Enciclopèdia Catalana.
- Rochemont, M. and P. Culicover. 1990. *English focus constructions and the theory of grammar*. Cambridge: Cambridge University Press.

- Rump, H. H. and R. Collier. 1996. Focus conditions and the prominence of pitch-accented syllables. *Language and Speech* 39 (1): 1-17.
- Salcioli, V. 1988a. Estudio fonético-experimental de la entonación interrogativa catalana. *Estudios de Fonética Experimental* 3: 37-70. Universitat de Barcelona.
- 1988b. *La entonación: estudio fonético-experimental de la entonación interrogativa catalana*. PhD dissertation. Universitat de Barcelona.
- Savino, M. 1999. Bari Italian ToBI. Paper presented at the 14<sup>th</sup> International Conference of Phonetic Sciences Satellite Meeting. Workshop on "Intonation: Models and ToBI labeling".
- Schmerling, S. F. 1976. *Aspects of English sentence stress*. Austin: University of Texas Press.
- Selkirk, E. 1980. The role of prosodic categories in English word stress. *Linguistic Inquiry* 2: 563-605.
1984. *Phonology and Syntax: the relation between sound and structure*. Cambridge MA: MIT Press.
- Silverman, K. and J. Pierrehumbert. 1990. The timing of prenuclear high accents in English. In J. Kingston and M. E. Beckman (eds.). *Papers in Laboratory Phonology I*. Cambridge: Cambridge University Press. 72-106.
- Silverman, K., M. E. Beckman, J. Pitrelli, M. Ostendorf, C. Wightman, P. Price, J. Pierrehumbert and J. Hirschberg. 1992. ToBI: a standard for labeling English prosody. *Proceedings of the Second International Conference on Spoken Language Processing* 2: 867-870. Banff, Canada.
- Solé, M. J. 1984. Experimentos sobre la percepción del acento. *Estudios de fonética experimental* 1. Universitat de Barcelona. 135-242.
- Sosa, J. M. 1991. Fonética y fonología de la entonación del español hispanoamericano. PhD dissertation. University of Massachusetts, Amherst.
1999. *La entonación del español*. Barcelona: Cátedra.
- Steele, S. A. and M. J. Altom. 1986. Nuclear stress: changes in vowel duration and F0 peak location. *Technical Memorandum*. AT and T Bell Laboratories.
- Tench, P. 1996. *The intonation systems of English*. London: Cassell.
- Thorsen, N. *See also* Grønnum.

1980. Intonation contours and stress group patterns in declarative sentences of varying length in ASC Danish. *Annual report of the Institute of Phonetics* 15: 13-47. University of Copenhagen.
1983. Two issues in the prosody of Standard Danish. In A. Cutler and D. R. Ladd (eds.). *Prosody: Models and measurements*. Berlin: Springer. 27-38.
- Toledo, G. 1989. Señales prosódicas del foco. *Revista Argentina de Lingüística* 5 (1-2): 205-230.
- Trager, G. L. and H. L. Smith. 1951. *An outline of English structure*. Norman, OK: Battenburg Press.
- Vaissière, J. 1983. Language-independent prosodic features. In A. Cutler and D. R. Ladd (eds.). *Prosody: Models and measurements*. Berlin: Springer. 53-66.
- Välismaa-Blum, R. 1993. Intonation: a distinctive parameter in grammatical constructions. *Phonetica* 50: 124-137.
- Vallduví, E. 1990. L'estructura informacional de l'oració i la dislocació en català. *Actes del Sisè col.loqui d'estudis catalans a Nord-Amèrica*. Barcelona: Publicacions de l'Abadia de Montserrat. 155-172.
1991. The role of plasticity in the association of focus and prominence. *Proceedings of the Eastern States Conference in Linguistics (ESCOL) 7*: 295-306.
- 1994a. Detachment in Catalan and information packaging. *Journal of Pragmatics* 22: 573-601.
- 1994b. Updates, files and focus-ground. In P. Bosch and R. van der Sandt (eds.) *Focus and natural language processing*, vol. 3: 649-658. (IBM Working Papers of the Institute for Logic and Linguistics 8. Heidelberg: IBM).
1995. Structural properties of information packaging in Catalan. In K. Kiss (ed.). *Discourse configurational languages*. New York, Oxford: Oxford University Press. 122-152.
- Vallduví, E. and R. Zacharski. 1994. Accenting phenomena, association with focus and the recursiveness of focus-ground. In P. Dekker and M. Stokhof (eds.) *Proceedings of the Ninth Amsterdam Colloquium*. 683-702. ILLC, Amsterdam.
- van den Berg, R., C. Gussenhoven and T. Rietveld. 1992. Downstep in Dutch: implications for a model. In G. J. Docherty and D. R. Ladd (eds.). *Papers*



- in *Laboratory Phonology II*. Cambridge: Cambridge University Press. 335-358.
- Vella, A. 1995. Prosodic structure and intonation in Maltese: influence on Maltese English. PhD dissertation. Edinburgh University.
- Venditti, J. J. 1995. Japanese ToBI labeling guidelines. [[http://ling.ohio-state.edu/Phonetics/J\\_ToBI/jToBI\\_homepage.html](http://ling.ohio-state.edu/Phonetics/J_ToBI/jToBI_homepage.html)].
1999. The J\_ToBI model of Japanese intonation. Paper presented at the 14<sup>th</sup> *International Conference of Phonetic Sciences Satellite Meeting*. Workshop on "Intonation: Models and ToBI labeling".
- Venditti, J. J., S. Jun, and M. E. Beckman. 1996. Prosodic cues to syntactic and other linguistic structures in Japanese, Korean and English. In J. L. Morgan and K. Demuth (eds.). *Signal to syntax: bootstrapping from speech to grammar in early acquisition IV: Speech and the acquisition of phrase structure*. Hillsdale: Lawrence Erlbaum. 287-311.
- Veny, J. 1993. *Els parlars catalans*. Palma de Mallorca: Editorial Moll.
- Virgili-Blanquet, V. 1971. Notas sobre la entonación catalana. *Archivium* 21. Universidad de Oviedo. 359-377.
- Wells, J. C. 1990. Syllabification and allophony. In S. Ramsaran (ed.). *Studies in the pronunciation of English: a commemorative volume in honour of A. C. Gimson*. London: Routledge. 76-86.
2000. *Longman pronunciation dictionary*. Harlow: Pearson Education Limited.
- Wells, R. 1945. The pitch phonemes of English. *Language* 21: 27-40.
- Wells, W. 1986. An experimental approach to the interpretation of focus in spoken English. In C. Johns-Lewis (ed.). *Intonation in discourse*. London: Croom Helm. 53-75.
- Wheeler, M. 1974. *Some rules in a generative phonology of Modern Catalan*. PhD dissertation. University of Oxford.
1979. *Phonology of Catalan*. Oxford: Blackwell.
- Wichmann, A., J. House and T. Rietveld. 1999. Discourse constraints on peak timing in English: experimental evidence. *Proceedings of the 14<sup>th</sup> International Conference of Phonetic Sciences*. San Francisco. 1765-1768.
- Zubizarreta, M. L. 1998. *Prosody, focus and word order*. Cambridge, AM: MIT Press.

## Appendix 1

This appendix contains the sentences designed for the production test (and also used in the perception test). The column on the right presents the questions posed by the researcher in the production test. The column on the left provides the list of expected responses. For the perception test, sentences on the left column are the stimuli presented to the listeners and sentences on the right column are the questions that the listeners were expected to identify.

Sentences are organised in the following way. Sentences are divided into target sentences, distractors and trial sentences. Within the target sentences, structures are grouped according to the focal domain: subject focus and verb focus. Within each focal domain, sentences are divided according to the kind of question that triggered narrow focus: contrast vs identification. Finally, sentences in standard font are those used in the reading task. Sentences in italics are the expected responses of a semi-spontaneous task.

### 1.1 Target sentences

#### 1.1.1 Narrow focus on the subject

##### 1.1.1.1 Narrow focus triggered by contrast

JO menjava melmelada	En Xavier menjava melmelada?
L'EVA guarda monedes romanes	En Lluís guarda monedes romanes?
LA ROSA llegia llibres	En Jordi llegia llibres?
LA NÚRIA vol una germana menuda	La Carme vol una germana menuda?
MALES MODES malmeten els joves	La vida malmet els joves?
VÀRIES JOIES eren lluminoses	El braçalet era lluminós?
ALGUNES MINYONES guarden mala lluna	Els cambrers guarden mala lluna?
LES ROSES VERMELLES eren oloroses	Les flors eren oloroses?
<i>L'OU bullia a l'olla</i>	<i>L'all bullia a l'olla?</i>
<i>L'ILLA brillava amb la llum de l'albada</i>	<i>La platja brillava amb la llum de l'albada?</i>
<i>L'ISIDRE du una gavardina vermella</i>	<i>En Ricard du una gavardina vermella?</i>
<i>LA MARE menja arengades</i>	<i>El pare menja arengades?</i>
<i>L'OLLA NEGRA guardava mongetes</i>	<i>La cassola guardava mongetes?</i>
<i>LES MEVES AMIGUES eren bones nenes</i>	<i>Les noies eren bones nenes?</i>
<i>LA NENA MORENA venia nines noves</i>	<i>La noia venia nines noves?</i>
<i>L'ALA NOVA dóna al jardí</i>	<i>L'edifici dóna al jardí?</i>

### 1.1.1.2 Narrow focus triggered by identification

L'AU venia de l'illa	Quin animal venia de l'illa?
L'HOME venia llimones madures	Qui venia llimones madures?
LA MARI du una sivella daurada	Qui du una sivella daurada?
L'EMÍLIA vol amanida	Qui vol amanida?
NOIES JOVES ballen munyeires	Qui balla munyeires?
VÀRIES NENES lligaven els globus	Qui lligava els globus?
LA MEVA GERMANA balla balades vieneses	Qui balla balades vieneses?
ALGUNES HABANERES eren alegres	Quines cançons eren alegres?

<i>ELL anava a Girona</i>	<i>Qui anava a Girona?</i>
<i>L'ANNA vivia a Vilavella</i>	<i>Qui vivia a Vilavella?</i>
<i>LA NÚVIA rentava la faldilla</i>	<i>Qui rentava la faldilla?</i>
<i>L'AMÈLIA duia robes negres</i>	<i>Qui duia robes negres?</i>
<i>L'HOME MAGRE menjava vedella</i>	<i>Qui menjava vedella?</i>
<i>L'AMO MURRI viu a Badalona</i>	<i>Qui viu a Badalona?</i>
<i>LA MINYONA MENUDA netejava la nevera</i>	<i>Qui netejava la nevera?</i>
<i>LA GALLINA RODONA vivia a la granja</i>	<i>Quin animal vivia a la granja?</i>

### 1.1.2 Narrow focus on the verb

#### 1.1.2.1 Narrow focus triggered by contrast

En Juli DIU bajanades	En Juli escolta bajanades?
La Mila NEGA la maionesa	La Mila lliga la maionesa?
Els homes LLIMAVEN l'armari	Els homes vernissaven l'armari?
En Joan DOMINA l'àlgebra	En Joan aprèn l'àlgebra?
Les noies VÈNEN DE VEURE l'examen	Les noies acaben de fer l'examen?
Les ovelles VOLEN JEURE a l'herba	Les ovelles estan jugant a l'herba?
En Raimon VOLIA ARRUÏNAR la germana	En Raimon podia ajudar la germana?
La Lluïsa HAVIA D'ARRIBAR diumenge	La Lluïsa volia marxar diumenge?

<i>La Glòria VE de Vilanova</i>	<i>La Glòria va a Vilanova?</i>
<i>La Rosa REGA els geranis</i>	<i>La Rosa talla els geranis?</i>
<i>En Joan ANAVA al museu</i>	<i>En Joan venia del museu?</i>
<i>La Mireia REMENA l'olla</i>	<i>La Mireia fa bullir l'olla?</i>
<i>Les noies VOLEN REBRE la reina</i>	<i>Les noies han vist la reina?</i>
<i>Les mares VEUEN RIURE les nenes</i>	<i>Les mares fan plorar les nenes?</i>
<i>En Juli HAURIA DE DEMANAR l'hora</i>	<i>En Juli sap l'hora?</i>
<i>Les nenes HAURIEN DE DONAR l'enhora bona</i>	<i>Les nenes rebien l'enhora bona?</i>

### 1.1.2.2 Narrow focus triggered by identification

En Jaume MOU la galleda	Què fa en Jaume amb la galleda?
En Jordi MIRA les mones	Què fa en Jordi amb les mones?

La Neus ANIMA els alumnes  
 En Jordi GUARDAVA medalles  
 Els joves VOLEN VIURE la vida  
 L'àvia VA A RAURE a la masia  
 La Laura VOLIA DIVIDIR l'herència  
 La mare HAVIA DE MENJAR verdura  
 La Marina BEU llimonada  
 L'Emili MANA els obrers  
 La Remei MIMA les nenes  
 En Ramon OLORA la benzina  
 L'Elena MIRA DE MOURE el mobiliari  
 Els homes VOLEN VEURE aiguardent  
 L'Elena VOLIA MUNYIR l'ovella  
 La Ramona VOLIA MULLAR la roba

Com tracta la Neus els alumnes?  
 Què feia en Jordi amb les medalles?  
 Com es plantegen els joves la vida?  
 Què fa l'àvia a la masia?  
 Què intentava fer la Laura amb l'herència?  
 Què feia la mare amb la verdura?  
 Què fa la Marina amb la llimonada?  
 Com tracta l'Emili els obrers?  
 Com cuidava la Remei les nenes?  
 Què fa en Ramon amb la benzina?  
 Què fa l'Elena amb el mobiliari?  
 Què fan els homes amb l'aiguardent?  
 Què feia l'Elena amb l'ovella?  
 Què feia la Ramona amb la roba?

## 1.2 Distractors

En Pere té vuit anys  
 La carn m'agrada poc feta  
 La Remei vol comprar un quadre  
 La Carme té tres fills i una filla  
 Demà em vull llevar a les set  
 Els diumenges vaig a pescar  
 La Raquel arriba dilluns  
 L'ajuntament està al carrer major  
 La meva amiga és de Portugal  
 Per sopar vull escudella  
 La Caballé cantarà tres cançons  
 La festa major és l'onze de maig  
 En Ferran ha rebut cinc cartes  
 El meu germà estudia a Barcelona  
 El concert comença a les nou  
 En Rafel té un Alfa Romeo  
 La Laura s'ha casat tres cops  
 Aquest bolso val cinc mil pessetes  
 Al pastís hi poso tres ous  
 De vacances aniré a Cuba  
 L'Ignasi ha estudiat econòmiques  
 Els meus pares viuen a Sabadell  
 L'Elies vindrà a les vuit  
 En Carles pot venir dijous  
 La Marta viu al tercer pis  
 El meu apartament té dos banys  
 La fira és el quinze de setembre  
 El meu cosí es diu Miquel  
 Al casament vindran tres-cents convidats  
 L'àvia fa cigrons per dinar  
 En Toni compra l'Avui  
 El meu escriptor preferit és Carner

Quants anys té en Pere?  
 Com t'agrada la carn?  
 Què vol comprar la Remei?  
 Quants fills té la Carme?  
 A quina hora et vols llevar demà?  
 Què fas els diumenges?  
 Quin dia arriba la Raquel?  
 On està l'ajuntament?  
 D'on és la teva amiga?  
 Què vols per sopar?  
 Quantes cançons cantarà la Caballé?  
 Quan és la festa major?  
 Quantes cartes ha rebut en Ferran?  
 On estudia el teu germà?  
 A quina hora comença el concert?  
 Quin cotxe té en Rafel?  
 Quants cops s'ha casat la Laura?  
 Quant val aquest bolso?  
 Quants ous poses al pastís?  
 On aniràs de vacances?  
 Què ha estudiat l'Ignasi?  
 On viuen els teus pares?  
 A quina hora vindrà l'Elies?  
 Quin dia pot venir en Carles?  
 A quin pis viu la Marta?  
 Quants banys té el teu apartament?  
 Quant és la fira?  
 Com es diu el teu cosí?  
 Quants convidats vindran al casament?  
 Què fa l'àvia per dinar?  
 Quin diari compra en Toni?  
 Quin és el teu escriptor preferit?

El llaç del gat és blau  
 La nena morena porta un globus  
 El gos gros porta un os  
 Les pomes són vermelles  
 En Xavier va a l'esglèsia  
 En Miquel rega cinc flors  
 La papallona és blava, vermella i groga  
 Al quadre hi ha un núvol  
 L'Anna renta la roba al riu  
 La casa té tres finestres  
 Les cortines són verdes  
 Sobre la catifa hi ha tres animals  
 Des de la finestra es veuen les muntanyes  
 El davantal de la Neus és vermell  
 La Roser escombra  
 La Cristina treu la pols del gerro  
 La faldilla de la nena és verda i blanca  
 A la paret queden dos quadres  
 Als peus de la Mireia hi ha un gat  
 En Manel dorm  
 El vestit de la Núria és taronja  
 Al penjador hi ha dos barrets  
 A la finestra hi ha dues flors  
 L'all i la ceba estan sobre la taula  
 Al cel hi ha tres ocells  
 La palmera té tres cocos  
 El vestit de la dona és verd  
 Els àngels estan sobre un núvol  
 Està plovent  
 La nena rossa porta un estel  
 A la cistella hi ha una pera  
 El quadre està sobre la taula blava

De quin color és el llaç del gat?  
 Quants globus porta la nena morena?  
 Què fa el gos gros?  
 De quin color són les pomes?  
 On va en Xavier?  
 Quantes flors rega en Miquel?  
 De quins colors és la papallona?  
 Quants núvols hi ha al quadre?  
 On renta la roba l'Anna?  
 Quantes finestres té la casa?  
 De quin color són les cortines?  
 Quants animals hi ha sobre la catifa?  
 Què es veu des de la finestra?  
 De quin color és el davantal de la Neus?  
 Què fa la Roser?  
 D'on treu la pols la Cristina?  
 De quins colors és la faldilla de la nena?  
 Quants quadres queden a la paret?  
 Què hi ha als peus de la Mireia?  
 Què fa en Manel?  
 De quin color és el vestit de la Núria?  
 Quants barrets hi ha al penjador?  
 Quantes flors hi ha a la finestra?  
 On estan l'all i la ceba?  
 Quants ocells hi ha al cel?  
 Quants cocos té la palmera?  
 De quin color és el vestit de la dona?  
 On estan els àngels?  
 Quin temps fa?  
 Què porta la nena rossa?  
 Quantes peres hi ha a la cistella?  
 On està el quadre?

### 1.3 Trial sentences

En Pere té tres germans i una germana  
 En Jordi sap cantar molt bé  
 Els Reis vull que em portin una camisa  
 La Maria vol sortir demà

Quants germans té en Pere?  
 En Jordi sap llegir molt bé?  
 Què vols que et portin els Reis?  
 Qui vol sortir demà?

En Jaume és pescador  
 Les botes d'en Jaume són verdes  
 L'Olga pren el sol  
 La tovallola de l'Olga és lila

En Pere és pescador?  
 De quin color són les botes d'en Jaume?  
 Qui pren el sol?  
 De quin color és la tovallola de l'Olga?

#### 1.4 Examples of sketches used in the semi-spontaneous task

*Triggering question*

Qui duia robes negres?  
"Who was wearing black clothes?"

*Expected response*

L'AMÈLIA duia robes negres  
"AMELIA was wearing black clothes"

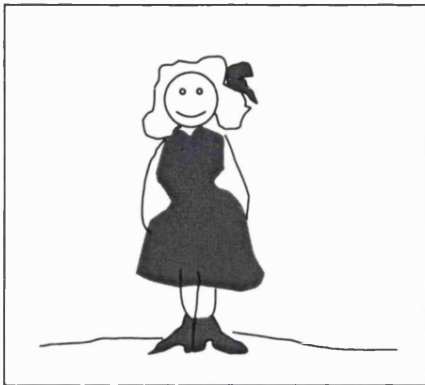
*Triggering question*

La Rosa talla els geranis?  
"Is Rosa cutting the geraniums?"

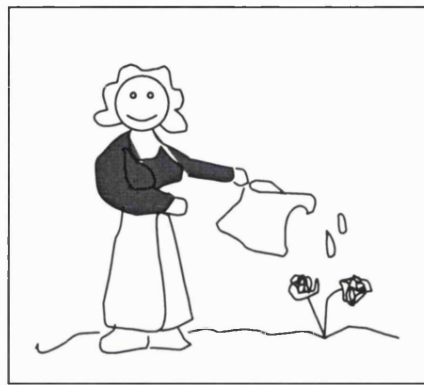
*Expected response*

La Rosa REGA els geranis  
"Rosa is WATERING the geraniums"

L' AMÈLIA



LA ROSA



## Appendix 2

Appendix 2 is divided into three sections. The first section contains the list of single-stressed structures (lexically stressed words) produced by six speakers in the second recording session. Words are nouns and verbs which have different stress patterns: oxytones (words with stress on the last syllable), paroxytones (words with stress on the penultimate syllable) and proparoxytones (words with stress on the antepenultimate syllable). The second section includes sentences where proparoxytones appear in initial (subject), medial (verb) and final (object) sentence positions. Finally, section 2.3 includes a few sentences with exactly the same segmental material expect for the first word, which consists of items with different stress patterns: oxytones, paroxytones and proparoxytones.

### 2.1 Isolated words

#### 2.1.1 Nouns

##### Proparoxytones

la làmina  
la Melanie  
la mínima  
l'ànima  
L'Àngela  
L'àguila

##### Paroxytones

la Mila  
la nena  
la mare  
la nina  
la Rosa  
la lluna

##### Oxytones

en Ramon  
el senyor  
la muller  
la mamà  
la llavor  
el baró

#### 2.1.2 Verbs

##### Proparoxytones

demana-li  
manej-a-li  
menja-li  
mana-li  
mima-li  
roba-li

##### Paroxytones

demana  
manej-a  
menja-ven  
mana  
mima  
roba

##### Oxytones

demanarà  
manejarà  
menjarà  
manarà  
mimarà  
robarà

### 2.2 Proparoxytones

#### 2.2.1 Initial position (subject)

La làmina lluirà amb el sol  
La Melanie mira les novel·les  
El número resulta massa gran  
La mínima baixarà demà

La nòmina no serà més alta  
 La bòveda donava al carrer  
 La màquina lligava l'herba  
 La Mònica menjava mongetes  
 La màniga no es podrà allargar  
 El nòmada menjava al carrer  
 L'ànima demanava descans  
 L'òliba dormia vora l'arbre  
 L'Àngela manejava l'ordinador  
 L'àrbitre demanava l'hora  
 L'íntegre no es podrà calcular  
 L'àguila volava sobre el mar

### 2.2.2 Medial position (verb)<sup>1</sup>

(És necessari que) en Jaume sàpiga la veritat  
 (És necessari que) els llibres càpiguen a l'armari  
 (És necessari que) la noia sàpiga la sol.lució  
 (És necessari que) la nena càpiga dins la vànova  
 (És necessari que) els homes sàpiguen el camí  
 (És necessari que) els jugadors càpiguen a l'estadi  
 Ell va dir diga-li la veritat  
 Ell va dir amaga-li la veritat  
 Ell va dir mira-li la mà  
 Ell va dir nega-li la mà  
 Ell va dir mima-li la nena  
 Ell va dir dóna-li la mandarina

### 2.2.3 Final position (object)

La nena mira ...

la nòmina / la bòveda / la màniga / la làmina / el número / la mínima / l'íntegre /  
 l'òliba / l'àguila / l'ànima / la Melanie / la Mònica / l'Àngela / la màquina / el nòmada  
 / l'àrbitre

## 2.3 Oxytones, paroxytones and proparoxytones in initial position

La <u>là</u> mina	} lluirà amb el sol	El <u>nú</u> mero	} resulta massa gran
La <u>ll</u> una		El <u>nom</u> bre	
La <u>llav</u> or		El <u>cam</u> ió	

<sup>1</sup> In Catalan there are not many verbs with stress on the antepenultimate syllable. Hence, only two kinds of structures have been included in the data: the subjunctive forms of *saber*(*sàpiga/sàpiguen*) and *cabre* (*càpiga/càpiguen*) and a few imperative forms followed by a pronoun (*li*).



La Melanie }  
La Mila } mira les novel·les  
En Ramon }

La mínima }  
La nena } baixarà demà  
La Roser }

La nòmina }  
La dona } no serà més alta  
La muller }

La bòveda }  
La tribuna } donava al carrer  
El saló }

La màquina }  
Els homes } lligava/en l'herba  
El baró }

La Mònica }  
La mare } menjava mongetes  
La mamà }

La màniga }  
El gorro } no es podrà allargar  
El jersei }

El nòmada }  
La nena } menjava al carrer  
El senyor }

L'ànima }  
La nina } demanava descans  
La muller }

L'Àngela }  
La Rosa } manejava l'ordinador  
En Raimon }

L'àguila }  
La lluna } volava sobre el mar  
L'avió }

L'àrbitre }  
El nino } demanava l'hora  
El jugador }

L'òliba }  
El lladre } dormia vora l'arbre  
El senyor }

L'íntegre }  
La suma } no es podrà calcular  
L'operació }

## **Appendix 3**

This appendix shows two kinds of materials. First, section 3.1 contains sentences designed to analyse narrow focus on the object and the triggering questions. For each sentence, there are two kinds of triggering questions: 1) questions that prompt identification and 2) questions that prompt contrast. Second, section 3.2 includes sentences with narrow focus on the subject and on the verb consisting of proparoxytones and oxytones.

### **3.1 Narrow focus on the object**

La Rosa llegia LLIBRES	Què llegia la Rosa? La Rosa llegia revistes?
En Juli hauria de demanar L'HORA	Què hauria de demanar en Juli? En Juli hauria de demanar el dia?
La Mireia remena L'OLLA	Què remena la Mireia? La Mireia remena la cassola?
L'au venia de L'ILLA	D'on venia l'au? L'au venia del continent?
L'Emília vol AMANIDA	Què vol l'Emília? L'Emília vol sopa?
La Glòria ve de VILANOVA	D'on ve la Glòria? La Glòria ve de Barcelona?
La Marina beu LLIMONADA	Què veu la Marina? La Marina beu cervesa?
Els homes llimaven L'ARMARI	Què llimaven els homes Els homes llimaven la taula?
La nena morena venia NINES NOVES	Què venia la nena morena? La nena morena venia juguets?
L'Amèlia duia ROBES NEGRES	Què duia l'Amèlia? L'Amèlia duia faldilles?
Les meves amigues eren BONES NENES	Com eren les teves amigues? Les teves amigues eren antipàtiques?

Algunes minyones guarden MALA LLUNA	Què guarden algunes minyones? Algunes minyones guarden rancúnia?
L'home venia LLIMONES MADURES	Què venia l'home? L'home venia taronges?
La meva germana balla BALADES VIENESES	Què balla la teva germana? La teva germana balla valsos?
L'Eva guarda MONEDES ROMANES	Què guarda l'Eva? L'Eva guarda segells?
La Núria vol una GERMANA MENUDA	Què vol la Núria? La Núria vol un germà?

### 3.2 Proparoxytones and oxytones in subject and verb focus

#### 3.2.1 Paroxytones

LA LÀMINA lluirà amb el sol	El quadre lluirà amb el sol?
LA MELANIE mira les novel.les	Qui mira les novel.les?
EL NÚMERO resulta massa gran	La xifra resulta massa gran?
EL NÒMADA menjava al carrer	Qui menjava al carrer?
LA MÍNIMA baixarà demà	La màxima baixarà demà?
LA BÒVEDA donava al carrer	Què donava al carrer?
L'ÀNIMA demanava descans	El cos demanava descans?
L'ÀNGELA manejava l'ordinador	Qui manejava l'ordinador?
L'ÀGUILA volava sobre el mar	La gavina volava sobre el mar?
L'ÀRBITRE demanava l'hora	Qui demanaven l'hora?
L'ÒLIBA dormia vora l'arbre	El mussol dormia sobre l'arbre?
L'ÍNTEGRE no es podia calcular	Què no es podia calcular?
(És necessari que)...	(És necessari que)...
En Jaume SÀPIGA la veritat	En Jaume digui la veritat?
Els llibres CÀPIGUEN a l'armari	Els llibres estiguin a l'armari?
La nena CÀPIGA dins la vànova	La nena dormi dins la vànova?
(És necessari que)...	(Què és necessari que)...
La noia SÀPIGA la sol.lució	Faci la noia amb la sol.lució?
Els jugadors CÀPIGUEN a l'estadi	Facin els jugadors a l'estadi?
Els homes SÀPIGUEN el camí	Facin els homes amb el camí?

#### 3.2.2 Oxytones

UN arribarà demà	Tres arribaran demà?
L'ANY havia acabat malament	Què havia acabat malament?
L'ALL era olorós	La ceba era olorosa?
L'ULL havia sagnat	Què havia sagnat?

NOU miraven el mar  
L'OM era verd

La nena DEMANARÀ la pregunta  
La mare MANEJARÀ l'ordinador  
El nen MENJARÀ l'amanida  
La noia MIMARÀ els alumnes  
En Jordi MANARÀ els obrers  
El lladre ROBARÀ la maleta

Deu miraven el mar?  
Quin arbre era verd?

La nena respondrà la pregunta?  
Què farà la mare amb l'ordinador?  
El nen farà l'amanida?  
Què farà la noia amb els alumnes?  
En Jordi pagarà els obrers?  
Què farà el lladre amb la maleta?

## **Appendix 4**

This appendix contains the materials designed for the analysis of English. Section 4.1 contains a list of isolated words. Words are divided in proparoxytones, paroxytones and oxytones. Second, the list of sentences produced with a broad focus reading is presented in section 4.2. Finally, the sentences with narrow focus on the subject, verb and object are included in section 4.3 along with their triggering questions (narrow focus triggered by contrast and by identification).

### **4.1 Isolated words**

#### **Proparoxytones**

Melanie  
Marjorie  
Emily  
Marilyn  
Jeremy  
the melody  
the gardener  
the lullaby  
the marmalade  
the memory

#### **Paroxytones**

Mary  
Millie  
Molly  
Mini  
Julie  
the lawyer  
the mother  
the lady  
the nanny  
the mummy

#### **Oxytones**

Jo  
Lee  
Ray  
John  
Ron  
the law  
the balloon  
the meal  
the name  
the moon

### **4.2 Broad focus sentences**

Jo relies on my money  
Lee remembered the melody  
Ray resigned from his job  
John divided the legacy  
The law has to be modified  
The balloon was moving upwards  
The meal was absolutely delicious  
The name was written at the door

Mary learns modern languages  
Millie lives in Ireland  
Molly made a marvellous dinner  
Mini loves lemon marmalade  
The lawyer's reading the journal  
My mother loves babies  
The lady ordered an orange juice  
The nanny's ironing the linen

Melanie will win the award  
Marjorie married an Indian marine  
Emily damaged the window  
Marilyn nurses the baby  
The melody was amazing  
The gardener used to live in London  
The lullaby brings good memories  
The marmalade is on the fridge  
The gardener mows the lawn

The boys need a new game  
Norma runs every morning  
The bluebells die with the rain  
Marjorie adores movies  
The girls believe in fairies  
Brenda removes the garbage  
The jam remained in the jar

Josephine needed an explanation  
My neighbour married the model  
The general murdered the villain  
Melanie ordered an orange juice  
Ron amended the bill  
My brother remembers the rhymes  
The boy delivered the magazines  
Jonathan erases the message

Melanie's menacing the hostages  
Grandma's ironing the linen  
The rain damages the harvest  
The journalist balances the disorder  
The minister moderates the audience  
Jeremy manages his company  
Julie memorised the names  
John normalised the numbers

Ron's paying bills  
The journalist's ordering wine  
Jeremy's wearing jeans  
Marilyn's boiling eggs  
The nominee won the award  
Amelia married a marine  
The jam remained in the jar  
The bluebells die with the rain

My mother loves babies  
Nelly adores movies  
John married Mary  
The lady's wearing diamonds

Gary borrowed the money  
Julie memorised the numbers  
The nanny's ironing the linen  
The general murdered the villain

John married Melanie  
Jimmy loves gardening  
The barber adores Marilyn  
Mary learns languages  
The lady ordered the marmalade  
The boy delivered the magazines  
The minister remembers the melody  
The nanny sings a lullaby

### 4.3 Narrow focus sentences and triggering questions

#### 4.3.1 Narrow focus on the subject

##### 4.3.1.1 Narrow focus triggered by contrast

RAY resigned from his job  
JOHN divided the legacy  
THE BALLOON was moving upwards  
THE NAME was written at the door

Did Peter resign from his job?  
Did Michael divide the legacy?  
Was the plane moving upwards?  
Was the address written at the door?

MOLLY made a marvellous dinner  
MINI loves lemon marmalade  
MY MOTHER loves babies  
THE NANNY's ironing the linen

Did Helen make a marvellous dinner?  
Does Karen love lemon marmalade?  
Does your sister love babies?  
Is the butler ironing the linen?

MELANIE will win the award  
EMILY damaged the window  
THE MELODY was amazing  
THE MARMALADE is on the fridge

Will Mary win the award?  
Did Judith damage the window?  
Was the novel amazing?  
Is the jam on the fridge?

##### 4.3.1.2 Narrow focus triggered by identification

JO relies on my money  
LEE remembered the melody  
THE LAW has to be modified  
THE MEAL was absolutely delicious

Who relies on my money?  
Who remembered the melody?  
What has to be modified?  
What was absolutely delicious?

MARY learns modern languages  
MILLIE lives in Ireland  
THE LAWYER's reading the journal  
THE LADY ordered an orange juice

Who learns modern languages?  
Who lives in Ireland?  
Who's reading the journal?  
Who ordered an orange juice?

MARJORIE married an Indian marine

Who married an Indian marine?

MARILYN nurses the baby	Who nurses the baby?
THE GARDENER used to live in London	Who used to live in London?
THE LULLABY brings good memories	What brings good memories?

### 4.3.2 Narrow focus on the verb

#### 4.3.2.1 Narrow focus triggered by contrast

The boys NEED a new game	Do the boys use a new game?
The bluebells DIE with the rain	Do the bluebells grow with the rain?
The girls BELIEVE in fairies	Do the girls dislike fairies?
The jam REMAINED in the jar	Did the jam come out of the jar?

Josephine NEEDED an explanation	Did Josephine want an explanation?
Melanie ORDERED an orange juice	Did Melanie buy an orange juice?
My brother REMEMBERS the rhymes	Does your brother learn the rhymes?
The boy DELIVERED the magazines	Did the boy keep the magazines?

The rain DAMAGES the harvest	Does the rain benefit the harvest?
The journalist BALANCES the disorder	Does the journalist cause the disorder?
The minister MODERATES the audience	Does the minister speak to the audience?
Jeremy MANAGES his company	Does Jeremy work in his company?

#### 4.3.2.2 Narrow focus triggered by identification

The gardener MOWS the lawn	What does the gardener do with the lawn?
Norma RUNS every morning	What does Norma do every morning?
Marjorie ADORES movies	What does Marjorie do with movies?
Brenda REMOVES the garbage	What does Brenda do with the garbage?

My neighbour MARRIED the model	What did your neighbour do with the model?
The general MURDERED the villain	What did the general do with the villain?
Ron AMENDED the bill	What did Ron do with the bill?
Jonathan ERASES the message	What does Jonathan do with the message?

Melanie's MENACING the hostages	What's Melanie doing with the hostages?
Grandma's IRONING the linen	What's grandma doing with the linen?
Julie MEMORISED the names	What did Julie do with the names?
John NORMALISED the numbers	What did John do with the numbers?

### 4.3.3 Narrow focus on the object

#### 4.3.3.1 Narrow focus triggered by contrast

The journalist's ordering WINE	Is the journalist ordering beer?
Marilyn's boiling EGGS	Is Marilyn boiling vegetables?
Amelia married A MARINE	Did Amelia married an officer?



The bluebells die with THE RAIN

Do the bluebells die with the sun?

Nelly adores MOVIES

Does Nelly adore novels?

John married MARY

Did John marry Susan?

Julie memorised THE NUMBERS

Did Julie memorise the letters?

The general murdered THE VILLAIN

Did the general murdered the thief?

Jimmy loves GARDENING

Does Jimmy love fishing?

Mary learns LANGUAGES

Does Mary learn music?

The boy delivered THE MAGAZINES

Did the boy deliver the newspapers?

The nanny sings A LULLABY

Does the nanny sing a carol?

#### 4.3.3.2 Narrow focus triggered by identification

Ron's paying BILLS

What's Ron paying?

Jeremy's wearing JEANS

What's Jeremy wearing?

The nominee won THE AWARD

What did the nominee win?

The jam remained in THE JAR

Where did the jam remain?

My mother loves BABIES

What does your mother love?

The lady's wearing DIAMONDS

What's the lady wearing?

Gary borrowed THE MONEY

What did Gary borrow?

The nanny's ironing THE LINEN

What's the nanny ironing?

John married MELANIE

Who did John marry?

The barber adores MARILYN

Who does the barber adore?

The lady ordered THE MARMALADE

What did the lady order?

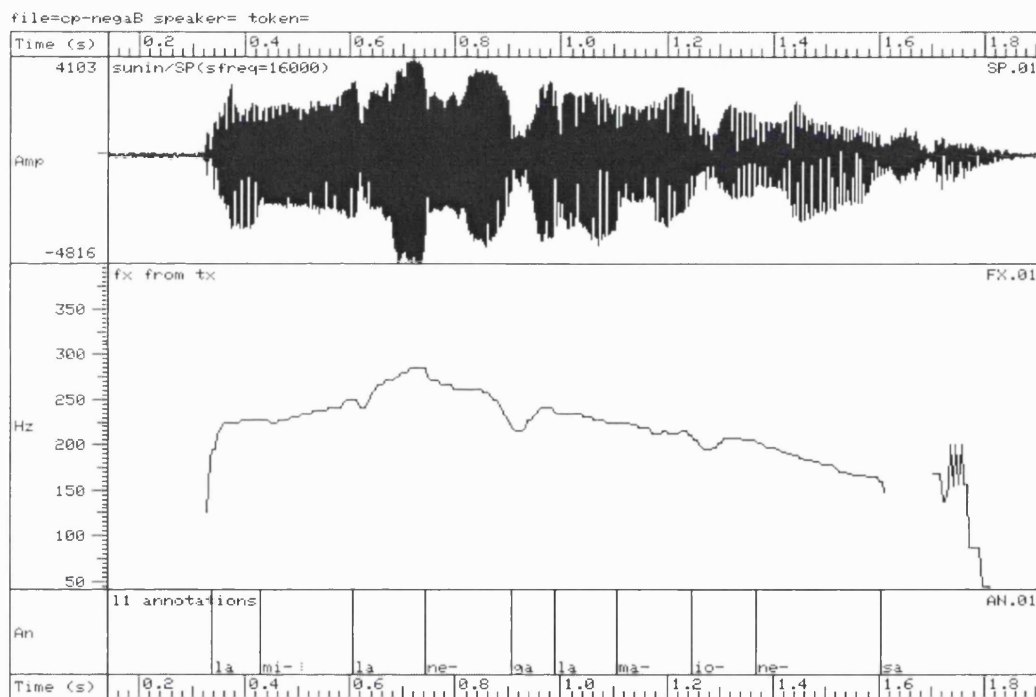
The minister remembers THE MELODY

What does the minister remember?

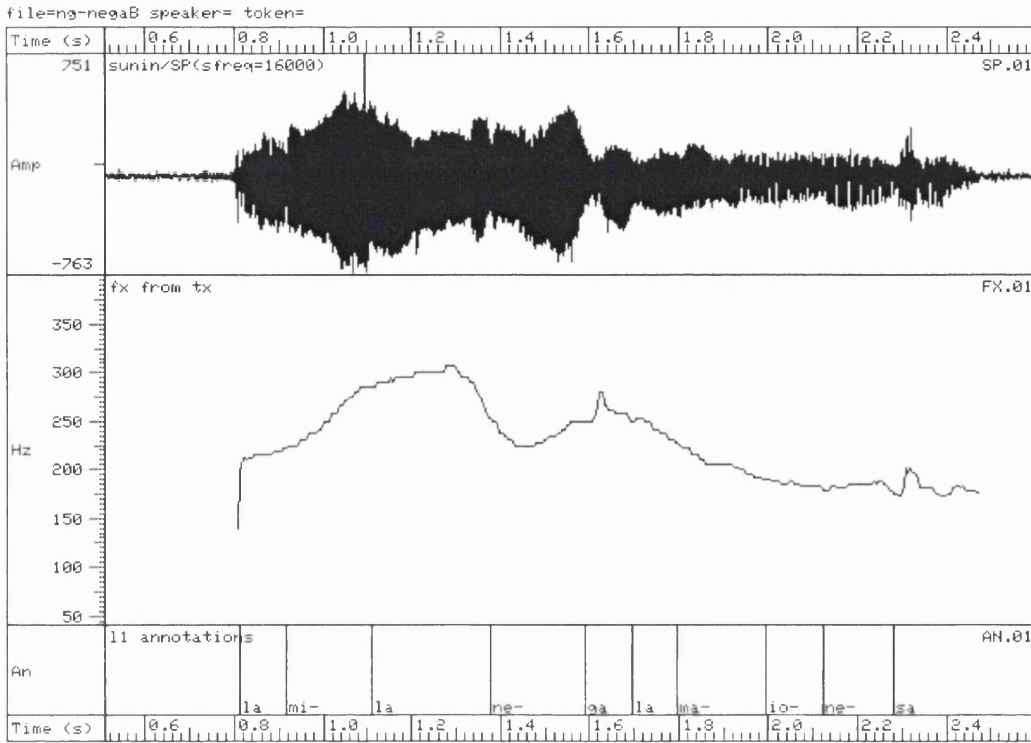
## Appendix 5

This appendix contains the figures of the speech waveforms and F0 traces corresponding to the sketched examples of chapters 3-6.

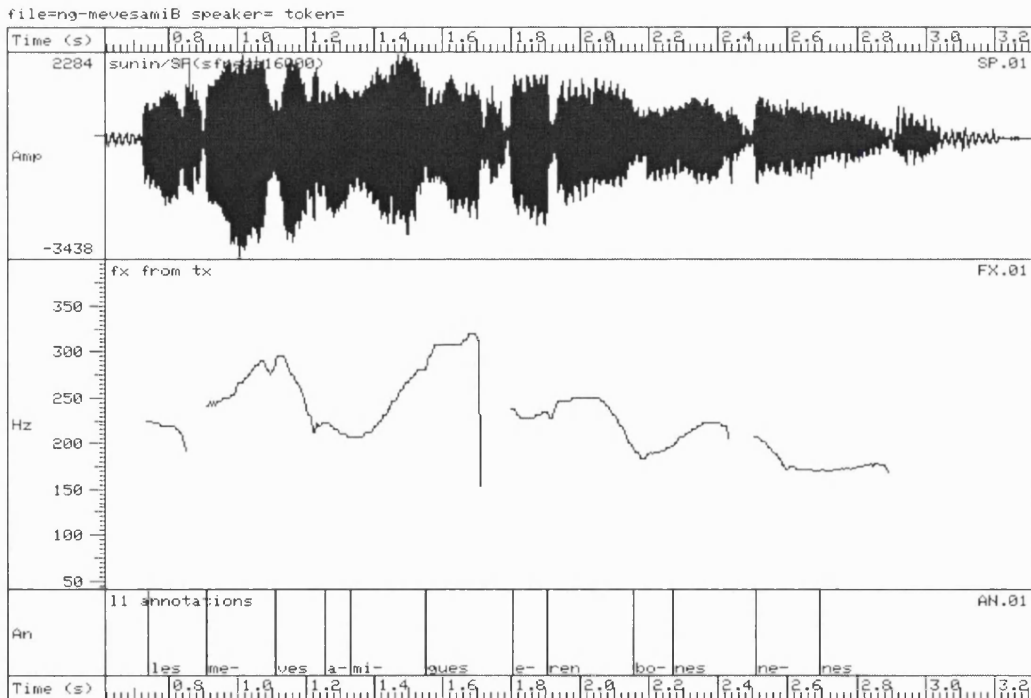
### 5.1. Figures chapter 3



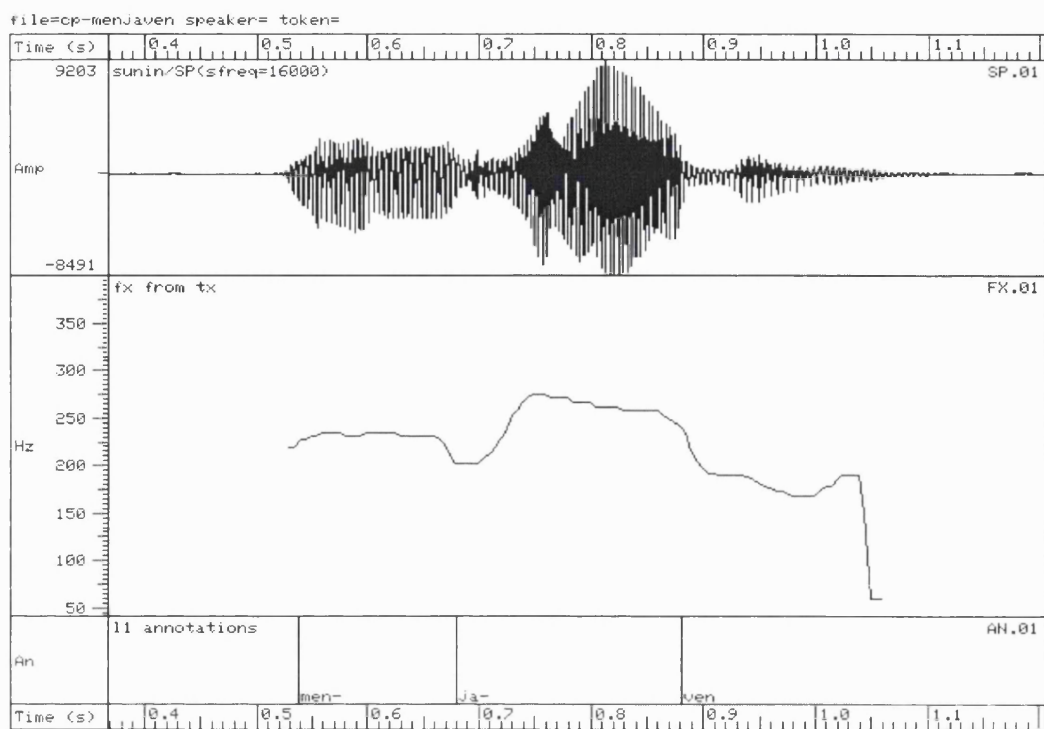
A.1. Figure for example (3.4). *La Mila nega la maionesa* (speaker CP).



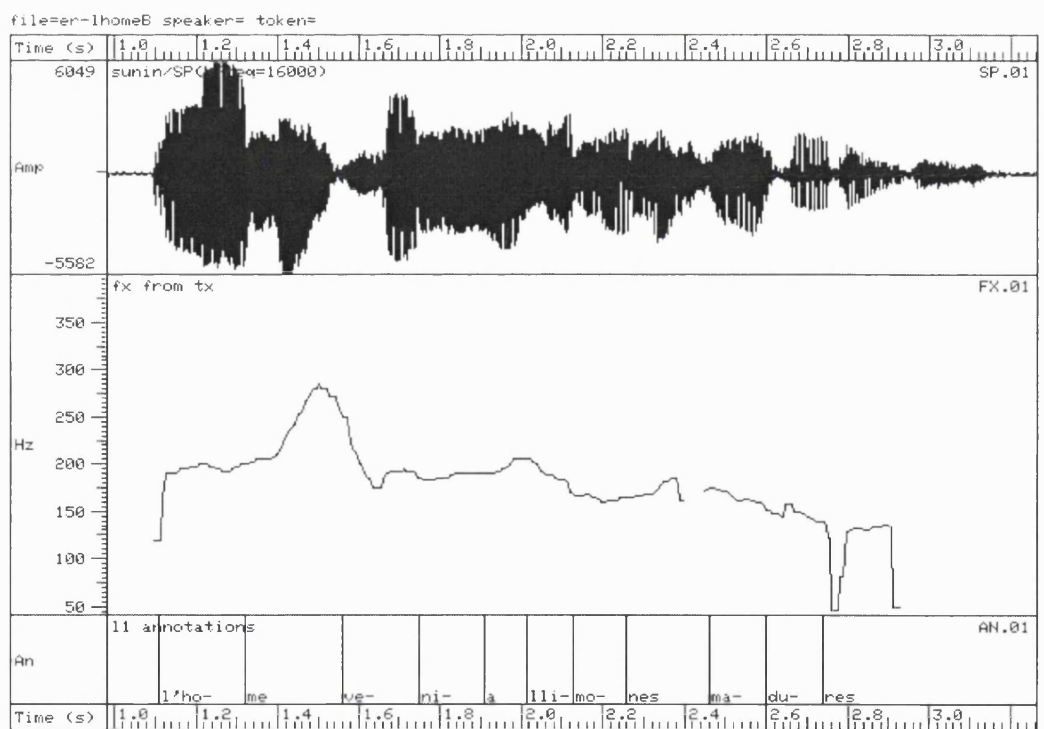
A.2. Figure for example (3.5). *La Mila nega la maionesa* (speaker NG).



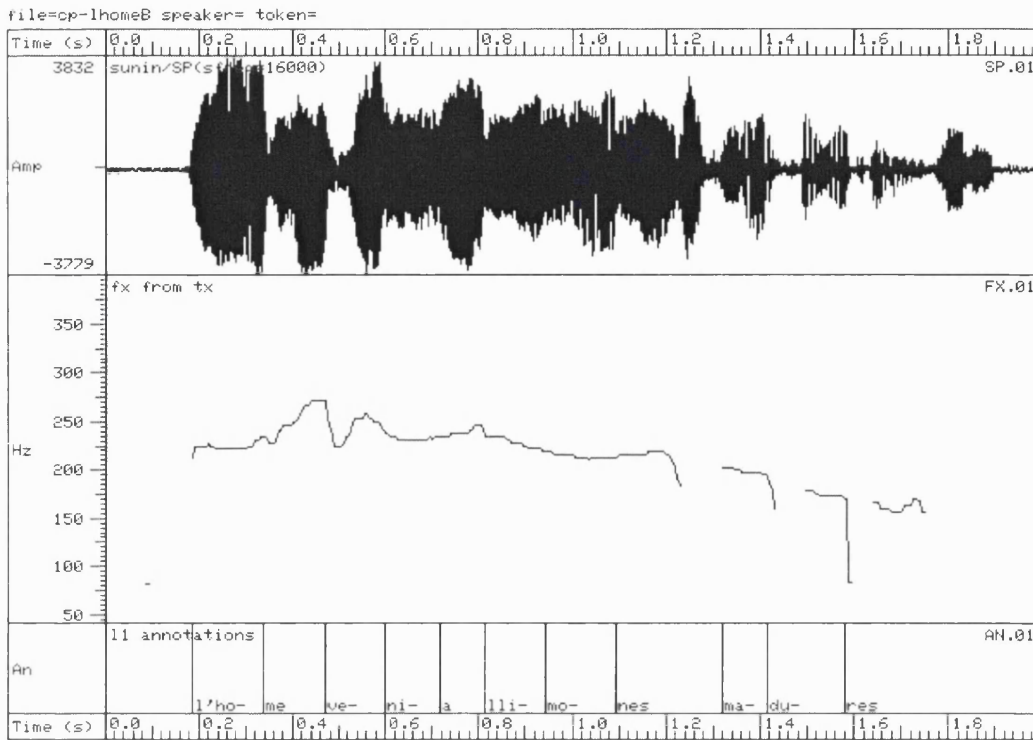
A.3. Figure for example (3.6). *Les meves amigues eren bones nenes* (speaker NG).



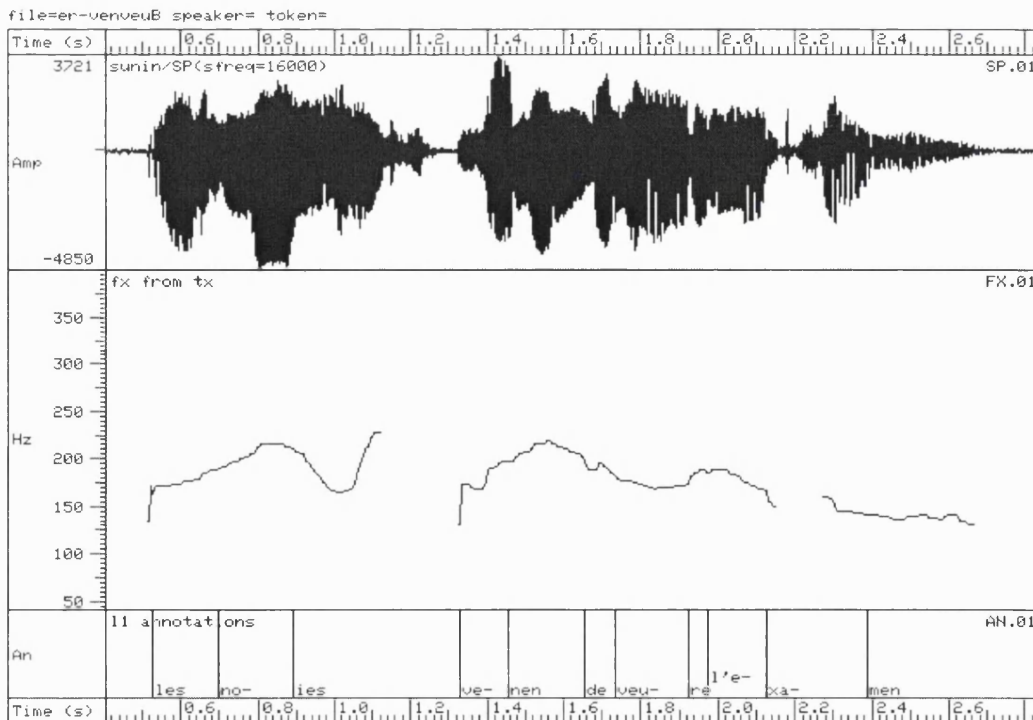
A.4. Figure for example (3.7). *Menjaven* (speaker CP).



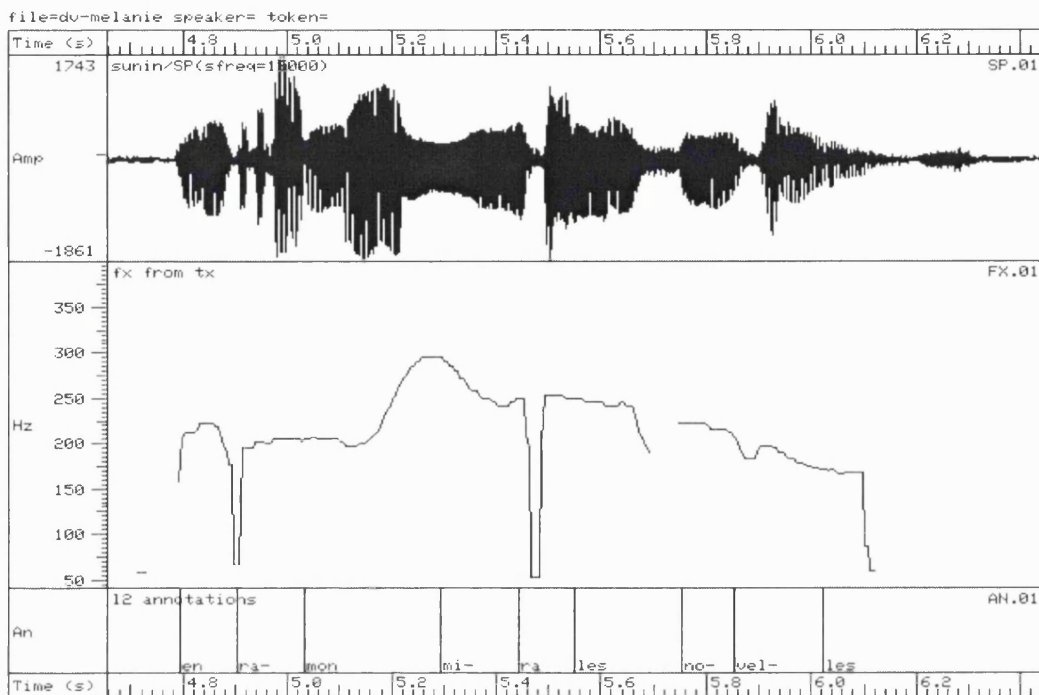
A.5. Figure for example (3.13a). *L'home venia llimones madures* (speaker ER).



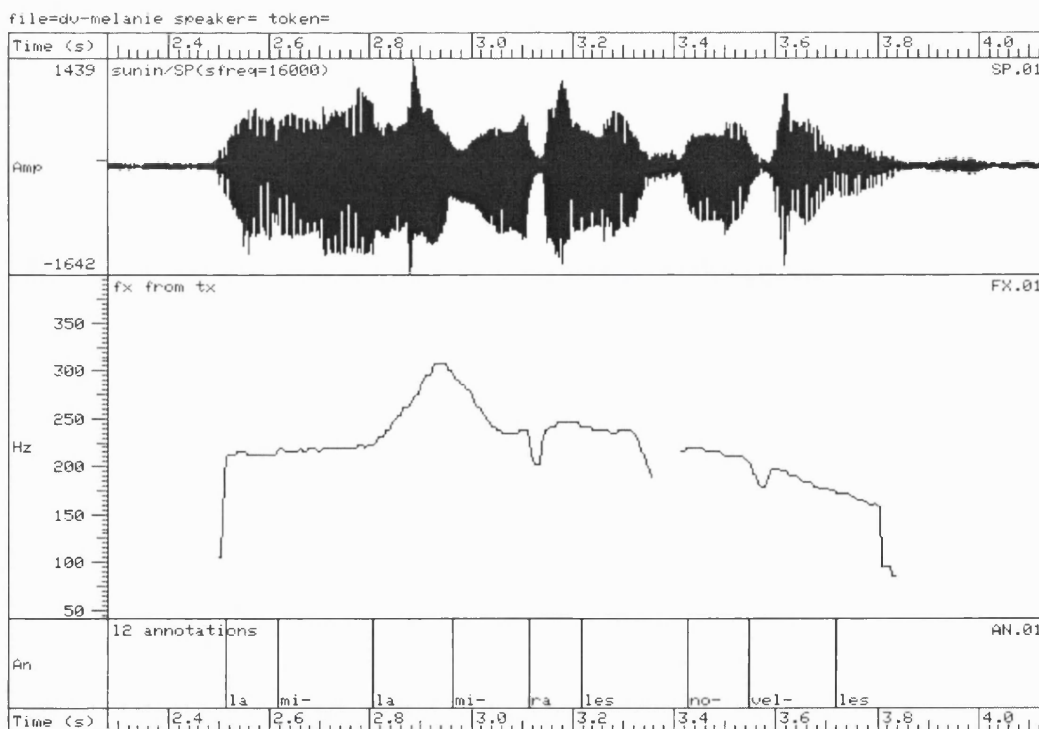
A.6. Figure for example (3.13b). *L'home venia llimones madures* (speaker CP).



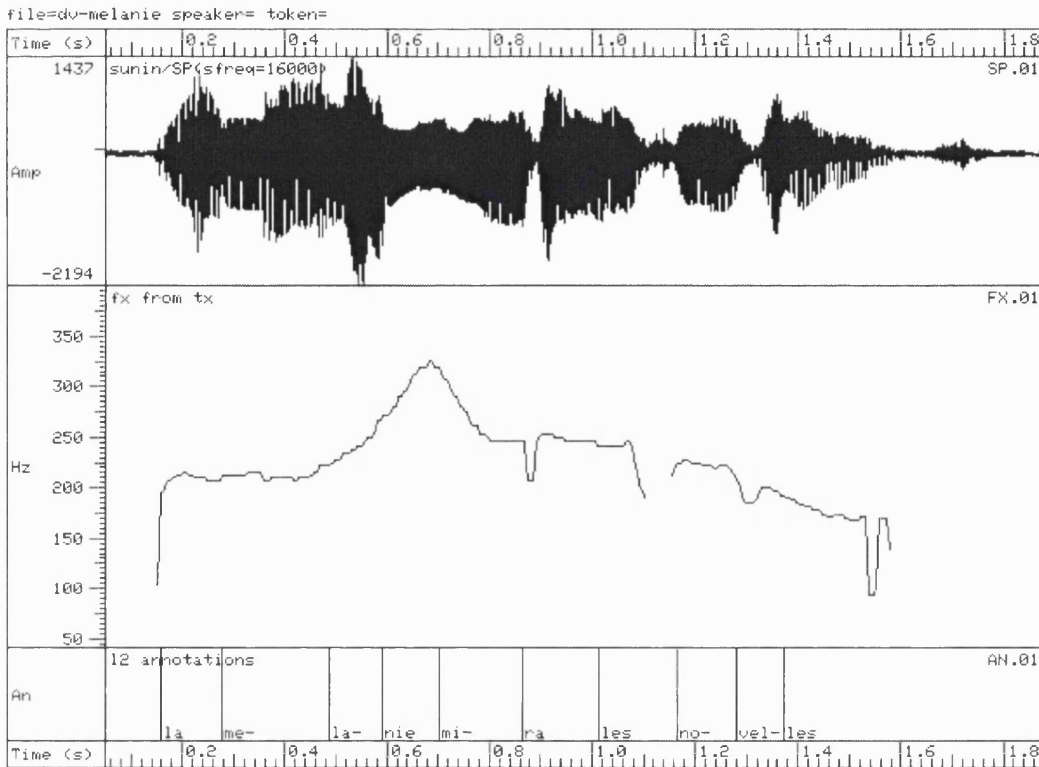
A.7. Figure for example (3.14). *Les noies vénen de veure l'examen* (speaker ER).



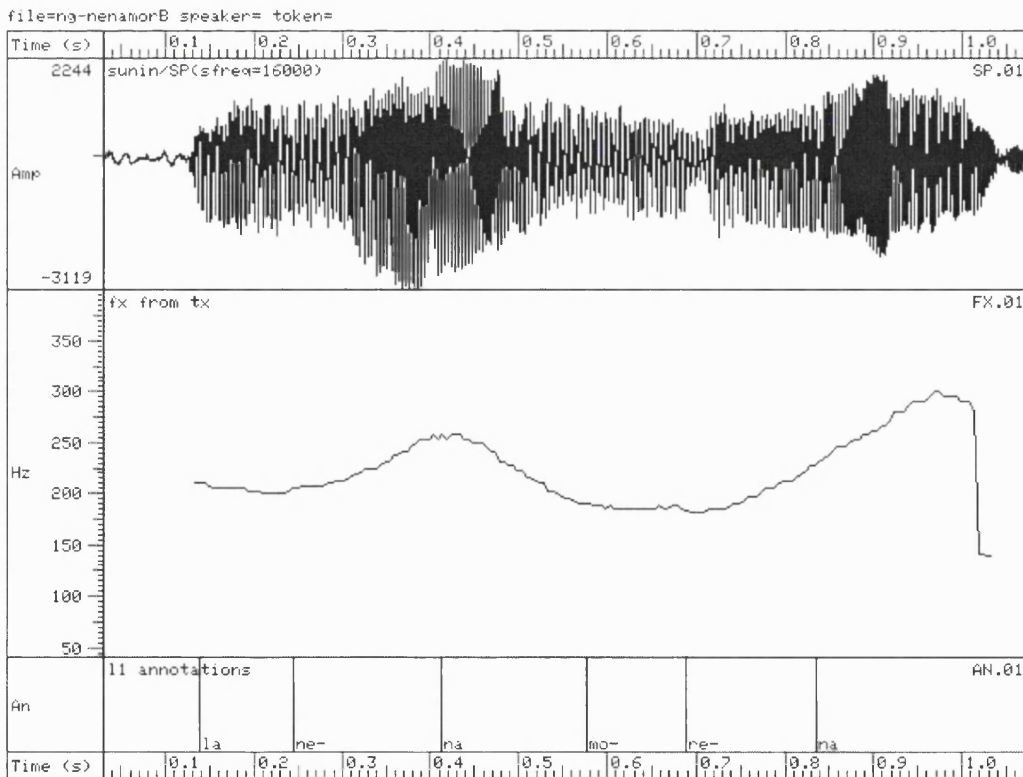
A.8. Figure for example (3.19). *En Ramon mira les novel.les* (speaker DV).



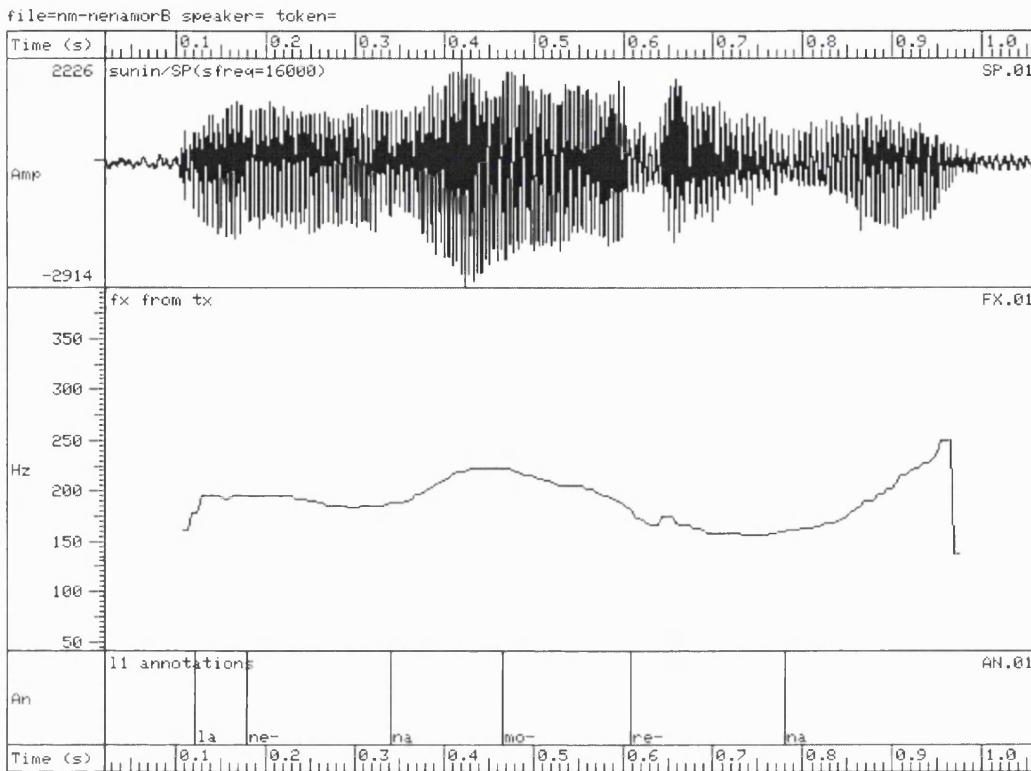
A.9. Figure for example (3.20). *La Mila mira les novel.les* (speaker DV).



A.10. Figure for example (3.21). *La Melanie mira les noyel.les* (speaker DV).

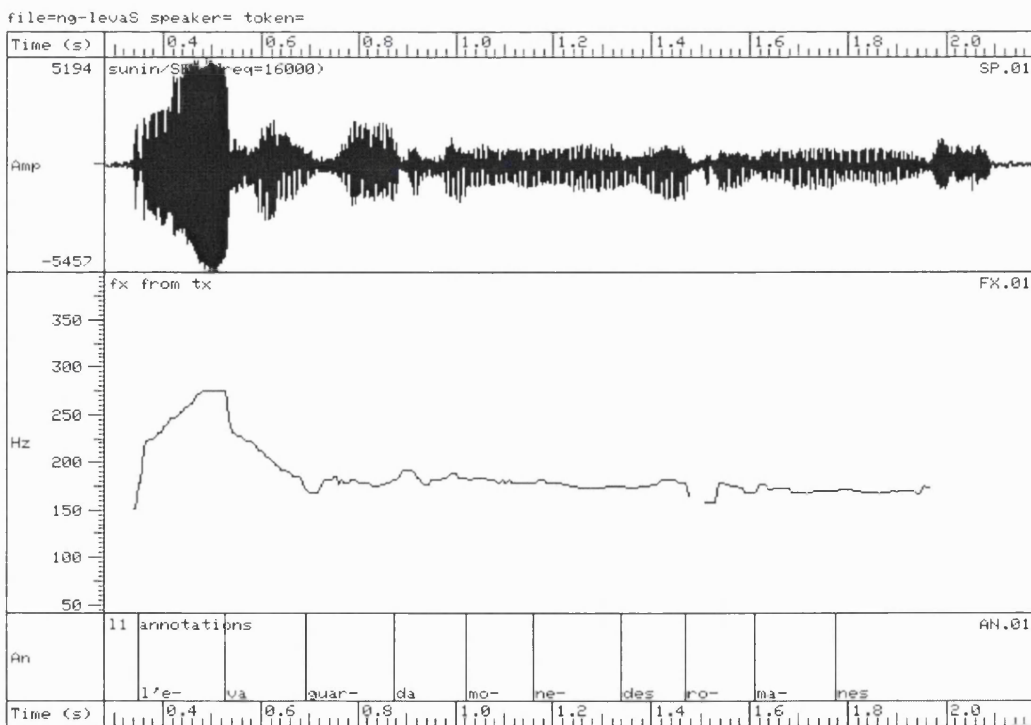


A.11. Figure for example (3.23a). *La nena morena ...* (speaker NG).



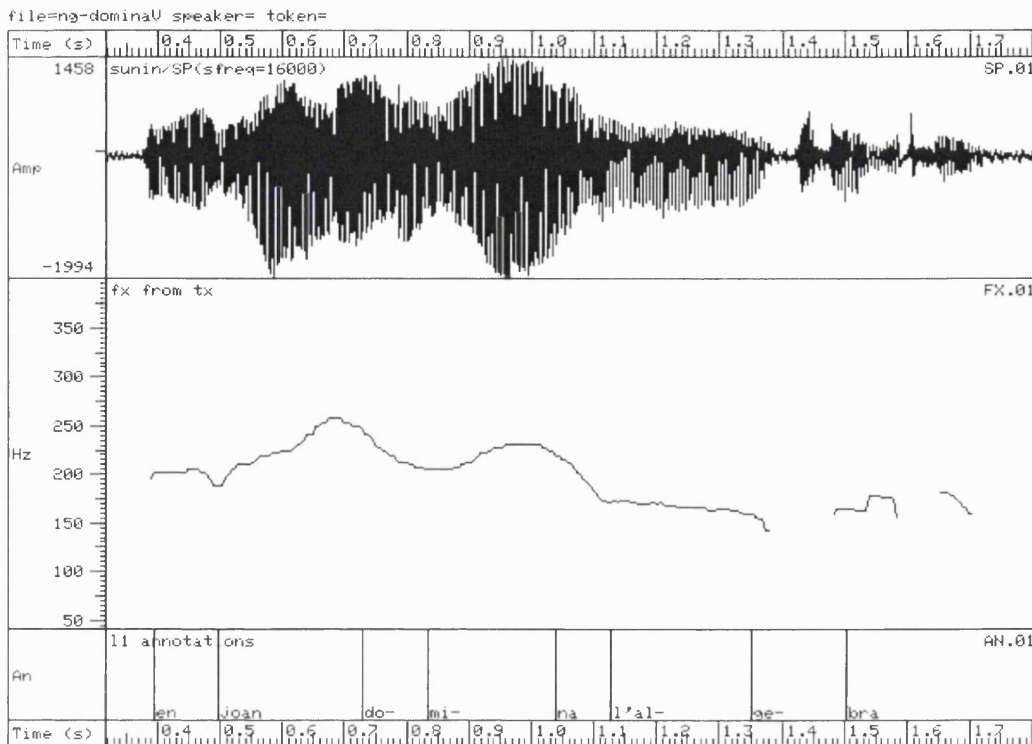
A.12. Figure for example (3.23b). *La nena morena ...* (speaker NM).

## 5.2 Figures chapter 4

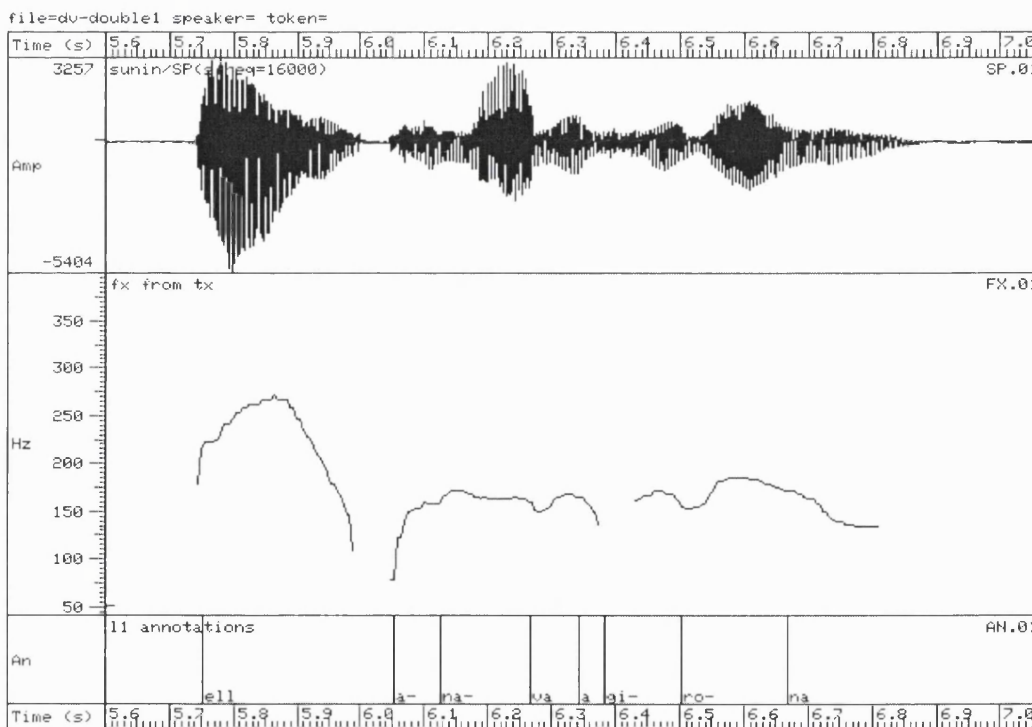


A.13. Figure for example (4.6). *L'EVA guarda monedas romanes* (speaker NG).

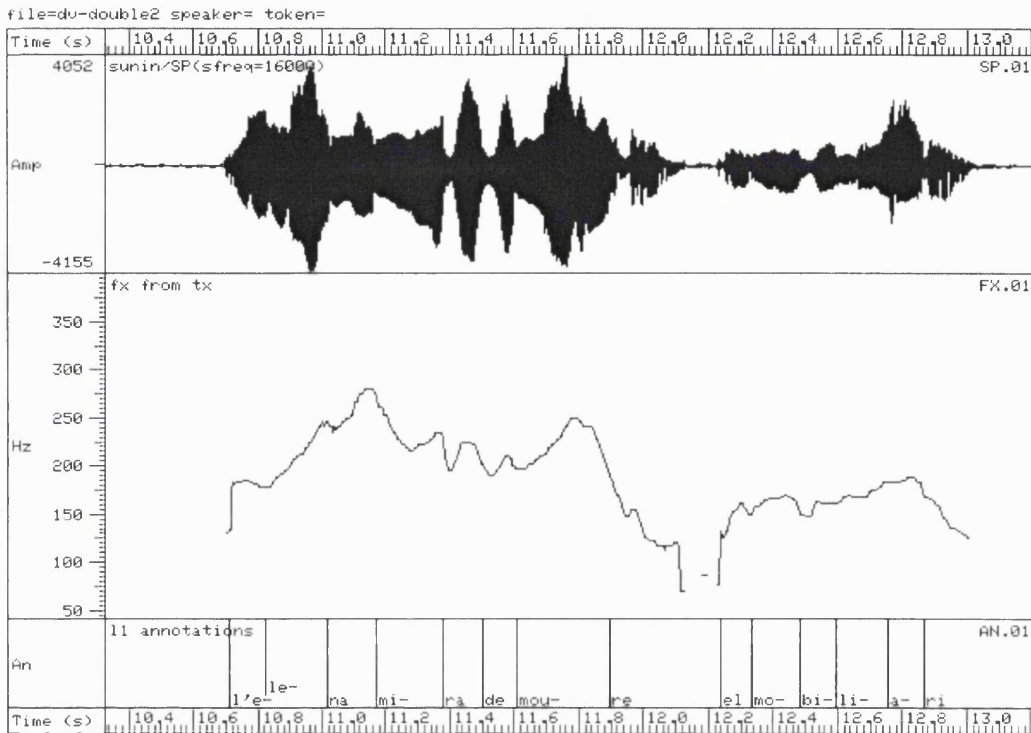




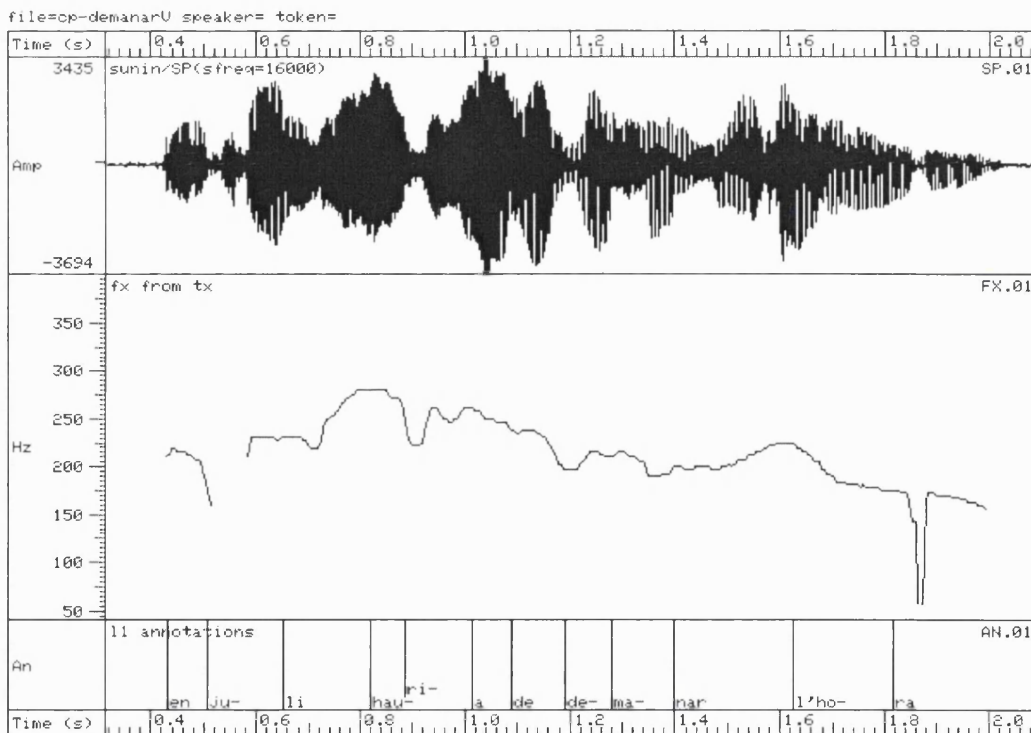
A.14. Figure for example (4.7). *En Joan DOMINA l'àlgebra* (speaker NG).



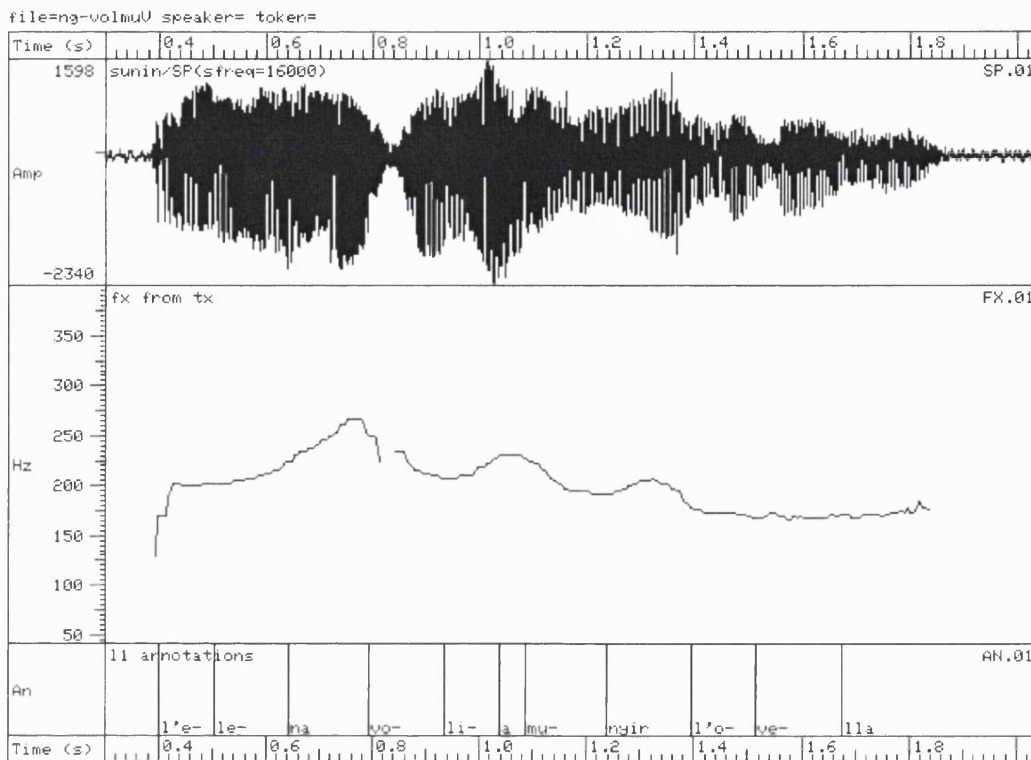
A.15. Figure for examples (4.8) and (4.25). *ELL anava a Girona* (speaker DV).



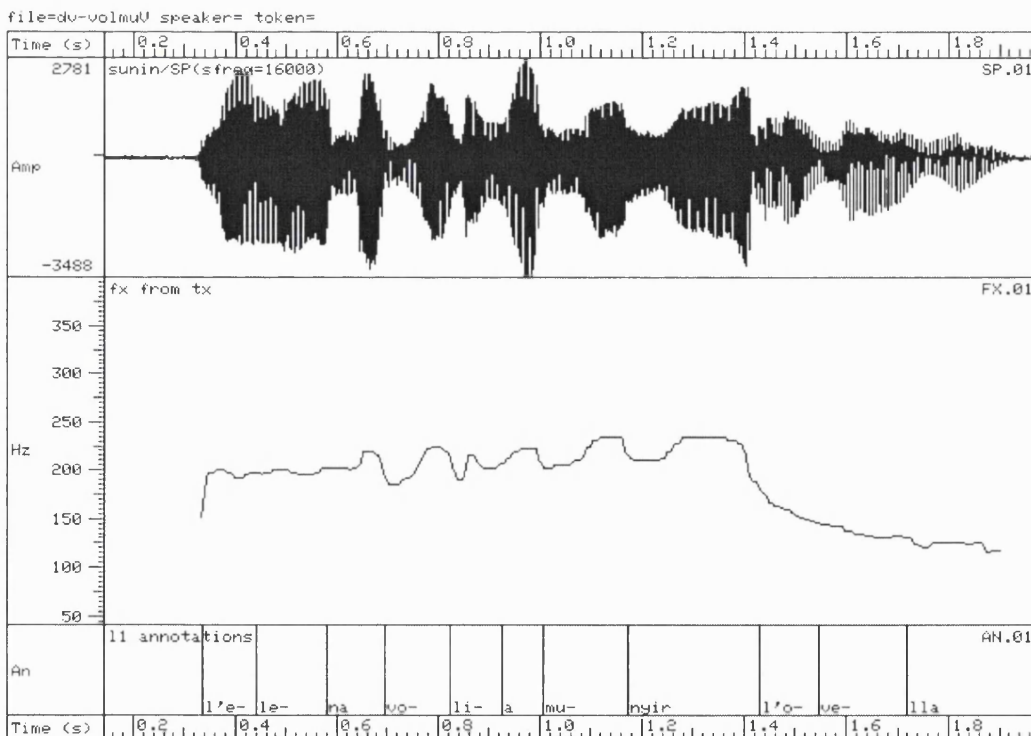
A.16. Figure for examples (4.9) and (4.26). *L'Elena mira de MOURE el mobiliari* (speaker DV).



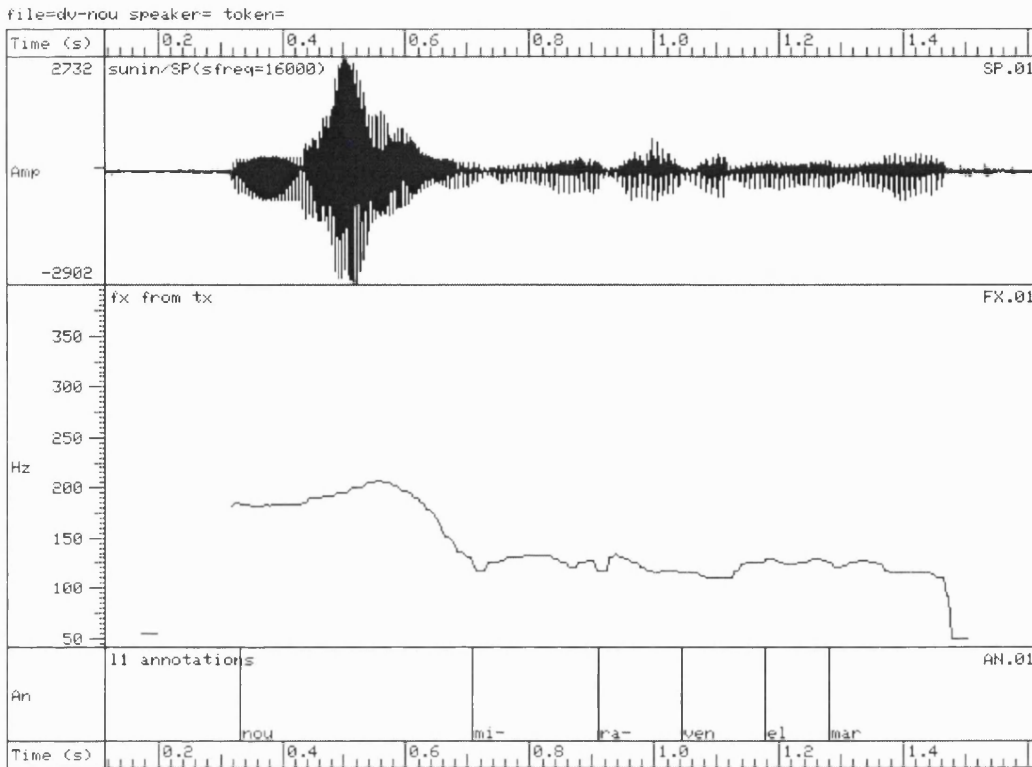
A.17. Figure for example (4.10). *En Juli hauria de DEMANAR l'hora* (speaker CP).



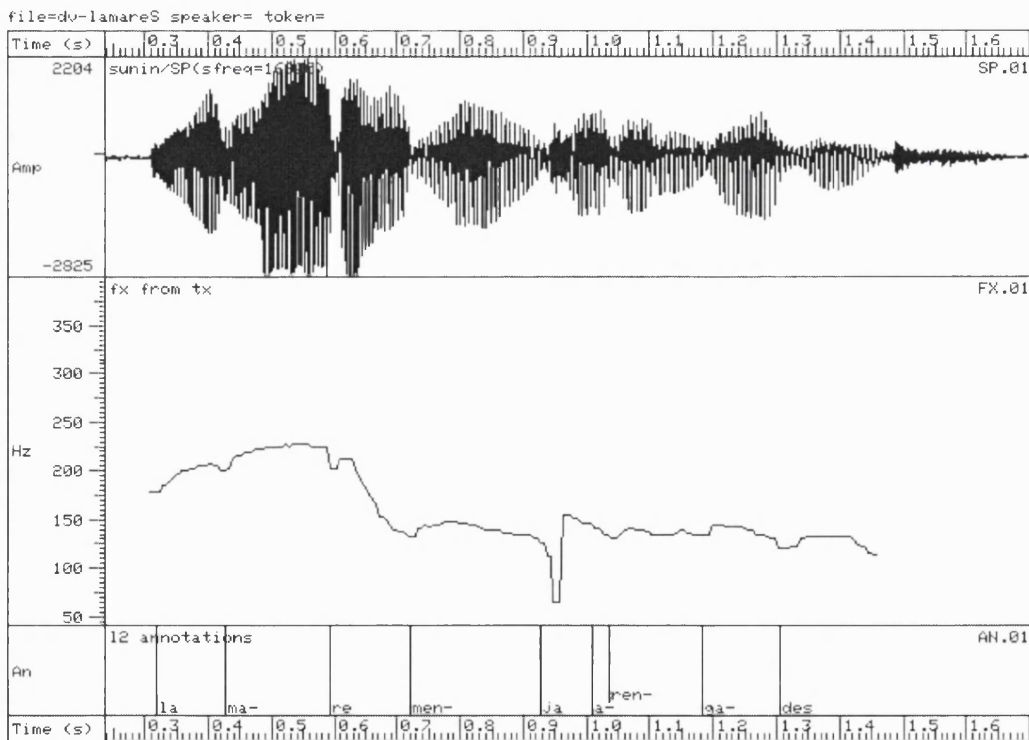
A.18. Figure for example (4.11). *L'Elena volia MUNYIR l'ovella* (speaker NG).



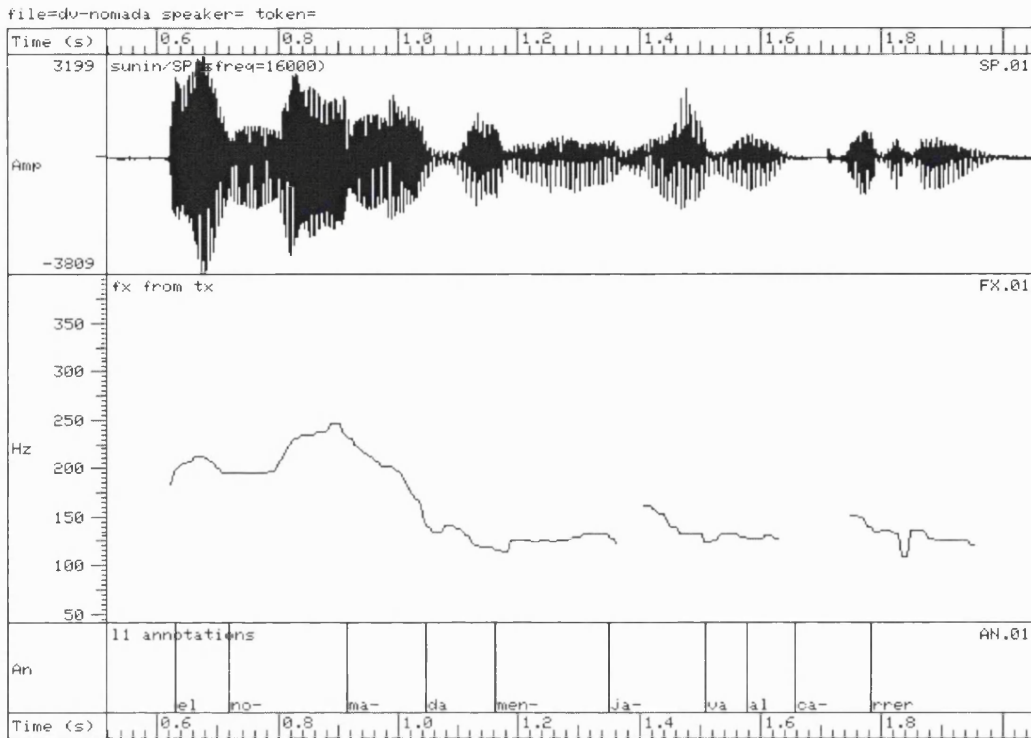
A.19. Figure for example (4.12). *L'Elena volia MUNYIR l'ovella* (speaker DV).



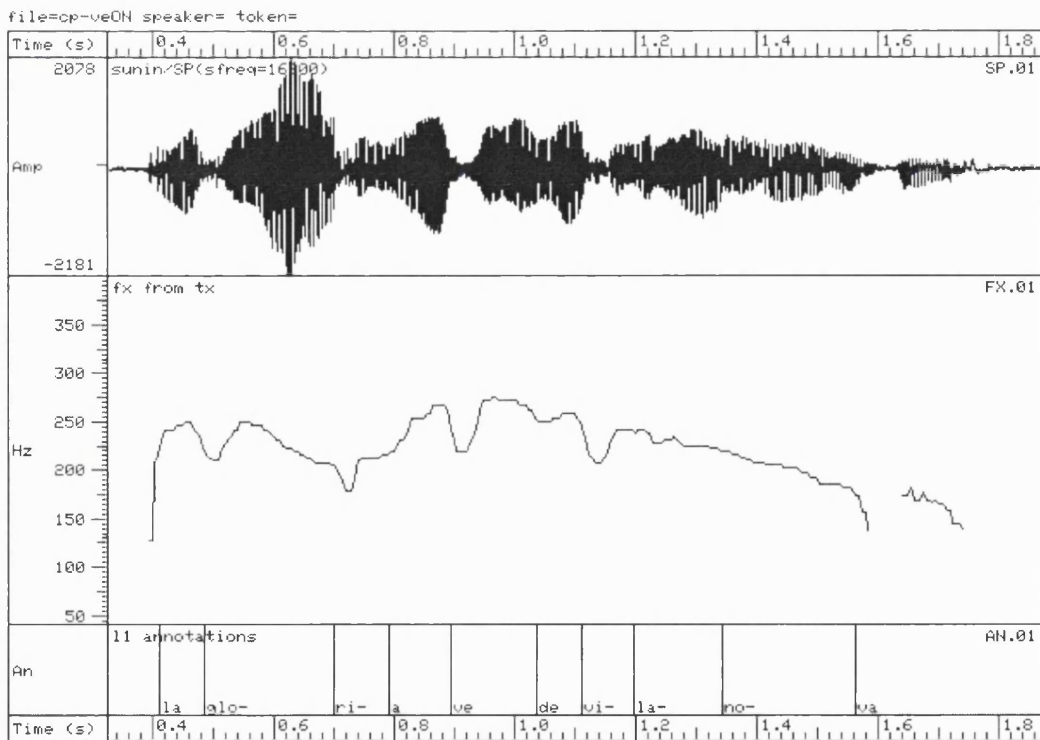
A.20. Figure for example (4.14). *NOU miraven el mar* (speaker DV).



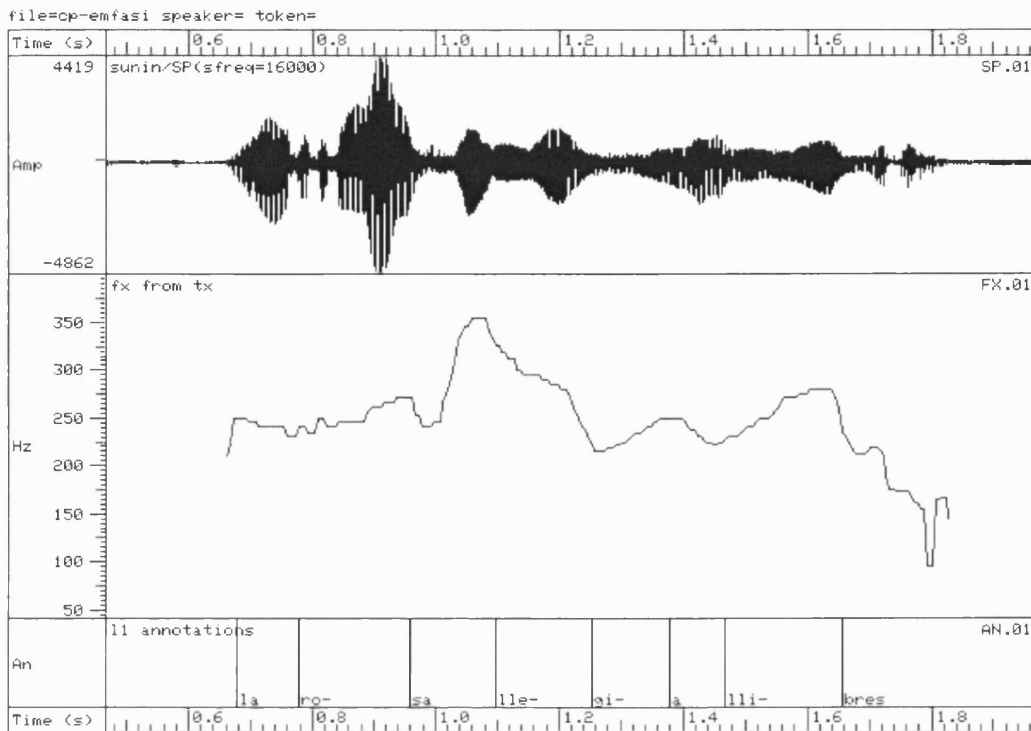
A.21. Figure for example (4.15). *La MARE menja arengades* (speaker DV).



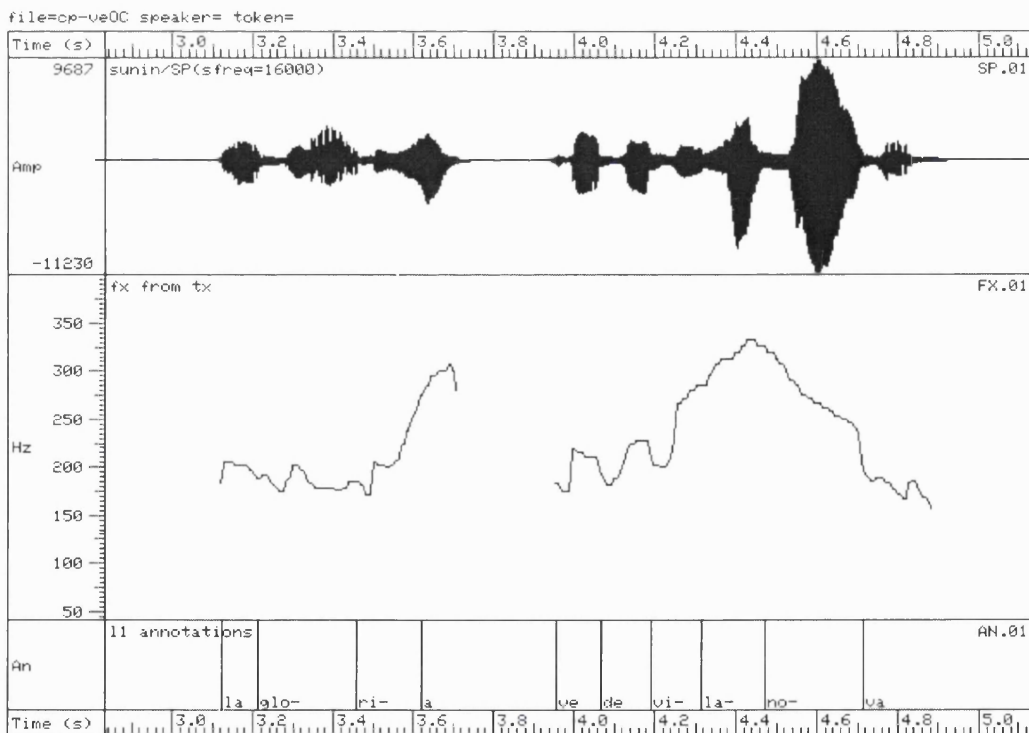
A.22. Figure for example (4.16). *El NÒMADA menjava al carrer* (speaker DV).



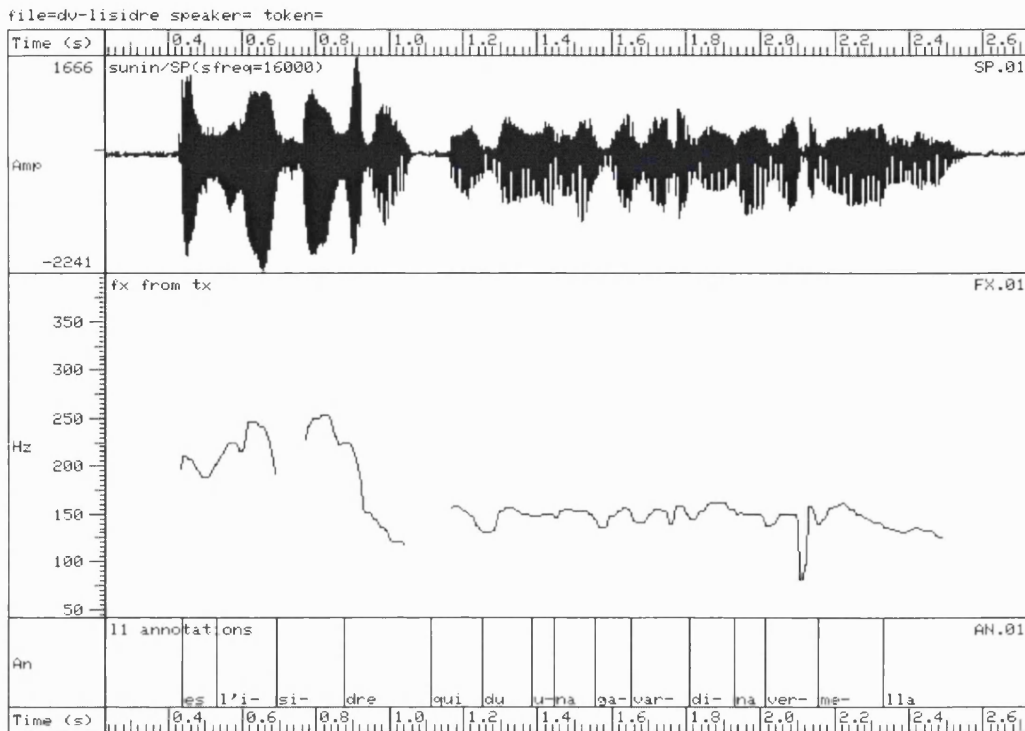
A.23. Figure for example (4.17). *La Glòria ve de VILANOVA* (speaker CP).



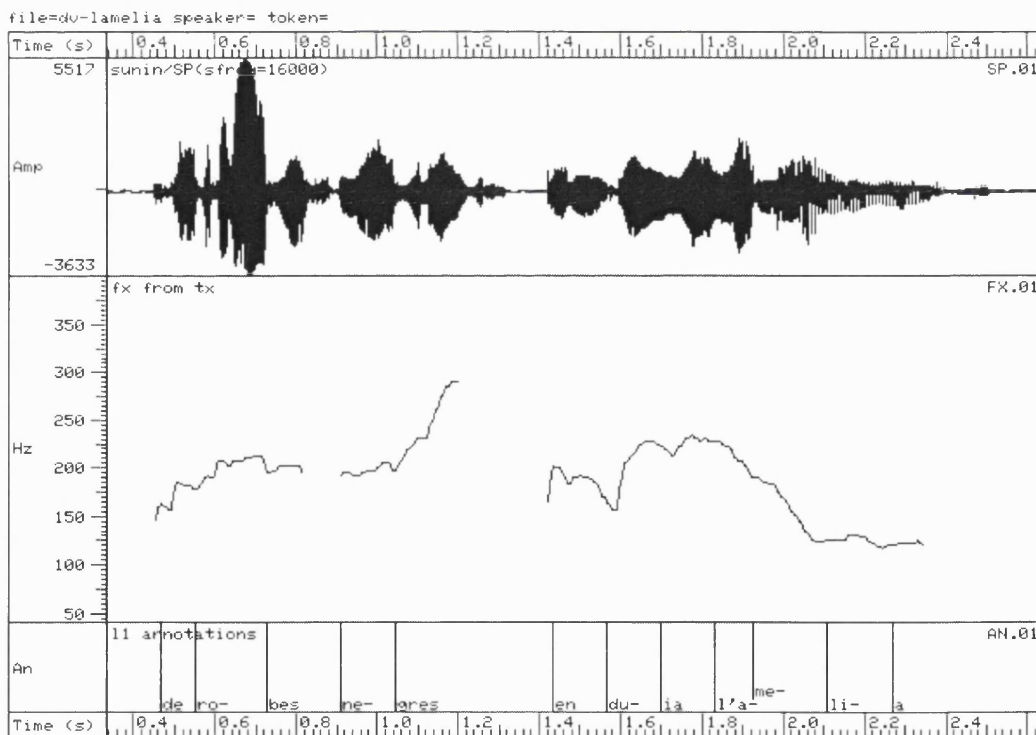
A.24. Figure for example (4.18). *La Rosa llegia LLIBRES* (speaker CP).



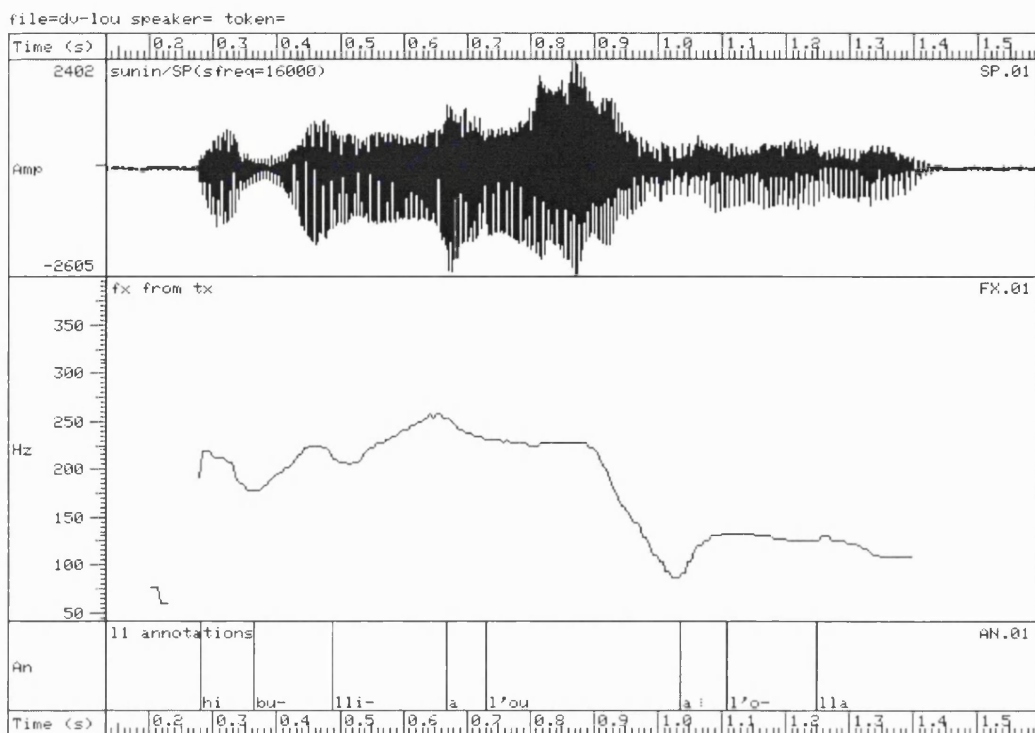
A.25. Figure for example (4.19). *La Glòria ve de VILANOVA* (speaker CP).



A.26. Figure for example (4.28). *És l'ISIDRE, qui du una gavadina vermella* (speaker DV).

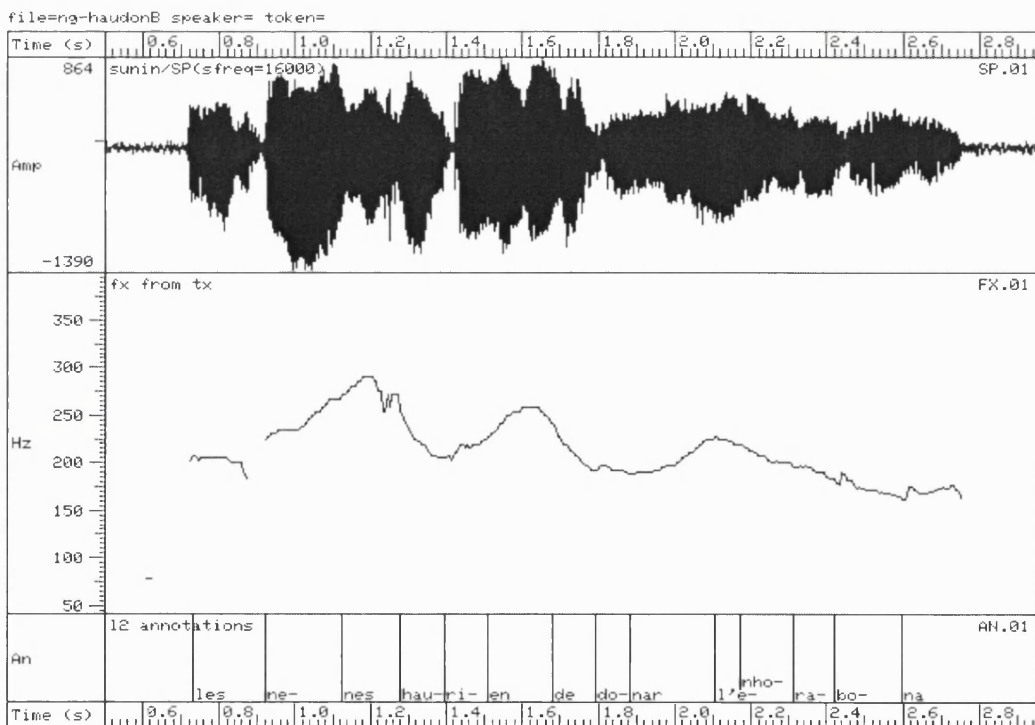


A.27. Figure for example (4.29). *De robes negres, en duia l'AMÈLIA* (speaker DV).



A.28. Figure for example (4.30). *Hi bullia l'OU, a l'olla* (speaker DV).

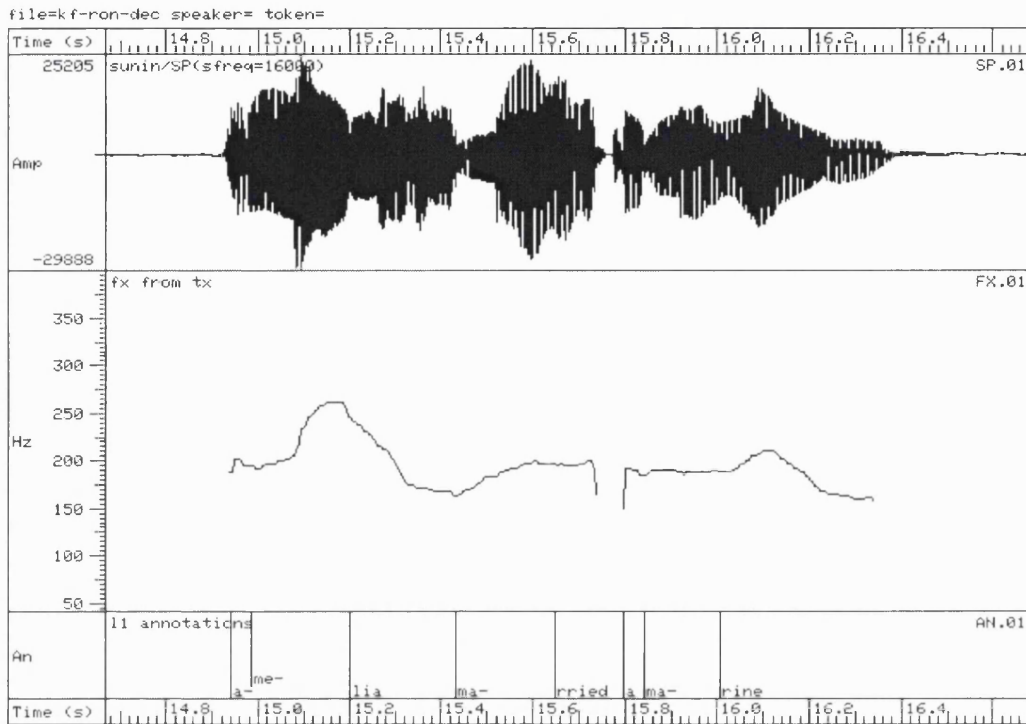
### 5.3 Figures chapter 5



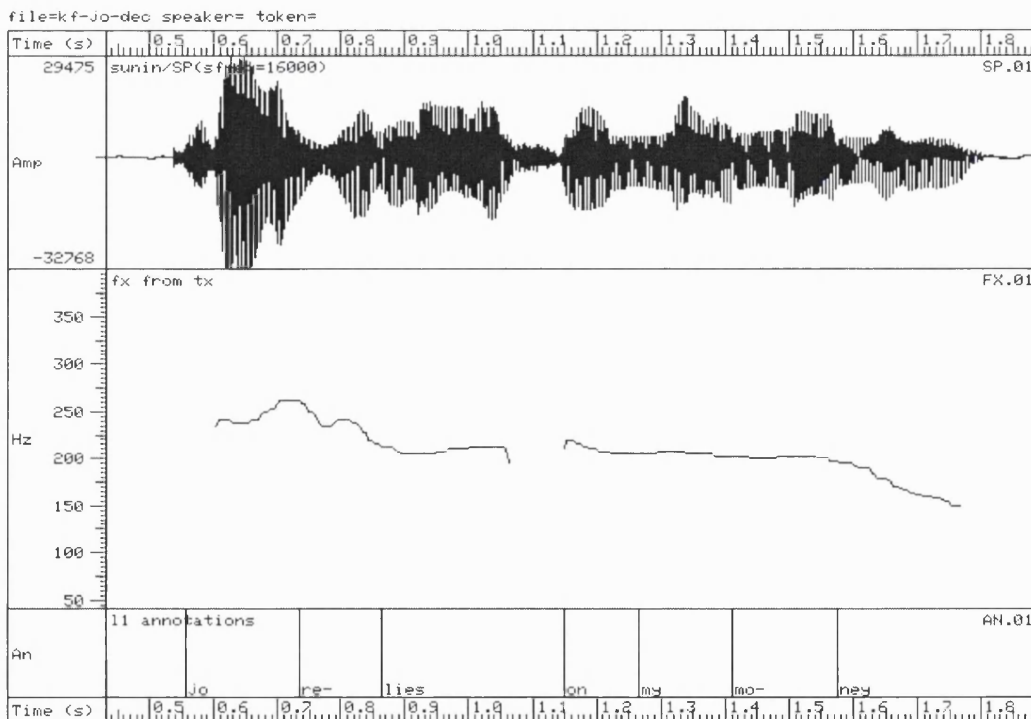
A.29. Figure for example (5.1). *Les nenes haurien de donar l'enhorabona* (speaker NG).



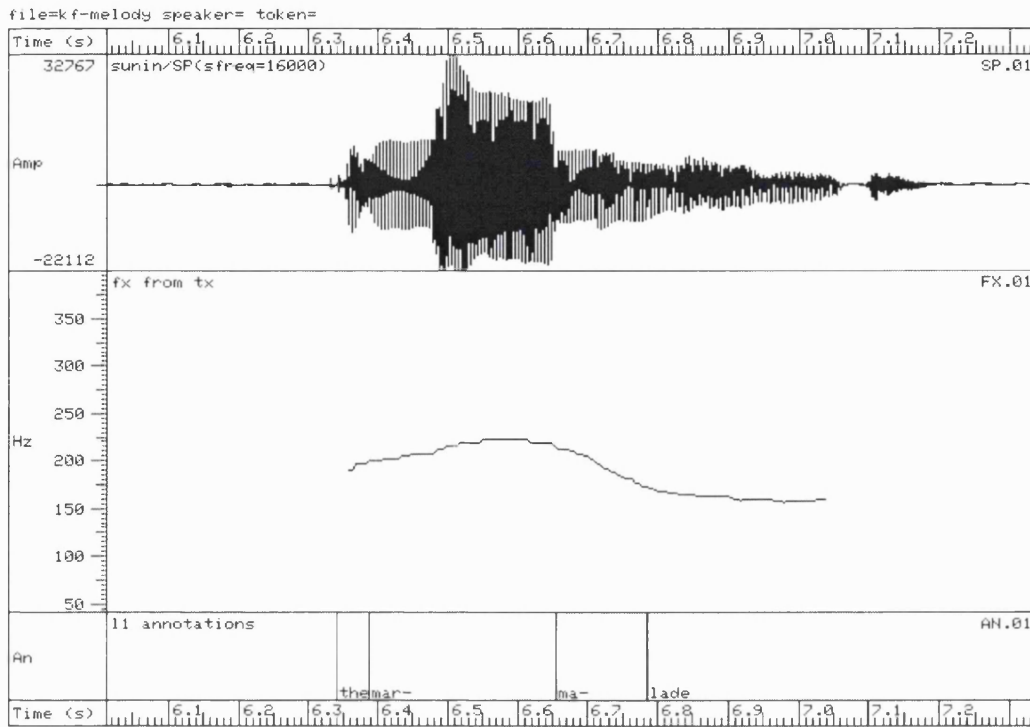
## 5.4 Figures chapter 6



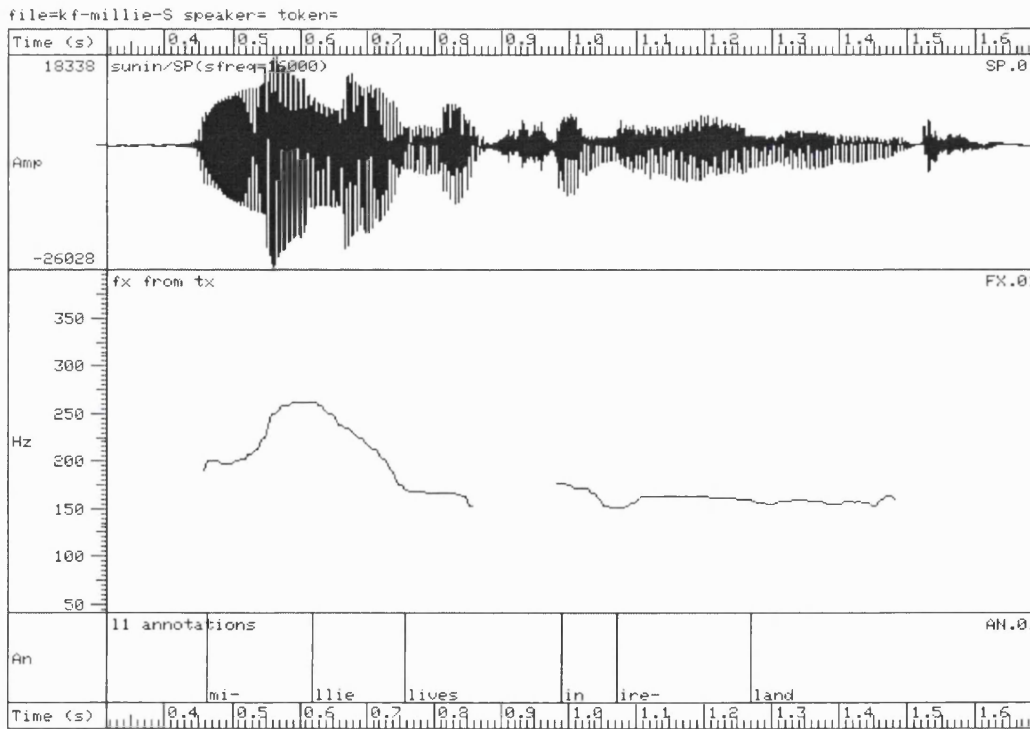
A.30. Figure for example (6.4). *Amelia married a marine* (speaker KF).



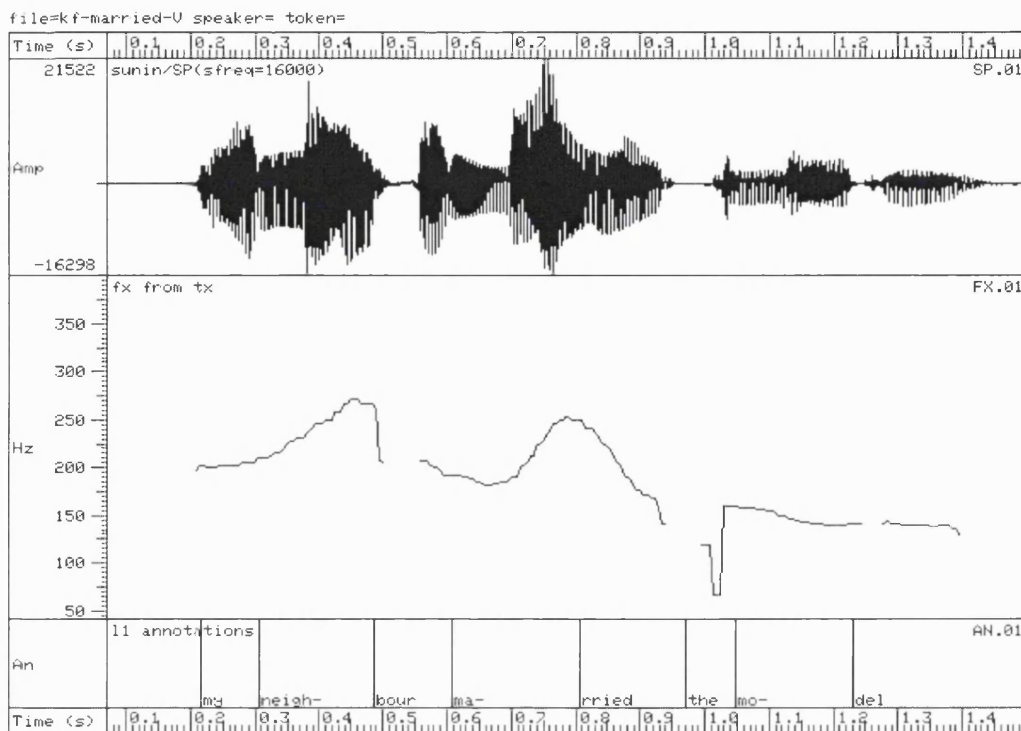
A.31. Figure for example (6.5). *Jo relies on my money* (speaker KF).



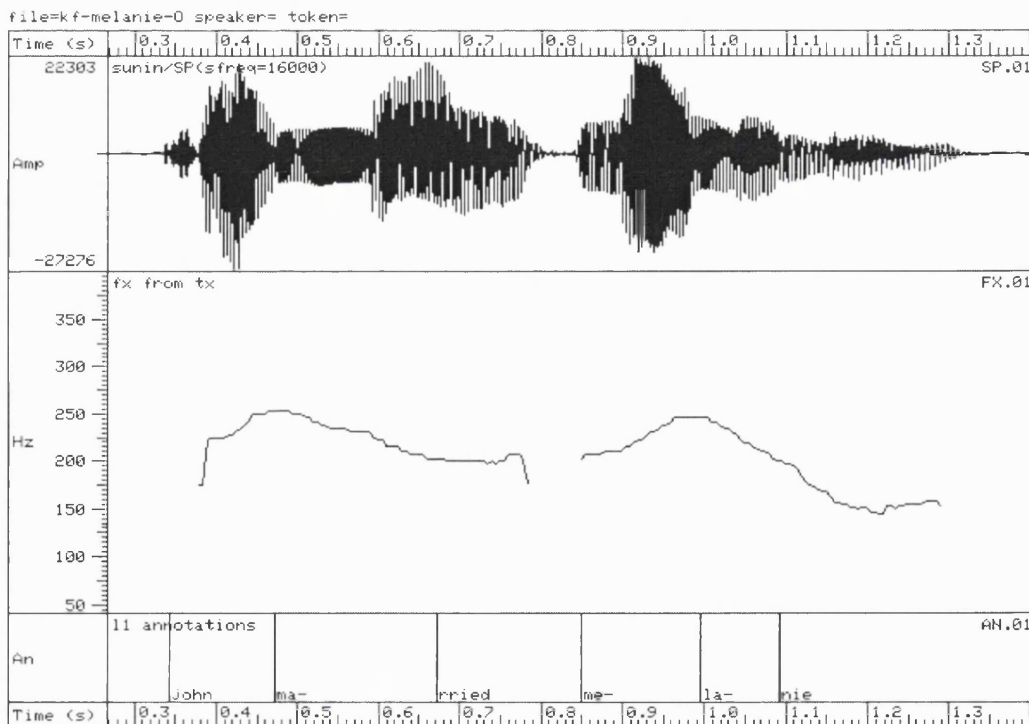
A.32. Figure for example (6.6). *The marmalade* (speaker KF).



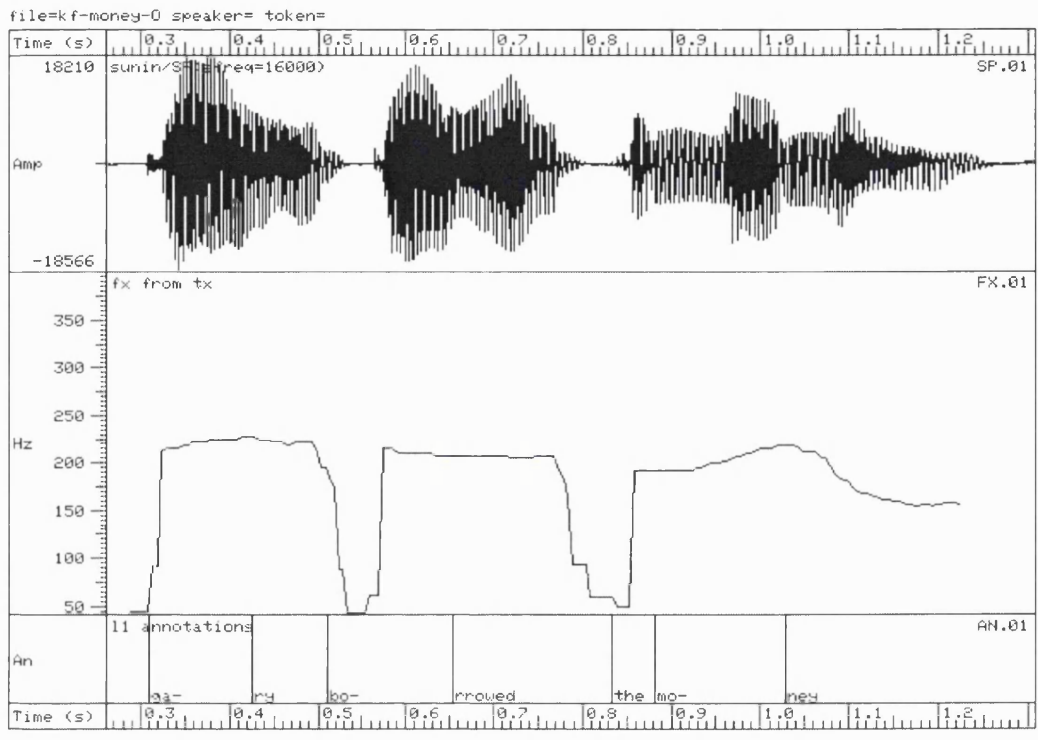
A.33. Figure for example (6.12). *MILLIE lives in Ireland* (speaker KF).



A.34. Figure for example (6.13). *The neighbour MARRIED the model* (speaker KF).



A.35. Figure for example (6.14). *John married MELANIE* (speaker KF).



A.36. Figure for example (6.15). *Gary borrowed the MONEY* (speaker KF).