	Professional	Non-professional
News	Recordings from radio bulletins	Readings of newspaper news Readings of extracts from the radio news
Commentaries	Recordings from TV news	Readings of newspaper commentaries and editorials Readings of extracts from the TV news

Table 2.1. summary of the contents of the collected speech corpus

The collection of this corpus is described in the next section. Its contents were classified as described in section 2.3. The 'non-professional' set was used to build four different sub-corpora:

1) a tonic group sub-corpus
2) an intonation group sub-corpus
3) a sentence sub-corpus
4) a paragraph sub-corpus

The description of these sub-corpora is presented in chapters 4-7. The inventory of elements included in each sub-corpus can be found in the appendices.

2.2. Collection of the speech corpus

2.2.1. Collection of the 'professional' set

The recordings that make up the 'professional' set came from two sources:

a) the recordings of four *Radio Nacional de España* news bulletins, recorded at the radio station studies.

b) the recordings of eight speeches extracted from two television news programs of the *Tele 5* Spanish TV station, recorded on domestic video and re-recorded (only the audio channel) in audio tapes.

The use of both types of sources was due to the fact that the radio news bulletins only contained news readings, but not commentaries. The *Tele 5*

news bulletins included some contributions of invited speakers who commented the news.

The contents of all these recordings were orthographically transcribed in ASCII files.

A more detailed summary of the contents of this 'Professional' set is presented in table 2.2.

	News	Commentaries
Recordings	• Diario de la Tarde RNE (27.2.92) • España a las 8 RNE (28.2.92) • Diario de la Tarde RNE (28.2.92) • España a las 8 RNE (2.3.92)	• A. Ramiro 1 (TELE5) • A. Ramiro 1 (TELE5) • C. Tomás 1 (TELE5) • C. Tomás 2 (TELE5) • F. Jáuregui 1 (TELE5) • F. Jáuregui 2 (TELE5) • J. J. Santos 1 (TELE5) • J. J. Santos 2 (TELE5)
Speakers	44 (23 male/21 female)	4 (3 male/1 female)

Table 2.2. Summary of the contents of the 'Professional' set.

As has been indicated before, this set has not been finally analysed for this work, due to clear differences in the reading style between professional and non-professional speakers found during previous analyses. For this reason, this material in not commented in more detail here, nor have the transcribed texts been included in the appendices.

2.2.2. Collection of the 'non-professional' set

The contents of the 'non-professional' set are summarized in table 2.3.

	News	Commentaries
Recordings	• Noticia EL OBSERVADOR/1 • Noticia EL OBSERVADOR/2 • Noticia EL PAÍS/1 • Noticia EL PAÍS/2 • Noticia LA VANGUARDIA/1 • Noticia LA VANGUARDIA/2 • Noticia radio/1 • Noticia radio/2 • Noticia radio/3 • Noticia radio/4	• Artículo opinión EL OBSERVADOR • Artículo opinión EL PAÍS • Artículo opinión LA VANGUARDIA • Artículo opinión EL OBSERVADOR 2 • Opinión TV/1 • OpiniónTV/2
Speakers	2 (1 male/1 female)	

Table 2.3. Summary of the contents of the 'Non-professional' set.

2.2.2.1. Definition of the reading material

A set of newspaper articles was chosen to be read by the selected speakers. These articles were extracted from three different newspapers published (at that moment) in Barcelona: *El Observador*, *El País*, and *La Vanguardia*. The selected articles referred to facts and events that also appeared in the radio and TV news bulletins. Two different subsets were built, the first one containing the news articles and the second one containing the commentaries articles. The first subset contained 6 different articles, 2 from each newspaper. The second one included 4 articles, 2 short ones extracted from *El Observador*, and other 2, longer than the previous ones, extracted from *La Vanguardia* and *El País*.

In order to allow a direct comparison between 'professional' and 'non-professional' recordings, some transcriptions of the radio and TV bulletins were also included in the reading material. As in the case of the newspaper articles, a 'news' subset and a 'commentaries' subset were organized. The first one included 4 different excerpts from the radio bulletins, and the second one 2 transcriptions of speeches from the *Tele 5* news bulletins.

The text of the articles included in this set can be found in appendix 1.

2.2.2.2. Selection of the speakers

Several conditions were established for the selection of the non-professional set of speakers:

a) language:

- they had to speak Spanish as mother tongue;
- they had to speak the same dialect of Spanish, but not with very marked (segmental or suprasegmental) features;
- they had to use Spanish in most of the communicative situations of their every-day life.

b) age: they had to be in a range between 20-40 years.
c) social level: they had to have an homogeneous, medium-high, cultural background, (secondary school or university studies).
d) reading skill: they had to be able to read a text in a understandable way.

A short selection test was performed in order to determine the more adequate readers for the final recording. 14 different readers were recorded while reading a short article extracted from the selected reading material. The recordings of all the readers were evaluated by three different listeners, who gave a score for each one of them. In this evaluation the ability of the speakers to read in a understandable way and to avoid monotonous reading were mainly considered. The best scored male and female readers were chosen for the recording of the material.

The selected speakers were:

1) AR, male, aged 35, postgraduate student at the *Departament de Filologia Espanyola* of the *Universitat Autònoma de Barcelona*, born in Huelva but having lived from his childhood in Barcelona, Spanish as mother tongue and most used language, and with some experience as an actor.

2) ME, female, aged 29, born in Barcelona, where she has lived since then, Spanish as mother tongue, although uses Catalan quite frequently in her work as a member of the administration staff at the *Facultat de Lletres* of the *Universitat Autònoma*.

Both speak Spanish without (noticeable) dialectal or Catalan interferences, at least from the intonation point of view.

2.2.2.3. Recording procedure

The speech corpus was recorded in the semianecoic room of the Phonetics Laboratory of the *Universitat Autònoma de Barcelona*, both in Revox and cassette tapes. The speakers were told to read as if they were reading aloud the text to a public audience, as in a radio bulletin. For each speaker, the texts were recorded in several sessions. The texts were recorded with no interruptions. If the reading contained too many reading mistakes, speakers were asked to read again the whole text.

2.3. Classification of the material

The goal of this classification task was to obtain a series of 'prosodic databases', which would contain the information about the characteristics of the contents of the collected recordings. Two different analyses were in fact carried out:

a) a first analysis included all the collected material, and aimed at defining the number and the characteristics of the sentences and paragraphs contained in it.
b) a second analysis was only carried out on a group of sentences extracted from the 'non-professional' set, and it was intended to establish the number and characteristics of their tonic groups and intonation groups.

In both cases, prosodic databases were built using the information obtained in these analyses. These two analyses, and the prosodic databases obtained in each case are described in the following subsections.

2.3.1. Paragraphs and sentences

2.3.1.1. Analysis of the speech corpus

The paragraphs and sentences of the corpus were analysed in order to obtain the following information:

a) number of sentences and paragraphs;
b) length of sentences and paragraphs (in syllables);
c) sentence type of the sentences.
d) number of sentences contained in each paragraph

The number of syllables of paragraphs and sentences was obtained by means of an automatic transcription and syllabification system developed at the

Departament de Teoria del Senyal i Comunicacions of the *Escola Tècnica Superior d'Enginyers de Telecomunicacions* of the *Universitat Politècnica de Catalunya.* This means that the result of this counting is the number of theoretical syllables that should appear in the texts, without taking into account possible deletions, mispronunciations or resyllabifications during reading.

The results of this analysis for the 'non-professional' set are summarized in table 2.4.

	'Non-professional' set
News	Paragraphs: 50 (x 2 speakers = 100) Sentences: 112 (x 2 speakers = 224) • Declarative: 112 (x 2 = 224) • 'Yes-No' questions: 0 • 'Wh-' questions: 0 • 'Relative questions': 0 • Exclamative: 1 (x 2 = 2)
Commentaries	Paragraphs: 39 (x 2 speakers = 78) Sentences: 179 (x 2 speakers = 358) • Declarative: 140 (x 2 = 280) • 'Yes-No' questions: 12 (x 2 = 24) • 'Wh-' questions: 20 (x 2 = 40) • 'Relative questions': 4 (x 2 = 8) • Exclamative: 10 (x 2 = 20)

Table 2.4. Summary of the analysis of the paragraphs and sentences in the 'non-professional' set

2.3.1.2. Building the prosodic databases

The collected information was stored in 8 different databases, 4 for the 'professional' set and 4 for the 'non-professional' set. They are listed in table 2.5, as well as the information included in each case.

Units	Databases	Information included for each unit
Paragraphs	• News/Non-professional • News/Professional • Commentaries/Non-professional • Commentaries/Professional	• Orthographic transcription • Location in the texts • Number of syllables • Number of sentences
Sentences	• News/Non-professional • News/Professional • Commentaries/Non-professional • Commentaries/Professional	• Orthographic transcription • Location in the texts • Number of syllables • Sentence type

Table 2.5. Summary of the contents of the paragraph and sentence prosodic databases.

2.3.2. Intonation groups and tonic groups

2.3.2.1. Analysis of the speech corpus

The analysis presented here was performed on the set of sentences selected for the sentence sub-corpus. This corpus contains sentences extracted only from the 'non-professional' material. The complete inventory of these sentences can be found in appendix 6.

In the analysis of the intonation groups the following information was obtained:

1) phonic groups contained in each sentence: as has been indicated in chapter 1, phonic groups have been considered here to be always equivalent to intonation groups; the set of sentences was analysed acoustically and perceptually in order to locate the pauses defining phonic group boundaries.

2) number of syllables in each group: in this case, the counting was carried out by hand from the text.

3) position of the phonic group in the sentence.

4) syntactic boundary appearing at the end of each group.

In the analysis of the tonic groups, the following information was obtained:

1) phonic groups were divided in tonic groups according to the definition established in chapter 1.

2) the number of syllables in each group, obtained by hand as in the case of the phonic groups.

3) the position of the stressed syllable within the group

4) the position of the tonic group within the phonic group

Table 2.6. summarizes the results of this analysis:

		Non-Professional
News	Phonic groups	150
	Stress groups	549
Commentaries	Phonic groups	250
	Stress groups	877

Table 2.6. Summary of the analysis of phonic and tonic groups

2.3.2.2. Building the prosodic databases

As in the case of sentences and paragraphs, several databases (2 for the intonation groups and 2 for the tonic groups) were built using the obtained information. Their contents are summarized in table 2.7.

Units	Databases	Information included for each unit
Intonation groups	• News/Non-professional • Commentaries/Non-professional	• Orthographic text • Location • Number of syllables • Sentence type • Position in the sentence
Tonic groups	• News/Non-professional • Commentaries/Non-professional	• Orthographic text • Location • Number of syllables • Position of the stressed syllable • Sentence type • Position in the intonation group • Position in the sentence

Table 2.7. Summary of the contents of the intonation group and tonic group prosodic databases.

2.4. General summary

In this chapter, the corpus collected for the experimental part of this work has been described. A set of recordings of text readings from professional and non-professional speakers was initially collected to be analysed. However, results of some preliminary analyses showed that the intonation of the professional and non-professional sets were different enough to be considered separately. Only the non-professional set was finally considered for analysis.

This final analysis has been carried out on four different sub-corpora selected from the non-professional material. In these sub-corpora, one for each of the analysis levels defined in chapter 1, the chosen variables have been controlled as much as possible. The description of the definition process and of the contents of each sub-corpus has been left for the chapters describing the analysis of the different levels.

Chapter 3

STYLIZATION: A PHONETIC APPROACH TO THE REPRESENTATION OF INTONATION

The first step in a phonetic study of intonation is the representation of the intonation contours, by means of a transcription or stylization procedure. This chapter describes the procedure used in this work for the stylization of intonation contours. The method presented here offers a representation of the intonation contours similar to the 'close copy stylizations' of the IPO ('t Hart *et al.*, 1990) or the MOMEL system from Aix (Hirst & Espesser, 1991; Hirst *et al.*, 1991), but there are also important differences, as discussed later. The procedure followed here includes two steps, a first one of automatic stylization, and a second one of 'post-processing' of the output of the automatic system. Both steps are described in more detail in the next two sections.

3.1. Stylization of F0 contours

3.1.1. Description of the stylization system

The stylization procedure used for this work has been developed in collaboration with a group of engineers at *Enginyeria La Salle* of the *Universitat Ramon Llull* of Barcelona. It is an automatic system that defines a series of **inflection points** in the F0 contour and calculates the complete stylized contour by linking these inflection points with straight lines. The system uses 'islas melódicas' ('melodic islands'), or portions of an intonation contour between two unvoiced segments, as the domain of the stylization process. For this reason, unvoiced portions of the contour are retained in

the stylized version as 'zero' inflection points. More details about the stylization procedure can be found in Jiménez (1994).

The automatic definition of the inflection points is controlled by the user by means of a **threshold value** ('deviation error percentage') that determines the degree of dissimilarity between the original F0 contour and its corresponding stylized counterpart. This variable also controls the number of inflection points in a contour, in such a way that a lower error threshold means more inflection points, and a more approximate copy of the original contour; inversely, a higher error threshold calculates contours with less inflection points but less similar to the original one.

The stylization system has been integrated in **'Pitch'**, a Windows application that allows the acoustic and perceptual analysis of intonation. The program includes, among others, the following functions:

1) **signal recording/playing**: using a SoundBlaster™ card, it is possible to record and play utterances.
2) **F0 tracking**: the tool includes a F0 tracking system, based on the SIFT algorithm (Markel, 1972).
3) **F0 stylization**
4) **contour edition:** the tool allows correction of the detection errors by using the mouse and to edit the values of the stylized contours to modify their shape or to create new intonation patterns.
5) **energy calculation**: the system includes an energy normalization procedure.
6) **LPC synthesis**: it is also possible to synthesize utterances with original and stylized F0 contours using an LPC procedure.

A more detailed description of the program is given in Martínez (1995).

Figure 3.1. offers an example of the representations obtained by means of 'Pitch', displaying the speech wave of a sample utterance, its original intonation contour, and the corresponding stylized version superimposed to it.
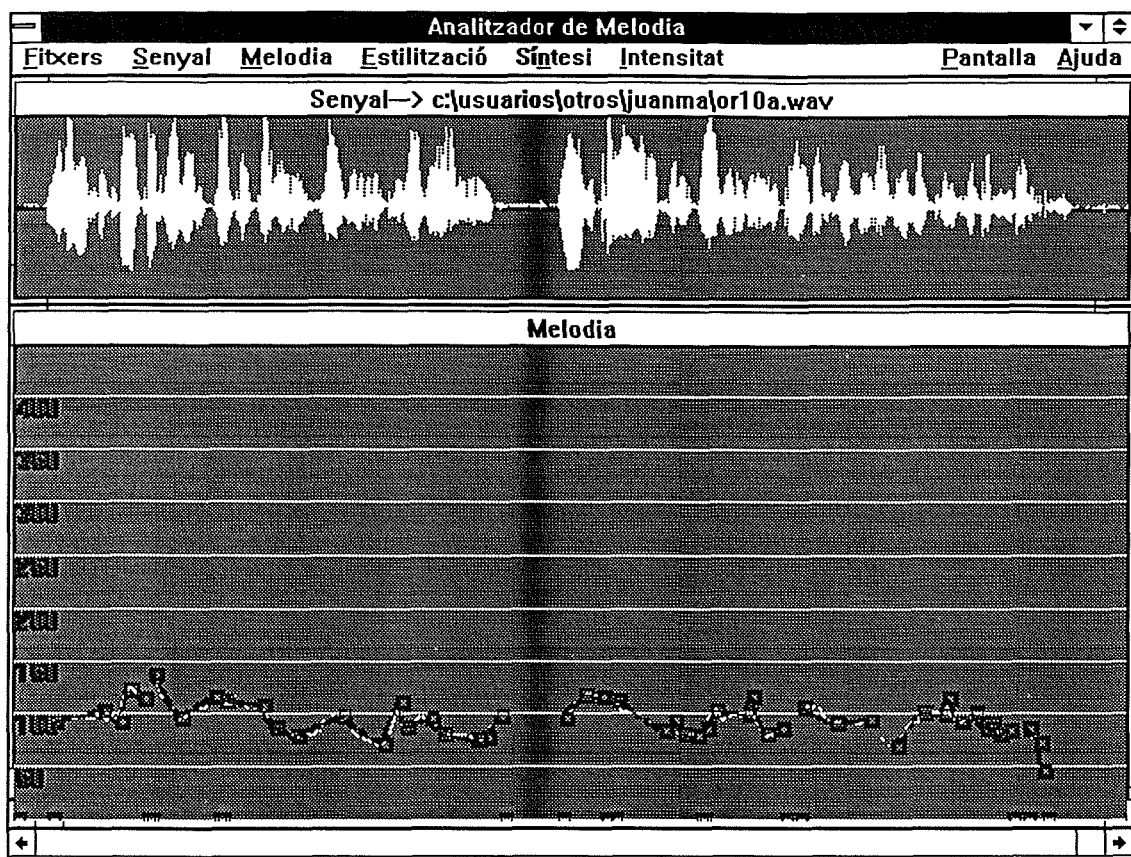
Figure 3.1. Example of original F0 contour and its corresponding stylized version obtained with 'Pitch'

3.1.2. The concept of 'perceptually equivalent stylizations'

This automatic stylization system can then produce different stylized versions of the same original contours, depending on the value of the defined threshold. The next step is to determine the most adequate threshold value for the calculation of the stylized contours of this analysis.

It was established that the stylized contours obtained with the chosen threshold value had to meet simultaneously two conditions:

a) they had to be perceptually equivalent to the original version, and
b) they had to contain the minimal possible number of inflection points

The stylized contours defined in this way are called henceforth **'perceptually equivalent stylizations'**.

3.1.2.1. Definition of the threshold

To establish the defined threshold value, a series of perceptual experiments were carried out. The results of two different tests are presented in the following sections.

3.1.2.1.1. Jiménez (1994)

In this work, a test was designed to determine the degree of perceptual equivalence between an original contour and several versions of contours stylized with a different threshold value (henceforth, T) to the original. The goal was to establish the highest value for this factor that would retain the relevant perceptual information.

Five different values for T were considered in the test: 4, 8, 12, 16 and 20. Six different sentences were used, taken from the speech material analysed in Garrido (1991). These sentences were between 5 and 7 syllables long, represented different sentence types (declarative, Yes-No question, Wh-question, 'echo-question', exclamative and imperative), and all of them contained only voiced segments. The 6 sentences had been uttered by the same male speaker. The recordings of these 6 sentences where used to generate, using the application described above, synthesized versions of each one with the original F0 contour and the 5 different versions of the corresponding stylized contours. The test stimuli were created by combining pairs for each sentence the synthesized version with the original F0 contour and one of the 5 different 'stylized' versions. This gave a total of 30 test pairs.

The pairs were presented grouped by sentence, but the order of presentation within each group was random. The whole test was presented twice, each one with a different presentation order (Order 1 and Order 2).

The subjects of the test were 21 undergraduate and postgraduate students of Spanish Philology at the *Universitat Autònoma de Barcelona*. The subjects were asked to judge the perceptual equivalence of the intonation contours in both utterances of each pair, by answering the following question:

> "¿*Las frases que componen cada una de las parejas son percibidas con la misma entonación? No se pretende evaluar si la entonación es correcta sino únicamente la diferencia entre ellas*"

All the subjects of the test, students of Spanish Philology, had some knowledge of the notion of intonation.

The total number of answers for each of the five defined values of T was 252 (6 sentences x 21 listeners x 2 series).

The results are presented in table 3.1, and plotted in figure 3.2.

	Answers 'different'									
	T=4		T=8		T=12		T=16		T=20	
	N	%	N	%	N	%	N	%	N	%
Order 1	13	10,32	15	11,9	18	14,29	30	23,81	42	33,33
Order 2	19	15,08	15	11,9	23	18,25	26	20,63	62	49,21
Total	32	12,69	30	11,90	41	16,27	56	22,22	104	41,27

Table 3.1. Results of the Jiménez (1994) test
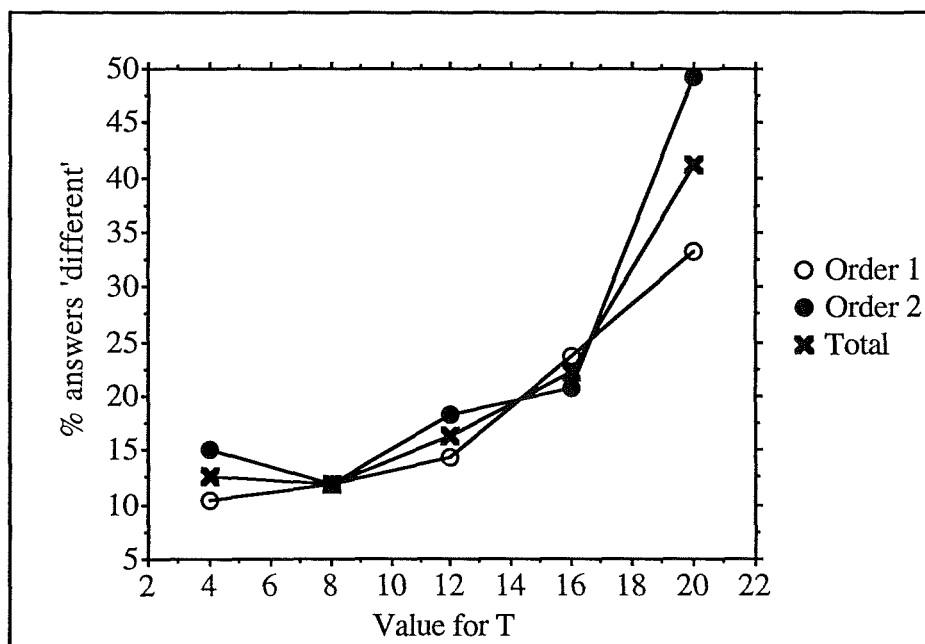
Figure 3.2. Graphical representation of the results of the Jiménez (1994) test

According to these results, the optimal value for T as threshold value to obtain 'perceptually equivalent contours' was set to 8. Two facts were considered for the choice of this value:

a) The percentage of answers 'different' in the case of T=8 is the lowest of all the considered conditions, even lower than in the case of T=4. It can be considered then that '8' is the optimal value to obtain 'perceptually equivalent contours' containing the minimum number of inflection points.

b) The obtained results were compared with those presented in De Pijper (1983) for a similar test with the 'close-copy stylization' technique. De Pijper accepts as 'perceptually equivalent' sentences with a 13.2% of

responses 'different'. The results obtained for T = 8 are already below the percentage used as threshold by De Pijper.

3.1.2.1.2. Estruch *et al.* (1995)

The goal of this second test was to check whether the stylized contours obtained with T=8 and the original ones were really perceived as equal or different, even with sentences of a different language, Catalan in this case.

The test stimuli were 40 pairs of Catalan utterances, obtained from a previous analysis comparing the intonation of one-phonic-group sentences in Catalan and Spanish (Estruch *et al.*, 1994). These stimuli were synthesized using the same analysis and synthesis tool as in the previous case, and distributed in the following way:

   a) 29 pairs containing versions of the same sentence with the original intonation contour, in one case, and with stylized contour (T=8) in the other;
   b) 2 pairs with two versions of the same sentence and the stylized contour;
   c) 2 pairs containing two versions of the same sentence with its original contour;
   d) 11 pairs containing versions of the same sentence with two different stylized contours, and representing different sentence types.

A tape with these stimuli was prepared and presented to 37 students of the *Universitat Autònoma de Barcelona*. The subjects were asked to decide, as in the previous experiment, if the intonation contour in both components of the pair was equal or different. The subjects were encouraged to answer in all cases.

The results of this test are summarized in the following table:

	% 'equal'	% 'different'	% no answer
original-stylized	94.03	5.55	0.42
original-original	97.29	1.35	1.35
stylized-stylized	94.59	5.4	0
different contours	21,23	78.37	0.38

Table 3.2. Results of the Estruch *et al.* (1995) test

It may be concluded from these results that the selected stylization procedure gives stylized contours that are considered by the listeners as (almost) equivalent to the original ones.

3.1.2.2. 'Perceptually equivalent contours' and other stylization procedures

The system presented here is similar to other procedures, mainly in the way that they decompose intonation contours into a series of target points linked by lines, and that the stylized result is intended to be perceptually indistinguishable from the original. IPO 'close-copy' stylizations and the Aix MOMEL system are the most known examples of this type of stylization procedures.

3.1.2.2.1. 'Close-copy stylizations'

Some differences can be pointed out between the system defined here and the IPO 'close-copy stylization':

a) 'Close-copy stylizations' have to be obtained by hand after several steps of edition, resynthesis and perceptual comparison between the original and the stylized versions. The stylization procedure used here, however, is automatic, and it runs in a single step.

b) 'Close-copy stylizations' do not retain unvoiced segments in the resulting stylization contour. These segments, however, are present in the contours obtained with the system used here.

c) The type of perceptual constraint established in this stylization system is different from the one used in the IPO stylization procedure. In the 'close-copy stylization', the resulting stylized contours contain **only** the perceptually relevant inflection points, because the stylization procedure is based on perceptual criteria. The stylization procedure defined here, however, ensures that the obtained 'perceptually equivalent contours' retain all the perceptually relevant information contained in the corresponding original contour. However, due to the fact that the stylization procedure is based on acoustic criteria, it does not guarantee that the number of obtained inflection points are exactly the minimal ones necessary to obtain a perceptually equivalent contour. In other words, it is possible that some non-relevant inflection points can be also found in the 'perceptually equivalent contours' obtained with this procedure. In fact, the analysis of the stylized contours obtained in this way has revealed that some variations related to micromelodic phenomena, that are not perceptually relevant, are still retained. As is described in the next section, these movements were detected manually and excluded from the next steps of the analysis.

3.1.2.2.2. MOMEL

MOMEL and the system presented here have in common that they both obtain the stylized contours in an automatic way. Accordingly, they both need some type of post-processing to correct the errors of the automatic procedure. However, there are also some significant differences:

a) In the MOMEL system, target points are linked with parabolas instead of straight lines, as is the case of the IPO 'close-copies' and the procedure presented here.

b) As in the case of the IPO 'close-copies', unvoiced segments are not retained in a stylized contour obtained by means of MOMEL. In the case of the stylization procedure presented here, they are still present in the resulting stylized contour.

3.1.3. Stylization procedure

The material selected for the analysis at each level was processed using the defined stylization procedure. Target sentences were digitized at 10 KHz using a MacADIOS™ card and the MacSpeech Lab II™ speech analysis software for Macintosh (GW Instruments, 1988). The 'Pitch' program was used for the F0 estimation and stylization tasks. The threshold value was set to 8, according to the results of the tests described above. Figures 3.3 and 3.4 present the representation obtained for the same utterance after F0 estimation (figure 3.3) and stylization (figure 3.4) by means of the 'Pitch' program.
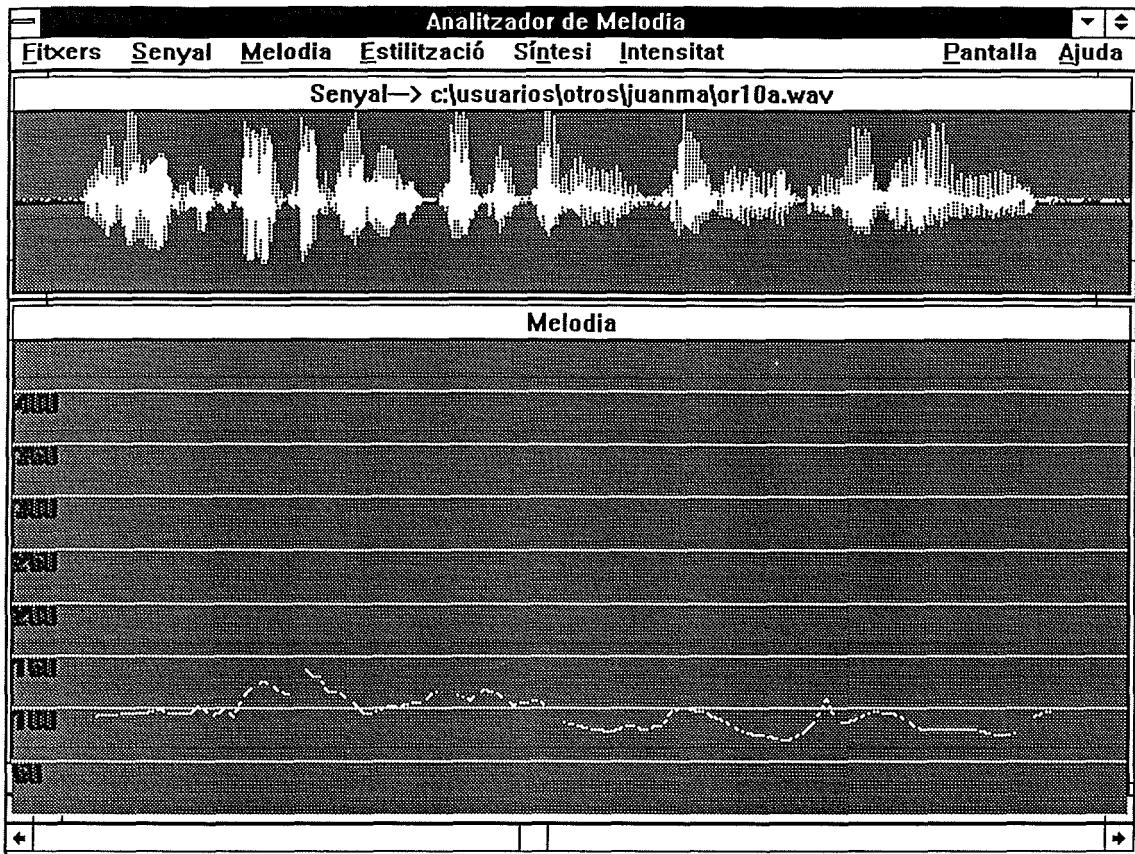
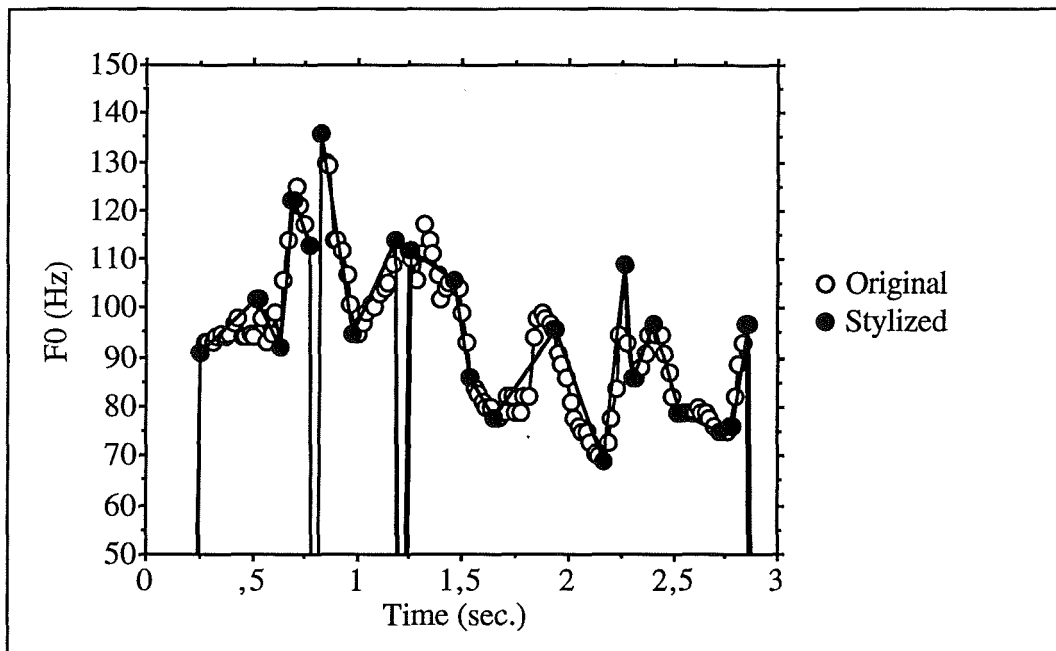Figure 3.3. Waveform and original intonation contour of the phonic group
'*La organización terrorista ha protagonizado en lo que va de año*' (speaker AR),
obtained with 'Pitch' application.
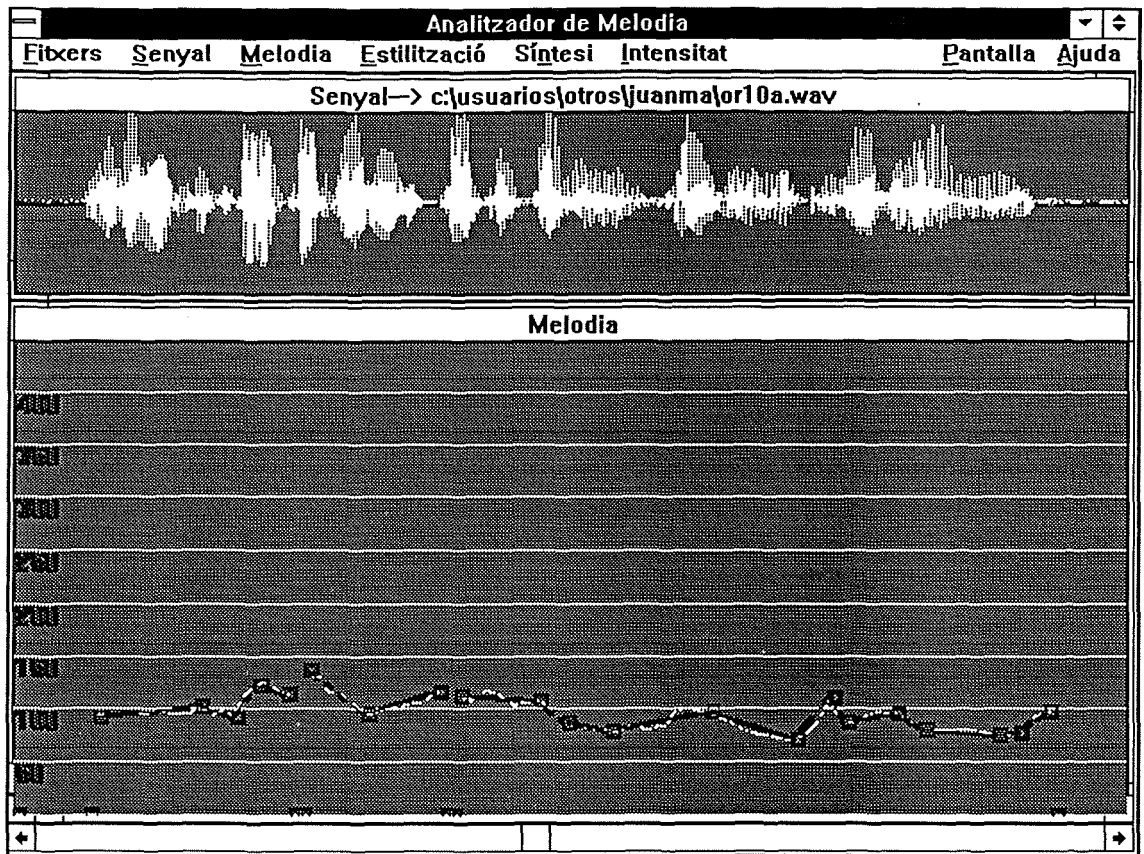
Figure 3.4. Waveform, original intonation contour and stylized contour (superimposed to the original contour) of the phonic group '*La organización terrorista ha protagonizado en lo que va de año*' (speaker AR), obtained with 'Pitch' application.

The values corresponding to the different inflection points (time and frequency) of the resulting stylized contours were first stored in ASCII files, and then converted to Statview™ format for further processing. Statview™ is a data analysis and presentation program for Macintosh that has been used in this work for the treatment and display of the obtained stylized contours. An example of representation of an intonation contour obtained with this program is included in figure 3.5.
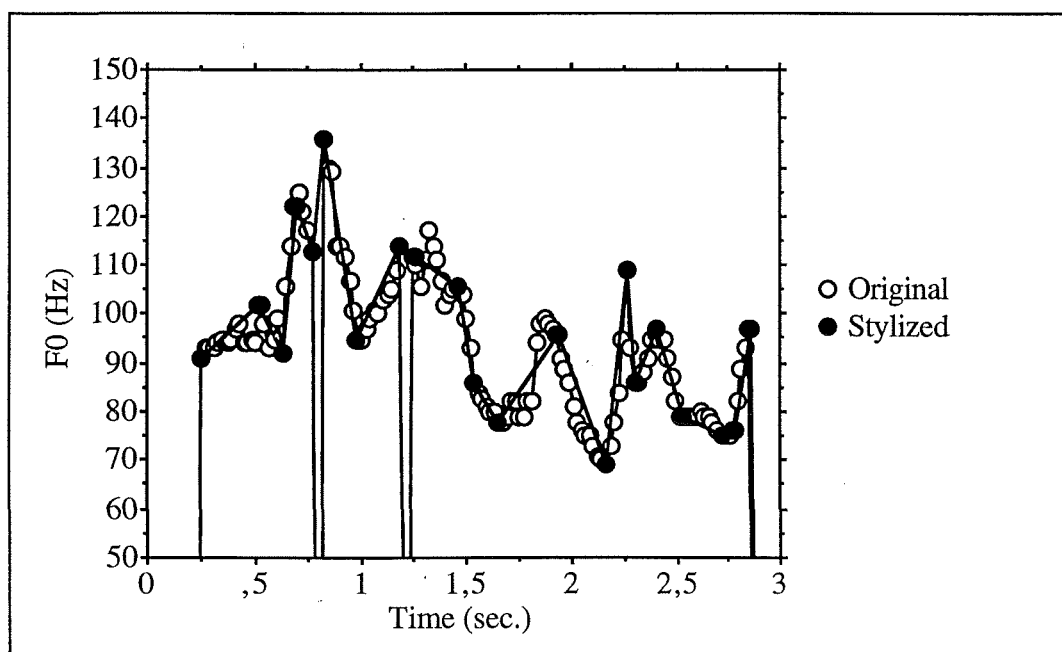
Figure 3.5. Original and stylized contour of the phonic group 'La organización terrorista ha protagonizado en lo que va de año' (speaker AR) displayed by means of Statview™.

3.2. 'Post-processing' of the stylized contours

The 'perceptually equivalent contours' obtained using the procedure described above still retained some unnecessary information:

a) errors of the pitch tracker, which triggered 'spurious' inflection points in the stylized version;

b) inflection points related to micromelodic variations.

The resulting stylized contours had then to be 'post-processed' in order to recognize the errors and mark the variations. The characteristics of this 'post-processing' step are described in the next sections.

3.2.1. Fixing errors of the pitch tracker

The output of any available pitch tracker presents some estimation errors (Rabiner et al., 1976; Hess, 1983). It is necessary then to check manually its output in order to detect them. In this work, fixing of pitch estimation errors was carried out in two different tasks:

a) the first one, previous to the stylization procedure, was carried out using 'Pitch' edition facilities; clear errors (for example, points at which the F0 value was half or the double of the value of the surrounding

points) were removed from the F0 contour before the calculation of the 'perceptually equivalent contour'.

b) a the second one, after stylization, compared the stylized contours with the output of a different pitch tracker (the Rabiner-Schafer autocorrelation procedure included in the SoundScope™ software package; see GW Instruments, 1992, for details) in order to find out less evident errors that were not fixed in the first step. In this case, the inflection points related to these errors were marked and removed.

3.2.2. Avoiding unvoiced 'zeroes'

As mentioned before, the stylization procedure considers as 'zeroes' those portions of the intonation contour that do not show periodic vibration of the vocal folds, as in the case of the unvoiced segments. Then, inflection points with 'zero' values appear in the resulting stylized contours. These inflection points have been also eliminated in the 'post-processed' version of the contours, in such a way that stylized contours are then a continuous line from the beginning to the end of a phonic group. Figure 3.6 shows the same stylized contour of figure 3.5 after applying this 'zero deletion' process.

Figure 3.6. Original and stylized contour of the phonic group 'La organización terrorista ha protagonizado en lo que va de año' (speaker AR) after 'zero' deleting.

3.2.3. Identifying micromelodic inflections

The analysis of the inflection points in the contours related to the influence of the segments that make up the utterance is outside the scope of this work (see section 1.1.). For this reason, these inflection points have been labelled as 'micromelodic' in order to exclude them from the analysis.

To detect micromelodic variations, the stylized contour was compared with the intonation contour obtained using the F0 tracker of SoundScope™. The inflection points that appeared to be related to micromelodic movements were deleted from the stylized contour.

Three different types of micromelodic variations have been found in the analysed stylized contours. They are described briefly in the following sections, but a detailed analysis is postponed for a further study.

3.2.3.1. Micromelodic variations at the beginning and the end of the utterances

One of the types of micromelodic variations detected in the analysed material are the lowering and rising movements appearing at the beginning or end of a phonic group. Figures 3.7 and 3.8 show two examples of these movements. The first one is an example of an inflection point lower than the expected at the beginning of the utterance. The second one shows the presence of some inflection points higher than the preceding ones at the end of the utterance. These variations may be caused by alterations in the normal vibration of the vocal folds due to the beginning or ending of their vibrating period.
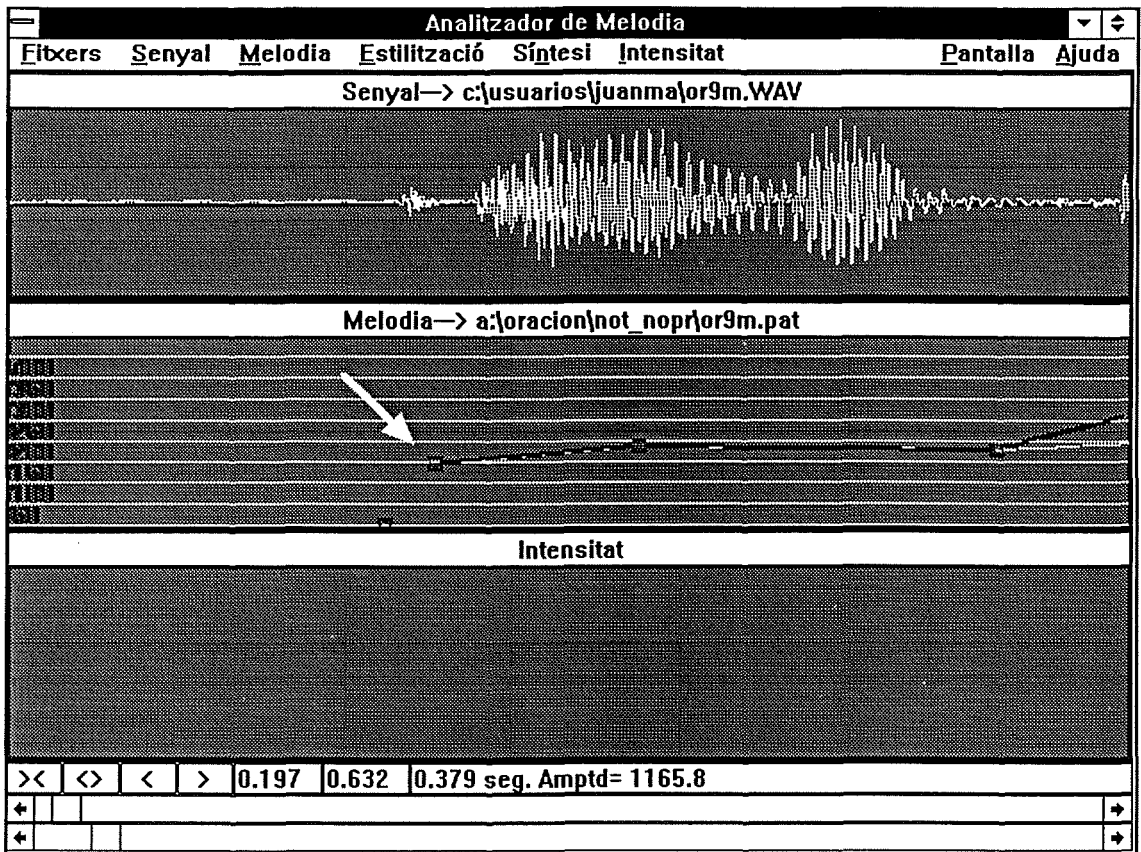
Figure 3.7. Initial part of the stylized contour corresponding to the phonic group '*El documento reitera*' (Or. 9/Mar., 'news' set). The arrow indicates the initial inflection point lower than the usual due to the micromelodic influence.
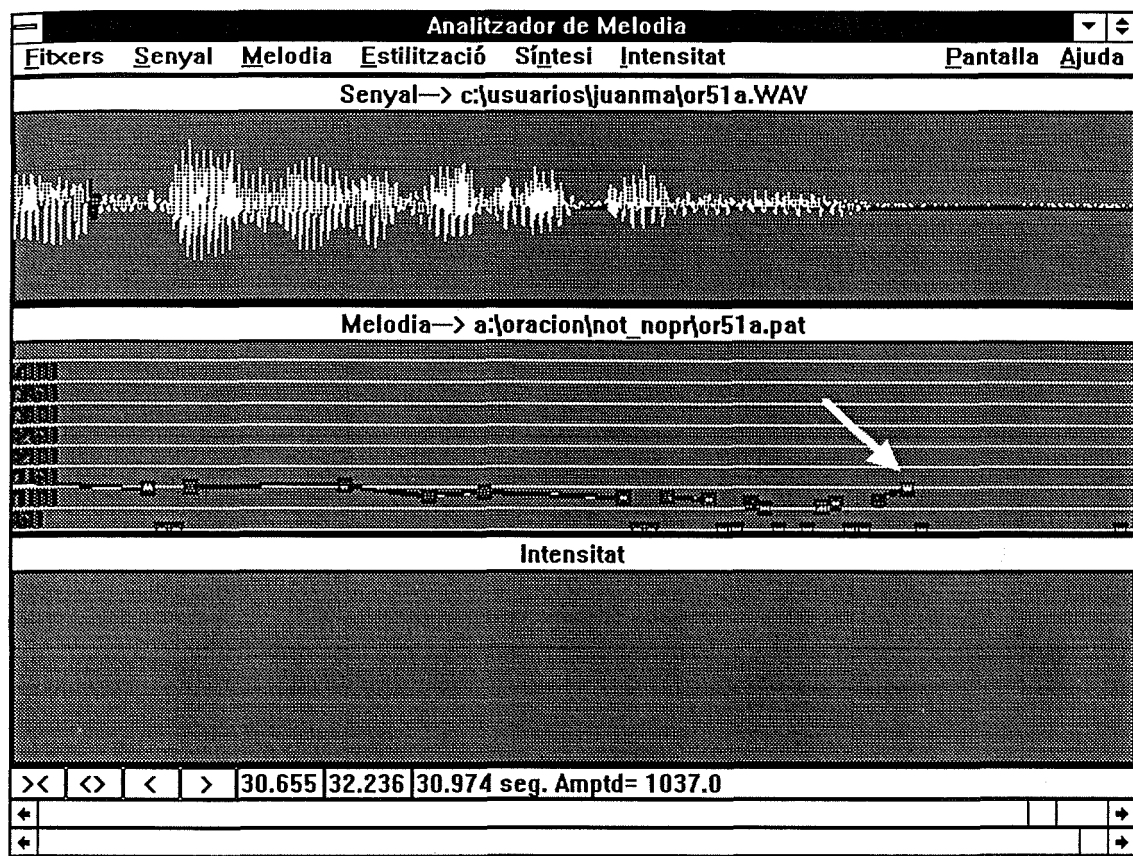
Figure 3.8. Final part of the stylized contour corresponding to the phonic group '*que deberá desembocar en un acuerdo de libre cambio.*' (Or. 51/Ant., 'news' set). The arrow indicates the final inflection point higher than the usual due to the micromelodic influence.

3.2.3.2. Micromelodic variations in the vicinity of unvoiced consonants

The beginning or end of voicing in the vicinity of an unvoiced consonant is also a place where inflection points related to micromelodic influence have been detected. For Spanish, Gili (1924) stated that the beginning of vowels following an unvoiced consonant normally have a slightly lower F0 compared to the center of the vowel: "*las cuerdas vocales al empezar a vibrar de nuevo producen para vencer su inercia un sonido más grave, que puede afectar aún a las vocales acentuadas*" (Gili, 1924, p. 177). These lowering movements are also reported by Navarro (1944): "*Las consonantes sordas y en especial las oclusivas, al interrumpir la línea de entonación, ocasionan cierto descenso en los puntos de contactos con las vocales contiguas*". (Navarro, 1944, p. 21).

Figures 3.9 and 3.10 are examples of 'micromelodic' inflection points before and after an unvoiced consonant in the analysed corpus. Figure 3.9 shows an example of a rising contour at the end of the last vowel [e] of the word '*instituciones*', due to the influence of the final unvoiced fricative [s]. Figure 3.10 shows an example of inflection point at the beginning of a

vowel which follows an unvoiced stop, in this case the [t] in the word 'entre'. As in this case, these inflection points usually present an F0 level higher than excepted, due to the influence of the consonant.
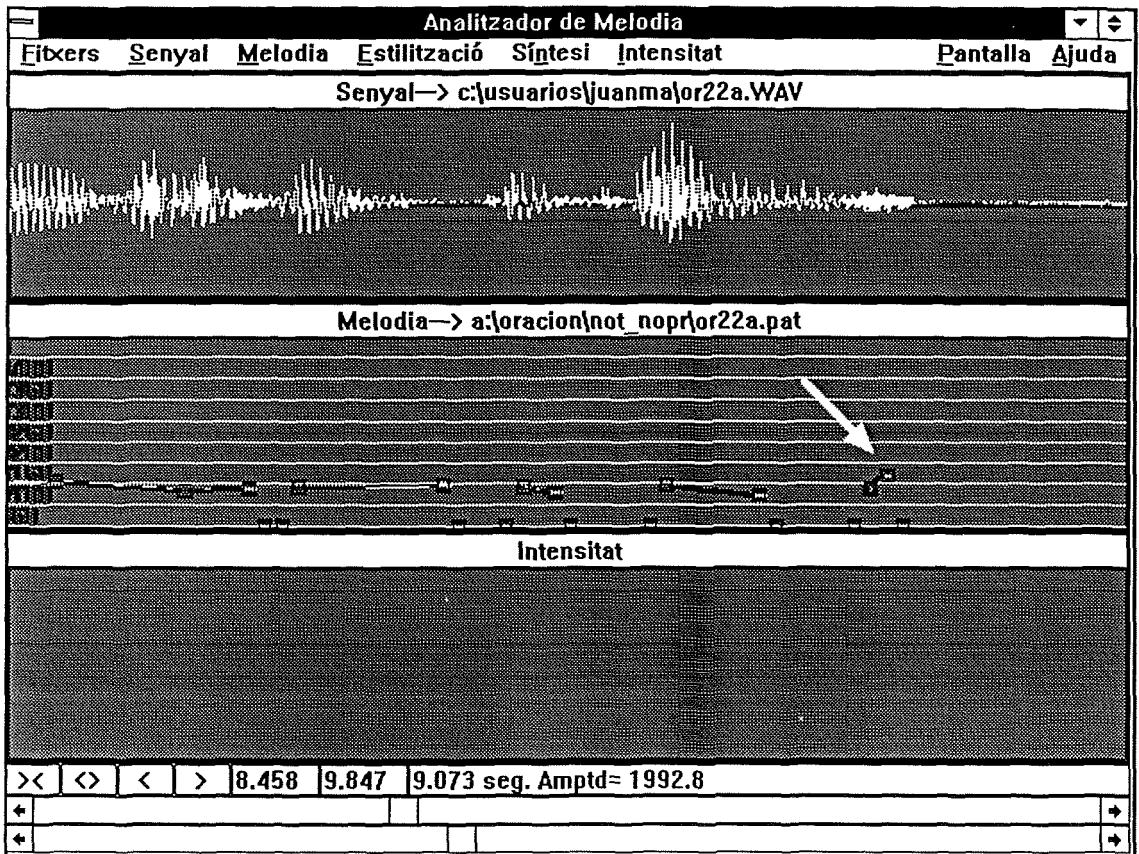

Figure 3.9. Final part of the stylized contour corresponding to the phonic group '*la deslegitimación de las instituciones*,' (Or. 22 Ant., 'news' set). The arrow indicates the inflection point affected by the influence of the final [s] of the word '*instituciones*'.
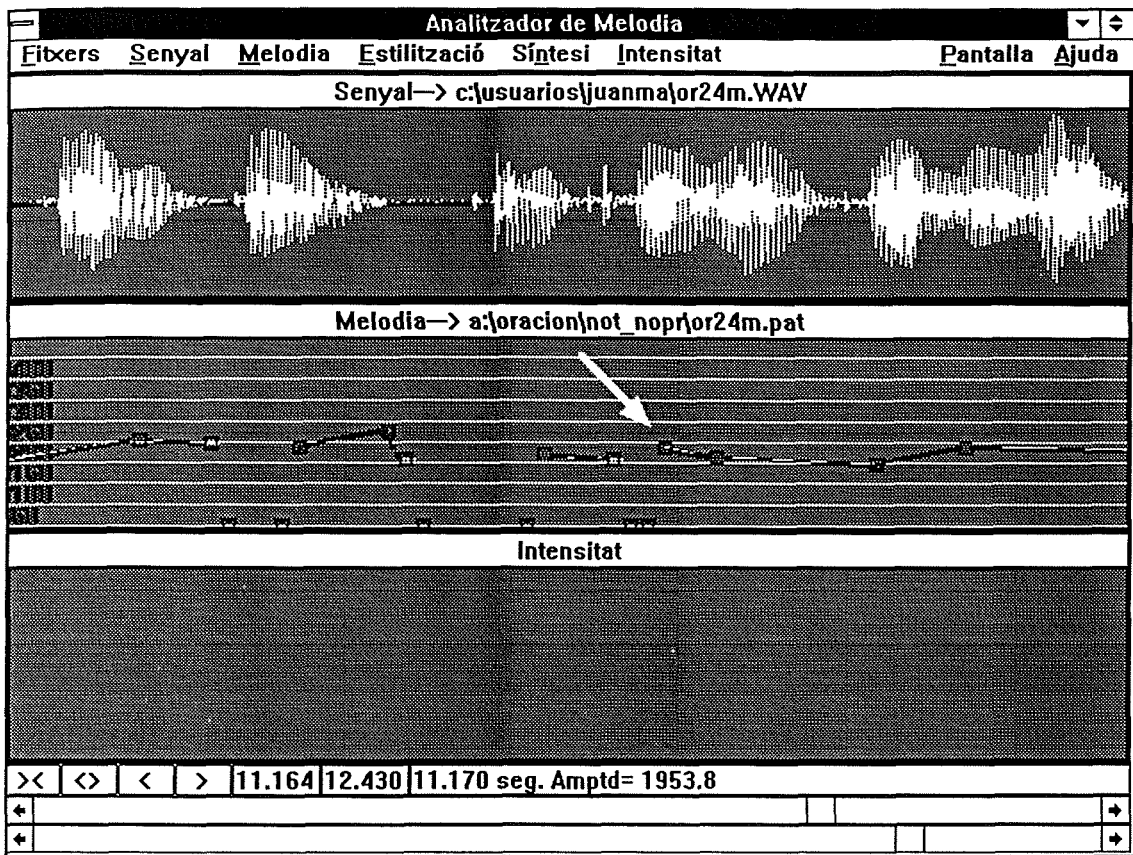
Figure 3.10. Partial view of the stylized contour corresponding to the phonic groups '*y la relación inexistente*' and '*entre González y Aznar*' (Or. 24 Mar., 'news' set). The arrow indicates the inflection point affected by the presence of [t] in the word '*entre*'

3.2.3.3. Micromelodic variations in voiced consonants

The lowering of intonation contours due to the presence of voiced consonants has been analysed already for languages other than Spanish in House & Fairbanks (1953), Léon & Martin (1970), Lea (1973), Boe (1973), Larreur & Boe (1973), and Di Cristo (1982). Gili (1924) reported this lowering effect for Spanish: "*En muchos casos la consonante sonora se nos presenta más baja que la vocal con que forma sílaba, con una diferencia de uno a dos semitonos, por lo común, y aun en algún caso extremo, de cuatro semitonos*" (p. 176). Navarro (1944) also mentions this phenomenon, but, as in the case of Gili, he does not describe it in detail: "*En las oclusivas sonoras b,d,g, la depresión del tono suele ser mayor que en sus fricativas correspondientes b,d,g.*" (Navarro, 1944, p. 21).

In general, these falls in the contours are not retained by the stylization procedure. In some cases, however, they are deep enough to create an inflection point, as in the case of figure 3.11. This example shows an inflection point which is slightly lower than the adjacent one, that coincides with the lateral consonant [l] in the sequence '*con el objeto*'.
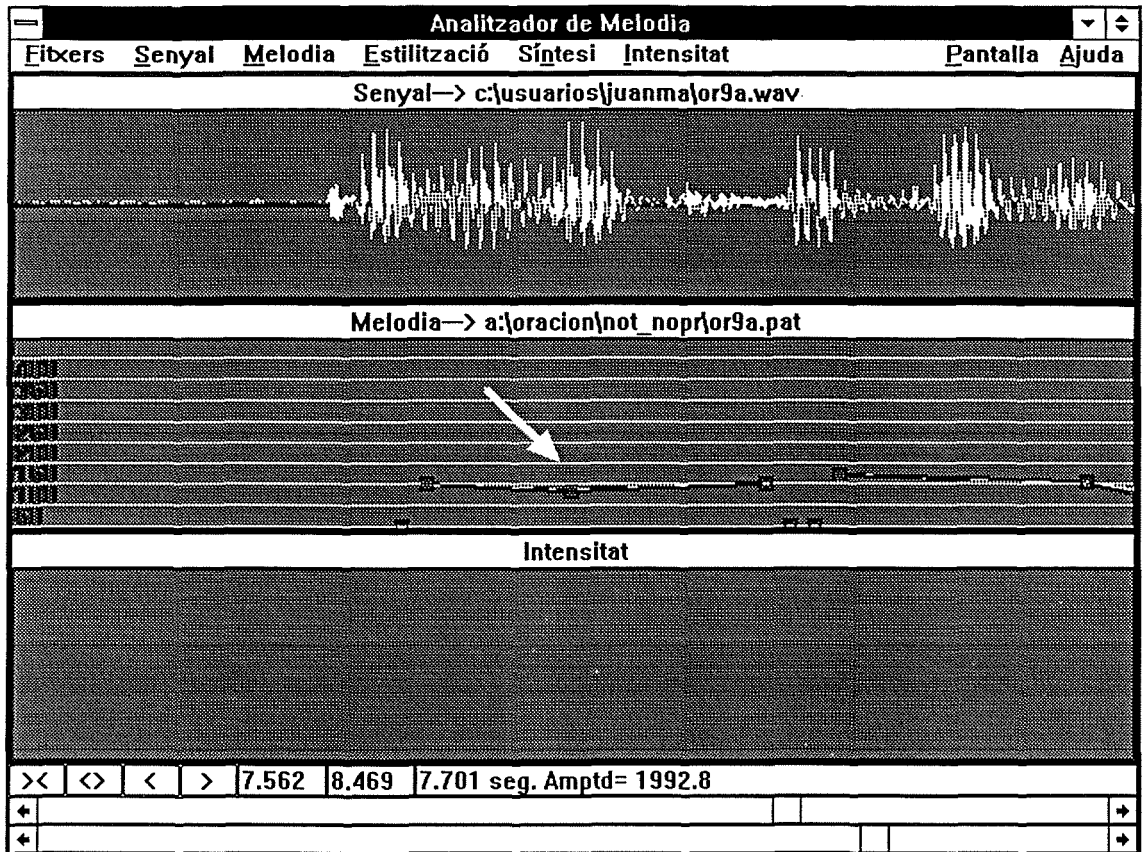
Figure 3.11. Initial part of the stylized contour corresponding to the phonic group 'con el objeto de evitar los atentados planeados por ETA.' (Or. 9/Ant., 'news' set). The arrow indicates the inflection point affected by the presence of the [l] in 'con el objeto.'

3.2.4. The representation of the post-processed stylized contours

Final post-processed stylized contours do not contain inflection points related to the presence of pitch estimation errors, unvoiced segments, or micromelodic variations. To obtain graphical representations of these post-processed contours, the Statview™ software was used. An example of post-processed contour obtained with this program is presented in figure 3.12.
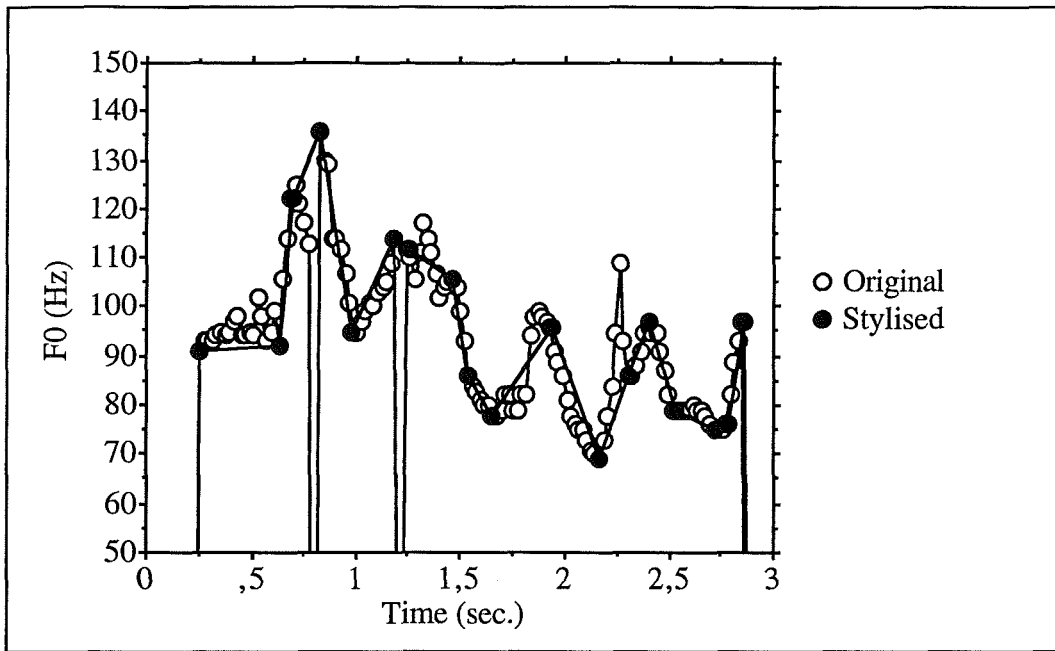
Figure 3.12. Original and stylized contour of the phonic group '*La organización terrorista ha protagonizado en lo que va de año*' (speaker AR) after deleting micromelodic variations.

3.3. General summary

In this chapter, the procedure used in this work to simplify the shape of the original intonation contours has been described. By this method, the original intonation contours have been converted to a series of inflection points linked by straight lines. This procedure consists of two different steps, a first one of automatic stylization of the contours, and a second one of post-processing of the stylized contours.

The **stylization** process attempts to obtain a first stylized representation of the contour perceptually indistinguishable from the original. This procedure is carried out automatically by means of the 'Pitch' program. The stylization procedure includes four different tasks:

1) F0 estimation.
2) Detection and correction, if possible, of the F0 estimation errors.
3) Calculation of the stylized contour.
4) storage in ASCII files.

After the stylization procedure, however, the stylized contours still retain some inflection points that can be considered as irrelevant for the analysis. These 'unnecessary' inflection points are excluded from the contour with a **post-processing** procedure. Post-processing is carried out by hand, using Statview™ and SoundScope™. The post-processing procedure is also divided in four different tasks:

1) Conversion of the ASCII files to Statview™ format.
2) Detection of F0 estimation errors, and deletion of the associated inflection points.
3) 'Zero' deletion.
4) Deletion of micromelodic variations.

The contour representations obtained by this method have been used for the analyses presented in the following chapters.

Chapter 4

SPANISH INTONATION CONTOURS
AT THE TONIC GROUP LEVEL

In this chapter the results of the analysis of intonation movements at the tonic group level are presented.

The review of the Spanish intonation analyses presented in section 1.2 has shown that the description of local intonation movements in Spanish is far from being complete. Only partial data is available, and no formal descriptions have been given. In this study, a formal description of these local movements in Spanish has been attempted.

Following the methodology proposed in chapter 1, the first task consists on the definition of a framework for the description of the contours at tonic group level. The proposed framework is similar to the one used by the IPO approach for languages such as Dutch and English, as described later.

Using this framework, the evolution of the intonation contours along a set of tonic groups has been described. The 'tonic group' has been the chosen domain for the analysis of local movements in this work. However, as has also been discussed in chapter 1, the tonic group is a rather controversial intonation unit in Spanish. Defined by Navarro as *"la parte del discurso que tiene por base prosódica un sólo acento espiratorio y por contenido ideológico un núcleo de significación no susceptible de divisiones más pequeñas"* (Navarro, 1944, p. 29), he does not consider that this domain is an intonation unit in Spanish. It has been used in later studies (Alcina & Blecua, 1975; Fant, 1984; López, 1993), but no arguments justifying its use have been given. In other descriptions, as Canellada & Kuhlmann (1987) or Quilis (1981, 1993) the tonic group is not considered. For these reasons, it

must be pointed out that this unit has been chosen here only as an 'analysis domain', with no 'a priori' assumptions about its validity as an intonation unit in Spanish. This validity is discussed here, in the light of the data obtained.

The results of this description have also been used to define a set of typical concatenations of movements (or 'patterns') appearing in the analysed corpus. Finally, the relationship between these patterns and different variables has been analysed.

According to this, the goals of this study can be summarized in the following way:

1) to propose and apply a formal framework for the description of the local movements.
2) to define intonation patterns for local movements in Spanish using this framework.
3) to analyse the validity of the tonic group as an intonation domain in Spanish.
4) to study the relationship between the defined patterns and three linguistic phenomena that can determine their shape and use: stress, sentence type and syntactic structure.

The following sections describe the procedure and results of this study. First, the tonic groups sub-corpus is described (section 4.1). Section 4.2 presents the proposed descriptive framework. Section 4.3 describes the analysis procedure of the tonic group contours. Questions related to the consideration of the tonic group as an intonation unit in Spanish are reviewed in section 4.4. Section 4.5. presents the proposed intonation movements and their main features. Section 4.6 deals with the problem of the relationship between the appearance of the different patterns and phonetic factors, such as the length of the tonic groups. Finally, section 4.7 analyses some aspects related to the influence of linguistic variables, as stress, sentence type and syntactic structure, in the use of these patterns. Section 4.8 summarizes the results and conclusions of this chapter.

4.1. Definition of the tonic groups sub-corpus

The analysed sub-corpus is a set of tonic groups chosen from the material described in chapter 2. The procedure of definition and selection of this sub-corpus is described in the following sections.

4.1.1. Variables

As outlined in chapter 2, several variables have been taken into account for the definition of the tonic group sub-corpus. These variables are:

1) Position of the tonic group within the phonic group (initial, medial, final and initial-final):

Navarro's (1944) description distinguishes three parts in an intonation group: initial, medial, and final parts. According to his description, the location of the group implies a different shape of the contour of the group. This distinction of three types of local movements according to their location in the utterance is also present, in fact, in the IPO model (see section 1.3.5.3). According to this variable, four different types of tonic groups have been considered, following Fant (1984): initial (at the beginning of a phonic group), final (at the end of a phonic group), medial (neither in initial nor in final position), and initial-final (tonic group forming a single phonic group).

2) Number of syllables in the tonic group:

As stated in chapter 1, classical descriptions of Spanish local movements have generally focused their attention in the location of the peaks, but very few data can be found about the shape of the contours between peaks. Accordingly, if the complete shape of the movements is going to be described, the number of syllables in a tonic group becomes an important variable to consider. Tonic groups with different number of syllables were then included in the sub-corpus.

3) Position of the stressed syllable within the group:

It has been reported in the literature that the position of the stressed syllable tends to determine the location of the peaks in the intonation contours (Navarro, 1944, p. 21 & ff., for example; see section 1.2.3.1 for details). Tonic groups with the stressed syllable in different positions within the group have been included in the sub-corpus to analyse this phenomenon.

4) Position of the tonic group within the sentence (initial or final):

It was discussed before (section 1.2.3.3, for example) that, in Spanish and other languages, special local movements at the beginning or end of sentences can be used to indicate, for example, sentence type. For this reason, examples of sentence-initial and sentence-final tonic groups have been included in the sub-corpus.

5) **Type of the sentence containing the group:**

As has been pointed out in chapter 1, the literature has reported that sentence type can determine the shape of local movements at the beginning or end of the sentences. Accordingly, tonic groups extracted from sentences with different sentence types have been included.

4.1.2. Selection and organization of the material

The first task in the selection and organization of the sub-corpus was to establish the range of values allowed for each of the five variables described in section 4.1.1. To do this, some analyses had to be performed on the corpus material. For each variable, the final ranges of values were the following:

1) 'Position of the tonic group in the phonic group'.

Four positions were considered: initial, medial, final and initial-final.

2) 'Number of syllables'.

To determine the possible values of this variable, the range of number of syllables in the analysed set of tonic groups was obtained. It was observed that there were tonic groups containing from 1 to 8 syllables. Examples of these 8 types of groups were included in the sub-corpus.

3) 'Position of the stressed syllable'.

The analysis of the tonic groups revealed that most groups had their stressed syllable in the last and penultimate position, and in some cases, in the syllable before the penultimate. Only a few examples of stressed with the main stress placed before these positions were found. However, all of them have been included in this analysis.

4) 'Position of the tonic group within the sentence'.

In this case, three positions have been considered: initial, medial and final.

5) 'Sentence type'.

Four different types have been included in the analysis: declarative, yes-no questions, wh- questions, and exclamative sentences.

The sub-corpus was built considering the combination of all the values for each of the variables. Unfortunately, it was not possible to find out examples of all possible combinations, but when possible, two tokens of each case were chosen. Each tonic group was analysed for the two recorded speakers. From the 'News' subset, 57 groups were selected, and 148 in the case of the 'Commentaries' subset. This gave a total of (57+148)*2 speakers = 410 analysed tonic groups.

A complete inventory of the tonic groups included in this sub-corpus, and two tables summarizing its contents can be found in appendix 2.

4.2. Definition of the framework

As has been pointed out in chapter 1, intonation studies have used two types of frameworks for the description of local intonational movements: levels and contours. In this case, a contour approach has been preferred. The method presented here is very close to the IPO method.

4.2.1. P, M and V levels

As explained in chapter 3, the intonation contours obtained after stylization and post-processing may be viewed as series of inflection points linked by straight lines disregarding micromelodic influence. An example of this type of resulting representation can be seen in Figure 4.1.
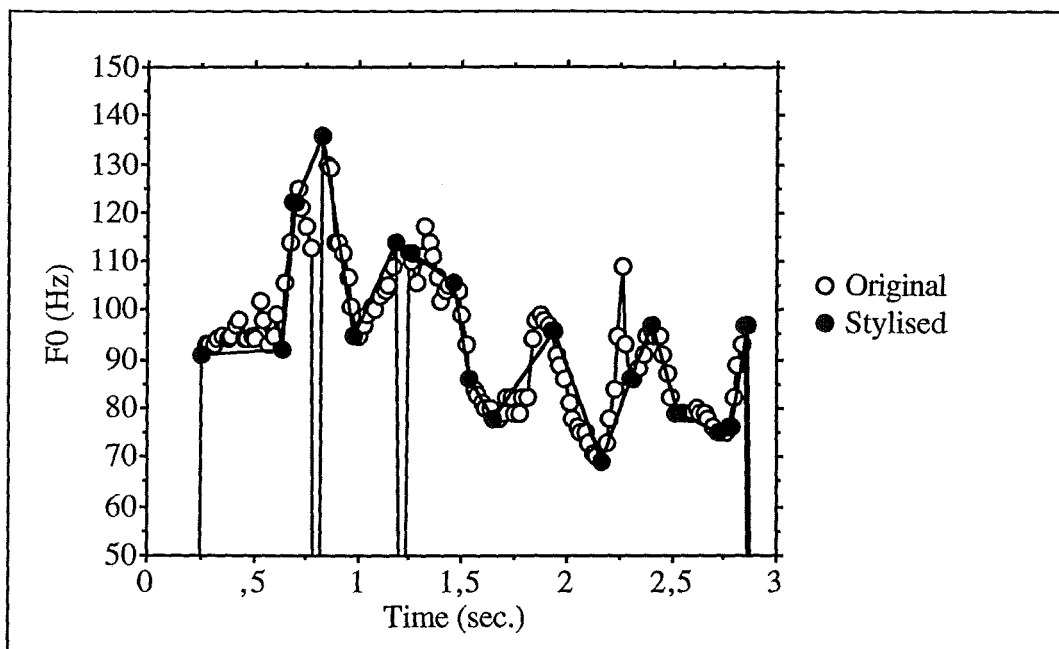
Figure 4.1. Original contour and stylized and post-processed representation (dark line) of the contour in the phonic group "*La organización terrorista ha protagonizado en lo que va de año*", read by speaker AR.

Visual inspection of the resulting representations allows to hypothesize the existence of three different levels in the intonation contours: a high one, labelled as P (Peak), that would correspond to the inflection points coinciding with 'full' accent peaks; a low one, labelled as V (Valley), that would be identified with those inflection points appearing at full lowerings or minima in the intonation contour; and a medium one, labelled as M (Mid), for those inflection points coinciding with 'half' accent peaks or half lowerings. Examples of each of these levels can be found in Figure 4.2.
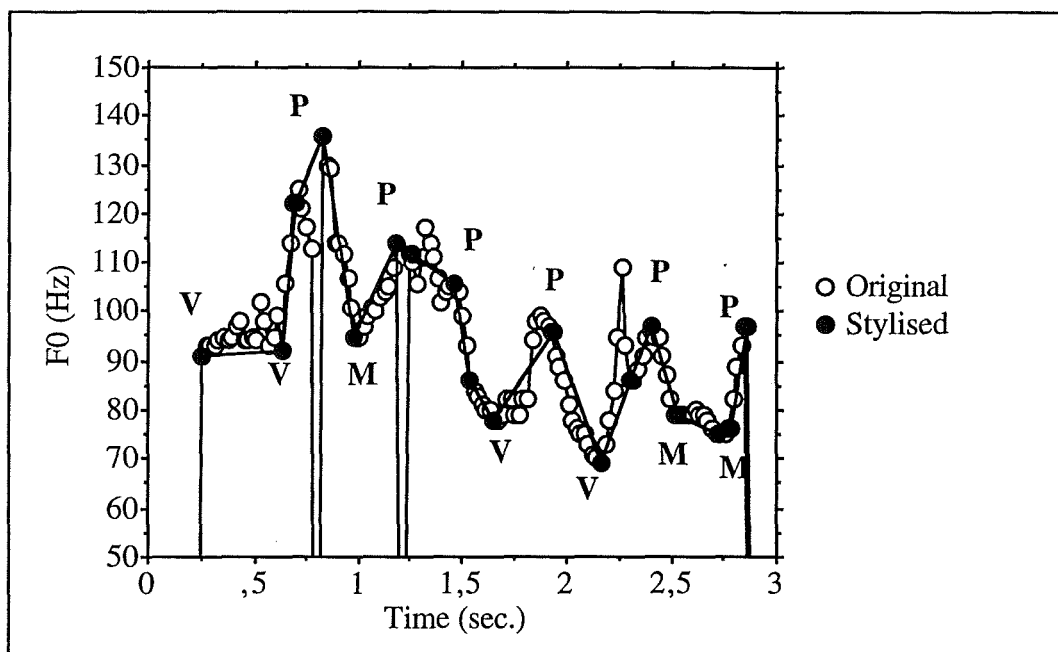
Figure 4.2. Annotated version of the contour presented in Figure 4.1.

Only those inflection points representing a clear change in the direction of the contour have an associated label. It has to be pointed out that the labelling of the inflection points assumes some kind of falling tendency along the contours. For this reason, some points that are apparently at the same level have been assigned to different categories. The existence of this falling tendency is analysed in more detail in chapter 5.

Perhaps the most controversial aspect of this proposal lies on the assumption of the existence of an M level. However, some phenomena found during the analysis of the tonic groups contours seem to support this idea:

a) when the first syllable of the group is stressed, it may appear that the beginning of the first inflection point of the phonic group is higher than usual, as in the example presented in figure 4.3:
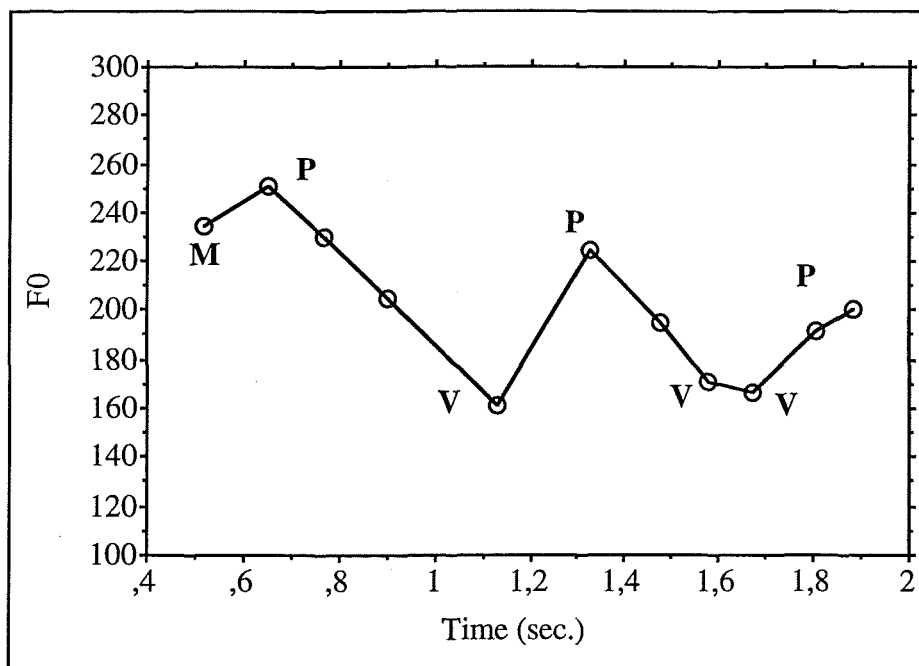
Figure 4.3. Stylized, post-processed and annotated version of the phonic group '*Poco después de la firma del pacto,*' uttered by speaker ME.

b) the level of the first 'valley' in a phonic group is frequently higher than the rest of valleys of the group, as in the contour of figure 4.4:
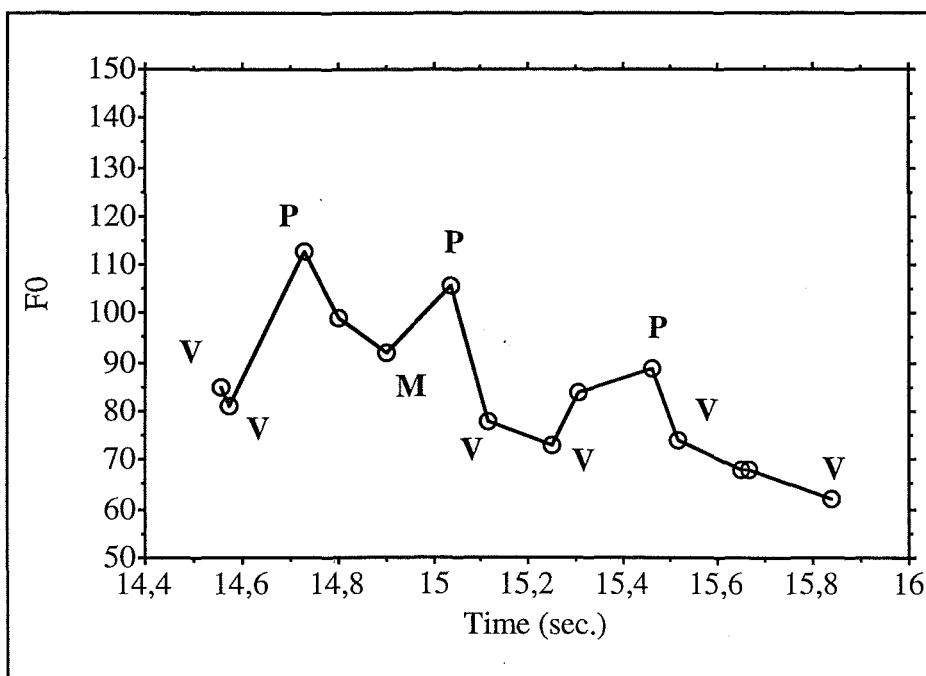
Figure 4.4. Stylized, post-processed and annotated version of the phonic group '*que sirven para encubrir cosas,*', uttered by speaker AR.

c) in some cases, a 'peak' inflection point seems to be lower than those in the rest of the group, as in figure 4.5:
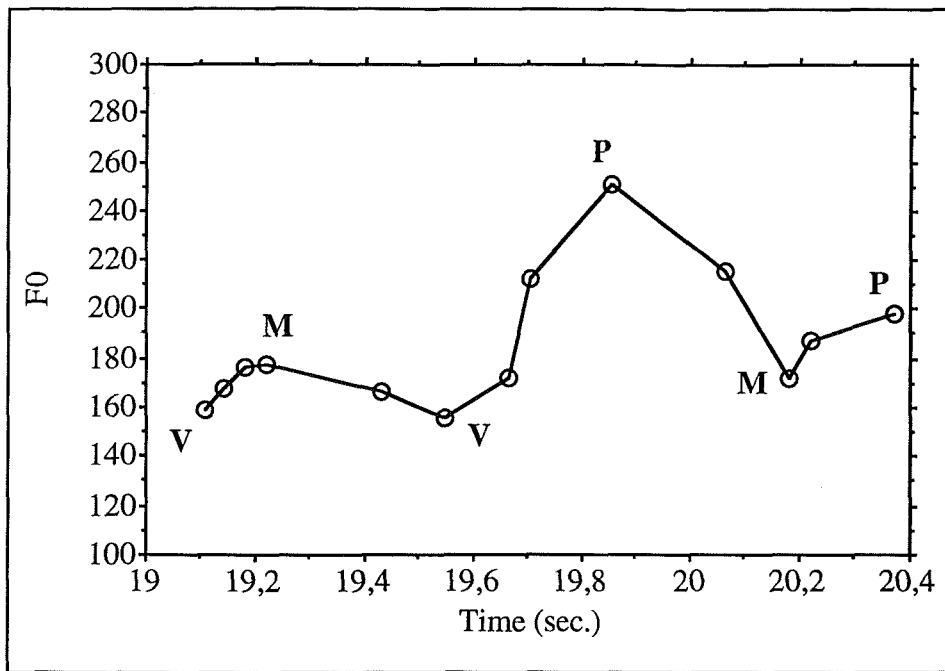
Figure 4.5. Stylized, post-processed and annotated version of the phonic group '*que, como el coeficiente de caja,*', uttered by speaker ME.

d) finally, figure 4.6 illustrates the case of some contours that show inflection points whose height seems to be between P and V levels:
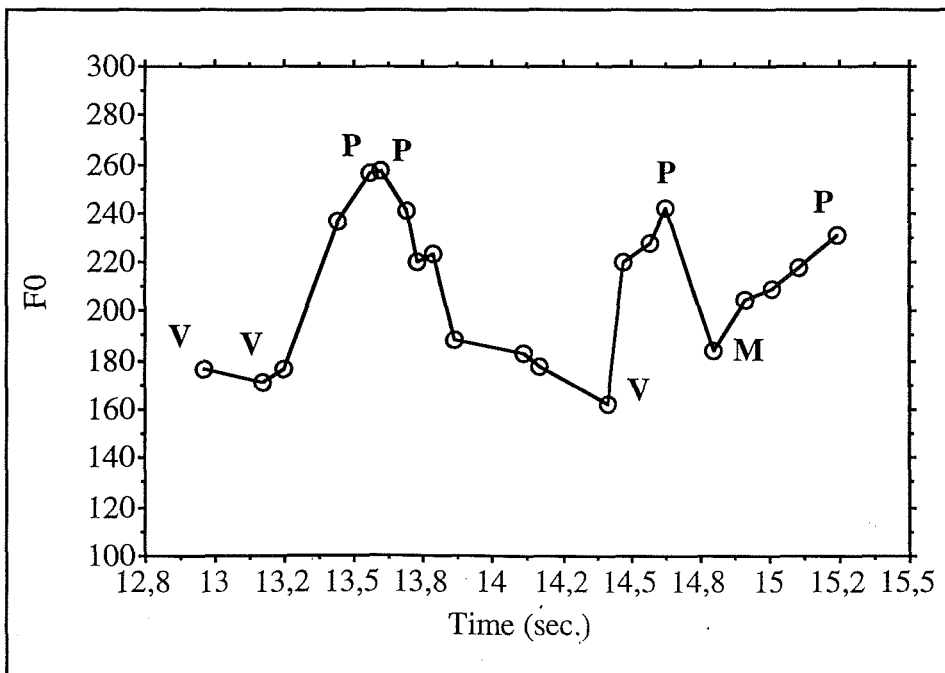
Figure 4.6. Stylized, post-processed and annotated version of the phonic group '*de mantenerse el coeficiente en el cinco por ciento,*', uttered by speaker ME.

In all these examples, the existence of an intermediate level seems to be a good solution for the annotation of these points.

The use of three different levels to represent intonation contours has already been proposed for other languages, as in the case of the IPO description and the INTSINT transcription system.

4.2.2. Intonation movements

In addition to their intrinsic level, P,. M or V, inflection points can also be defined according to the **intonation movement** preceding and following it. An intonation movement is the line linking two inflection points of any level. Three different movements have been established:

1) **Rising**: the direction of the intonation movement is rising.
2) **Level**: the direction of the movement is level, or slightly falling, following the general tendency of the intonation contour to fall.
3) **Falling**: the direction of the movement is clearly falling.

These intonation movements are similar to the 'pitch movements' defined in the IPO model (see section 1.3.5.3. for details). The movements presented here could also be described in terms of intonational features, as has been done for pitch movements. However, this type of definition has not been attempted yet.

4.2.3. Intonation contours as series of levels and movements

An inflection point can then be defined as a combination of two movements, the preceding and the following one, and one level. The inflection point presented in figure 4.7, for example, could be defined as 'Falling - V - Rising':
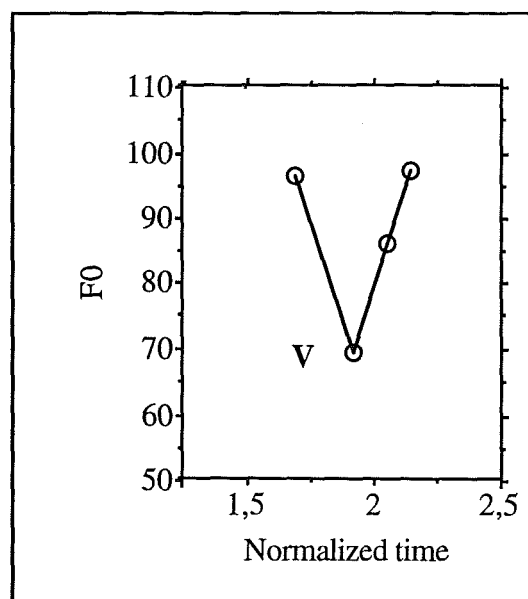
Figure 4.7. Intonation movements and inflection point
corresponding to the 'Falling-V-Rising' representation

Accordingly, the intonation contours can be considered as series of inflection points labelled with one of these three values, 'P', 'M' or 'V', and movements labelled with one of the values 'Rising', 'Level' or 'Falling'. And, finally, the intonation patterns of local movements can be equally viewed as series of movements and inflection points. The description of these patterns presented in section 4.5 is based on these assumptions.

4.3. Analysis of the tonic group contours

The stylized and post-processed F0 contours of the selected tonic groups were analysed and annotated using the described framework. The resulting transcribed version was used to analyse the evolution of the intonation contours along tonic groups in Spanish. This study involved two main steps:

a) **classification of tonic groups according to their contour shape**:

The different tonic groups were classified according to the contour shape they presented. This shape has been defined as the combination of levels and movements appearing along each tonic group. 30 different shapes were found in the case of the initial tonic groups, 61 in the case of medial groups, 59 in the case of final groups, and 9 in the case of the initial-final tonic groups.

b) **analysis of the location of different inflection points within the tonic group**:

The location of the inflection points found in each tonic group was analysed and annotated, taking the stressed syllable as reference point. The location of the point within the syllable has also been labelled, but this information has not been taken into account in this work.

The results of these analyses were used to check the validity of the tonic group as an intonation domain in Spanish, and to define a set of local intonation patterns. The results are presented in the two following sections.

4.4. Tonic groups and the domain of the local intonation patterns

A first conclusion that can be extracted from the analysis of the shape of the intonation contours along tonic groups is that there is no exact matching between the tonic group and the domain of movements. The following figures offer some examples.
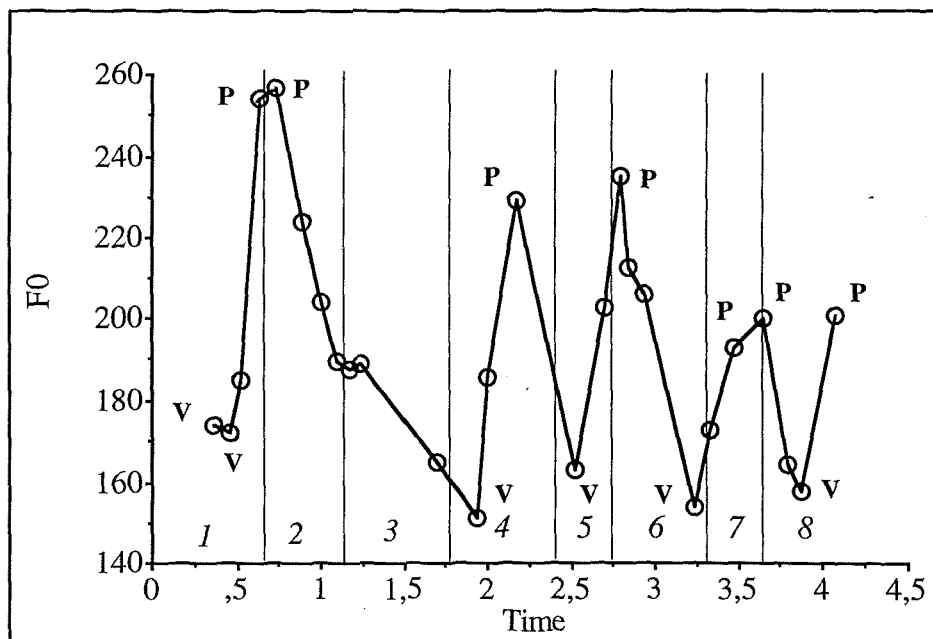
Figure 4.8. Stylized, post-processed and annotated contour corresponding to the phonic group '*La firma (1) del pacto (2) autonómico (3) entre los líderes (4) de los dos (5) principales (6) partidos (7) del país (8)*' uttered by speaker M.E.
Vertical lines represent tonic group boundaries.

It can also be concluded from this figure that there is no exact correspondence between the number of 'peaks' and the number of stressed syllables. In some groups, as the initial one (number 1, '*La firma*'), or some

internal groups, (number 4, '*entre los líderes*'; number 7, '*partidos*'), there is a peak at the end, or at least in the second part of the group. In other cases, however, only a falling movement appears along the tonic group, as in the case of groups number 2 ('*del pacto*') and 3 ('*autonómico*').

The boundary between tonic groups is not clearly marked with an intonational cue. In some cases, the peak coincides with the boundary (group 1); in other cases, however, the peak is before (group 4), or after (group 5) the boundary.
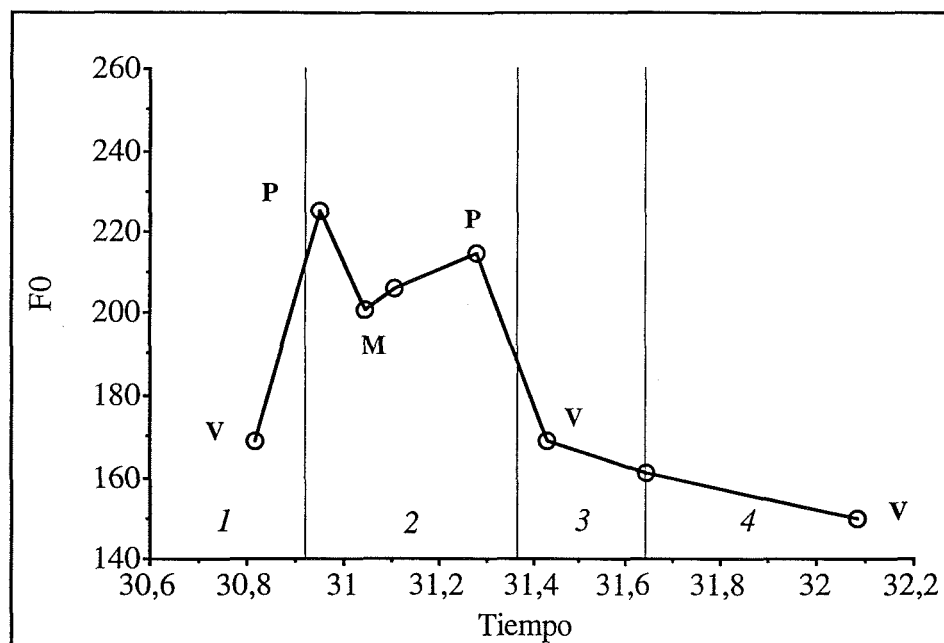
Figure 4.9. Stylized, post-processed and annotated contour corresponding to the phonic group '*en un (1) acuerdo (2) de libre (3) cambio (4).*' uttered by speaker M.E. Vertical lines represent tonic group boundaries.

The contour presented in figure 4.9 illustrates an additional case of mismatch between tonic group and intonation pattern. Groups number 2 ('*de libre*') and 3 ('*cambio*'), do not show any peak; they seem to be part of the same intonation pattern.
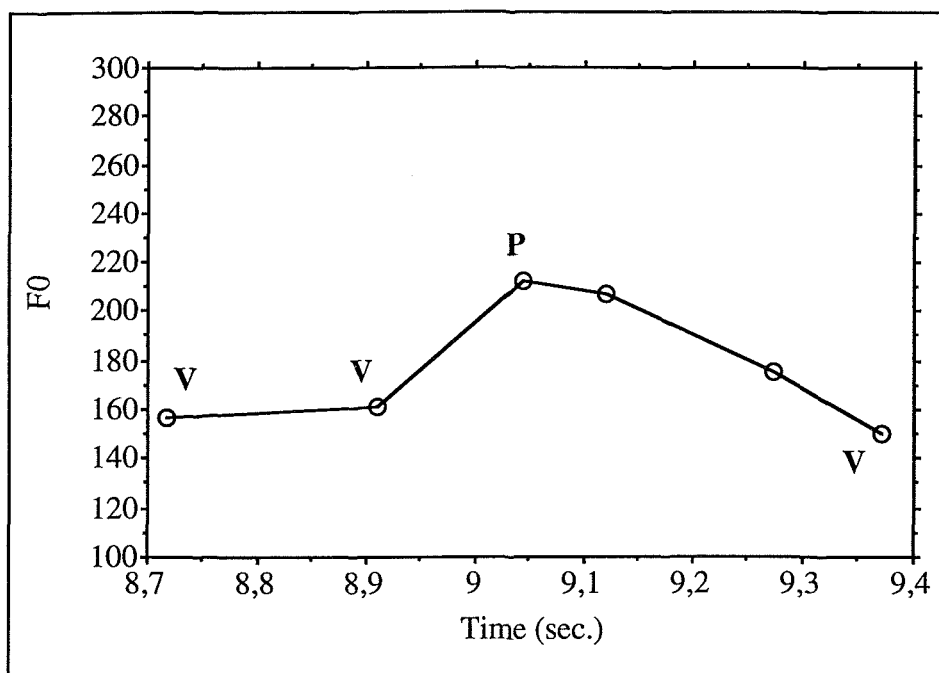
Figure 4.10. Stylized, post-processed and annotated contour corresponding to the phonic
group *'en los coches'* uttered by speaker M.E.

Finally, figure 4.10 shows an example of a single tonic group whose
intonation contour can be described in terms of two separate intonation
patterns. It presents an initial-final tonic group, whose contour can be
interpreted in terms of an initial rising pattern, until the P point, and a
second 'final' pattern, until the end of the contour. This combination allows
the description of the shape of the contours of the initial-final tonic groups
as the combinations of an initial and a final pattern.

The examples above do not favour the idea of considering the tonic group
as the domain for local intonation movements.

4.5. Defining local intonation patterns

According to these examples, it seems that local intonation patterns do not
have tonic groups as their natural domain, and it is better to define them
independently of their domain, only considering their shape as a classifying
factor. The contour shapes obtained in the study of the tonic groups were
re-analysed to define a set of intonation patterns that are presented in this
section.

A local pattern can be defined as any specific combination of movements
and levels that has been found to appear, irrespective of its frequency, in
the analysed material. Any combination of this type, even if it appears only
once in the material, has been included in this description as a 'pattern'.

According to this definition, patterns can then be represented exclusively as series of combinations and levels, without paying attention to the specific location of the inflection points in the utterance, or to the F0 value associated to each level. For example, the pattern offered in figure 4.11 can be described using the combination of levels and movements 'V - Level - V - Rising - P':
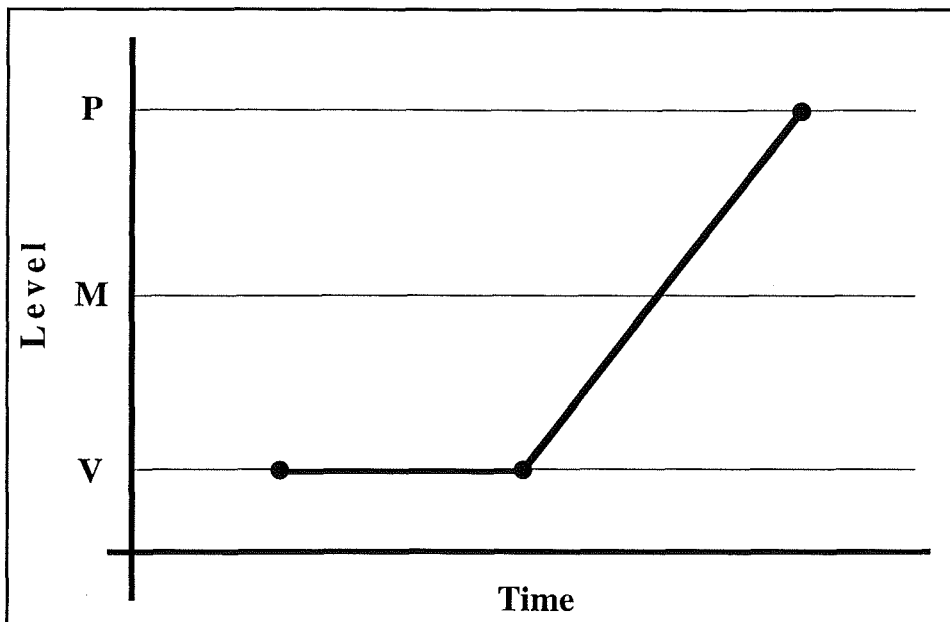

Figure 4.11. An example of representation of a local pattern. The pattern described here is the initial pattern number 7 (V - Level - V - Rising - P)

The boundary of the patterns has been established to coincide in all cases, more or less arbitrarily, with the 'peak' of the pattern, or, more precisely, with an inflection point 'P' or 'M' followed by a falling movement. The intonation contour of any phonic group can be divided then into intonation patterns by placing their boundary at the end of each peak, as can be seen in figure 4.12. This definition of intonation patterns is specially useful for intonation models oriented to TTS systems, which usually build contours by concatenation of patterns.
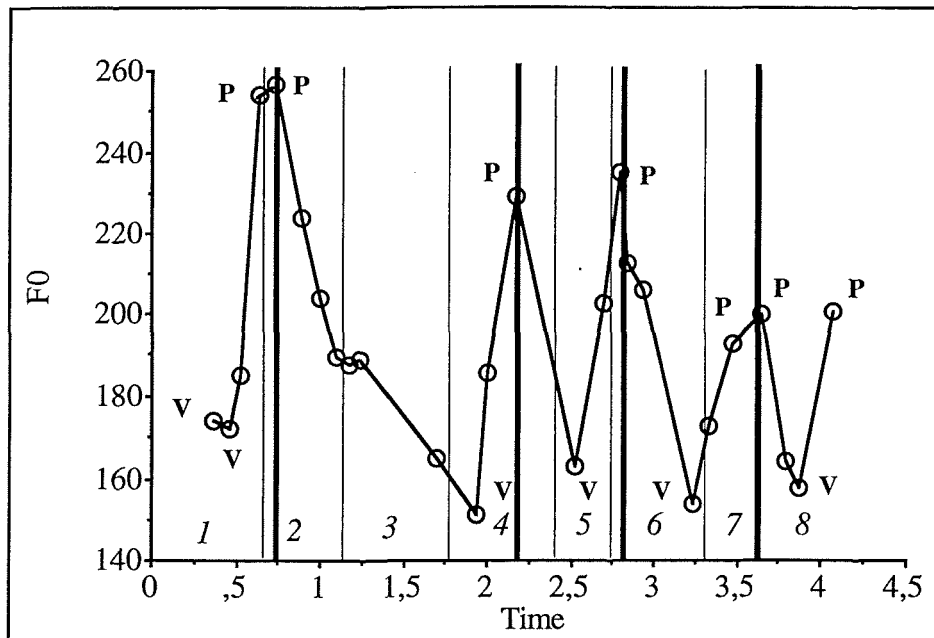
Figure 4.12. The same stylized, post-processed and annotated contour as in figure 4.8 (*'La firma del pacto autonómico entre los líderes de los dos principales partidos del país'*, speaker M.E.) segmented into intonation patterns. As in 4.8, narrow vertical lines represent tonic group boundaries; bold lines have been added to indicate the boundaries of the patterns.

A total of 66 different patterns has been found. However, the patterns appearing in each position of the phonic group show similar features so they can be grouped into several sets. The identified patterns were then divided into three main groups, depending on the position of the tonic group within the phonic group:

a) **Initial**, for patterns appearing at the beginning of a phonic group;

b) **Final**, for patterns appearing at the end of a phonic group;

c) **Medial**, for patterns located neither in initial nor in final position.

Within these three groups, some subdivisions have also been established. In the case of medial patterns, two subtypes have been considered:

a) **Post-Initial patterns**, or patterns that typically appear (only) immediately following an initial one;

b) **Internal patterns**, which include the patterns that may also appear in non-Post-Initial position.

The final patterns have been classified into three subtypes according to the shape of the contour:

a) **Falling patterns**, for those final patterns ending with a fall or a low-level movement;

b) **Rising patterns**, which include those patterns ending with a rise or a high or mid-level movement;

c) **Rise-fall patterns**, for those final patterns containing such type of movements.

Minor subclasses have been established within these types of final patterns. They are described in more detail in the following sections.

Table 4.1 shows the distribution of the patterns along the different defined subsets.

Initial	Medial		Final					
	Post-Initial	Internal	Falling	Rising		Rise-Fall		
				A	B	A	B	C
14	3	15	7	10	2	9	4	2

Table 4.1. Distribution of patterns by type

The patterns obtained in this work can be compared with previous descriptions of intonation, in Spanish and other languages. Some similarities can be established, for example, with the patterns presented in Fant (1984), or with the configurations of the IPO model.

As far as final patterns are concerned, the classification proposed here is slightly different from classical descriptions of final movements, as the one by Navarro (1944) presented in section 1.2.2.1.3. The classification suggested here is based not only on the direction of the last movement, as in the case of Navarro's 'tonemes', but also on the global shape of the pattern. In addition, final movements as 'level', 'half-rising' or 'half-falling', used in Navarro's description, have not been considered here. Accordingly, no distinction is made between 'sentence-final' patterns and 'non-sentence-final' patterns, established in Navarro's system by using the 'half-rising' and 'half-falling' movements.

In the following sections, a description of the different types of defined patterns is offered. The complete inventory of patterns included in each group, and a description of the location of the inflection points in these patterns, can be found in appendix 3 (section 1).

4.5.1. Initial patterns

As stated above, the group of initial patterns includes all those patterns that have been found to appear at the beginning of phonic groups. They show common features, both in their shape and in the location of their inflection points, that allow to include them in a single group.

A prototypical shape for initial patterns is a low start, V or M, followed by a rising movement that tends to coincide with the stressed syllable, and a high end, M or P, at the stressed syllable itself or in the syllable following the stressed one. The most frequent realization of this shape includes a 'level' movement before the rise, as can be seen in figure 4.13:
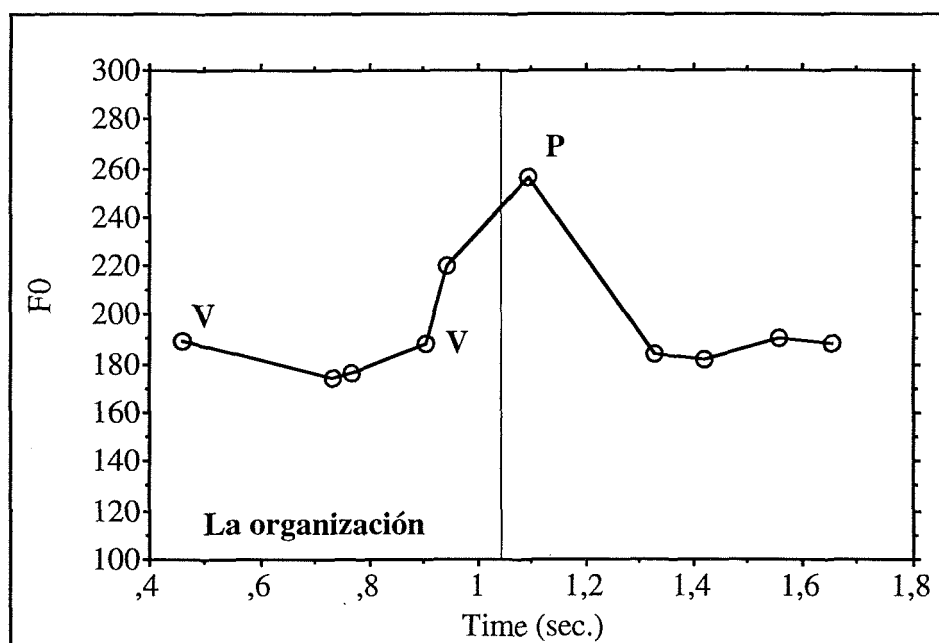
Figure 4.13. An example of intonation contour showing the initial pattern number 7 (V - Level - V - Rising - P). The example corresponds to the phonic group "*La organización terrorista*" uttered by speaker ME. The vertical line indicates the tonic group boundary.

A less frequent variant of this shape is pattern 3, which does not include the 'level' movement before the rise:
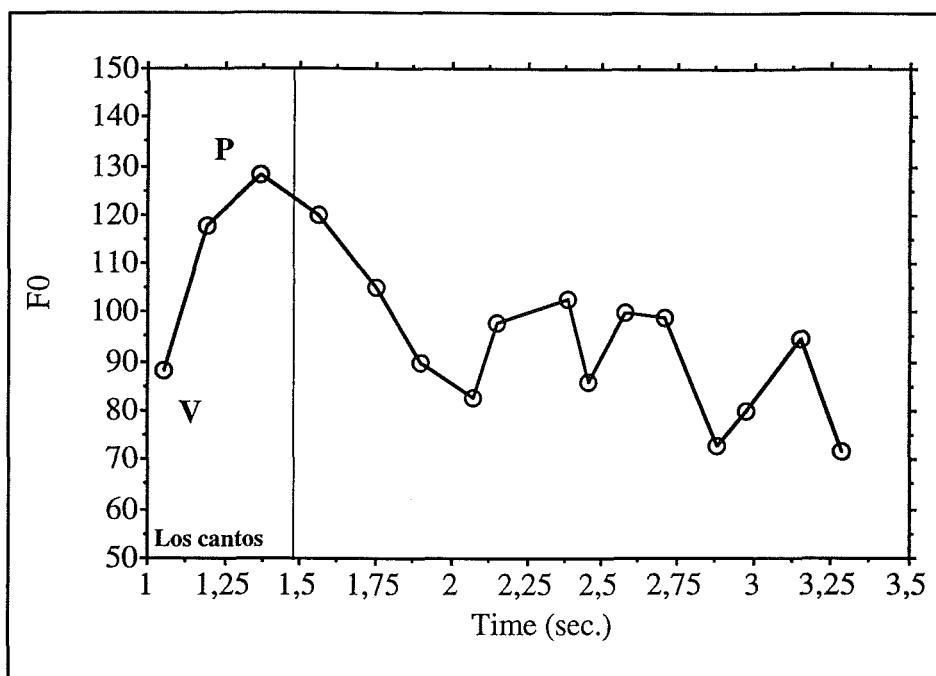
Figure 4.14. An example of intonation contour showing the initial pattern number 3 (V - Rising - P). It corresponds to the initial tonic group in the phonic group "*Los cantos de sirena en favor de un bloque de primera*," uttered by speaker AR.

Pattern number 12 appears also in a significant percentage of cases. It is characterized by the presence of a 'flat' peak, as can be seen in figure 4.15:
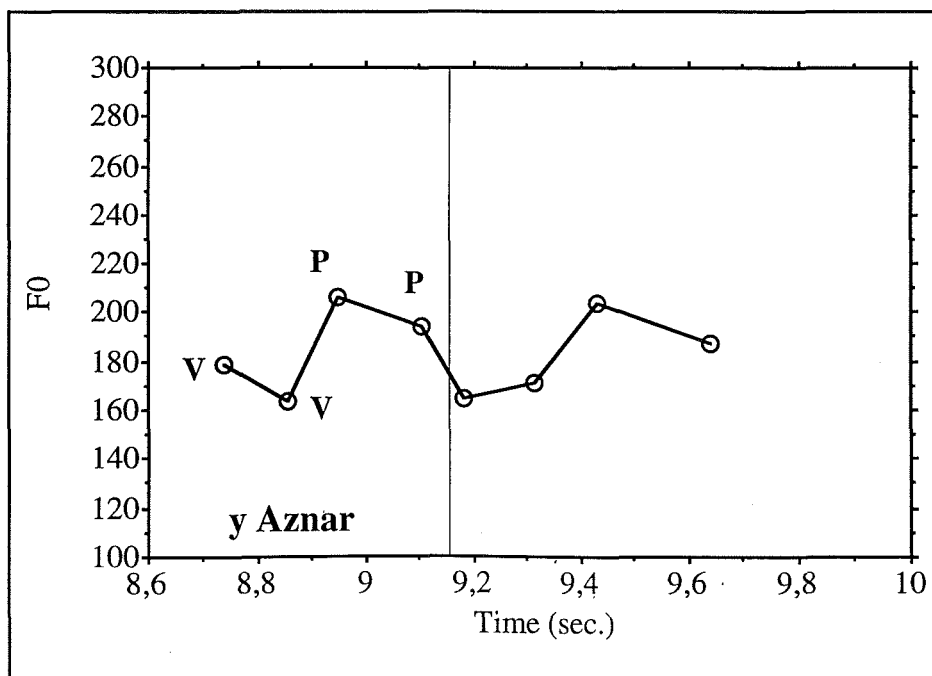

Figure 4.15. An example of intonation contour showing the initial pattern number 12 (V-Level - V - Rising - P - Level - P), extracted from the phonic group "*y Aznar recordaba*", uttered by speaker ME

In some cases, however, when the stressed syllable is the first syllable of the group, the pattern can start with in P or M level, and it is followed by a level movement. This type of shape is exemplified in figure 4.16:
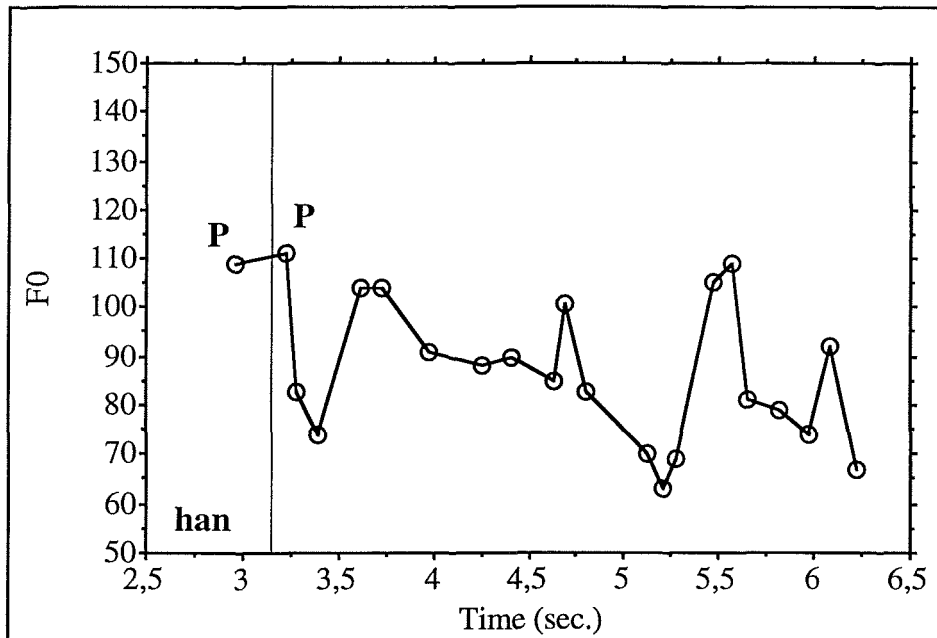
Figure 4.16. An example of intonation contour showing the initial pattern 2 (P - Level - P), extracted from the phonic group "*han hecho que algunos instrumentos financieros, como los cheques conformados*", uttered by speaker AR

The stressed syllable of the initial tonic group tends to be associated then with a P or M level, although the inflection point does not have to coincide specifically with the stressed syllable, as has been reported already in earlier studies (Navarro, 1944; Fant, 1984).

4.5.2. Medial patterns

Medial patterns have been divided in two different subtypes: the group of post-initial patterns, appearing exclusively after an initial one, and the group of internal patterns, that can appear in any medial position of a phonic group (that is, excluding only initial and final locations).

4.5.2.1. Post-Initial patterns

The main feature of this type of patterns is the fact that they usually extend over more than one tonic group. The initial point (P in all the analysed cases) coincides with the final point of the initial pattern, and is located then in the stressed syllable of the first tonic group, or in the syllable following it. A falling movement starts at this point, continues along the second tonic group of the phonic group, and even the third one in some cases, and ends

in a M or V level close to the stressed syllable of the third or fourth tonic group. Then a rising movement starts, finishing in a P or M level to connect with the first internal pattern of the group. When this pattern appears in a phonic group, then, the second stressed syllable of the group does not have a peak associated to it. An example of this type of patterns is given in figure 4.17.
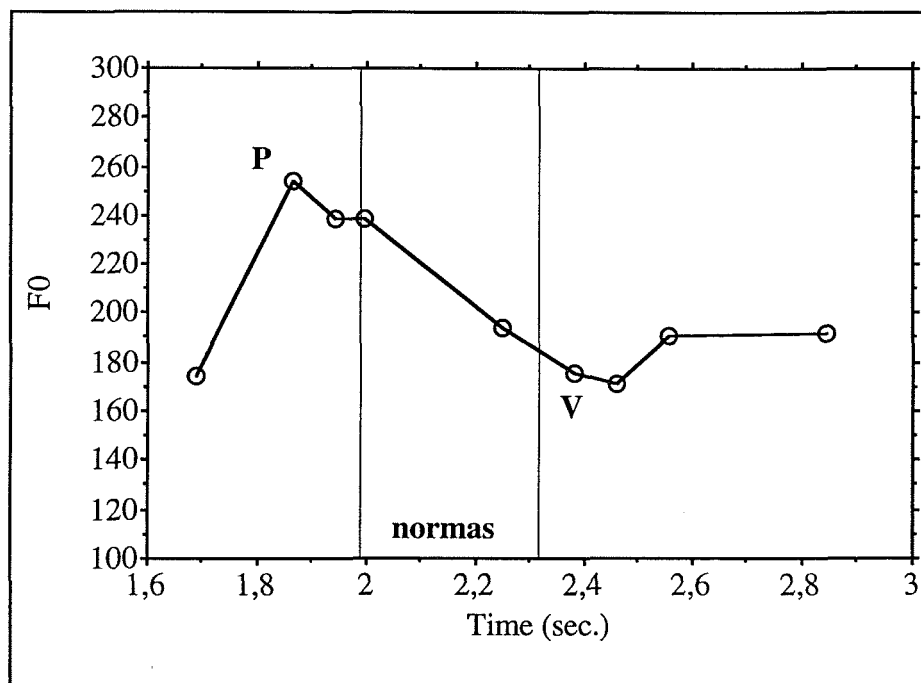

Figure 4.17. An example of intonation contour showing the post-initial pattern number 1 (P - Falling - V). It has been taken from the phonic group "*las nuevas normas contables*", uttered by speaker ME.

In some cases, however, a level movement can appear in the middle of the pattern, coinciding with the second stressed syllable of the phonic group. The stressed syllable is marked then with two M points, one before and another after the stressed syllable. This case is illustrated in figure 4.18.
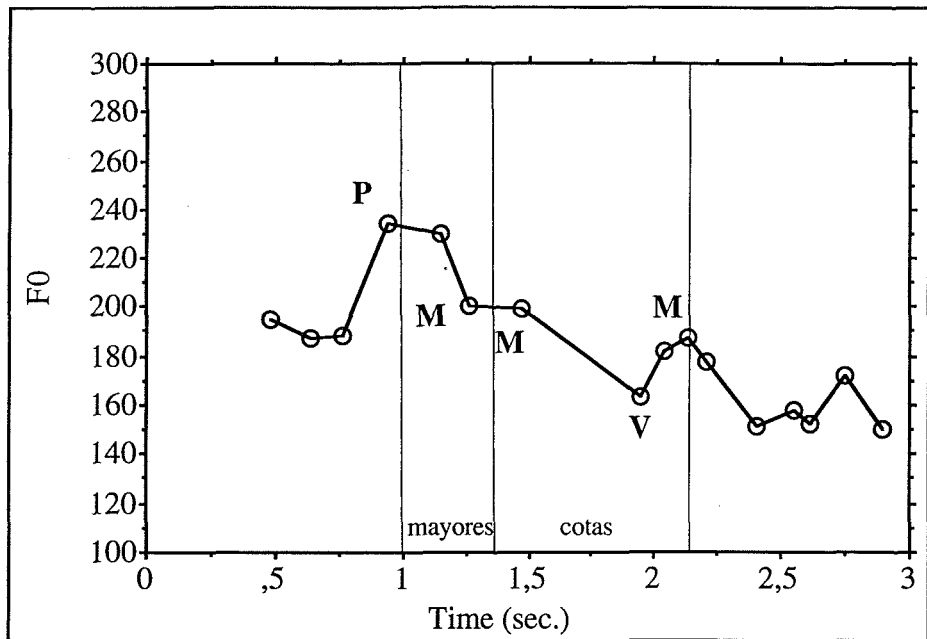
Figure 4.18. An example of intonation contour showing the post-initial pattern number 3 (P - Falling - M - Level - M), taken from the phonic group "¿O de obtener mayores cotas de bienestar para los catalanes" uttered by speaker ME.

4.5.2.2. Internal patterns

The most typical shape of internal patterns is presented in figure 4.19. This shape includes two movements, an initial fall and a final rise, linked by a low-level inflection point, V in this case, that tends to be located just before the beginning of the stressed syllable. The point corresponding to the end of the rising movement can be located in the stressed syllable or in a syllable following it. This 'peak delay' phenomenon was already reported by Garrido et al. (1993) or Prieto et al. (1994, 1995b), as indicated in chapter 1.
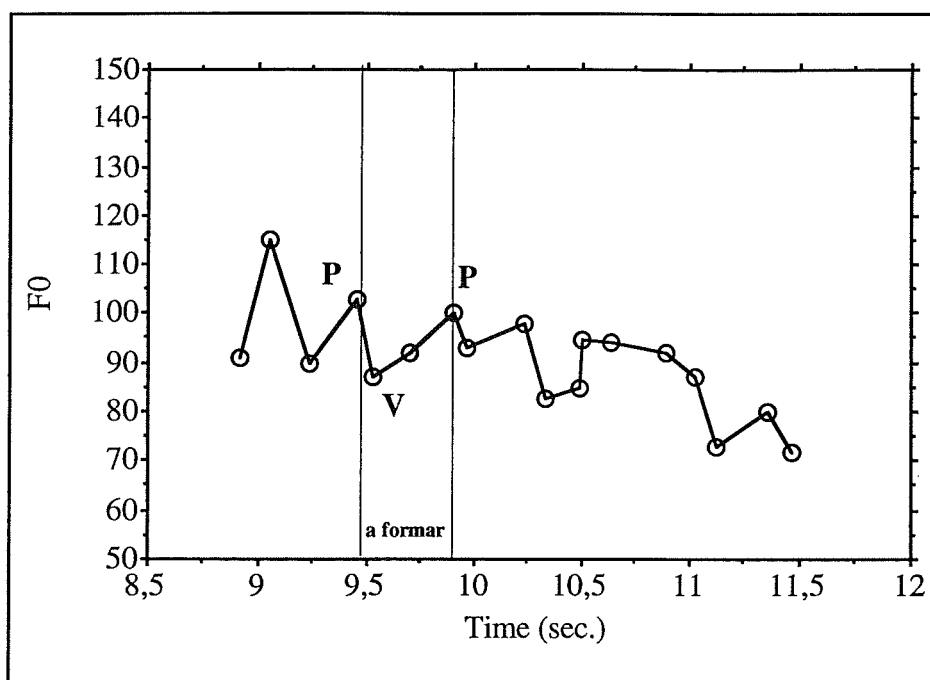
Figure 4.19. An example of intonation contour showing the internal pattern number 1 (P - Falling - V - Rising - P), included in the phonic group *"hayan pasado a formar parte de los depósitos de los bancos,"*, uttered by speaker AR.

The rest of the patterns included in this group can be described as variations of this basic pattern. Accordingly, they can be described as composed by two different parts, linked by a V or M inflection point:

a) a falling section, that generally links the end of the previous pattern, at P or M level, and the beginning of the stressed syllable;
b) a rising section, from the beginning of the stressed syllable to the end of the pattern, again at P or M level.

If some variations are introduced in the basic pattern, then other frequent patterns are obtained. For example, if the intermediate inflection is at M level instead of at V level, pattern 5 appears. This pattern is presented in figure 4.20.
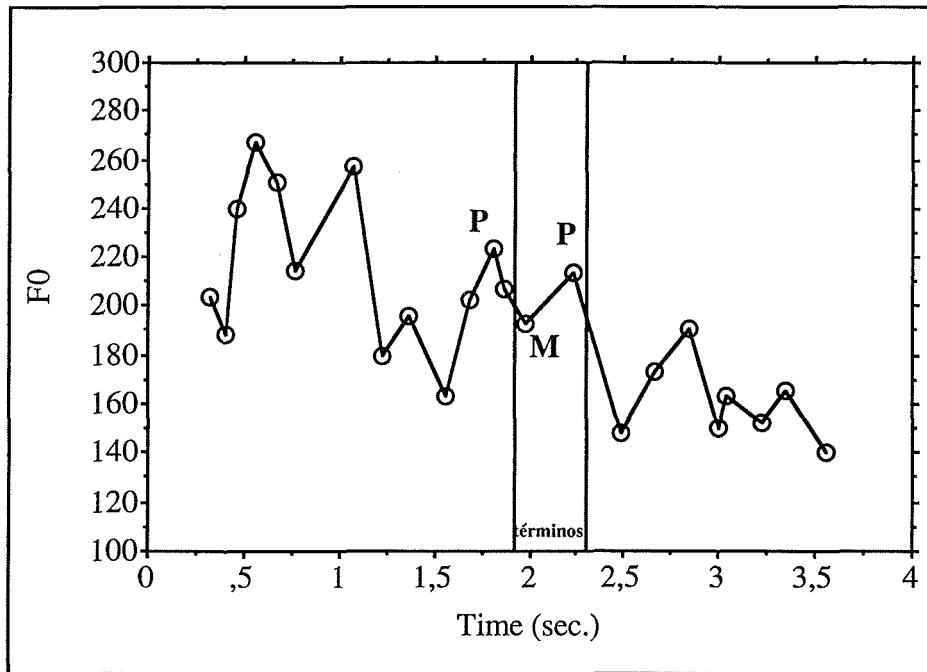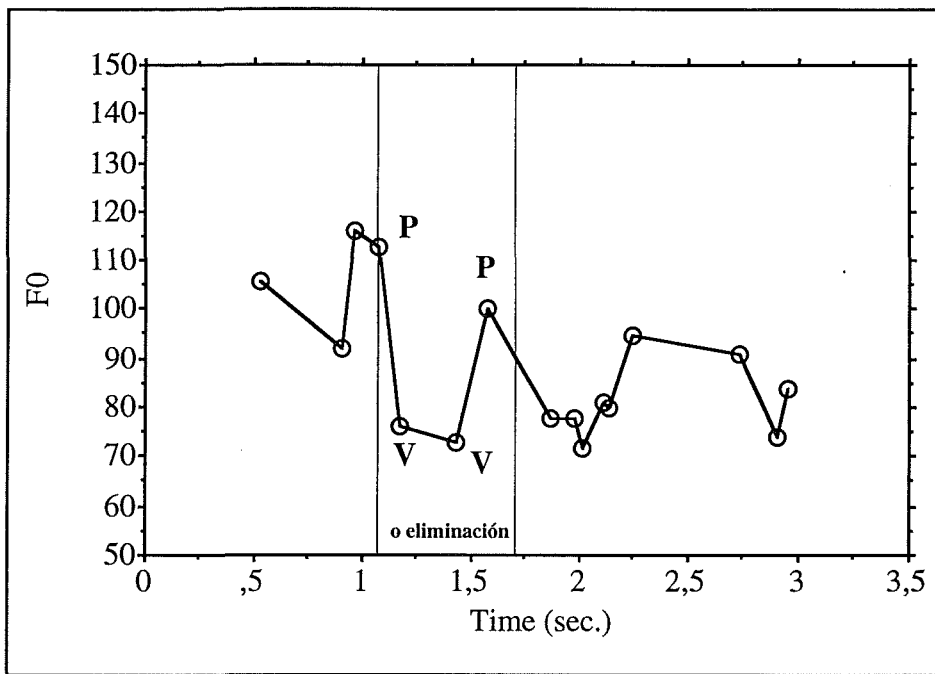
Figure 4.20. An example of intonation contour showing the internal pattern number 5
(P - Falling - M - Rising - P), extracted from the phonic group "*Se pueden seguir
empleando los mismos términos de la crítica especializada*", uttered by speaker ME.

If a level movement is introduced in the falling part, before the
intermediate inflection point, patterns 6 is obtained, also quite frequent in
the analysed material. This pattern is presented in figure 4.21. Pattern 7,
showing the same shape as pattern but with the 'level' movement at M
height, is also quite common.

Figure 4.21. An example of intonation contour showing the internal pattern number 6 (P - Falling -V - Level - V - Rising - P), extracted from the phonic group *"La reducción o eliminación de los coeficientes de caja"*, uttered by speaker AR.

If a level movement is introduced in the rising part of the pattern, a 'flat' peak is obtained, as was already found in the case of initial patterns. An example of pattern with this type of 'flat' peak is presented in figure 4.22.
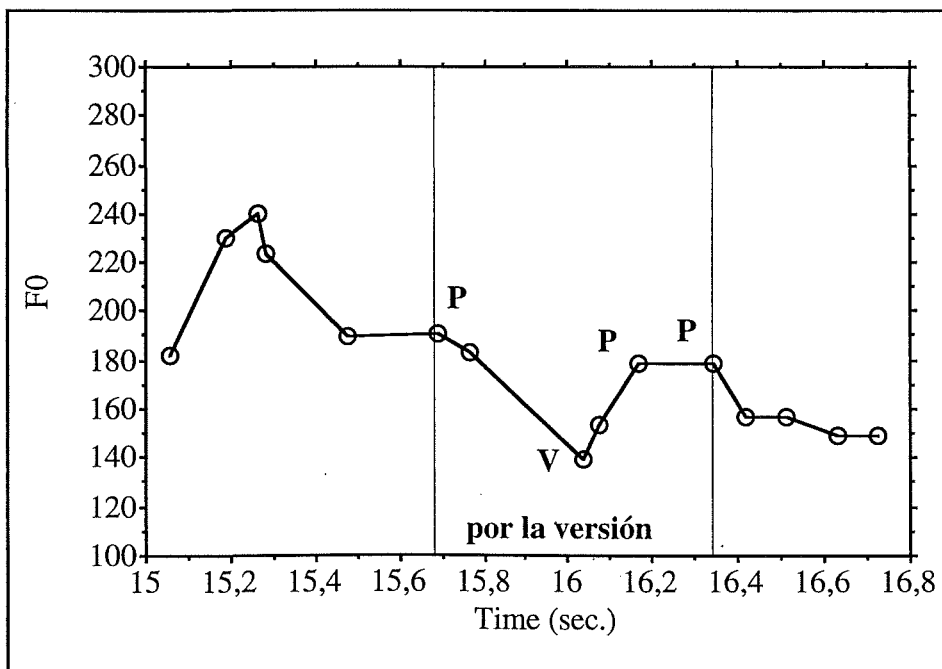

Figure 4.22. An example of intonation contour showing the internal pattern number 10 (P - Falling - V - Rising - P - Level - P), included in the phonic group *"y no confirmada por la versión oficial."*, uttered by speaker ME.

The combinations of these possible variants in the falling part, rising part, and intermediate point produce the different patterns found in this analysis.

4.5.3. Final patterns

As stated before, three different types of final patterns have been established:

1) Falling patterns
2) Rising patterns
3) Rise-Fall patterns

They are described in the following sections.

4.5.3.1. Falling patterns

Final falling patterns are characterized by having a global falling shape. They end with a falling or (low) level movement, and their last inflection point is usually at V level. These patterns typically start at a high level (P or M), at the end of the penultimate stressed syllable, and go down until the end of the contour. This fall can be realized in several ways, the most frequent one being represented in figure 4.23. In this case, the intermediate V point tends to be placed in the syllable before the stressed one or in the stressed syllable itself.
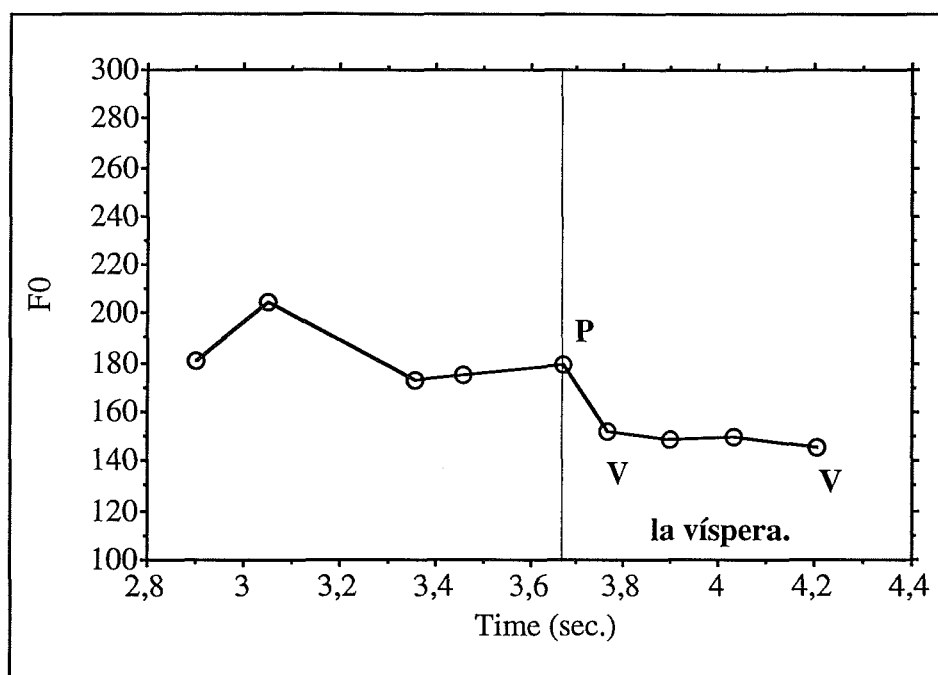
Figure 4.23. An example of intonation contour showing the final falling pattern number 3 (P - Falling - V - Level - V), embedded in the phonic group "*ambas delegaciones la víspera.*", uttered by speaker ME.

A frequent shape is also the one represented by pattern 1, schematized in figure 4.24. In this pattern, the stressed syllable of the last tonic group does not show any inflection point or movement associated with it. The falling movement starts in the penultimate stressed syllable of the group, or after it, and goes down directly to the end, at V level. Pattern 4 is a variant of this shape starting at M level instead of at P level.
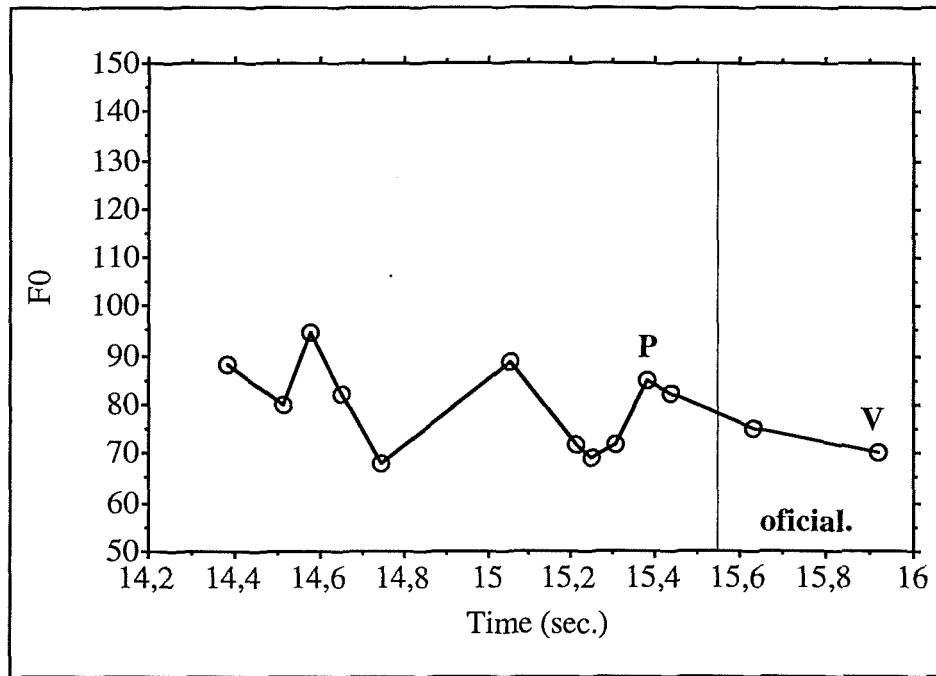
Figure 4.24. An example of intonation contour showing the final falling pattern number 1
(P - Falling - V). It is included in the phonic group "*y no confirmada por la versión
oficial.*", uttered by speaker AR.

Finally, a less frequent choice for falling patterns is the one represented by
figure 4.25. In this case, the first M point is 1-2 syllables before the stressed
one, and the second M can be placed just before the stressed syllable or at
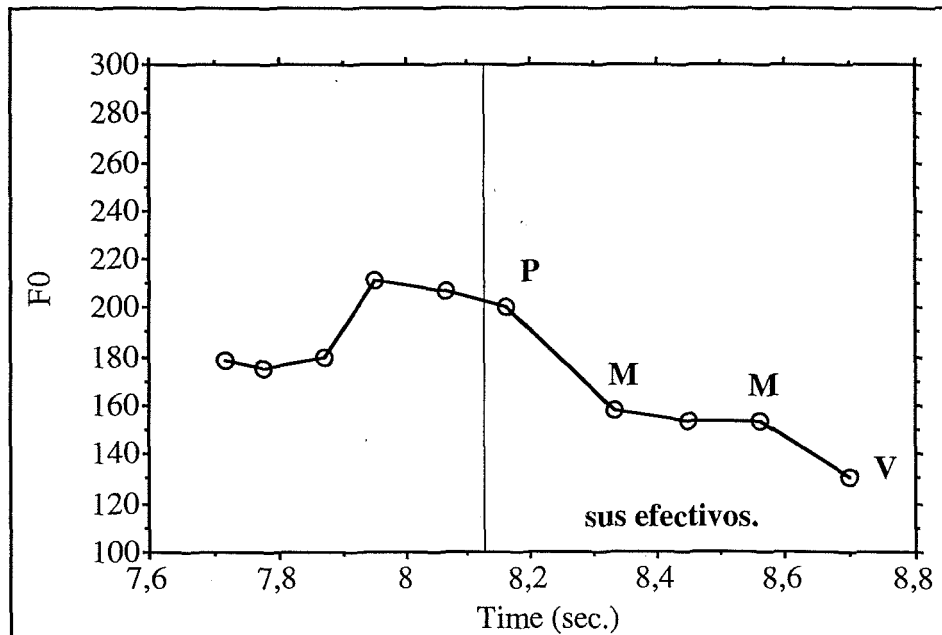the stressed syllable itself.

Figure 4.25. An example of intonation contour showing the final falling pattern 6
(P - Falling - M - Level - M - Falling -V). It is included in the phonic group "*y con todos
sus efectivos.*", uttered by speaker ME.

The rest of patterns included in this group are variants of these three basic shapes.

4.5.3.2. Rising patterns

Rising patterns are characterized by having a rising movement in the second half of the pattern. They can end with a rising or (high) level movement, and the final point is usually at P level. The basic shape for this type of patterns, and also the most frequent one, is the pattern presented in figure 4.26.
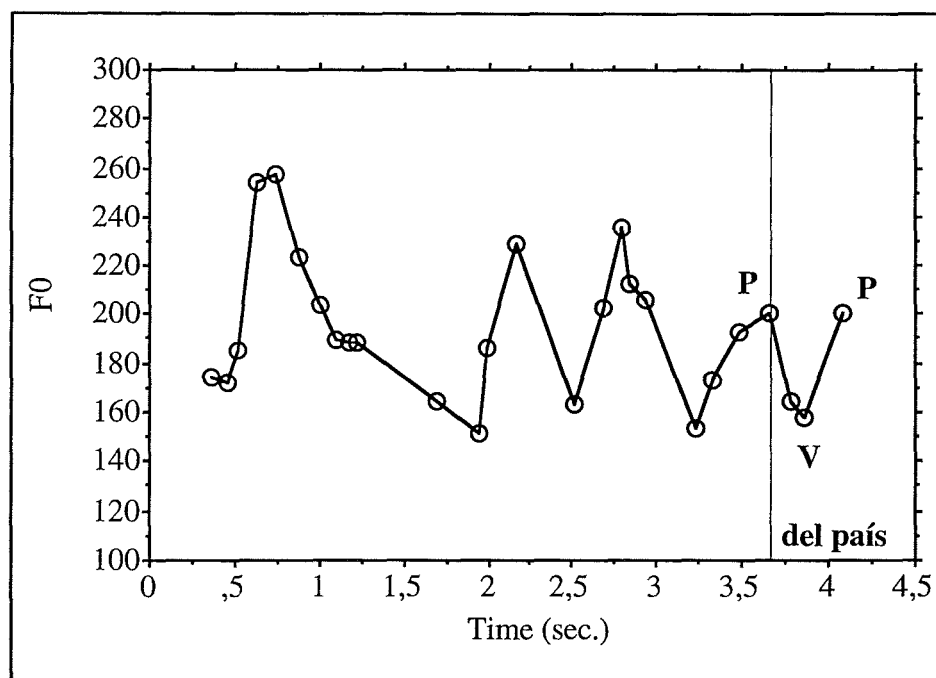

Figure 4.26. An example of intonation contour showing the final rising pattern (type A) number 1 (P - Falling - V - Rising - P), extracted from the phonic group "*La firma del pacto autonómico entre los líderes de los dos principales partidos del país*", uttered by speaker ME.

As in the case of internal patterns, two parts can be distinguished, a falling one and a rising one, divided by an intermediate P or M inflection point.

Also as in the case of the internal patterns, both parts of the pattern can present variations, that are represented by patterns appearing frequently in the analysed material. If the falling part includes a level movement before the intermediate point, the scheme represented by figure 4.27 is obtained:
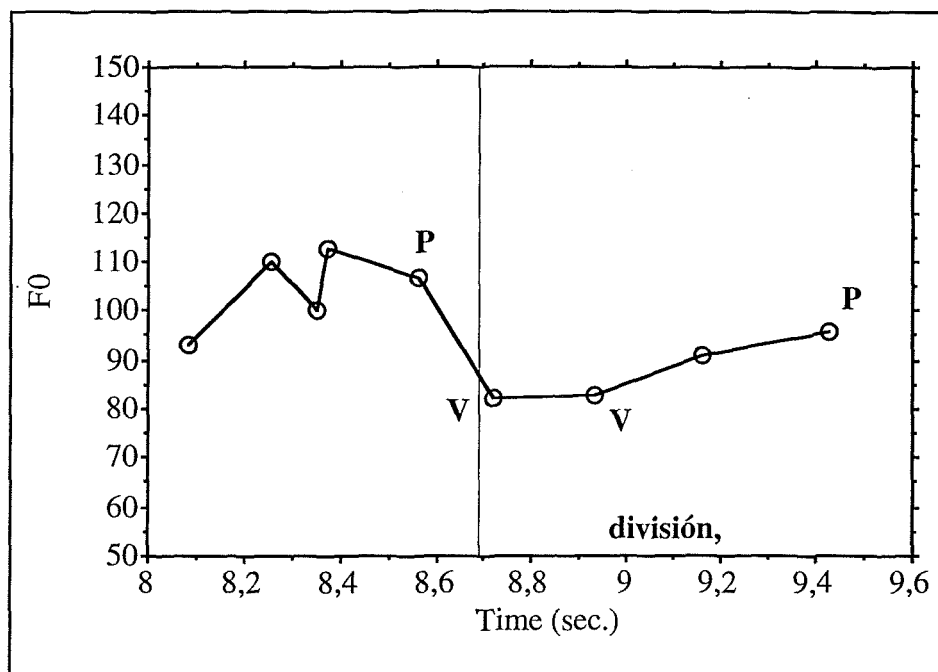
Figure 4.27. An example of intonation contour showing the final rising pattern (type A) number 4 (P - Falling - V - Level - V - Rising - P), included in the phonic group *"en una segunda división,"*, uttered by speaker AR.

If a level movement is introduced in the rising part, the pattern represented in figure 4.28 appears:
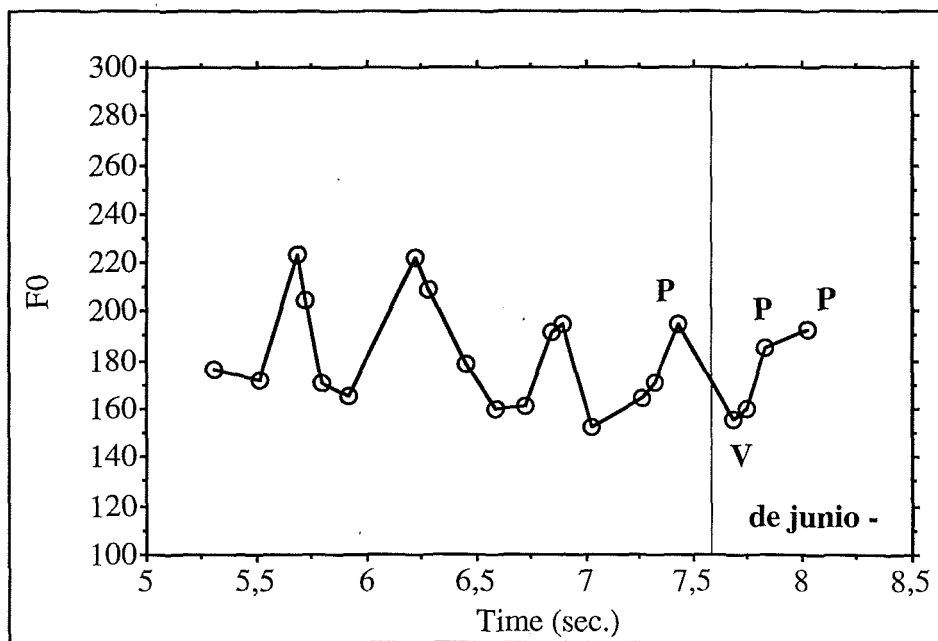

Figure 4.28. An example of intonation contour showing the final rising pattern (type A) number 8 (P - Falling - V - Rising - P - Level - P). It is included in the phonic group *"y Dinamarca como fruto del mandato del referéndum de junio -"*, uttered by speaker ME.

Finally, if the intermediate point is at M level instead of at V level, the pattern of figure 4.29 is obtained:
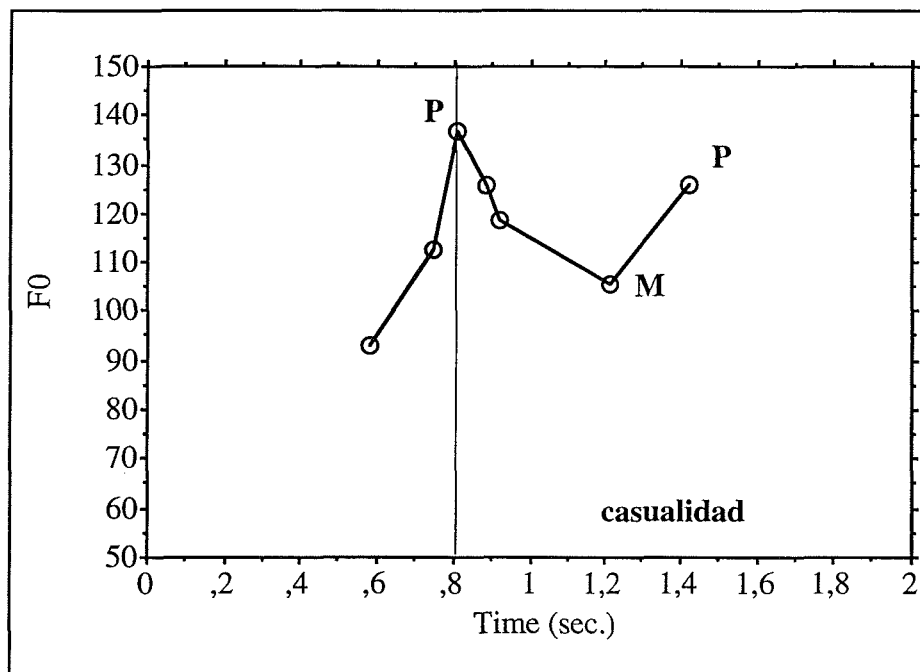
Figure 4.29. An example of intonation contour showing the final rising pattern (type A) number 2 (P - Falling - M - Rising - P), embedded in the phonic group "¡*Curiosa casualidad*", uttered by speaker AR.

The rest of the rising patterns found in this study are variations of the basic patterns described here.

Depending on the level of the last inflection point, rising patterns have been divided into two subtypes:

a) Type A, including patterns ending with a P inflection point. All the examples presented in these sections correspond to type A patterns.

b) Type B, for patterns ending with an M inflection point. This type of patterns is, however, much less frequent than type A patterns. An example of type B pattern can be found in figure 4.30.
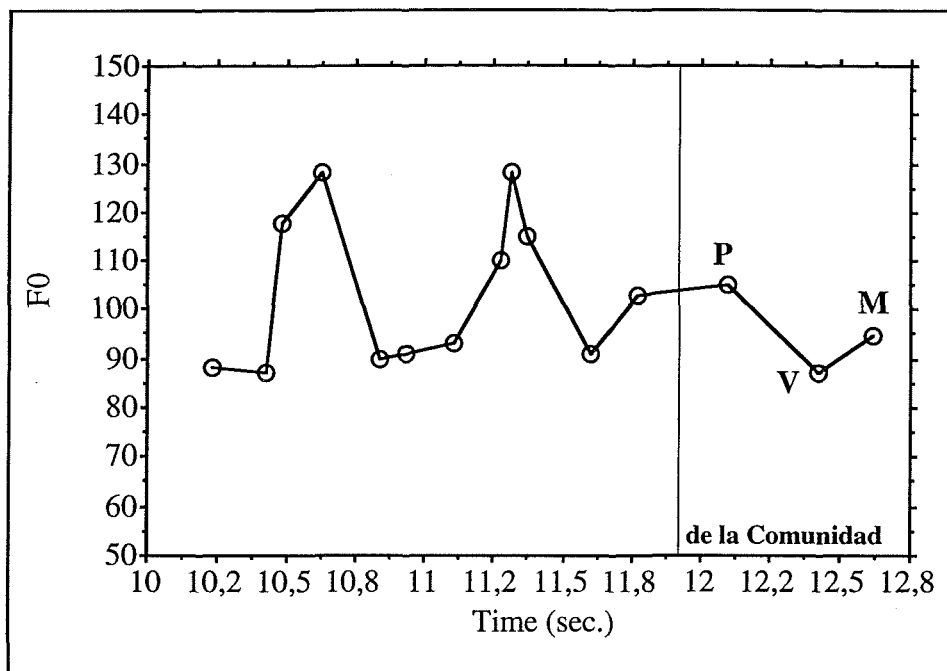
Figure 4.30. An example of intonation contour showing the final rising pattern (type B) number 1 (P - Falling - V - Rising - M). It is included in the phonic group "*se sucedan en la presidencia semestral de la Comunidad*", uttered by speaker AR.

4.5.3.3. Rise-Fall patterns

The presence of a special 'rise-fall' concatenation of movements is the feature common to all the patterns of this group. This 'rise-fall' pattern would be equivalent to the '*tonema circunflejo*' described in Navarro (1944) and reviewed in chapter 1. They are also divided in several special types, three in this case (A, B, and C), depending on the level of the inflection points appearing in the 'rise-fall' complex movement.

4.5.3.3.1. Type A

In type A, the inflection points located at the beginning and the end of the rise-fall movement are at the same level. Figure 4.31 shows the most frequent shape for this type of patterns:
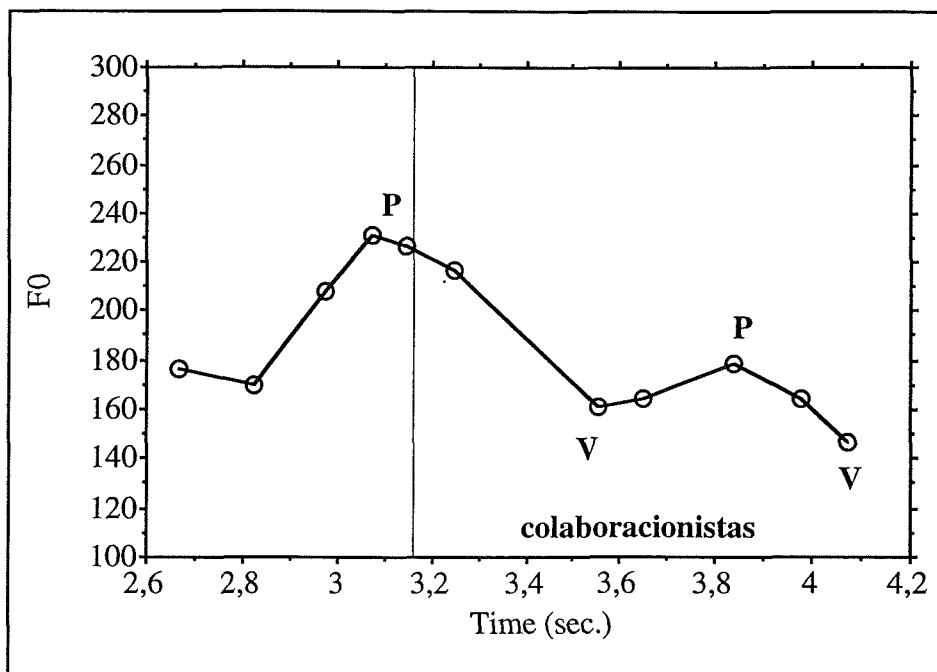
Figure 4.31. An example of intonation contour showing the final rise-fall pattern (type A) number 1 (P - Falling - V - Rising - P - Falling - V), extracted from the phonic group "*los ministros colaboracionistas*", uttered by speaker ME.

Some variations can be found from this basic pattern. The most frequent one is the insertion of a 'level' movement before the rise-fall scheme, as illustrated by figure 4.32:

Figure 4.32. An example of intonation contour showing the final rise-fall pattern (type A) number 4 (P - Falling - V - Level - V - Rising - P - Falling - V). It is included in the phonic group, "*censuraba al PP su actitud ante la corrupción,*", uttered by speaker AR.

The rest of patterns are less frequent, and they can be described as variants of the basic pattern.

4.5.3.3.2. Type B

Type B rise-fall patterns are characterized by the fact that the inflection point which starts the rise-fall scheme is always an M, while the end of the fall is always a V. The basic shape of these patterns is the one found in figure 4.33.

Figure 4.33. An example of intonation contour showing the final rise-fall pattern (type B) number 1 (P - Falling - M - Rising - P - Falling -V), included in the phonic group "*una escalada de atentados en los que han perdido la vida quince personas.*", uttered by speaker AR.

As in the case of type A patterns, a frequently found variant is the one that includes a level movement before the rise-fall scheme. This variant is illustrated in figure 4.34:
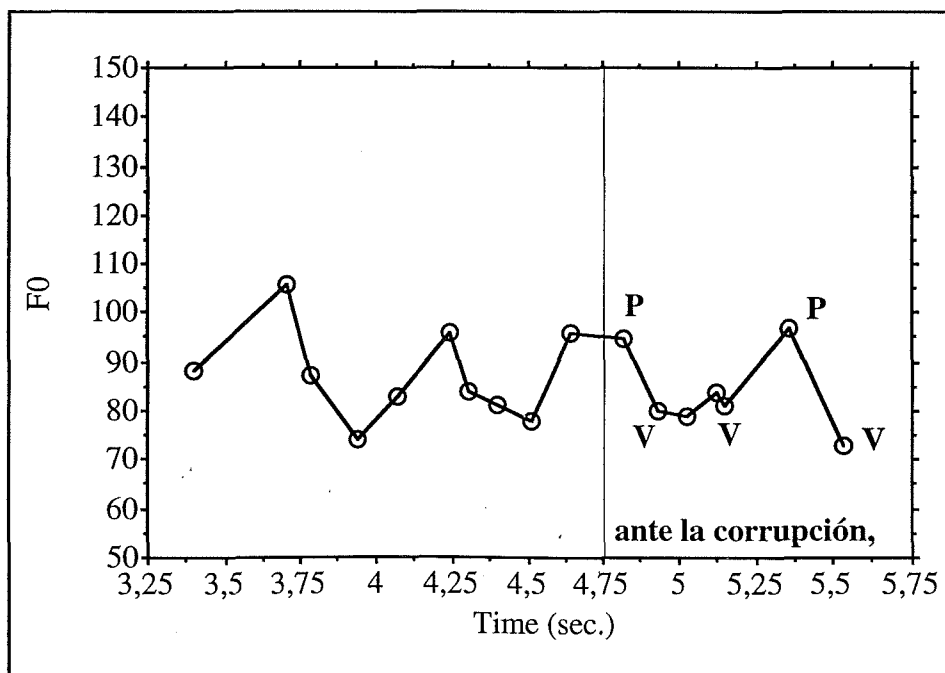
Figure 4.34. An example of intonation contour showing the final rise-fall pattern (type B) number 2 (P - Falling - M - Level - M - Rising - P - Falling - V), included in the phonic group "*¿Poseer una obra original de un artista?*", uttered by speaker AR.

4.5.3.3.3. Type C

Type C is characterized by showing a V level at the beginning of the rising movement and an M level at its end. Figure 4.35 illustrates this case, not very frequent in the analysed material:

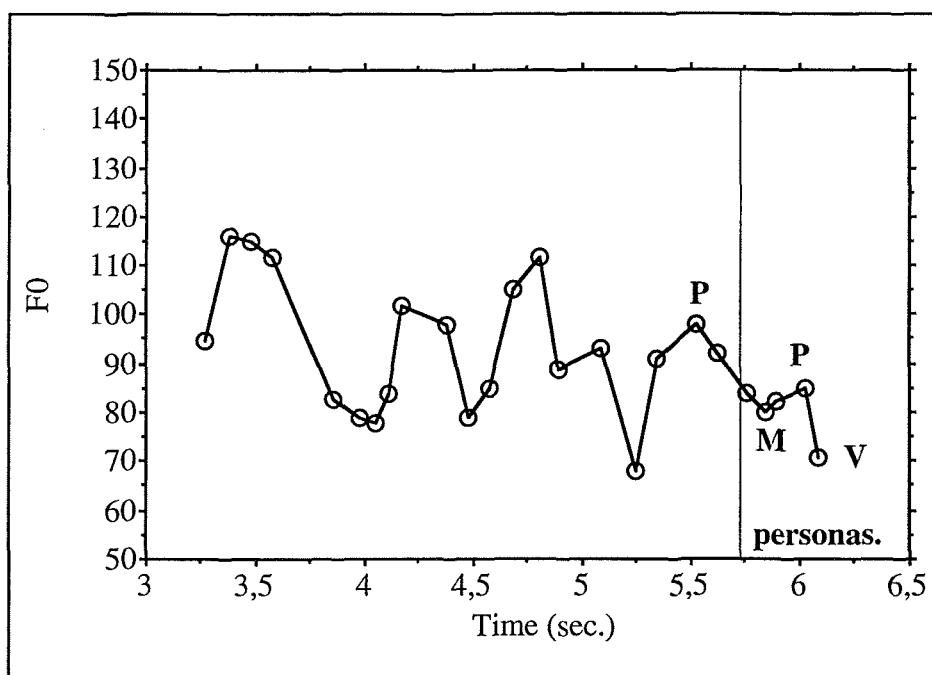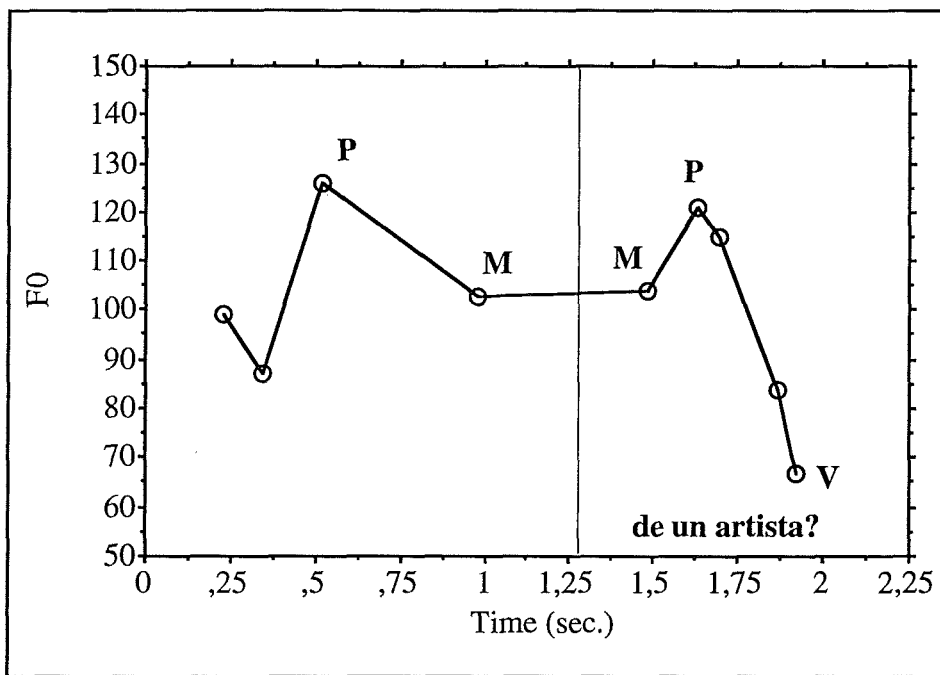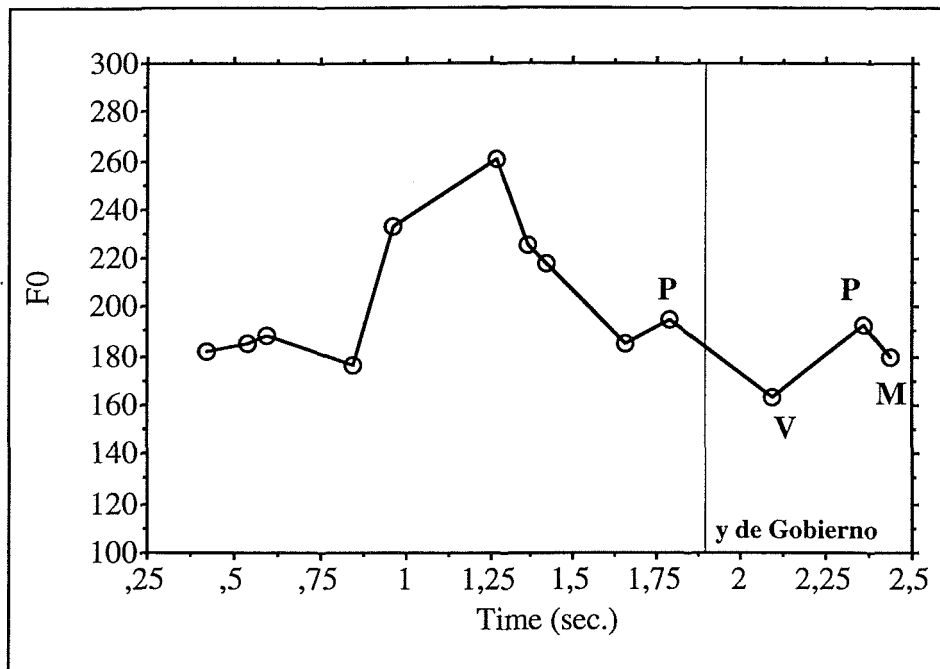Figure 4.35. An example of intonation contour showing the final rise-fall pattern (type C) number 1 (P - Falling - V - Rising - P - Falling - M). It has been taken from the phonic group "*De la reflexión de los jefes de Estado y de Gobierno*", uttered by speaker ME.

4.5.4. Summary

A set of patterns for local intonation movements has been defined in this section. They have been obtained from the analysis of the stylized contours of the selected tonic group sub-corpus. Three different types of patterns have been defined, according to their location in the phonic group. It has been shown that the different patterns in the same group share some characteristics and that they can be described in an unified way.

The shape of the patterns corresponding to each type can be expressed in the terms of a 'grammar' that defines the possible combination of movements and inflection points in the different patterns of a type. These 'grammars' are presented in appendix 3 (section 2) for each of the defined types. In the case of initial patterns, two different grammars have been established, one for the patterns showing a rise at the beginning, an a second one for those patterns which contain only a 'level' movement. Each grammar is accompanied by a table that summarizes the type and location of the corresponding inflection points.

4.6. Intonation patterns and phonetic factors: the case of tonic group length

After the definition of the patterns, the analysis of the use of these patterns according to different factors has been attempted. First the relationship between these patterns and the length of the tonic groups has been considered.

The distribution of the different patterns depending on the number of syllables of the tonic group has been analysed. The results of this analysis for some of the defined types of patterns are discussed in the following sections, although a detailed presentation of the results is left for appendix 3 (section 3).

4.6.1. Initial patterns

According to the collected data, it can be hypothesized that there seems to be some relationship between number of syllables in a tonic group and the number of movements in the corresponding pattern. In general, the longer the tonic group is, the higher is the number of movements in the chosen pattern:

a) One-syllable groups show one-movement patterns in 7 of the 8 analysed cases: pattern 2 (3 cases), pattern 4 (3 cases) and pattern 3 (1 case). Only in one case a larger pattern has been found (pattern 12).

b) In the case of the two-syllable groups, one-movement patterns are used in 6 of the 13 analysed cases (pattern 3, 3 cases; pattern 4, 3 cases). Two-movement patterns are used in 5 cases (pattern 7, 3 cases; pattern 6, 2 cases).

c) A two-movement pattern, number 7, is the most used pattern in medium-length (3-7 syllables) groups: 70 of the 116 analysed cases. A three-movement pattern, number 12, which contains a 'flat peak', appears as a less frequent alternative choice in these groups (13 cases).

d) In 8-9 syllable groups (only 6 analysed cases), three-movement patterns (pattern 12, 2 cases; pattern 13, 1 case) appear as frequently as two-movement patterns (pattern 7, 1 case; pattern 10, 2 cases).

Also, some relationship can be established between the presence of a one-syllable group in initial position and the level of the first inflection point of the corresponding pattern. Patterns used for one-syllable groups usually start with a P (pattern 2; 3 cases) or M level (pattern 4; 3 cases).

4.6.2. Internal patterns

One-syllable groups do not have an associated pattern to them; they 'share' the pattern corresponding to the previous or following group. For this reason, they are not included in the results presented in appendix 3 (section 3.2.2).

The analysis of the results reveals that there seems to be a tendency to use shorter patterns with shorter groups, as in the case of initial patterns:

a) Two-movement patterns are the most frequent ones in tonic groups ranging from 2 to 5 syllables: they are used in 34 of the 53 analysed cases (number 1, 15 cases; number 5, 9 cases; number 2, 7 cases; number 4, 3 cases) Three-movement patterns are used in 18 cases, less than the two-movement patterns (number 7, 6 cases; number 6, 5 cases; pattern 10, 5 cases; pattern 8, 1 case; pattern 11, 1 case).

b) In the case of long groups (6 syllables or more), three-movement patterns are the most frequent ones: they appear in 15 of the 27 analysed cases (number 6, 7 cases; number 7, 3 cases; number 13, 2 cases). Two-movement patterns are used in 11 cases (number 1, 7 cases).

The number of movements in the second part of the pattern (after the rise) does not seem to be dependent on group length. In other words, the use of patterns containing 'flat peaks' at the end, or 'rising-level' combinations, (numbers 10-15, 13 cases), does not seem to be related to group length. They are much less frequent than patterns containing 'pointed peaks' (patterns 1-9, 67 cases).

The number of movements in the first part of the pattern does seem to be dependent on group length. In 2-5 syllable groups, patterns containing only movement in their first part (numbers 1, 2, 3, 4, 5, 10, 11 and 12) appear in 39 of the 53 considered cases, while patterns containing two patterns in the first part (numbers, 6, 7, 8, 9, 13, 14 and 15) are used in the remaining 14 cases. However, in the case of 6-8 syllable groups, one-movement patterns are used only 13 times, while two-movement patterns are used in 14 cases.

4.6.3. Final patterns

4.6.3.1. Falling patterns

Pattern 3 (P-falling-V-level-V) is in general the most used one (30 of 56 considered cases). Also, as in the previous types, some relationship can be

suggested between length of the group and number of movements of the pattern:

a) One-movement patterns are mainly used with rather short groups: 11 of the 13 registered cases of these movements appear in groups ranging from 1 to 4 syllables (pattern 1, 7 cases; pattern 2, 4 cases).
b) Complex patterns (3-4 movements) are used in long groups: the 6 registered cases appear in groups ranging from 3 to 6 syllables (number 6, 5 cases; number 7, 1 case).

4.6.3.2. Rising patterns

In this case, the results do not indicate a clear relationship between length of the group and number of movements of the patterns, as can be inferred from the data of type A patterns:

a) In the case of groups ranging from 2 to 4 syllables, there is not a clear preference for any type of patterns: two-movement patterns appear in 8 of the 21 analysed cases (number 1, 5 cases; number 2, 2 cases; number 3, 1 case); three-movement patterns also appear in 8 cases (number 4, 2 cases; number 5, 4 cases; number 6, 2 cases); and longer patterns (numbers 7-9) appear in 5 cases.

c) In 5-7 syllable groups, there is not a clear preference for one type or another: two-movement patterns appear in 8 of the 18 cases found in this range (number 1, 5 cases; number 2, 2 cases; number 3, 1 case); three-movement patterns also appear in 8 cases (number 4, 5 cases; number 5, 3 cases); and longer patterns appear only in 2 cases.

As was reported in the case of the internal patterns, one-syllable groups do not have their own associated pattern: they are merged with the previous group, and they both share the same pattern.

As in previous cases, the use of patterns containing 'flat peaks' (numbers 8-10 of the type A set, 7 cases) does not seem to be related to length factors. They appear as a less frequent option for 'pointed peak' patterns (numbers 1-7 of type A, 33 cases).

4.6.3.3. Rise-Fall patterns

Some relationship between length of the group and number of movements of the pattern seems to appear again in this case. This is illustrated by the data obtained for type A patterns:

a) In 2-4 syllable groups, three-movement patterns are the most frequent ones: 7 of the 10 cases considered in this range (number 1, 2 cases; number 2, 3 cases; number 3, 1 case; number 9, 1 case). Patterns containing 4 or 5 movements appear only in 3 cases (number 4, 1 case; number 7, 1 case; number 8, 1 case).

b) In 5-7 syllable groups, patterns containing 4 or 5 movements are slightly more frequent than three-movement patterns: they appear in 8 of the 15 considered cases (number 4, 5 cases; number 5, 1 case; number 6, 1 case; number 8, 1 case); three-movement patterns appear in 7 cases (number 1, 5 cases; number 3, 2 cases).

As in the case of internal patterns, this influence seems to be concentrated in the first part of the patterns, before the rise-fall combination: longer groups tend to show patterns containing two movements (falling-level) in this first part, while shorter groups show in general only one falling movement. If type A data are analysed, it can be observed that patterns showing only one movement in their first part are much more frequent (numbers 1, 2, 3 and 6; 6 of 9 considered cases) than patterns showing two movements (numbers 4, 5, 7 and 8) in 2-4 syllable groups. However, the number of cases in both sets is more balanced in 5-7 syllable groups: 8 cases of patterns showing one movement, and 7 cases of patterns showing two movements.

4.6.4. Summary

In general, the results of this analysis suggest that it exists some type of relationship between group length and shape of the patterns. The data seem also to indicate that this influence is concentrated in the first part of the patterns: patterns having a single 'falling' movement in their first part can be used in general with groups of any length, although they are more likely to appear with short groups; inversely, patterns showing 'falling-level' combinations at the beginning are more likely to occur with long tonic groups, but they can also be used in short groups. The limit for this 'short-long' distinction seems to be placed around the 4-5 syllable boundary.

An analysis of the data in terms of distance between stressed syllables instead of tonic group length could be necessary to extract more precise conclusions about this question.

4.7. Relating intonation patterns to linguistic functions

4.7.1. Local F0 movements and stress location

As was discussed in chapter 1, the location of lexical stress is one of the factors that determine the shape of local intonation patterns, although no exhaustive analysis of its influence is available yet. Some questions related to this relationship still remain unclear. However, these factors are very important in the definition of a grammar of use of local intonation patterns. In this work, two of these questions have been considered in more detail:

a) the problem of the 'peak shift' or 'peak delay:'

As has already been reported in previous studies, and it has also been noted above (sections 4.5.1 and 4.5.2.2), there is a tendency in Spanish to find intonation peaks shifted with respect to the stressed syllable. The data obtained in this study have been analysed in order to give some additional information about this phenomenon.

b) the relationship between stress placement and final patterns:

Final intonation patterns have been proposed in previous descriptions of Spanish intonation. However, less attention has been paid to the question of the variation of these patterns in relation to the location of the stressed syllable at the end of the group. This question has been considered using the data obtained in this work.

The results of the analysis of these two questions are presented in the following sections.

4.7.1.1. Peak shift in initial and internal patterns

The descriptions of non-final local patterns in Spanish available at this moment (see section 1.2.3.1.) seem to indicate that the position of the stressed syllable determines the beginning of the rising part. Fant (1984) proposes a model that places the beginning of the rise in the boundary between the stressed syllable and the preceding one. Prieto et al. (1995b) have also reported the location of the beginning of the rise in the syllable onset. The data obtained in the present study seem to confirm these statements (see sections 4.5.1 and 4.5.2.2).

However, as far as the location of the end of the rising movement (the 'peak') is concerned, there seems to be less agreement in the literature. Fant (1984) considers that end of the rising movement coincides with the end of