

BRITISH ENGLISH DOWNTRENDS: DOWNSTEP OR DECLINATION?

EVA ESTEBAS VILAPLANA
UNED. Madrid

ABSTRACT

The aim of this paper is to shed some more light on the issue of whether the pitch or F0 downtrend observed in British English declarative sentences should be analysed as the result of an overall *declination* effect, triggered by a physiological or phonetic mechanism, or as an intended, phonologically controlled, *downstep* phenomenon. Previous studies on English intonation disagree on the interpretation of such downtrends. In this paper the scaling of pitch accents is analysed as an effect of the temporal distance (or number of unaccented syllables) between accents. 72 declarative sentences are examined. Sentences consisted of two or three accents. The number of intervening syllables between accents varied between 1 and 8. The hypothesis is that if the number of intervening syllables or temporal distance has an effect on the scaling of accents, then a declination effect should be postulated. On the other hand, if pitch accents have the same scaling despite differences in the number of intervening syllables, then F0 downtrends should be interpreted as the result of downstep. The results show that the scaling of peaks is not affected by the number of intervening syllables between accents and hence pitch downtrends are interpreted as the result of a linguistically controlled downstep rather than a gradual declination effect.

0. INTRODUCTION

It has been observed cross-linguistically that pitch (or F0)¹ tends to gradually decline over the course of an utterance (Ladd 1984). This is illustrated in Figure 1, which shows an English declarative sentence produced with four High (H*) tones or pitch accents. These tones are not scaled at the same level but each H* is lower than the preceding one. According to the Autosegmental-Metrical approach of intonational analysis (Pierrehumbert (1980); Beckman & Pierrehumbert (1986); Pierrehumbert & Beckman (1988); Beckman & Hirschberg (1994); Ladd (1996); among many others), the phonological entity H* stands for an F0 rise within the limits of the accented syllable. An ! before H* indicates that this F0 peak is lower than the preceding one. L- and L% account for the low F0 contour at the end of the utterance.

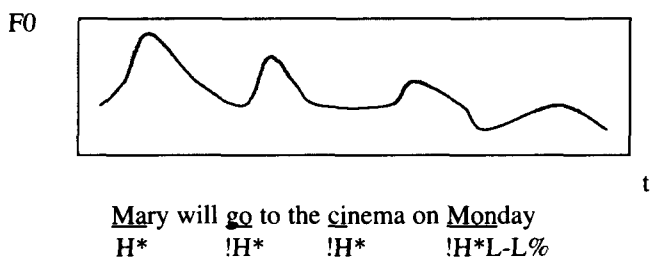


Figure 1. Schematised F0 contour for the sentence «Mary will go to the cinema on Monday». Stressed syllables are underlined. The shaded areas show the limits of the accented syllables.

The interpretation of such a downtrend varies in the literature. Two major views have been proposed for the description of F0 downtrends, referred to by among others Ladd (1983) and Nolan (1995) as the *Contour Interaction* model and the *Tone Sequence* model (henceforth CI and TS respectively). Within the CI model, the downward trend of pitch has been analysed as a global effect, conceived as a component of the overall contour. Thus, F0 downtrends were believed to be triggered by an automatic physiological mechanism (Fujisaki 1983, 1988 for Japanese; Thorsen 1980 for Danish; Bruce 1977 for Swedish; Lieberman 1967, Liberman 1975, Cooper and Sorensen 1981, Lieberman and Blumstein 1990 for English; among others). Followers of the TS model, however, treated this F0 downtrend as a local, phonologically controlled mechanism, which affects accents individually and in relation to previous accents. Pierrehumbert (1980) and Liberman and Pierrehumbert (1984) discovered that the value of each F0 peak in a descending contour could be expressed as a constant proportion of the previous peak. They found that time-dependent lowering was almost absent in their data and that pitch descent could be characterised as an accent-by-accent decay, which they termed *downstep* (or *catathesis* in Pierrehumbert and Beckman 1988), eliminating the need for a global declination component. Additionally, they found out that the final pitch accent in a sentence undergoes a more drastic lowering in F0 than expected by the application of the downstep rule and proposed to account for it by using a lowering constant. The existence of

¹ In this paper, the terms «pitch» and «F0» will be used interchangeably. Fundamental frequency (F0) is the acoustic correlate of the perceptual sensation of pitch.

downstep and *final lowering* has been observed in many other languages apart from American English (Japanese: Beckman and Pierrehumbert 1986, Poser 1984; Mexican Spanish: Prieto et al 1996; German: Möbius 1993, Grabe 1998, to appear, Central Catalan: Estebas-Vilaplana, 2000).

However, some studies on intonation (Pierrehumbert and Beckman 1988, Poser 1984, Fujisaki 1983, 1988) propose that the analysis of pitch downtrends as a sequence of localised phonological events does not exclude the existence of a global declination effect and suggest that both phenomena, declination and downstep, can coexist in the same utterance. Also Grabe (to appear) suggests that what Liberman and Pierrehumbert (1984) interpreted as final lowering in American English sentences seems to be an effect of declination in British English utterances, since the greater F₀ reduction observed on the utterance-final peak can be explained by a greater time interval between peaks. In the studies that claim the existence of the two mechanisms, declination is considered a residue downtrend after all other predictable lowering mechanisms have applied (see Prieto 2003 for more details on these downward trends).

In this paper, the nature of F₀ downtrends in British English will be examined in relation to the temporal distance (or number of unaccented syllables) between accents. In particular, we will investigate whether the F₀ downtrend observed in the neutral declarative sentences produced by a native British English speaker responds to a global, time-dependent declination effect or to a local, linguistically-controlled downstep mechanism. This will be done by comparing the F₀ maximum of consecutive peaks separated by different temporal distances (or number of intervening syllables between accents). If declination is active, a greater amount of F₀ lowering is expected as the temporal distance between H* pitch accents increases. Alternatively, if the height of peaks is stable and not affected by temporal distance, then the pitch downtrend would be the result of a controlled downstep or accent-by-accent decay.

1. EXPERIMENTAL DESIGN

The data examined in this study consisted of 72 Subject-Verb-Object English declarative sentences (e.g. *Amelia married a marine*). The majority of sentences consisted of 3 stresses (88% of sentences) and only a few of them had 4 stresses (12% of sentences). The number of unstressed syllables between stresses varied between 1 and 4. Thus, our initial aim was to compare the scaling of F₀ peaks in two conditions: i) 1-2 intervening syllables and ii) 3-4 intervening syllables. However, as we will see in section 3, not all stresses became accented². Sentences with 3 stresses tended to be produced with 2 pitch accents. Sentences with 4 stresses were produced with 3 pitch accents. Depending on which syllables the speaker decided to accentuate, the number of unaccented syllables between accents varied between 1 and 8 syllables. Thus, three different conditions were finally analysed: i) 1-2 intervening syllables, ii) 3-4 intervening syllables, and iii) 5-8 intervening syllables. Sentences were designed with the maximum number of voiced segments possible to avoid interrupted F₀ contours. See the Appendix for the list of sentences used in the experiment.

² Following Cruttenden (1986), in this paper we assume that *stressed* syllables are those produced with a rhythmic beat but no pitch movement and *accented* syllables are those syllables which show a pitch movement.

The informant was a 30 year-old female speaker of a Southern (London) variety of British English. Sentences were gathered by means of a reading activity. The speaker was given a list of structures and was asked to read them in a neutral way, as if they were answers to the question «what happens?». The data were recorded in the anechoic room of the Department of Phonetics and Linguistics of University College London. For each sentence, speech and laryngeal (Lx) signals were obtained. Speech was recorded to a B&K sound level meter of the type 2231, which was fitted with a 4165 microphone. The microphone was placed at about twenty centimetres from the speaker. The laryngeal signal was obtained by means of a laryngograph processor.

2. ANALYSIS OF THE DATA

An acoustic analysis of the data was performed by means of the Speech Filing System (SFS) program, which allowed us to obtain F0 traces from the laryngeal signal. To transfer the speech and laryngeal signals obtained in the recordings into the computer, the two signals were played on a Dennon DN-770R tape recorder connected into a Sun Sparc-10 computer running SFS. Acquisition of the signals was done at 16 KHz sampling rate, following the routines of the program. For the acoustic analysis a time-aligned inspection of the speech waveform and the F0 contour was carried out. An example of the sentence *Amelia married a marine* with three stresses (underlined) but produced with two pitch accents (on the first and last stressed syllables) is presented in Figure 2 below. The upper box shows the speech waveform, the middle box the F0 trace and finally the lower box signals the beginning of each syllable.

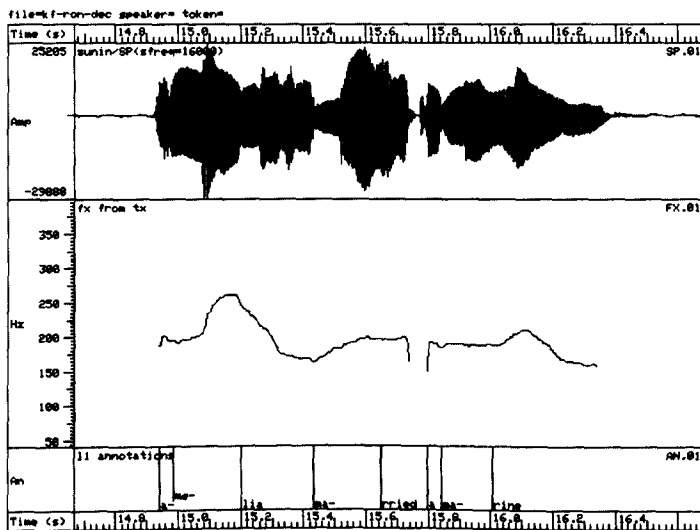


Figure 2. SFS display screen showing: 1) the speech waveform, 2) the F0 contour and 3) the beginning of each syllable for the sentence *Amelia married a marine*.

In order to examine the nature of F0 downtrends in the productions, the following points were labelled in the F0 contours:

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

3. RESULTS

The results showed that not all stressed syllables became accented. Sentences with 3 stresses tended to get two pitch accents (on the first stressed syllable and on the last stressed syllable, as illustrated in Figure 2 above). Sentences with 4 stresses had 3 accents (usually the second stressed syllable did not become accented). The schematised mean F0 contours of utterances with a different number of stresses (and accents) are presented in Figure 3. The L and H letters indicate the mean values of F0 minimum and F0 maximum obtained in the analysis respectively. The number of stresses and accents (in brackets) is specified on the graph.

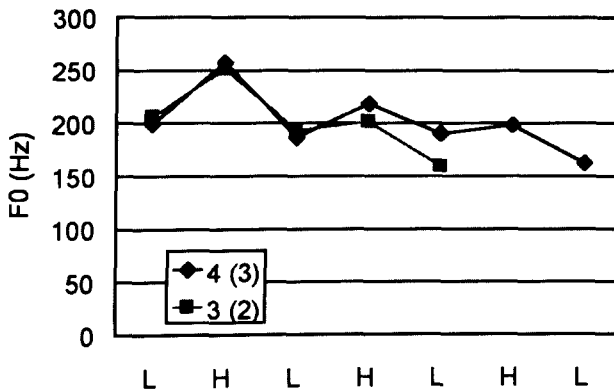


Figure 3. Schematised mean F0 contours for neutral declarative sentences with different number of stresses and accents (in brackets).

As expected, the results in Figure 3 show that British English neutral declaratives are produced with a sequence of High accents and that each accent is much lower than the preceding one.

3.1. Utterance initial and final F0 values and pitch range

Before analysing the nature of British English F0 downtrends and deciding whether the pitch drop in each accent is the result of a phonologically controlled downstep or an overall declination effect, it was considered necessary to see whether differences in sentence length had an effect on the utterance initial and final F0 values and on the pitch range with which sentences were produced. Whereas in some studies (Thorsen 1980, Cooper and Sorensen 1981), utterance length seems to have an effect on the F0 range of the first peak (and subsequently of all other peaks), in other studies (Lieberman and Pierrehumbert 1984, Prieto et al 1996, Estebas-Vilaplana 2000) sentence length does not influence the degree of prominence in sentences. Since in our data sentences with a different number of stresses (and hence with different length) were gathered, we first needed to prove that pitch accents are actually comparable and hence that utterance length does not affect pitch range. In order to do so, utterance-initial and utterance-final F0 values were calculated as well as the scope of the first F0 rise (measured as the difference in Hz from the lowest point at the beginning of a contour to the highest F0 value of the first peak). If the F0 excursion between the initial F0 value and the first peak is the same in sentences with different length, this will show that utterances with a different number of stresses are produced with the same pitch range and hence accents are comparable.

The mean F0 values (in Hz) in utterance initial and utterance final positions for sentences with different length (3 and 4 stresses) are presented in Table 1.

Number of stresses	Utterance initial mean F0 value (Hz)	Utterance final mean F0 value (Hz)
3	205.9	158.9
4	198.8	162

Table 1. Mean F0 values (Hz) in utterance initial and utterance final positions for sentences with different length (3 and 4 stresses).

As presented in Table 1, the mean F0 value in utterance initial position is slightly higher in sentences with three stresses than in those with four stresses. However, the results of a t-test performed on the data showed no significant differences in utterance-initial F0 values for sentences of different length at 1 per cent significance level ($p=0.07$, $t=1.66$). Similarly, utterance-final F0 values did not present significant differences with respect to sentence length ($p=0.16$, $t=1.66$). Thus, the number of stresses per sentence (and subsequently the sentence length) does not seem to affect the initial and final F0 values.

In addition, the F0 range (difference in Hz from the lowest F0 value at the beginning of an utterance to the highest F0 point of the first peak) was also rather constant in sentences with different number of stresses. This is shown in Table 2 which displays the mean F0 range in utterances with three and four stresses and the results of t-tests comparing the F0 range values. The results of the t-tests show that there are no significant differences in the F0 range with respect to the number of stresses. Thus, the number of stresses per sentence (and subsequently the sentence length) does not seem to

affect the degree of prominence or pitch range with which sentences were produced. These results agree with those of Liberman and Pierrehumbert (1984) for American English and prove that pitch accents are comparable irrespective of sentence length.

Stresses	F0 range	p	t
3	46.9	0.09	1.66
4	55		

Table 2. Mean F0 range in utterances with three and four stresses and results of t-tests comparing the F0 range values.

3. 2. Effects of temporal distance between peaks

One of the strategies used by some researchers (Prieto et al 1996, Estebas-Vilaplana 2000, Grabe to appear, among others) to detect whether declination is active in downtrend F0 contours is the analysis of the temporal distance (examined in terms of the number of intervening unaccented syllables) between pitch accents. If declination is present, the F0 reduction is expected to be bigger as the time interval (or number of intervening syllables) between peaks increases. Otherwise, if there is no declination, the F0 reduction is expected to have a constant value irrespective of the time interval between peaks.

In order to see whether a declination component was operating in the British English sentences, the values of H1 and H2 were compared in sentences with two accents but with a different number of intervening syllables between them. Sentences were divided into three groups according to the number of unaccented syllables between accents: i) 1-2 syllables, ii) 3-4 syllables and iii) 5-8 syllables. Before analysing the effect of syllable distance on the scaling of two peaks, we wanted to confirm that differences in the number of syllables involved differences in real duration. Hence, the temporal distance between H1 and H2 was measured for the three groups of syllables. The mean duration between H1 and H2 according to the number of intervening syllables is presented in Table 3.

Number of syllables		
1-2	3-4	5-8
0.259	0.361	0.595

Table 3. Mean duration (in seconds) between peaks with 1-2, 3-4 and 5-8 intervening syllables.

As expected, the bigger the number of intervening syllables between H1 and H2, the longer the temporal duration. T-tests comparing the results showed that the mean duration is significantly different ($p < 0.001$, $t = 10.9$) in all cases, indicating that the bigger the number of intervening syllables the longer the time to produce them.

Once it had been confirmed that differences in the number of syllables involved differences in time, we compared the mean peak height of H2 preceded by 1-2, 3-4 and

5-8 unaccented syllables as well as the amount of pitch drop (F0 mean difference) between H1 and H2. The results plotted in Figure 4 show the mean F0 values for H1 and H2 with different number of intervening syllables. Table 4 presents the mean F0 values of H2 according to the number of intervening syllables. Finally, Table 5 displays the mean F0 difference between H1 and H2 with 1-2, 3-4 and 5-8 intervening unaccented syllables.

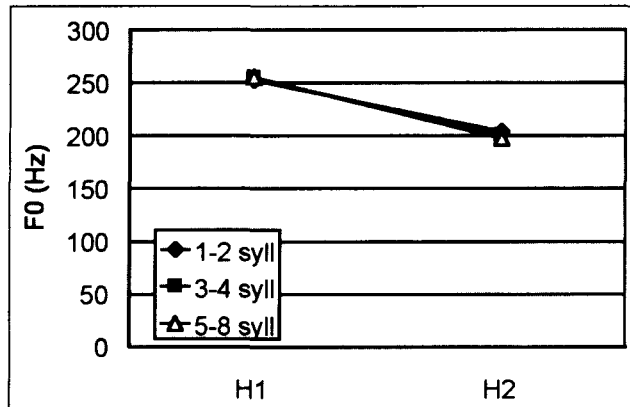


Figure 4. Mean F0 values for H1 and H2 corresponding to an increase in the number of intervening syllables (1-2, 3-4 or 5-8).

1-2 syllables	3-4 syllables	5-8 syllables
203.7	200.3	197

Table 4. Mean F0 values of H2 (Hz) with 1-2, 3-4 and 5-8 intervening syllables.

1-2 syllables	3-4 syllables	5-8 syllables
48.5	54.1	58.2

Table 5. Mean F0 difference between H1 and H2 (Hz) with 1-2, 3-4 and 5-8 intervening syllables.

The results presented in Table 4 show that the mean F0 values of H2 slightly decrease as the number of intervening syllables increases. However, the statistical analysis performed on the data (ANOVA) shows that these differences are not significant ($p=0.16$ and $F=2.3$). The results presented in Table 5 are similar to those observed in Table 4. Even though the mean F0 difference between H1 and H2 seems to increase as the number of syllables increases, the statistical analysis (ANOVA) shows that these differences are not significant ($p=0.13$ and $F=2.11$). These findings indicate that the scaling of H2 is very little affected by the number of intervening unaccented syllables between H1 and H2, proving the lack of a declination effect in British English neutral

declaratives. This suggests that downtrend contours in this language originate from the repeating occurrence of downstep accents, rather than from a global declination component. These results agree with those of Liberman and Pierrehumbert (1984) for American English who treated pitch downtrends as the result of a linguistically controlled downstep rather than a gradual declination effect and support the proposal of the TS view, which analyses pitch downtrends as the result of an accent-by-accent decay.

3.3. Final Lowering

One final issue derived from the data is the question whether the last pitch accent undergoes final lowering or not. Liberman and Pierrehumbert (1984) carried out several experiments, which showed that the final accent of a declarative is lower in F0 range than predicted by the location of the immediately preceding accent. However, recent work on British English (Grabe 1998, Grabe to appear) has shown that, unlike American English, Southern Standard British English does not have final lowering. The results of her data demonstrate that the last pitch accent was affected by the amount of segmental material between accents (steps between successive accents increased as more material was inserted) suggesting the presence of declination. Obviously, to be able to support one of these two views more data and more speakers would be needed. However, from the observations gathered so far, final lowering does seem to operate in the data. This is illustrated in Figure 5, which plots 1) the mean peak values of H1 and H2 in sentences with three stresses and two pitch accents (produced with a different number of intervening syllables between accents), and 2) the mean peak values of H1, H2 and H3 in sentences with four stresses and three pitch accents. The results show that H2 (or final accent) in sentences with three stresses but two accents (3,2) is much lower than the second H2 (non-final accent) in utterances with four stresses but three accents (4,3), indicating the presence of final lowering on the last accent in (3,2) utterances.

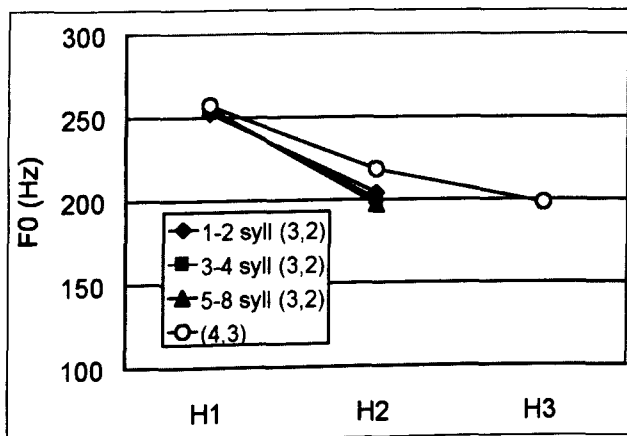


Figure 5. Mean F0 values for H1 and H2 in sentences with 3 stresses but 2 accents (produced with a different number of unaccented syllables between accents) and mean F0 values for H1, H2 and H3 in sentences with 4 stresses but 3 accents.

Overall, the results of F0 downtrends in the English data confirm the findings of Liberman and Pierrehumbert (1984), Beckman and Pierrehumbert (1986), and Pierrehumbert and Beckman (1988) for American English. English declaratives are produced with a series of local downstepping accents. The last accent of the declarative shows a greater amount of lowering than the predicted by the downstep rule. This differs from the results of Grabe (1998, to appear) who shows that in British English final lowering is absent.

4. CONCLUSION

In this paper the nature of pitch or F0 downtrends in British English declarative sentences has been analysed as an effect of the temporal distance between pitch accents. The predictions were that if the scaling of pitch accents is affected by the temporal distance (or number of unaccented syllables) between accents, then a declination effect should be postulated. On the other hand, if pitch accents had the same scaling despite differences on the number of intervening syllables, then F0 downtrends should be interpreted as the result of an intended, phonologically controlled, downstep phenomenon. 72 declarative sentences produced with two or three pitch accents and with a different number of intervening syllables between accents (from 1 till 8) were examined. The results showed that the scaling of F0 peaks is not affected by the number of intervening syllables between accents and hence pitch downtrends in British English should be interpreted as the result of a linguistically controlled downstep mechanism rather than a gradual declination effect. These results support the findings of the TS model, which interprets F0 downtrends as a local, accent-by-accent phenomenon rather than a global component of the utterance.

REFERENCES

- BECKMAN, Mary & Julia HIRSCHBERG (1994): *The ToBI Annotation Conventions. Ms. and Accompanying Speech Materials*, Ohio State University.
- BECKMAN, Mary & Janet PIERREHUMBERT (1986): «Intonational Structure in Japanese and English», *Phonology Yearbook* 3, p. 255-310.
- BRUCE, Gosta (1977): *Swedish Word Accents in Sentence Perspective*, Lund, Gleerup.
- COOPER, W. F. & J. M. SORESENSEN (1981): *Fundamental Frequency in Sentence Production*, Heidelberg, Springer.
- CRUTTENDEN, Alan (1986): *Intonation*, Cambridge, Cambridge University Press.
- ESTEBAS-VILAPLANA, Eva (2000): «Peak F0 downtrends in Central Catalan neutral declaratives», *Speech, Hearing and Language: work in progress*, vol. 12, London, University College London, p. 16-41.
- FUJISAKI, Hiroya (1983): Dynamic characteristics of voice fundamental frequency of speech and singing. In P. F. MacNeilage (ed.), *The production of speech*, New York and Berlin, Springer-Verlag, p. 39-55.

- (1988): «A note on the Physiological and Physical Basis for the Phrase and Accent Components in the Voice Fundamental Frequency Contour.» In Osamu Fujimura (ed.), *Vocal Physiology: Voice Production, Mechanisms and Functions*, New York, Raven, p. 347-355.
- GRABE, Esther (1998): *Comparative Intonational Phonology: English and German*, Wageningen, Ponsen and Looijen bv.
- (To appear): «Fundamental frequency downtrends in English and German», *Language and Speech*.
- LADD, Robert (1983): «Peak Features and Overall Slope», In Anne Cutler & Robert Ladd (eds.), *Prosody: Models and Measurements*, Berlin, Springer, p. 39-52.
- (1984): «Declination: a Review and Some Hypotheses», *Phonology Yearbook* 1, p. 53-74.
- (1996): *Intonational Phonology*, Cambridge, Cambridge University Press.
- LIEBERMAN, Philip (1967): *Intonation, Perception, and Language*, Cambridge MA, MIT Press.
- LIEBERMAN, Philip & Sheila BLUMSTEIN (1990): *Speech Physiology, Speech Perception and Acoustic Phonetics*, Cambridge MA, Cambridge University Press.
- LIEBERMAN, Mark (1975): *The intonational system of English*, PhD thesis, MIT.
- LIEBERMAN, Mark & Janet PIERREHUMBERT (1984): «Intonational Invariance under Changes in Pitch Range and Length», In M. Aronoff and R. Oerhle (eds.), *Language Sound Structure*, Cambridge MA, MIT Press, p. 157-233.
- MÖBIUS, Bernd (1993): *Ein Quantitatives Modell der Deutschen Intonation. Analyse und Synthese von Grundfrequenzverläufe*, Niemeyer, Tübingen.
- NOLAN, Francis (1995): «The Effect of Emphasis on Declination in English Intonation». In Jack Windsor Lewis (ed.), *Studies in General and English Phonetics. Essays in Honour of Professor J. D. O'Connor*, London, Routledge, p. 241-254.
- PIERREHUMBERT, Janet (1980): *The Phonology and Phonetics of English Intonation*, PhD dissertation. Bloomington, Indiana University Press.
- PIERREHUMBERT, Janet & Mary BECKMAN (1988): *Japanese Tone Structure*, Cambridge MA, MIT Press.
- POSER, William (1984): *The Phonetics and Phonology of Tone and Intonation in Japanese*, PhD dissertation, MIT.
- PRIETO, Pilar (2003): «Teorías Lingüísticas de la Entonación», In Pilar Prieto (ed.), *Teorías de la Entonación*, Barcelona, Ariel Lingüística, p. 13-33.
- PRIETO, Pilar, Chilin SHIH, & Holly NIBERT (1996): «Pitch Downtrend in Spanish», *Journal of Phonetics* 24, p. 445-473.
- THORSEN, Nina (1980): «Intonation Contours and Stress Group Patterns in Declarative Sentences of Varying Length in ASC Danish», *Annual Report of the Institute of Phonetics* 15, University of Copenhagen, p. 13-47.

APPENDIX

Jo relies on my money
 Ray resigned from his job
 The law has to be modified
 The meal was absolutely delicious
 Mary learns modern languages

Lee remembered the melody
 John divided the legacy
 The balloon was moving upwards
 The name was written at the door
 Millie lives in Ireland

Molly made a marvellous dinner
The lawyer's reading the journal
The lady ordered an orange juice
Melanie will win the award
Emily damaged the window
The melody was amazing
The lullaby brings good memories
The gardener mows the lawn
Norma runs every morning
Marjorie adores movies
Brenda removes the garbage
Josephine needed an explanation
The general murdered the villain
Ron amended the bill
The boy delivered the magazines
Melanie's menacing the hostages
The rain damages the harvest
The minister moderates the audience
Julie memorised the names
Ron's paying bills
Jeremy's wearing jeans
The nominee won the award
The jam remained in the jar
My mother loves babies
John married Mary
Gary borrowed the money
The nanny's ironing the linen
John married Melanie
The barber adores Marilyn
lady ordered the marmalade
The minister remembers the melody

Mini loves lemon marmalade
My mother loves babies
The nanny's ironing the linen
Marjorie married an Indian marine
Marilyn nurses the baby
The gardener used to live in London
The marmalade's on the fridge
The boys need a new game
The bluebells die with the rain
The girls believe in fairies
The jam remained in the jar
My neighbour married the model
Melanie ordered an orange juice
My brother remembers the rhymes
Jonathan erases the message
Grandma's ironing the linen
The journalist balances the disorder
Jeremy manages his company
John normalised the numbers
The journalist's ordering wine
Marilyn's boiling eggs
Amelia married a marine
The bluebells die with the rain
Nelly adores movies
The lady's wearing diamonds
Julie memorised the numbers
The general murdered the villain
Jimmy loves gardening
Mary learns languages
The boy delivered the magazines
The nanny sings a lullaby