TIMING OF F0 PEAKS AND PEAK LAG

Francis Nolan and Kimberley Farrar
Department of Linguistics, University of Cambridge, UK

ABSTRACT

In this paper we discuss the phenomenon of 'peak lag', where H* tones are not aligned with the accented syllable. The timing of f0 peaks has been shown to depend on a number of factors, such as the segmental structure and foot-size. This paper presents the results of a study using data collected in four locations in the British Isles[1]. Whilst confirming some of the results found in previous studies of this phenomenon, our findings indicate that there is a further factor contributing to peak lag, namely the presence of preceding unstressed syllables, or anacrusis. The results also highlight differences between varieties of British English with respect to the occurrence of peak lag.

1. INTRODUCTION

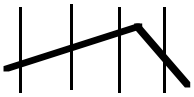
Pitch accents are associated with a particular stressed syllable, the 'accented syllable'. However, the timing of the corresponding f0 events with respect to that syllable is not constant, and the alignment of the peaks and troughs in the f0 contour is thus subject to variation. This variation in alignment is illustrated for example in a forthcoming paper by Wichmann et al [1], who examine the location of peaks to see where they occur during the accented syllable. They measure the location of the f0 peak by expressing it as a percentage of the duration of the accented syllable. Their results show that certain factors lead to a later alignment of the peak. Where such factors are *not* present in their data, the mean location of the peak is 73%, but where they are present, this mean value is considerably higher, at 98% [1:8]. Therefore in this condition peaks occur very close to the right hand edge of the accented syllable.

In such cases, the difference in alignment is relatively small; although the peak is displaced further rightwards, it is still temporally aligned with the accented syllable. It is merely much closer to the right edge of this syllable, than to the left. However, displacement is often far greater than this. When the displacement is particularly large, H may be aligned not with the accented syllable, but rather with a following unstressed syllable, a phenomenon we shall call 'peak lag'. Furthermore, f0 peaks may be aligned not even with the immediately adjacent syllable, but delayed still further, giving rise to 'extreme peak lag'. These two possibilities are illustrated in Figures 1 and 2:

peak is displaced one syllable to the right of accented syllable

accented syllable

Figure 1. Peak lag.



peak is displaced two syllables to the right of accented syllable

accented syllable

Figure 2. Extreme lag.

Differing alignment does not however necessarily entail a different phonological *association*. Examples such as those shown in the diagrams above may function in exactly the same way, and in such cases are best described as different *phonetic* implementations of the same *phonological* category. Thus, the H of an H*+L pitch accent can occur at any point within the accented syllable, or even outside that syllable, but the pitch accent is still H*+L, if it functions as this accent usually does in the dialect in question.

Thus there is a considerable degree of variation permitted in the surface patterns associated with pitch accents. In order to benefit from this knowledge and contribute to our labelling ability, we need to understand how different phonetic patterns are motivated. Although the use of a particular phonological specification does not uniquely determine the phonetic realization, it is also clear that this realization is not simply randomly chosen, but motivated by other factors [2:68].

Previous studies of peak alignment have suggested what some of these factors might be. It has been shown for instance that segmental factors play a role; intrinsically short vowels in the accented syllable give rise to earlier peaks for example [3]. The position of the pitch accent within the utterance also affects alignment, with peak lag found more often at the beginning of utterances [2:55]. Position in the discourse as a whole is important, and it has been shown that pitch accents at the beginning of topic initial utterances are more prone to peak lag [1:6]. The influence of prosodic factors has also been discussed elsewhere [4]. However, in studies like [4], it is the prosodic structure of the foot, which is examined. Since in English the accented syllable forms the head of the foot, no regard was given to what precedes this syllable. An experiment was therefore devised to test whether preceding prosodic context plays a role in determining peak alignment in the intonational onset.

2. METHOD OF EXPERIMENT

The data used was collected as part of an ESRC funded study of British English intonation[1]. The corpus arising from this project will contain directly comparable speech data collected from 16-18 year olds in seven different British English speaking locations. At the time of this experiment, data from four dialects was available for analysis: Cambridge, Leeds, Newcastle and Belfast (a total of 47 speakers). One task that speakers were asked to perform was to read a passage of English, namely a version of the Cinderella story. From this 10 utterances were chosen for analysis. In all the utterances an H* or H*+L pitch accent was associated with the first accented syllable (for details

of the IViE transcription system used here, please see [5]). However, utterances were varied according to the position of this syllable; some had no unstressed syllables preceding (i.e. no anacrusis), others had one or two unstressed syllables preceding (anacrusis). The utterances used are listed in Table 1.

No.	Text	Size of anacrusis
1	Once upon a time there was a girl called Cinderella	0
2	But everyone called her Cinders	1
3	Lily and Rosa were very unfriendly	0
4	The ball would be held in the Royal Palace	1
5	Lily and Rosa thought this was divine	0
6	They dreamed of wedding bells	1
7	Every girl in the land was willing to try on the slipper	0
8	But the Prince insisted that *all* girls must try the slipper	2
9	Then the girl looked at her old rags	0
10	After the ball, the Prince was resolved to find the beauty who had stolen his heart	0

Table 1. Utterances used and length of preceding prosodic context (accented syllables of interest are underlined).

Files were made containing each utterance from each speaker (using Xwaves). All cases where reading errors occurred, or where the f0 trace did not allow reliable measurements to be made, were rejected, leaving a total of 436 files. For each file the position of the peak associated with the first accented syllable was then noted. Peak position was classified as either 'peak in', where the peak is aligned with the accented syllable, 'peak lag', where the peak is outside the accented syllable, and located in the adjacent syllable to the right (see Figure 1), and 'extreme lag', where the peak is lagged still further, occurring two or three syllables to the right of the accented syllable (see Figure 2).

3. PEAK LAG AND ANACRUSIS

3.1. RESULTS

Utterances with and without anacrusis were separated out, and the location of peaks in the two groups compared. The results are shown in Figure 3:

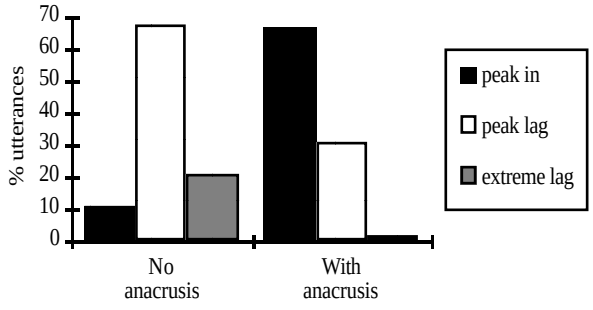


Figure 3. Peak lag and anacrusis.

A chi-square test was carried out on these figures, assessing the null hypothesis that the two samples (+ and – anacrusis) are from the same population. The observed value of X^2 is 150.95 Reference to critical values shows that X^2 150.95 has a probability of occurrence when H_0 is true of less than 0.001. Since the observed value exceeds the critical value 13.82, the null hypothesis can be rejected with greater than 99.9% certainty.

To summarise, only 11.07% of utterances with no anacrusis had the peak occurring within the accented syllable, compared to 66.67% of utterances with anacrusis. Peak lag is therefore more frequent when there is no anacrusis.

Examples of the patterns found in the data are given in Figures 4 and 5. Figure 4 shows the f0 trace for the beginning of utterance 3 (no anacrusis), as read by a male speaker from Cambridge. The peak associated with the H* pitch accent does not occur until the unstressed syllable following the accented syllable. This is therefore an example of peak lag, the category into which 67.94% of examples without anacrusis fall.

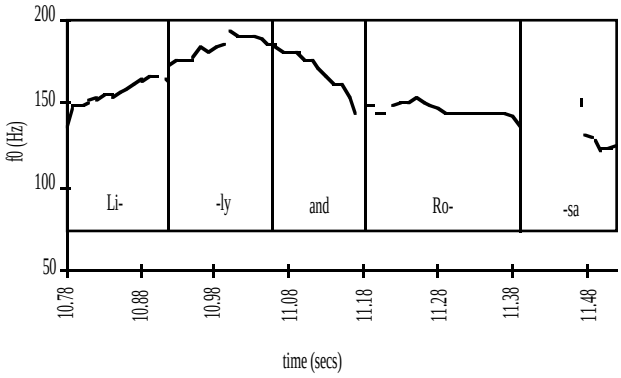


Figure 4. Utterance with no anacrusis showing peak lag.

Figure 5 shows the f0 trace for the beginning of utterance 6 (+ anacrusis) from the same speaker. The peak associated with the H*+L pitch accent is here aligned with the accented syllable, and the example is therefore categorised as 'peak in'. This is the category into which 66.67% of examples with anacrusis fall.

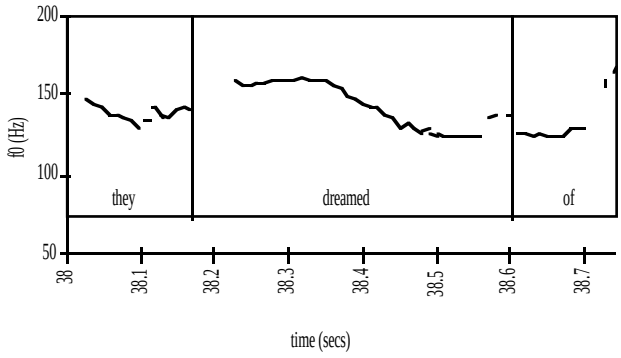


Figure 5. Utterance with anacrusis and peak in.

The results certainly indicate that peak alignment is indeed affected by preceding prosodic context. One factor which might be distorting these results is the position of each utterance in the discourse as a whole. Some of the utterances with no anacrusis

occur at the beginning of paragraphs. Paragraph-initial utterances have been shown to have a greater frequency of peak lag on their intonational onsets because they are often topic-initial, and discourse position is important in determining peak alignment [1:11]. It is possible therefore that the greater frequency of lag found for utterances with no anacrusis may well be partly due to discourse factors. In order to test this, utterances were further classified as 'para-initial' or 'non-initial', giving the figures in Table 2.

Location of utterance	Anacrusis?	peak in	peak lag	extreme lag	Total
Para-initial	No	22	110	38	170
Non-initial	No	7	68	17	92
	Yes	116	54	4	174

Table 2. Discourse position and peak lag.

Based on these figures it is possible to compare the *non-initial* utterances with and without anacrusis. Any possible impact of topic-initiality is therefore likely to be ruled out. The results of such a comparison are as follows: 7.61% of utterances with no anacrusis have the peak aligned with the accented syllable, compared to 66.67% of utterances with anacrusis. A chi-square test was carried out, and the null hypothesis that the two non-initial utterance sets were from the same population could be rejected with greater than 99.9% certainty.

3.2. DISCUSSION

There are a number of reasons why this link between preceding prosodic context and peak lag might exist. One possibility is that peak lag is physiologically determined. In order to align an f0 peak in an intonational onset, the speaker must reach the peak in a relatively short space of time. Although the 'aim' might be to align the peak with the accented syllable, performing such an f0 excursion is time-consuming. One might therefore expect peaks to be aligned more reliably if there are preceding unstressed syllables, on which f0 can already be rising. On the other hand, if there is no material prior to the accented syllable on which this excursion can be made, then it is possible that peak f0 will not be attained until after the accented syllable, hence peak lag will occur.

If peak lag does indeed have a physiological basis, then we would expect that as the size of the f0 excursion increases, so will the amount of time it takes the speaker to reach that peak. In other words, it will take you longer to climb further. In order to test whether this is the case, further measurements were made in each file. Firstly the f0 values at the beginning of the utterance and at the peak were measured. Using these two values, a range between the two was calculated, thereby giving a value for the f0 excursion size. Secondly, the duration (secs) from the beginning of the utterance to the peak was measured. Since informants had different speaking speeds, this value was normalised. The length of a specified chunk of the utterance for each speaker was measured, and the average length of this chunk across all speakers calculated. The duration start-peak was then normalised using the following calculation:

$$\text{duration start-peak} \quad X \quad \frac{\text{average length of chunk}}{\text{speaker's length of chunk}}$$

Using a Pearson Product Moment Correlation Test the two values of excursion size and normalised duration start-peak were compared. If the hypothesis of a physiological motivation for peak lag is correct, then we would expect a coefficient approaching +1. The actual value obtained was 0.1807. This militates against accepting a physiological account of peak lag; put simply, it does not necessarily take you longer to climb further.

It is clearly possible then to achieve a high peak in a relatively short space of time. However another possibility is that although this is physiologically possible, it is perhaps simply not a *priority* for the speaker. It may be the case that to perceive a peak, it is desirable to have a considerable rise up to that peak. Peaks are marked out as important points in the string by the fact that there is a slope on either side of them. Therefore the speaker's priority might be to render the peak more perceptually salient, by emphasizing the rise from first f0 to peak f0. One means of achieving this is to make the slope up to the peak a shallow one. This might override considerations of peak alignment, and therefore give rise to peak lag. This is particularly likely in cases where there is no anacrusis on which the rise can already be occurring, as shown in Figure 6.

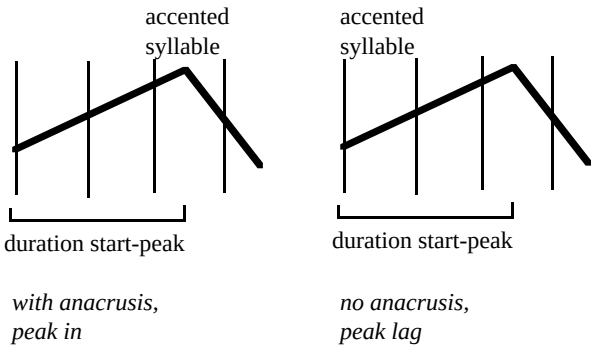


Figure 6. Peak lag as a result of slow rise to peak f0.

Here the rise is invariant, the only difference is the presence or absence of anacrusis.

If this hypothesis is true, then it explains why speakers *can* attain a high peak in a short period of time; it is simply the case that perceptual considerations may dominate.

4. PEAK LAG AND DIALECTAL DIFFERENCES

Data from four different locations in the British Isles was used in this experiment. It was therefore possible to investigate whether the locations differ with respect to the occurrence of peak lag. Firstly, utterances with and without anacrusis were compared in each dialect to see whether the apparent link between anacrusis and peak lag is true for all the locations. The results are shown in Figure 7.

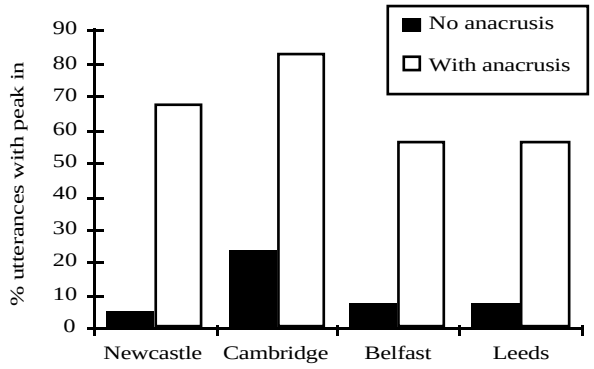


Figure 7. Anacrusis and peak alignment in different dialects.

Although some variation is evident, it is clear that anacrusis is an important factor in all locations. This variation, however, is an indication that there are differences between the dialects with respect to peak alignment.

In order to see whether the varieties do indeed behave differently, the numbers of utterances with peak in, peak lag and extreme lag were counted for each location. The results are shown in Figure 8.

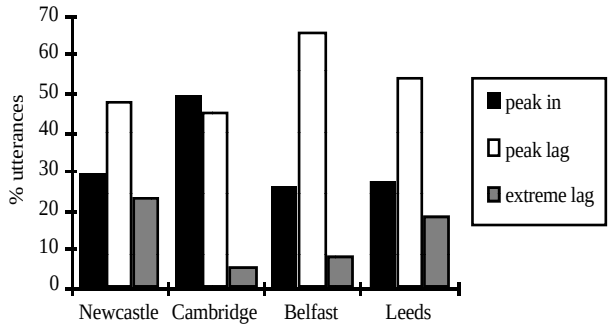


Figure 8. Peak alignment in different dialects.

A chi-square test was carried out, and the null hypothesis that the data from the four different locations came from the same population could be rejected with greater than 99.9% certainty. Several things are evident from the comparison. Firstly peak lag is less frequent overall in Cambridge. There are more utterances which fall into the 'peak in' category than in the data from other locations. Secondly, peak lag is more frequent in general in Belfast than elsewhere, but extreme lag is found less here. Extreme lag meanwhile is most common in data from Newcastle and Leeds.

These results provide further evidence against a physiological explanation for peak lag; if a Cambridge speaker is able to align the peak within the accented syllable, there is no clear physical reason why a Belfast speaker cannot.

5. SUMMARY AND CONCLUSION

In this paper we have presented results which indicate a link between the occurrence of peak lag and the presence or absence of anacrusis. This link does not appear to have a physiological basis; it is not the case that speakers simply cannot attain peak height in time when there is no anacrusis on which to 'take a run-up'. A more likely explanation is that it is perceptually advantageous to have a slow rise to peak height, in order to render that peak more salient. The desire to maximise the salience of the peak may override any considerations of alignment with the accented syllable, and therefore peak lag results in some cases.

Peak lag is also clearly more prevalent in data from some locations than others. Peaks are more frequenly subject to lag in the data from Belfast, Newcastle and Leeds, than in the data from Cambridge. That category of 'extreme lag' was found to be most common in data from Newcastle and Leeds.

The results of this experiment serve to emphasize the need for different levels of intonation analysis. Whilst all the pitch accents examined here involved an H*, there was a great deal of variation in the location of the corresponding f0 peak. A phonological level of analysis allows recognition of the fact that patterns with and without peak lag may function in the same way. A phonetic level of analysis meanwhile, allows for phonetic differences in peak alignment to be noted. Furthermore, the results supplement previous work on this phenomenon, by showing that preceding prosodic context and geographical location need to be added to the list of factors influencing the phonetic implementation of pitch accents.

NOTES

1. ESRC grant R000237145, 'Intonational Variation in the British Isles'. Award holders F. Nolan and E. Grabe; research associate K.J. Farrar, University of Cambridge, 1997-2000.

REFERENCES

[1] Wichmann, Anne, House, Jill and Rietveld, Toni. Forthcoming. Discourse constraints on peak timing. In Botinis, A. (ed.) *Intonation: Analysis, Modelling and Technology*. Cambridge, UK: Cambridge University Press.

[2] Ladd, D. Robert. 1996. *Intonational Phonology*. Cambridge, UK: Cambridge University Press.

[3] Steele, S. 1986. Nuclear accent f0 peak location: effects of rate, vowel, and number of following syllables. *JASA* Supplement 1, 80, s51

[4] Silverman, Kim E.A. and Pierrehumbert, Janet B. 1990. The timing of prenuclear high accents in English. In Kingston, John and Beckman, Mary E. (eds.) *Papers in Laboratory Phonology, Vol. 1: Between the Grammar and Physics of Speech.* Cambridge, UK: Cambridge University Press, 72-106.

[5] Grabe, E., Nolan, F. and Farrar, K.J. 1998. IViE - A Comparative Transcription System for Intonational Variation in English. In Proceedings of ICSLP5, 30th November - 4th December 1998, Sydney Australia.

[6] Cruttenden, A. 1997. *Intonation*. Cambridge, UK: Cambridge University Press.

[7] Lowry, O. 1997. *Intonation Patterns in N. Irish English.* unpub. MPhil thesis, University of Cambridge.

[8] Ladd, D. Robert. 1983. Phonological features of intonational peaks. *Language*, 59, 721-759.