

Speech, Hearing and Language: work in progress

Volume 11

The intermediate phrase in Central Catalan declaratives: a case for questioning the representation of downstep

Eva ESTEBAS I VILAPLANA and John A. MAIDMENT



**Department of Phonetics and Linguistics
UNIVERSITY COLLEGE LONDON**

The intermediate phrase in Central Catalan declaratives: a case for questioning the representation of downstep

Eva ESTEBAS I VILAPLANA and John A. MAIDMENT

Abstract

This paper examines two aspects of the intonation of S(ubject) V(erb) O(bject) Central Catalan declaratives produced in reading speech. First, it deals with the identification of an intermediate level of prosodic phrasing in Central Catalan declaratives. Second, it analyses the immediate implications of this intermediate phrase on the phonological representation of the F₀ contours and, in particular, on the interpretation of downstep. Three different cues are used for the identification of the prosodic boundaries: a pause, a local F₀ fall, and the lengthening of the boundary syllable. In all sentences, an intermediate level of prosodic structure, marked with a H- phrase accent, is observed between the subject and the verb. The tonal representation of the sentences is determined through both an auditory analysis and an acoustic analysis of the data. A pitch reset is observed at the beginning of the second intermediate phrase, which starts with a drastic lowering on the peak of the first pitch accent. Evidence for treating this lowering as an intended downstep movement is provided with the analysis of speaking rate differences.

1. Introduction

1.1 Theoretical framework

The data examined in this study are analysed within a metrical-autosegmental framework of intonational analysis. In particular, we follow the approach proposed by Pierrehumbert (1980), along with Ladd's (1983) first criticisms to the model and some of the subsequent modifications (Beckman and Pierrehumbert 1986). This approach is chosen since it has been proved to be a model with a wide application¹. In this section, the basic ideas of Pierrehumbert's phonological theory for the description of English F₀ contours are presented. Two issues are described in greater detail, namely, the identification of an intermediate level of prosodic phrasing and the phenomenon of downstep.

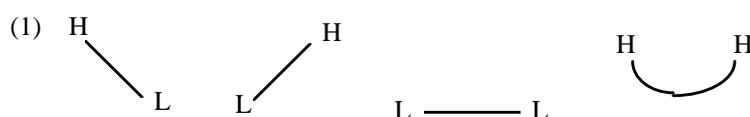
Pierrehumbert's theory involves the identification of three kinds of tonal events or phonological categories in the F₀ contours of English utterances: pitch accents, phrase accents and boundary tones. Pitch accents are produced in relation to the rhythmic pattern of the phrase. They are associated to metrically strong (stressed) syllables, which become accented by virtue of this association. Boundary tones and phrase accents are not lined up to the metrical structure of the utterance. Boundary tones are associated to the edges of the intonation phrase. In early versions of the theory, phrase accents are described as free-standing tones which account for the F₀ contour between the last pitch accent and the rightmost boundary tone. Later, phrase accents are used to

¹ So far, Pierrehumbert's theoretical framework has successfully described the F₀ contours of many languages, such as English (Pierrehumbert 1980), Japanese (Beckman and Pierrehumbert 1986, Pierrehumbert and Beckman 1988), Spanish (Sosa 1991), Italian (Grice 1992), German (Grice and Benz Müller 1995), and Catalan (Prieto 1995).

mark an intermediate level of prosodic phrasing (more details on this issue are provided later in this section). All these accent-types are characterised by means of two tone levels, H(igh) and L(ow).

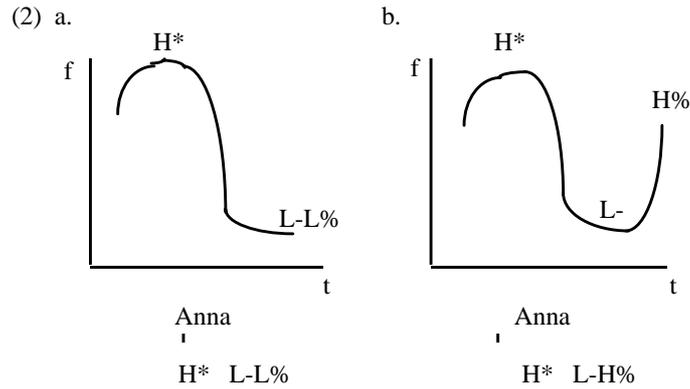
Pitch accents may consist of a single tone (H*, L*) or a bitonal tone (L*+H, L+H*, H*+L, H+L*)². Single H* is realised as a local peak in the F0 contour. L* involves a local trough. In bitonal pitch accents, the starred tone is associated to the accented syllable. The tones that lead or trail the starred one are usually realised in the syllables that precede or follow the accented tone. For example, L*+H consists of a low contour over the major part of the accented syllable followed by a sharp rise over the next unstressed syllable if there is one. Boundary tones are identified by means of the % symbol, i.e. H% and L%. Although they occur both at the initial edge and at the final edge of the intonation phrase, in initial position only H boundary tones are specified, since L% is considered the unmarked choice. Finally, phrase accents are represented as H- or L-.

In contrast to accented syllables, the intonational behaviour of unstressed syllables is not specified since it is not relevant phonologically. However, it can be predicted by means of interpolation rules, which account for the pitch transitions between two tones at the phonetic level. The interpolation between H and L tones or between two L tones involves a linear trajectory. The interpolation between two H tones consists of a sagging transition. This is shown in (1) below.

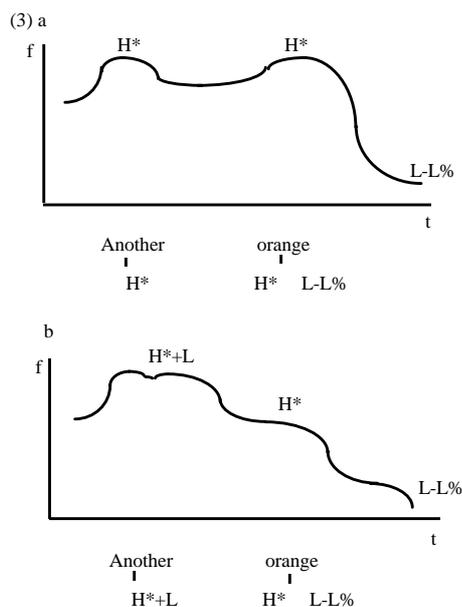


Examples of Pierrehumbert's analysis, taken from her first version (Pierrehumbert 1980), are reproduced in (2) and (3). (2a) and (2b) illustrate two different F0 contours for the word “Anna”. The metrical structure of this word consists of a strong (stressed) syllable (“An-”) followed by a weak (unstressed) syllable (“-na”). In both examples, the stressed syllable is associated to a H* pitch accent, characterised by a local peak in the F0 contour. The main difference between (2a) and (2b) lies in the behaviour of the following phrase accent and boundary tone. In (2a) both tonal events are low, indicating a fall in the F0 contour. (2b), on the other hand, shows a L- phrase accent followed by a H% boundary tone. This configuration accounts for the traditional fall-rise contour.

² In Pierrehumbert (1980) the H*+H bitonal accent was also proposed for English contours. In later versions (Beckman and Pierrehumbert 1986) this accent was dropped.



Examples (3a) and (3b) illustrate some differences between single and bitonal pitch accents. In (3a) the two stressed syllables of the phrase “another orange” are associated to a H* pitch accent. The transition between the two H tones shows a sag in the F0 contour, as predicted by interpolation rules. In (3b), on the other hand, the first stressed syllable is associated to a bitonal pitch accent (H*+L). The starred tone is aligned within the accented syllable, whereas the trailing tone is realised over the following unstressed syllable. If we compare the F0 contour corresponding to the first H* of (3a) with that of the bitonal accent of (3b), no big differences are observed since in both cases there is a peak followed by some kind of fall. The crucial point, however, is the realisation of the second pitch accent in the two utterances. Whereas in (3a) both H tones are scaled or realised around the same frequency, in (3b) the second H is realised at a much lower frequency than the starred H tone of the bitonal accent. The phenomenon whereby two peaks or H tones appear at different scaling levels within the same pitch range is known as downstep. According to Pierrehumbert, downstep in English is always triggered by the HLH tonal sequence including a bitonal accent, that is, H+L H or H L+H. Thus, in (3b) the lowering of the second H tone is explained by the presence of the trailing L in the bitonal pitch accent.

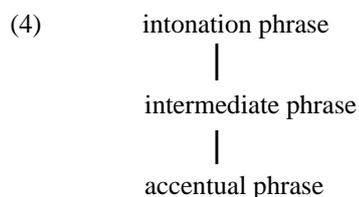


This analysis of downstep creates several problems, some of which are discussed by Pierrehumbert herself and by Ladd (1983) among others³. Basically, the main criticism lies in that the L of the H*+L is not realised as a low pitch but only acts as a downstep trigger for the following H tone. Ladd's alternative proposal to solve this problem involves treating downstep not as the phonetic consequence of a specific tonal configuration but as an independent phonological feature of the lowered tone itself. In Ladd's approach, the contours in (3a) and (3b), which in Pierrehumbert's model were represented with different sequences of tones, are treated as similar sequences that only differ by the choice of downstep. Thus, whereas (3a) is still analysed as H* H* L-L%, (3b) is analysed as H* !H* L-L%, (where ! indicates downstep). This entails the presence of two distinctive H tones in the intonational system of English, namely, a downstepped H (i.e. !H*), and a non-downstepped H (i.e. H*). These two separate entities potentially carry differences in meaning. The [+downstep] feature is applicable to all kinds of pitch accents.

Ladd's interpretation of downstep entails a reanalysis of the role of the H*+L pitch accent. By viewing downstep as a distinctive phonological feature rather than a phonetic consequence of bitonal accents, the H*+L accent has no role in the theory since it only functions as downstep triggerer of subsequent H tones. Ladd's alternative proposal is to use this tone to mark a final falling contour, such as the one illustrated in (2a) above, and consequently get rid of the phrase accent, which is a rather obscure concept in Pierrehumbert's original theory. This involves the reanalysis of the phrase final fall observed in (2a) as a H*+L pitch accent followed by a L% boundary tone and the final fall rise in (2b) as H*+L H%.

Beckman and Pierrehumbert (1986) completely rejected Ladd's proposals. First, they criticised Ladd's phonological view of downstep by claiming that in English there is no distinction between H* and !H* in certain phrase positions, such as at the beginning of the sentence. Thus, the idea of downstep being triggered by bitonality was retained. Second, Ladd's idea about the loss of the phrase accent was also abandoned due to the identification of an intermediate level of prosodic structure, which was said to be the domain of the phrase accent.

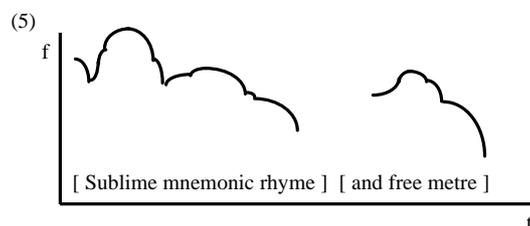
The identification of this new level of phrasing derived from the comparison of the intonational structure of English and Japanese. In Japanese, three different prosodic domains were observed: the intonation phrase, the intermediate phrase and the accentual phrase, which involved the following hierarchical structure.



³ Several studies (Grice 1992, 1995, Ladd 1983, 1993, among others) have criticised Pierrehumbert's idea of downstep. In this paper we will mainly deal with Ladd's (1983) approach.

The accentual phrase is the domain for a lexically defined pitch accent. This level of phrasing does not appear in languages such as English, in which accents are not specified by the lexicon but by the intonational system. The intermediate phrase is a larger unit of prosodic hierarchy, which may contain up to three accentual phrases in Japanese. Finally, the intonation phrase is the largest level of phrasing. It contains, at least, one intermediate phrase. Both the intermediate phrase and the intonation phrase are observed in English. Phonetically, the intermediate phrase and the intonation phrase are identified by similar cues, such as, the presence of a pause, a local F₀ fall or rise, or the lengthening of the boundary syllable. Phonologically, Beckman and Pierrehumbert proposed that the phrase accent should be used to mark the edge of the intermediate phrase, as the boundary tone signals the edge of the intonation phrase. With this proposal, Pierrehumbert's notation for the F₀ contours of (2) acquires a new theoretical dimension, since the L- and H- phrase accents are now associated to an intermediate phrase boundary instead of being free-standing tones.

With the identification of the intermediate level of phrasing, new ideas were put forward in relation to the phenomenon of downstep. Both in English and in Japanese downstep was observed to be blocked by the presence of an intermediate phrase boundary, which indicates that the intermediate phrase is the domain for downstep. An example from Beckman and Pierrehumbert (1986) is reproduced in (5) below for the phrase "sublime mnemonic rhyme and free metre". This sentence has an intermediate phrase boundary after the word "rhyme". A downstepped sequence of accents is observed within the domain of the first intermediate phrase. However, after the prosodic boundary downstep is blocked and the pitch is reset at the beginning of the second intermediate phrase.



So far, we have presented the main ideas of Pierrehumbert's model of intonational analysis along with some criticisms and subsequent modifications. We have seen that Pierrehumbert recognises two levels of intonational phrasing in English and accounts for the pitch contours of utterances by associating H and L tones to metrically accented syllables and to the edges of the prosodic domains. In this paper, this model will be used to analyse the intonational behaviour of declarative sentences in Central Catalan. Three main aspects will be taken into account: 1) the identification of different levels of prosodic structure, 2) the tonal configuration of the utterances, and 3) the relationship between downstep and the intermediate phrase boundary. Before presenting the data and the results obtained in our analysis, we will briefly review some of the traditional descriptions on the intonation of Central Catalan declaratives.

1.2 *The intonation of Central Catalan declaratives*

The intonational behaviour of declarative sentences in Central Catalan has been examined in a number of studies (Virgili Blanquet 1971, Recasens 1977, 1993, Bonet 1984, and Prieto 1995 among others). All these analyses include the intonational inspection of other kinds of syntactic structures, such as interrogatives and imperatives, along with declaratives in Central Catalan. Most of these studies describe the intonational patterns of the different sentences on a perceptual basis. Virgili Blanquet and Bonet compare some of their perceptual judgements with spectrographic representations of the utterances but no pitch trace analyses are provided. Prieto is the first study that presents an acoustic analysis of the pitch contours and interprets the intonational patterns within Pierrehumbert's model of intonational analysis. The other analyses do not use any particular framework of intonational representation and work out their own proposals for the modelling of the data. This makes the comparison between these studies rather difficult.

In the next two sections, we will reproduce some of the main ideas on the intonational behaviour of declaratives reported in the studies already mentioned. Two main issues will be taken into account, namely, the identification of levels of prosodic structure (intonation and intermediate phrases) and the phonological representation of the pitch contours in Central Catalan declaratives. Although most of these analyses account for the behaviour of both simple and complex declaratives, here we will only present the generalisations observed for simple sentences.

1.2.1 *The identification of levels of prosodic structure*

Except for Prieto (1995), all the studies on Catalan intonation reviewed in this paper deal with the division of declaratives into different tone units. Although in most cases no clear definition of "tone unit" is provided, here we will interpret it as a prosodic level which is similar to Pierrehumbert's notion of "intermediate phrase". In most of the analyses of Catalan intonation a correlation between tone units and syntactic constituents is proposed. This means that a prosodic boundary will always coincide with a grammatical boundary. However, not all grammatical boundaries will necessarily have an intonational boundary, since the division into different tone units is optional and depends on many extra-linguistic factors. For example, in a S(subject) V(erb) O(bject) declarative structure of the kind presented in (6), three possible ways of intonational division can be observed.

- (6) S V O
 La minyona menuda netejava la nevera
 (The small maid cleaned the fridge)
- a. La minyona menuda / netejava la nevera
 - b. La minyona menuda netejava / la nevera
 - c. La minyona menuda / netejava / la nevera

It is important to point out that most of the studies on Catalan intonation agree that it is very unlikely that a sentence such as (6) would be uttered as a single tone unit.

Bonet (1984, p:11) explicitly says that there is always some kind of intonational division among the constituents of a simple declarative sentence in Catalan. Recasens (1993), in his turn, analyses all SVO declaratives with an intonational boundary between the subject and the verb (as in (6a) above). Although he does not explicitly define this behaviour as the most natural or regular pattern in Central Catalan declaratives, his continual use of examples consisting of two tone groups is rather revealing.

Despite the general feeling that Central Catalan declaratives are uttered with some kind of intonational boundary, very few observations have been made on how these boundaries are realised. Recasens (1993) suggests that simple declaratives do not show a pause as the boundary marker for its different tone units. According to him, pauses are only used in complex structures to demarcate sentence boundaries. However, no alternative cue is provided to identify tone units within simple structures.

So far, the observations gathered in these studies have shown that the division into tone units (intermediate phrases in Pierrehumbert's terminology) seems to be a crucial issue for the intonational analysis of simple declaratives in Central Catalan. However, more experimental research is needed in order to see whether there is a more natural or neutral pattern of prosodic phrasing in Central Catalan declaratives, as well as to identify the main cues used to demarcate the prosodic boundaries.

1.2.2 The intonational representation

As far as the intonational representation of declaratives in Central Catalan is concerned, one point of agreement appears in all the aforementioned studies, that is, the falling tone or low level at the end of the sentence. Recasens (1977) is the first study which clearly specifies this behaviour. Although in Virgili Blanquet (1971) the same intonational tendency is observed, no final generalisations are proposed. The intonational characteristics of the initial part/s of declarative sentences are not explicitly stated till Bonet (1984).

Bonet's model of intonational representation involves the distinction between two kinds of stressed syllables, which she classifies as "basic" and "non-basic". Basic syllables are those syllables which have both a lexical stress and a rhythmic stress. Lexical stress refers to those syllables which are marked as stressed in the lexicon or dictionary and which have a potential for accent in utterances. Rhythmic stress, in its turn, refers to those syllables that are prominent to keep the rhythmic pattern in utterances. Syllables that do not have a lexical stress might get a rhythmic stress to preserve rhythmic regularity within an utterance. In Bonet's terminology, syllables with only a rhythmic stress are called non-basic syllables. Non-basic syllables do not get an accent. Basic syllables, on the other hand, are accented. According to this distinction, Bonet proposes the following generalisations to account for the intonation of declaratives in Central Catalan. First, all basic syllables of the last constituent are low. Second, the last basic syllable of non-final constituents is also low. Finally, all basic syllables that do not belong to the last constituent and are not the last basic syllable of their own constituent are high. An example of Bonet's proposal is illustrated in (7) for the sentence presented in (6a) above, consisting of two tone units. Basic syllables are underlined.

(9) La netejava (she cleaned it)

H* L-L%

(10) Netejava la nevera (she cleaned the fridge)

H* L* L-L%

No generalisations are made for the intonational behaviour of declaratives with more than two lexical stresses. The only observation proposed in this study is that not all syllables with a lexical stress acquire an accent. The example she gives is reproduced in (11) below⁴.

(11) L+H* H* H* H* L-L*

La meva germana va arribar ahir a la nit en l'ultim avio d'Alemanya

(My sister arrived yesterday night on the last plane from Germany)

Although Prieto's study represents a step forward in the analysis of Catalan intonation due to its being both experimentally and theoretically-based, there are still problems in her proposal. First, Prieto ignores most of the observations made in Bonet (1984) or Recasens (1993) concerning the division into tone units. The fact that she never considers the possibility of having intermediate phrases within an intonation phrase, makes it more difficult for her to find generalisations about the intonational behaviour of multi-stressed sentences. Second, since her analysis is based on a very small number of sentences, some of the descriptions provided seem rather inconsistent with one another. For example, the intonational behaviour of the final pitch accent and edge tones in (11) is more similar to sentences with one lexical stress, as the one in (9), than to sentences with two lexical stresses, as the one in (10). Although we do not question the accuracy of this interpretation, we think that it should be possible to account for the neutral pronunciation of declaratives in Central Catalan in a similar way.

In this paper, SVO declarative sentences in Central Catalan are acoustically analysed and modelled according to Pierrehumbert's framework of intonational representation. We expect to find some generalisations on the identification of prosodic domains and on the phonological representation of the pitch contours. Finally, the relationship between downstep and the intermediate phrase boundary in Central Catalan declaratives is also examined.

⁴ Since no generalisations are given, it is impossible to accommodate Prieto's description to our examples.

2. Method

2.1 Collection of data

The data included 64 declarative affirmative sentences. They all had the same structure: S(subject) V(erb) O(bject). Each constituent (S, V, O) presented two kinds of stress patterns either with one or with two lexically stressed words (i.e. content words with stress marked in the lexicon). All sentences were designed with the greatest number of voiced sounds possible in order to avoid gaps in the pitch traces. Utterances were elicited by means of a reading task. Subjects were asked to read the sentences in a natural way. Six female native speakers of Central Catalan were recorded. They all had similar characteristics as far as age, geographical origin, social status and education are concerned. The speakers were identified as NG, NM, ER, CP, MC, and DV. Overall, 384 sentences were analysed.

2.2 Instrumentation

Subjects were recorded in a soundproof room. Speech and laryngeal signals were obtained for each of the sentences. For the recording of the speech signal a microphone and a Marantz tape recorder were used. The microphone was set at about twenty centimetres from the speaker. The laryngeal activity was monitored by means of a laryngograph. A pair of electrodes was placed on the subjects' neck to record the vocal fold vibration. An oscilloscope was also used to check the laryngeal activity.

2.3 Analysis of the data

The data were analysed by means of the Speech Filling System program, which allowed us to carry out a time-aligned inspection of the speech waveform and the pitch traces. In order to obtain the pitch traces, the laryngograph signal was first turned into excitation period measurements and finally converted to a fundamental frequency trace by means of the VTX and FX programs.

3. Results and discussion

3.1 The identification of intermediate phrases

The intonational behaviour of SVO declarative affirmative sentences in Central Catalan produced in reading speech presented three different patterns as far as the identification of intermediate phrases is concerned. Pattern 1 consisted of utterances with two clear intermediate phrases, which were identified by the presence of a pause or silence between the subject and the predicate accompanied by an abrupt F0 fall at the beginning of the second phrase. This behaviour is illustrated in Figure 1 for the sentence *Algunes minyones guarden mala lluna* "some maids have a bad temper" produced by speaker NG. A pause is observed between the subject (*algunes minyones*) and the predicate (*guarden mala lluna*). The second intonational group starts with an F0 fall.

The second kind of pattern also involves the appearance of two intermediate phrases. In this case, however, no pause is observed between the two constituents and the only cue that allows us to identify the two phrases is the abrupt F0 fall at the beginning of the second group. An example is provided in Figure 2 for the sentence *L'home venia llimones madures* "the man sold ripe lemons" produced by speaker ER. Despite the

absence of a silence between the subject (*l'home*) and the predicate (*venia llimones madures*), the abrupt lowering of F0 allows us to identify the beginning of the second intermediate phrase.

Finally, pattern 3 showed no clear division of intermediate phrases in the sentence, since neither a pause (followed by a local F0 fall) nor the F0 cue alone were observed between the subject and the predicate. This is illustrated in Figure 3 for the sentence *En Jaume mou la galleda* "James moves the bucket" produced by speaker MC. Here, the absence of a silence and an F0 fall seems to indicate that this sentence consists of a single intonational group.

The number and percentage of occurrences of these three patterns for each speaker is provided in Table 1. A clear-cut division is observed between the behaviour of the first three speakers versus the behaviour of the last three speakers. The first three speakers produced most of the sentences with a clear boundary between the subject and the predicate (patterns 1 and 2). Pattern 1, which involves the presence of both the silence and the F0 fall as boundary markers, was highly used by speakers NM and ER. Speaker NG produced most of the sentences with pattern 2, which demarcates the two intermediate phrases by means of the F0 fall cue only. Pattern 3, which shows no clear division of intonation groups within the sentence, was not used by these speakers. The second group of informants, on the other hand, seemed to prefer pattern 3 over patterns 1 and 2. This is mainly the case for speakers CP and MC. Speaker DV, in her turn, shows a fairly similar use of the three patterns in her production.

Speakers	Number of occurrences			Percentage of occurrences		
	Pattern 1	Pattern 2	Pattern 3	Pattern 1	Pattern 2	Pattern 3
NG	17	47	-	26.6	74.4	-
NM	53	7	3	84.1	11.1	4.8
ER	48	15	1	75.0	23.4	1.6
CP	-	3	61	-	4.7	95.3
MC	7	18	39	10.9	28.1	61.0
DV	17	19	28	26.2	29.7	43.7

Table 1. Number and percentage of occurrences of the three intonational patterns in SVO declarative sentences in Central Catalan for each speaker.

The three-way division observed in our data can be further classified into two main kinds of intonational behaviour as far as the identification of intermediate phrases is concerned. Type A involves utterances that exhibit two clear intermediate phrases, identified by both a pause and a lowering in pitch or by the pitch cue only (patterns 1 and 2). Type B consists of utterances produced with a single phrase, since neither of the two cues is present (pattern 3). The grouping of patterns 1 and 2 into a single category as opposed to pattern 3 is justified by the consistent occurrence of these first two patterns in the productions of speakers NG, NM and ER, who hardly ever use pattern 3. However, two questions arise from these results. First, how is it that, given the identical circumstances under which the elicitation of the data were carried out

different intonational behaviours are observed among the speakers? Second, how can we explain the fact that the same speaker (see DV or MC for example) sometimes produces sentences with two clear intermediate phrases and other times produces sentences with only one phrase?

As seen in section 1.2.1, some descriptions on Catalan intonation (Bonet 1984, Recasens 1993) have usually proposed the identification of two (or more) tone units in Catalan declarative sentences. This is in line with the behaviour observed in Type A utterances (patterns 1 and 2) but not in Type B (pattern 3). The description provided in previous works on Catalan intonation together with the above questions led us to analyse the intonation of Type B utterances in further detail. One main question was addressed: could Type B utterances also be interpreted as having two intermediate phrases like Type A?

In order to determine whether Type B utterances could also be interpreted as having two intermediate phrases, further evidence, in addition the pause and low pitch cues, was required. The criterion we examined was that of syllable duration proposed by Cruttenden (1986). He claimed that the final syllable of an intonation group is lengthened, regardless of whether it is stressed or unstressed. This means that boundary syllables are expected to be longer than comparable syllables (with similar stress pattern and segmental structure) elsewhere within the same utterance. So, in order to determine whether Type B sentences can also be interpreted as having two separate intermediate phrases, we compare the duration of the syllable before the putative boundary with the duration of other similar syllables within the same utterance. If the duration of the syllable before the putative intermediate phrase boundary is significantly longer than these other syllables, we can posit the presence of two intermediate phrases in these utterances. In order to check this hypothesis both Type A and Type B sentences were analysed. Type A sentences were mainly included to confirm the validity of the syllable duration criterion. Boundary syllables were expected to be consistently longer than other syllables in this group of sentences.

The following points were measured in each sentence:

1. The duration of the final syllable before the first intermediate phrase in Type A sentences and before the putative intermediate phrase in Type B sentences (overall 383 syllables were analysed);
2. The duration of two similar non boundary syllables within the same sentence (overall 766 syllables were analysed);
3. The duration of the silence in Type A sentences with a pause boundary marker;
4. The duration of the whole utterance;
5. The number of syllables per utterance.

Three criteria were used for the selection of non boundary syllables. First, the comparable syllables had to have the same degree of lexical stress as the target one. If the syllable before the (putative) first intermediate phrase was stressed, two other stressed syllables were measured. If the target was unstressed, similar comparable syllables were analysed. The percentage of unstressed syllables (84.38%) was much higher than that of stressed syllables (15.62%). Second, the closest possible segmental structure was taken into account, both in the number and in the quality of the

segments. Finally, syllables with a word position similar to the target one, namely, word-final syllables were selected over syllables in different word positions. The last syllable of the whole sentence was never included since it occurs immediately before a boundary and hence is potentially longer than the others. It was considered desirable to measure two non boundary syllables since the segmental structure of the syllables was not always identical. An average duration of the two measurements was obtained for comparison with the duration of the target syllable.

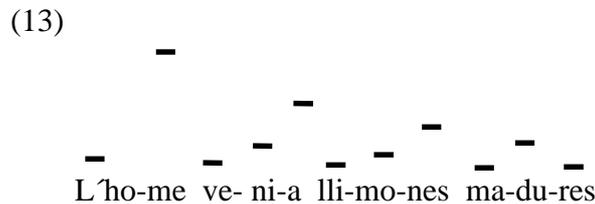
Once all the measurements were obtained and before the statistical analysis was carried out, the results were normalised to make inter-speaker comparisons possible. Normalisation was done by dividing the duration of a particular syllable (ds) by the average duration per syllable for the whole utterance (dav). To obtain the dav we divided the time of the whole utterance by the number of audible syllables in the utterance. Audible syllables were taken into account instead of written syllables since speakers varied in the number of syllables for the same sentence. The main source of variation derived from the pronunciation of [i]+vowel sequences as hiatus or diphthongs. Traditionally, these sequences have been analysed as hiatus. Recasens (1993), however, points out the increasing tendency by Catalan speakers to produce them as diphthongs. In our data, both realisations were observed.

In those sentences of Type A with a pause as boundary marker (pattern 1), the duration of the silence was subtracted from the duration of the whole sentence before obtaining the average syllable duration. Pauses were excluded for comparative purposes. In particular, we wanted to obtain a comparable dav in all cases, based on the duration of the segmental string only.

A normalised syllable duration was calculated 1) for the final syllable of the first intermediate phrase in Type A and for the final syllable before the putative boundary in Type B, and 2) for the average duration of the two non boundary syllables. A t-test was performed on the results so as to see whether the normalised duration of boundary syllables was significantly longer than the normalised duration of non boundary syllables. In Type A sentences differences in syllable duration were expected to be significant, since the target syllable appears in clear boundary position. In Type B sentences, significant differences between the two kinds of syllables would lead us to analyse this group of sentences in the same way as Type A, with two separate intermediate phrases. Non-significant differences in Type B sentences would suggest the absence of a boundary between the subject and the predicate and hence a different interpretation from Type A sentences.

The results are presented in Table 2 which shows: 1) the number of sentences analysed in Types A and B, 2) the mean value of normalised boundary syllable duration versus normalised non boundary syllable duration in Type A, 3) the mean value of normalised putative boundary syllable duration versus normalised non boundary syllable duration in Type B, 4) the t-value and 5) the probability of significance obtained in the t-test. The significance level used was $p < 0.01$.

llimones madures, we carried out the analysis illustrated in (13). The distance between lines is relative and was decided according to our auditory judgements.



After the auditory analysis of the utterances, we proceeded to perform an acoustic analysis of the data. This was done by means of the SFS program which allowed us to carry out a time-aligned inspection of two signals: speech waveform and Lx. For each sentence, we marked: 1) the onset (on) and the offset (of) of each accented syllable, and 2) the realisation of peaks (p) in the contour in relation to the limits of the accented syllables. Figure 4 shows the speech waveform and the F0 trace for the sentence in (13). The limits of the syllable and the peaks are marked.

The phonological analysis of the data was performed in line with Pierrehumbert's model of intonational representation (see section 1.1 for more details). So far, the results provided in this paper are still provisional since some of the utterances need further research.

At this stage of investigation, the tonal representation of Central Catalan declaratives has the following characteristics. The whole intonation phrase is divided into two intermediate phrases. The end of the first intermediate phrase is signalled by a high phrase accent (H-). A L- phrase accent marks the end of the second intermediate phrase. Based on this behaviour, we can state that phrase accents in Central Catalan coincide with the intermediate phrase. The whole intonation phrase ends with a low boundary tone (L%).

Usually there is a pitch accent associated to every stressed, and hence accented, syllable. This is illustrated in Figures 4 and 5 which reproduce the same sentences as in Figures 2 and 1 respectively. However, in some utterances (especially those produced at fast speech), not all stresses become accented. This can be seen in Figure 6 (the same as Figure 3), where only two of the three stresses are associated to a pitch accent. In all figures stresses are underlined.

Whenever there is a pitch accent, the tone representation presents the pattern in (14). The tones in brackets correspond to the initial pitch accent or accents of double stressed subjects and predicates. The whole intonation phrase is divided into two intermediate phrases. The end of the first intermediate phrase is signalled by a high phrase accent (H-). A L- phrase accent marks the end of the second intermediate phrase. The whole intonation phrase ends with a low boundary tone (L%).

$$(14) (H^*+L) L^*+H \quad H- \quad (L^*+H) H^* L-L\%$$

The last accented syllable of the first intermediate phrase is always associated to a L*+H tone. This accent involves a low contour in the major portion of the accented syllable followed by a rise which usually occurs in the following syllable if there is one. The same pitch accent in a similar phrasal position was reported in García-Lecumberri et al. (1997) for Peninsular Spanish. On the contrary, Prieto, et al. (1994, 1995) in their analysis of Mexican Spanish proposed for a similar contour a H* pitch accent with a peak delay. Although further research is expected to be done on this possibility, so far we consider it more appropriate the use of L*+H since it better describes the valleys attained in the F0 contour during the accented syllable.

The last accented syllable of the second intermediate phrase is associated to a H* pitch accent. This involves a peak in the F0 contour. Although in our data this peak sometimes reaches rather low levels, we still analyse it as a H* which, together with the following L-L%, accounts for the fall at the end of the intonation phrase. The drastic lowering of the last peak in an intonation phrase has been reported in other languages, such as Japanese (Pierrehumbert and Beckman, 1988), Danish (Thorsen, 1981), German (Möbius, 1993) and English (Lieberman and Pierrehumbert, 1984).

The first syllable of a double stressed subject domain is associated to a H*+L pitch accent. This tone starts with a peak over the accented syllable followed by an abrupt fall over the next syllables. This tone is clearly illustrated in the first word of Figure 5. Finally, if there is any pitch accent preceding the last tone of the second intermediate phrase, this accent is L*+H. This tone is the same as the one observed at the end of the first intermediate phrase.

So far, the intonational behaviour observed in SVO declaratives in Central Catalan partially agrees with the behaviour reported in other studies. As far as the representation of the first intermediate phrase is concerned, our findings confirm the final rise at the end of the first intermediate phrase suggested by Recasens (1993). Bonet's (1984) idea of this rise taking place only in certain circumstances is thrown into doubt, since a H- phrase accent consistently appears in all sentences. Bonet's idea of a L tone in the last syllable of the first intermediate phrase is rather consistent with our observations. Her proposal of a H tone over the accented syllable/s preceding the last accent is slightly modified. We propose a H*+L accent in this position. This pitch accent accounts for the immediate fall over the unstressed syllable/s following the accented one.

The intonational behaviour of the second intermediate phrase does not entirely agree with the proposals of other studies on Catalan intonation. Our representation mainly differs from these analysis in the way of accounting for the falling contour at the end of the utterance. Most of these studies propose a low tone (Bonet 1998) or a mid tone (Recasens 1993) over the last pitch accent. Our analysis, however, proposes a H* tone for the last accented syllable. The fact that this tone reaches rather low levels is explained by the phenomenon of final lowering observed in many languages. As for the initial syllables of the second intermediate phrase, our analysis partially agrees with the one proposed by Bonet (1984). She proposes a low tone for all the accented syllables of the second intermediate phrase and claims that the syllables surrounding these low accents are realised at a slightly higher pitch. This description coincides with the results observed in our data. The main problem with Bonet's analysis is the phonetic interpretation of the rising movement observed in the transition between two

L tones. This interpretation violates the universal rules of interpolation since the transition between two L tones can never involve a rise. We account for this rise in phonological terms by means of the L*+H. Our results also seem to agree with some of the descriptions proposed by Prieto (1995) for Catalan. However, her lack of generalisations makes comparisons with this study very difficult.

3.3 Downstep and the intermediate phrase boundary

In the last section of the results we will examine one final characteristic about the intonation of SVO Central Catalan declaratives concerning the phenomenon of downstep. Again, the observations presented in this part are provisional. Nevertheless, we hope to expose some of the problems encountered in the analysis of our data, as well as one preliminary solution and possible line of future research.

If we look at the F0 contour of Figures 4 and 5, we observe that the peaks of the second intermediate phrase are downstepped with respect to the previous one. That is each peak is produced at a much lower frequency than its antecedent. As proposed in Pierrehumbert (1980) and in Beckman and Pierrehumbert (1986), this behaviour can be explained by the presence of a bitonal accent which triggers the lowering down of the following tone. In the first intermediate phrase downstep is not observed since in our data there is only a maximum of two peaks in the first intermediate phrase and the second peak is always higher than the first one due to the presence of the H-.

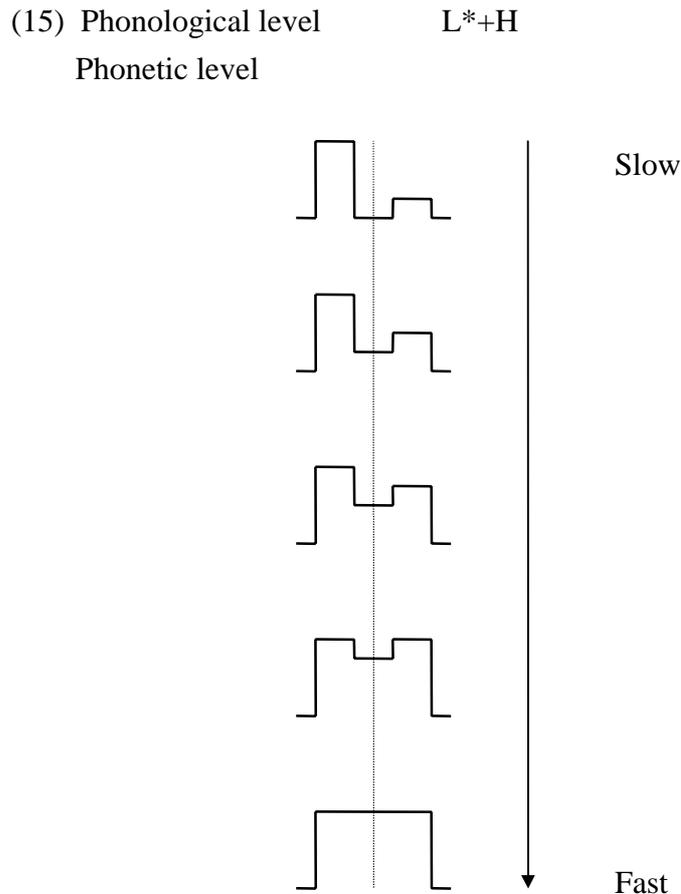
However, one question needs to be resolved and that is what is the status of the first peak of the second intermediate phrase. Is the first L*+H pitch accent of the second intermediate phrase downstepped or not? As observed in Figures 4 and 5, the first peak of the second intermediate phrase is produced at a much lower frequency than the preceding H tone (i.e. the H- marking the end of the first intermediate phrase). Although we can say that the lowering of the first peak of the second intermediate phrase is due to a pitch reset at the beginning of the new phrase, the lowering seems to us too drastic to exclusively involve a pitch reset and we think that the pitch reset is accompanied by an intended downstepped tone. Evidence for that derives from the fact that the first peak of the second intermediate phrase is much lower than the first peak of the first intermediate phrase, as observed in Figure 5. If the lowering of the first pitch accent of the second intermediate phrase was only due to a pitch reset, we would expect this peak to be at a similar frequency than the first peak of the first intermediate phrase since both peaks appear at phrase initial position. However, this is not the case since the first peak of the second intermediate phrase is much lower. This let us to hypothesise a downstepped tone in this position.

Based on these observations, we wanted to find out whether the first pitch accent of the second intermediate phrase is intendedly downstepped with respect to the preceding H- or it only shows a pitch reset. In order to test the hypothesis that the first pitch accent of the second intermediate phrase is downstepped, we analysed tempo differences. Since downstep is a phonological event, we expected it to be subject to speaking rate differences. That is, the slower the speaking rate, the higher the downstep and the faster the speaking rate the lower the downstep. Although speaking rate was not a variable included in this study, tempo effects were analysed since in some of the utterances the first pitch accent of the second intermediate phrase was not so clearly lowered due to a fast rate of speech. Figure 7, for instance, displays the F0 contour of the same utterance as in Figure 4 produced by another speaker. In this case,

the H of the bitonal accent in “-nia” is not lowered and appears with a similar scaling to the previous H tone.

So, in order to test the hypothesis that the first pitch accent of the second intermediate phrase of Central Catalan declaratives undergoes downstep, we compared the speaking rate of sequences with a clear lowering of the first peak of the second intermediate phrase with the speaking rate of sequences with not so clear lowering. If we can prove that the tempo of non-lowered structures is faster than that of lowered structures, we would be able to claim that the initial pitch accent of the second intermediate phrase is a downstepped tone, which shows different gradual realisations depending on speaking rate differences. This gradient behaviour would involve a continuum of forms, from the complete appearance of a downstepped L*+H tone in slow speech to its deletion in faster speech. This is schematised in (15) below. With the complete deletion of the tone we could even claim that in Figure 6 the first stressed syllable of the second intermediate phrase has an underlying L*+H pitch accent, its non-realisation being the consequence of a fast speaking rate.

In order to check whether tempo differences had an effect on the occurrence of the first peak of the second intermediate phrase, we compared the average syllable duration (dav) of sentences with a clearly lowered peak (henceforth +L) with that of sentences with a non-lowered peak (henceforth -L). If the average syllable duration of +L sentences proves to be significantly longer than that of -L sentences, we can conclude that the lack of lowering after the boundary in -L sentences is due to them being produced at a faster tempo. In other words, the shorter the dav, the quicker the speaking rate and the smaller the time for downstep to occur. A t-test was performed on the results so as to see whether the average syllable duration of +L sentences is significantly longer than the average syllable duration of -L sentences.



The results are presented in Table 3, which shows: 1) the average syllable duration for sentences with a clear lowering of the first pitch accent in the second intermediate phrase (+D) and the average syllable duration for sentences with no clear lowering (-L), 2) the t-value and 3) the probability of significance obtained in the t-test. The significance level used was $p < 0.01$.

+L sentences dav	-L sentences dav	t-value	p
0.1738	0.1553	8.4	<0.01

Table 3. Results of the comparison between lowered and non-lowered structures.

The results show that the average syllable duration of lowered structures is significantly longer than the average syllable duration of non-lowered structures ($p < 0.01$). These results seem to confirm our hypothesis that the first tone of the second intermediate phrase in Central Catalan declaratives is deliberately downstepped since it is subject to speaking rate realisational differences.

However, now that we have proved that the first pitch accent of the second intermediate phrase of Central Catalan is intentionally downstepped, some problems are encountered with the predictions of the theory. As pointed in section 1.1, Pierrehumbert (1980) and Beckman and Pierrehumbert (1986) treat downstep as a

phenomenon of phonetic realisation which is triggered by bitonal accents and blocked by the presence of an intermediate phrase boundary. According to this proposal, the tonal sequence L*+H H- L*+H observed in our data is expected to present the following characteristics: 1) H- is expected to be downstepped due to its being preceded by a bitonal pitch accent, and 2) the second L*+H tone is expected not to be downstepped due to the presence of the intermediate phrase boundary which is supposed to block the downstepping effect.

As observed so far, neither of these two predictions is fulfilled in our data, since the H marking the intermediate phrase boundary is not downstepped with respect of the preceding L*+H bitonal pitch accent and the first L*+H tone of the second intermediate phrase has proved to be downstepped even though it is immediately preceded by an intermediate phrase boundary. The behaviour observed in Central Catalan declaratives, therefore, questions two main issues on the interpretation of downstep in Pierrehumbert's theory, that is, its being triggered by bitonality and its being blocked by an intermediate phrase boundary.

An alternative solution seems to be required to explain the intonational behaviour observed in our data. At the moment, Ladd's view of downstep is being considered. According to Ladd, downstep can be treated as a distinctive feature of the lowered tone itself. If we follow this proposal, downstepped tones involve the selection of a separate phonological feature, i.e. [+downstep], which Ladd marks by means of the [!] symbol on the downstepped tone. So for example, a downstepped H* would be indicated as !H* and a downstepped L*+H as L*+!H.

So far, the results on Catalan declaratives seems to support Ladd's view of downstep as a phonological feature of the lowered tone itself and discard Beckman and Pierrehumbert's idea of downstep being triggered by bitonality. Thus, the neutral or default intonational pattern for SVO Central Catalan declaratives observed in our data can be said to have the tonal configuration presented in (16), where all pitch accents of the second intermediate phrase are downstepped. From our data, it can also be inferred that downstep in Central Catalan occurs within the intonation phrase, rather than within the intermediate phrase, since the appearance of a H- phrase accent does not block the realisation of downstep. More research on this issue is expected to be carried out in the near future.

(16) H*+L L*+H H- L*+!H !H* L-L%

4. Conclusions

In this paper we have analysed the intonational behaviour of simple SVO Central Catalan declaratives produced in reading speech. The results have been modelled within Pierrehumbert's theory of intonational analysis. Particular issues have been taken into account, namely, the division of Catalan declaratives into different prosodic units (intermediate phrase and intonation phrase), the tonal configuration of the F0 contours, and the analysis and interpretation of downstep. In order to identify the prosodic boundaries within an utterance, different cues have been examined, i.e., the presence of a pause, a local F0 fall, and the lengthening of pre-boundary syllables. An intermediate phrase boundary between the subject and the predicate is consistently

observed in all sentences. The neutral intonation of Central Catalan declaratives presents the following tonal configuration: (H*+L) L*+H H- (L*+!H) !H* L-L%. A pitch reset is observed at the beginning of the second intermediate phrase, which starts with a very drastic lowering on the peak of the first pitch accent. The effects of speaking rate on the realisation of this lowered peak support the analysis of this tone as intendedly downstepped. These findings provide evidence to treat downstep as a distinctive feature of the lowered tone itself, rather than as the consequence of bitonality. The domain of downstep in Central Catalan declaratives is shown to be the intonation phrase rather than the intermediate phrase.

Acknowledgements

Our thanks to Jill House and Mercedes Cabrera-Abreu for their insightful comments on an earlier version of this paper. Thanks also to Steve Nevard for his technical help and huge patience.

References

- Beckman, M.E. and J. Pierrehumbert. (1986). Intonational structure in Japanese and English. *Phonology Yearbook* 3, 255-310.
- Bonet, E. (1984). Aproximació a l'entonació del català. Tesi de llicenciatura. Universitat Autònoma de Barcelona.
- Cruttenden, A. (1986). *Intonation*. Cambridge: Cambridge University Press.
- García-Lecumberri, M.L., Cabrera-Abreu, M., and Maidment, J.A. (1997). The realisation of focal accent in English and Spanish: similarities and differences. *Intonation Theory, Models and Applications. Proceedings of ESCA Workshop*. September 18-20. Athens, Greece, pp. 141-144.
- Grice, M. (1992). The intonation of interrogation in Palermo Italian: implications for intonation theory. PhD dissertation, University College London. Published by Tübingen: Niemeyer, (1995).
- Grice, M. (1995). Leading tones and downstep in English. *Phonology* 12.2. 183-234.
- Grice, M. and R. Benzmüller. (1995). Transcription of German intonation using ToBI-Tones: the Saarbrücken System. *PHONUS (Research Report, Institute of Phonetics, University of the Saarland)* 1: 33-51.
- Ladd, D.R. (1983). Phonological features of intonational peaks. *Language* 59: 721-759.
- Ladd, D.R. (1993). In defense of a metrical theory of intonational downstep. In Hulst, van der H. and K. Snider, (eds.). *The phonology of tone: the representation of tonal register*. Berlin: Mouton de Gruyter.
- Liberman, M. and Pierrehumbert, J. (1984). Intonational invariance under changes in pitch range and length. In Aronoff, M. and Oehrle, R. (eds.) *Language sound structure*, pp. 157-233. Cambridge, MA: MIT Press.
- Möbius, B. (1993). *Ein quantitatives Modell der deutschen Intonation. Analyse und Synthese von Grundfrequenzverläufe*. Niemeyer: Tübingen.

- Pierrehumbert, J. (1980). *The Phonology and Phonetics of English Intonation*. PhD dissertation. Bloomington: Indiana University Press.
- Pierrehumbert, J. and Beckman, M. (1988). *Japanese Tone Structure*. Cambridge, MA: MIT Press.
- Prieto, P. (1995). Aproximació als contorns tonals del català central. *Caplletra* 19: 161-186.
- Prieto, P. van Santen, J. and Hirschberg, J. (1994). Patterns of F0 peak placement in Mexican Spanish. *Proceedings of the ESCA/IEEE Workshop on Speech Synthesis*, New Paltz, New York, September 12-15: 33-37.
- Prieto, P. van Santen, J. and Hirschberg, J. (1995). Tonal alignment patterns in Spanish. *Journal of Phonetics*, 23: 429-451.
- Recasens, D. (1977). Aproximació a les cadències tonals del català. *Anuario de Filología*, 3: 509-516.
- Recasens, D. (1993). *Fonètica i fonologia*. Barcelona: Enciclopèdia Catalana.
- Sosa, J.M. (1991). *Fonética y fonología de la entonación del español hispanoamericano*. PhD dissertation. Amherst, University of Massachusetts.
- Thorsen, N. (1981). Intonation contours and stress group patterns in declarative sentences of varying length in ASC Danish. –Supplementary data. *Annual Report of the Institute of Phonetics*, University of Copenhagen 14, 1-29.
- Virgili Blanquet, V. (1971). *Notas sobre la entonación catalana*. *Archivium*, XXI. Universidad de Oviedo. 359-377.

Appendix

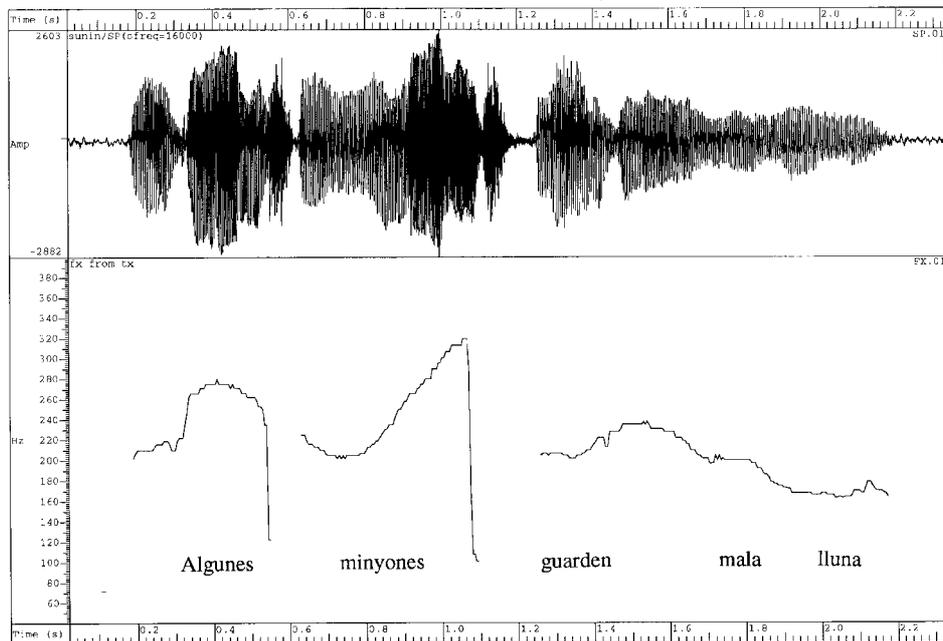


Figure 1. Speech waveform and F0 contour for the sentence *Algunes minyones guarden mala lluna* (some maids have a bad temper) produced by speaker NG. A pause and an abrupt F0 fall signal a prosodic boundary.

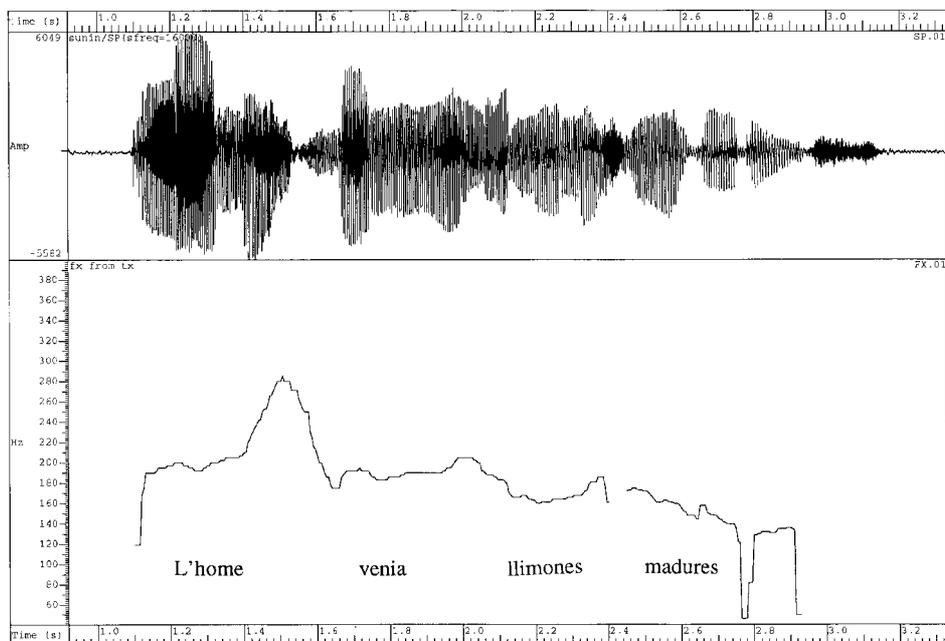


Figure 2. Speech waveform and F0 contour for the sentence *L'home venia llimones madures* (the man sold ripe lemons) produced by speaker ER. An abrupt F0 fall alone signals the prosodic boundary.

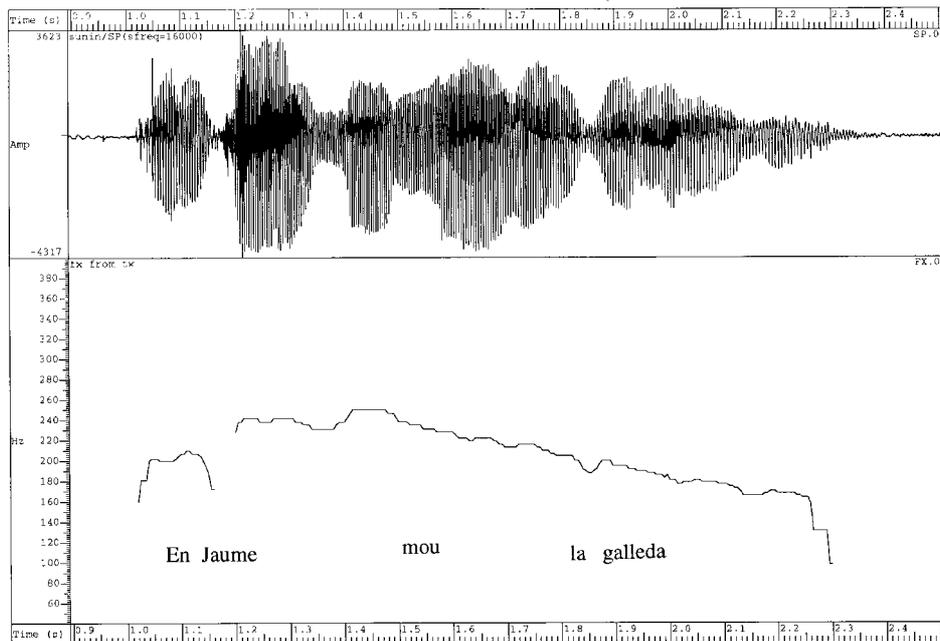


Figure 3. Speech waveform and F0 contour for the sentence *En Jaume mou la galleda* (James moves the bucket) produced by speaker MC. No pause or F0 fall are observed between the subject and the predicate.

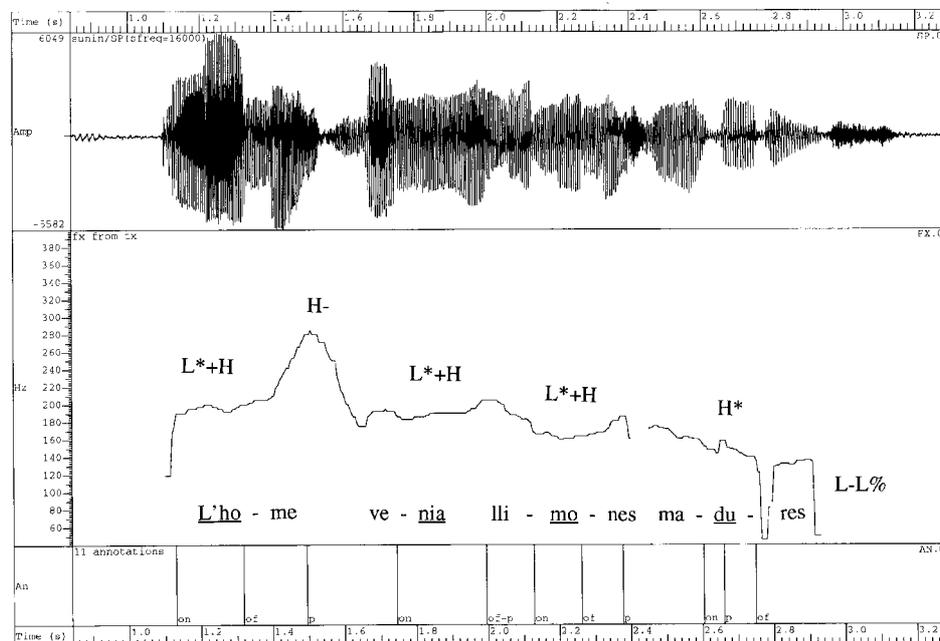


Figure 4. Tonal configuration for the same sentence as in Figure 2: *L'home venia llimones madures* (the man sold ripe lemons). The onset (on) and offset (of) of each accented syllable and the realisation of peaks (p) in relation to the accented syllables are signalled. Stressed syllables are underlined.

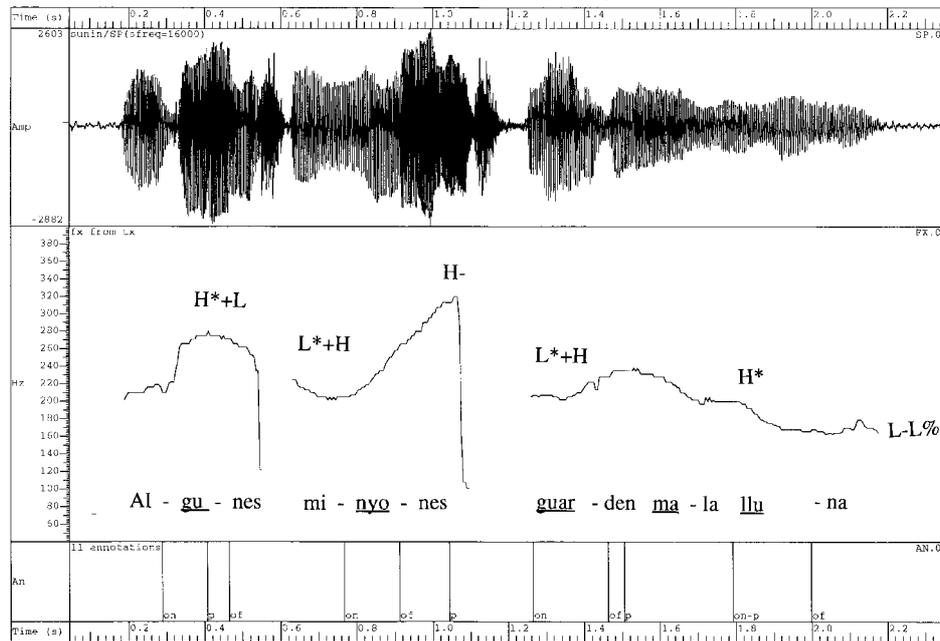


Figure 5. Tonal configuration for the same sentence as in Figure 1: *Algunes minyones guarden mala lluna* (some maids have a bad temper). The onset (on) and offset (of) of each accented syllable and the realisation of peaks (p) in relation to the accented syllables are signalled. Stressed syllables are underlined.

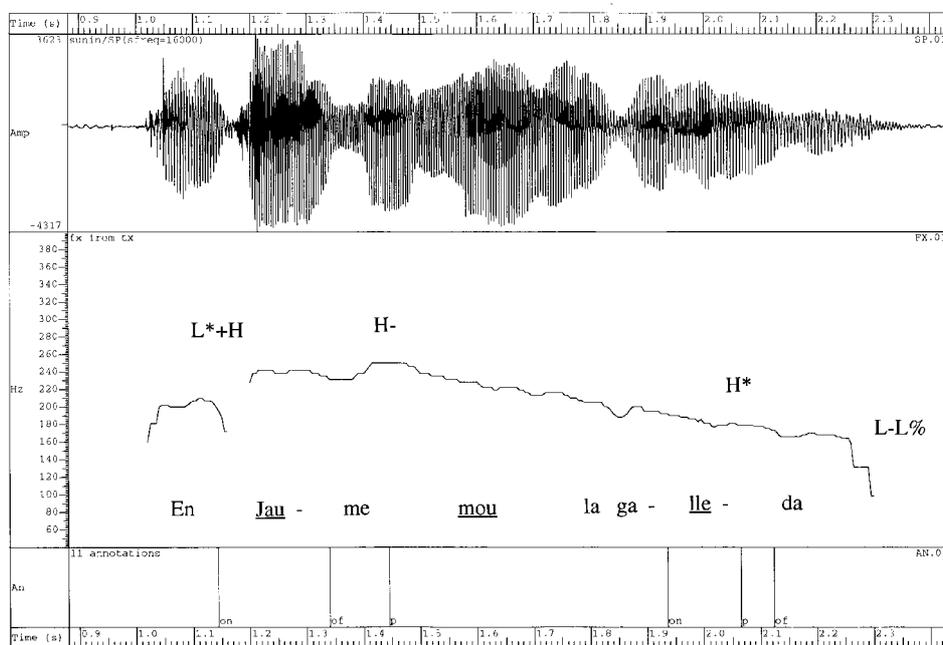


Figure 6. Tonal configuration for the same sentence as in Figure 3: *En Jaume mou la galleda* (James moves the bucket). The onset (on) and offset (of) of each accented syllable and the realisation of peaks (p) in relation to the accented syllables are signalled. Stressed syllables are underlined.

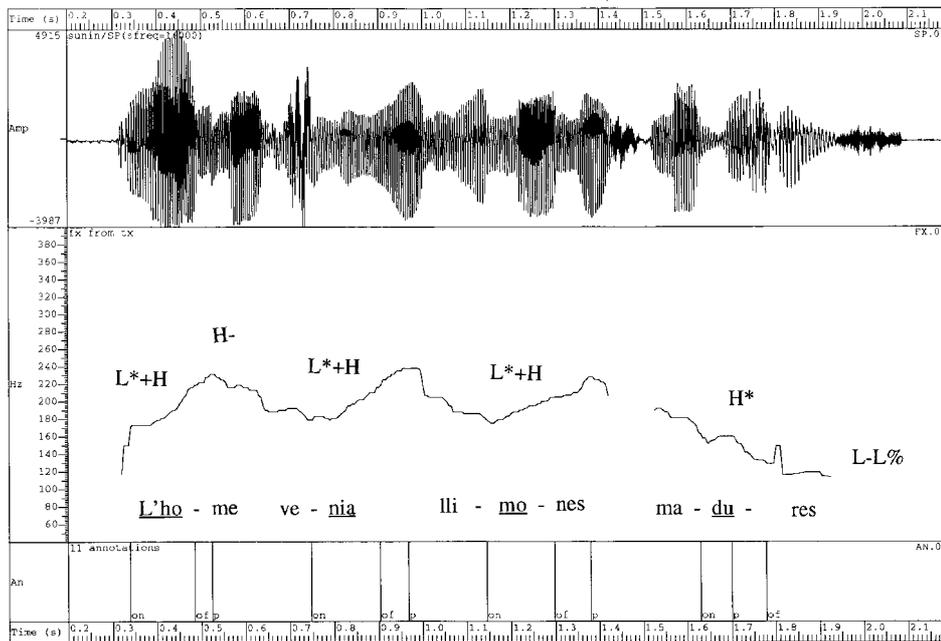


Figure 7. Speech waveform, F0 contour and tonal configuration for the sentence *L'home venia llimones maduras* (the man sold ripe lemons) produced by speaker DV. The onset (on) and offset (of) of each accented syllable and the realisation of peaks (p) in relation to the accented syllables are signalled. Stressed syllables are underlined.