Tonal Alignment Defined: the case of Southern Irish English

Raya Kalaldeh

Phonetics and Speech Science Laboratory, Trinity College Dublin

kalalder@tcd.ie

Abstract

This paper proposes to define tonal alignment features as either intrinsic; the default alignment, or extrinsic; the shifts away from the default alignment due to prosodic contextual factors. Intrinsic alignment is different for pre-nuclear (PN) and nuclear (N) accents. This distinction is illustrated for a variety of Irish English (IrE), Drogheda English (DroghE) where the PN and the N peaks of H* accents are intrinsically aligned at a time point 70% ~80% and 60% ~75% into the vowel of the accented syllable, respectively. Extrinsic alignment shifts of PN and N peaks are very small not exceeding the accented vowel boundaries.

Index Terms: intrinsic alignment, extrinsic alignment, Irish English, nuclear and pre-nuclear accents

1. Introduction

The study of tonal target alignment has been the focus of many studies for the past twenty years. Many factors have been reported to influence the timing of the peak [1]. Some of these factors include the boundaries (proximity to word/phrase boundary), rhythmical organization (e.g. stress clash), focus (prefocal, focal, or postfocal position), tonal environment (e.g. tonal crowding), pitch range (e.g. differences in degree of emphasis), global intonation (e.g. absence/presence of downdrift due to interrogative/declarative structure), segmental context (e.g. vowel length, syllable structure), and speaking rate (e.g. fast, normal, slow tempo). Many past and recent studies have investigated the influence of these and other factors in different dialects and languages [2-9].

Intrinsic alignment refers to the default alignment of the H (high) and L (low) f0 turning points of a pitch accent relative to the segmental string when there are no rhythmic or temporal constraints. This notion was partially alluded to as 'invariant features' [7]. Intrinsic alignment features are specific to each dialect and language. For example, in Modern Greek the pre-nuclear rise (L*+H) has the low of the rise aligned before the accented syllable and the high is aligned near the onset of the syllable following the accented syllable [2]. These systematic and consistent segmental anchors characterise the pre-nuclear rises in Modern Greek.

Extrinsic alignment refers to the shifts away from the default alignment which potentially may arise when there are constraints imposed by the rhythmic context, tonal environment (tonal crowding, proximity to phrase boundary) and other factors including degree of emphasis, pitch range, and tempo. Many contextual factors that could influence tonal alignment have been investigated in several studies [1, 6, 8, 9]. These factors can be prosodic such as the proximity of the word boundary, the rhythmic context, tonal environment, pitch range, and focus as well as speaking rate. In this paper, such factors are referred to as extrinsic since they do not alter the identity of the pitch accent but modify it as a requirement of the prosodic context. For example, it was found that in Inis Oírr Irish [10] the nuclear peak of a fall shifts rightwards as the number of following unstressed syllables increases. Although the nuclear peak is realised within the vowel of the following post-accentual syllable, when the nuclear syllable is followed by two unstressed syllables, this does not mask the identity of the pitch accent as a nuclear fall. Hence, these shifts in Inis Oírr Irish are extrinsically conditioned and are not intrinsic.

Differences between dialects and languages may be better expressed in terms of intrinsic and extrinsic alignment features. Languages and dialects may vary in terms of their intrinsic timing differences. For example, Southern German pre-nuclear rises are aligned later than those in Northern German in the same rhythmic context [3]. There are also cross dialect differences in extrinsic alignment. For example, it was found that in Inis Oírr Irish the peak of a fall shifts when the rhythmic structure is varied, whereas in Cois Fharraige Irish peak alignment of the fall is not influenced by the same variation in rhythmic structure [5]. These differences between the two varieties of Irish are essentially extrinsic ones (even though intrinsic differences also exist).

Moreover, most alignment studies can be grouped into an investigation of factors which control **intrinsic** alignment e.g. [3, 11-13], factors that cause **extrinsic** alignment features e.g. [5, 6, 8, 9], or both e.g. [2, 7].

In this paper, the two aspects of alignment just introduced are investigated in the IrE variety of DroghE. First, the optimal rhythmic context for intrinsic alignment for H* in both PN and N accents is investigated. Next, potential extrinsic shifts in alignment are investigated when the rhythmic context changes. The rhythmic context is varied by manipulating the number of unstressed syllables before and after the accented syllable (A). Therefore, the formula '$2^{+}A2^{+}$' is proposed and tested. The formula means that the accented syllable (A), when preceded and followed by at least two unstressed syllables should yield the default, intrinsic alignment for a given syllable, in so far as it allows for a 'full' realisation of the timing characteristics of the pitch accent. The type of the context proposed in the '$2^{+}A2^{+}$' formula has been partially suggested from results presented in [2, 5, 8, 14]. Then after defining the intrinsic alignment of H* in both PN and N accents, the potential extrinsic variation is examined when the '$2^{+}A2^{+}$' context is not available.

The town of Drogheda is shown in Figure 1. The variety of DroghE is described as a Southern variety of IrE that shares many intonational features with Dublin English [15]. Declaratives in DroghE are characterised by predominantly H* pre-nuclear and nuclear accents. Therefore, the investigated PN and N pitch accent in this paper is H* in declarative statements (commonly labelled as H*L in IViE [16]).

10.21437/Interspeech.2011-452

Figure 1: *Map of Ireland and the location of Drogheda*

2. Methods and Materials

2.1. Materials

The study is based on data from seven male adolescent speakers from Drogheda. For H* in both initial PN and N accents in this study, the following questions are addressed:

1. **When** do the PN and N intrinsic alignments of the H* accent in DroghE manifest?
2. **What** are the characteristics of the PN and N intrinsic alignments for such an accent?
3. **How much extrinsic variation** is there in DroghE if the rhythmic context is varied?

The rhythmic context for each PN and N accent is varied by changing the number of unstressed syllables 'preceding' the accented syllable from zero to three (0-3) and the number of unstressed syllables 'following' the accented syllable from zero to three (0-3). This gives four sentences for each of the four sets as shown in Table 1. If the size of unstressed material preceding the accented syllable was investigated, the size of unstressed material following the accented syllable is fixed to two unstressed syllables and vice versa. The manipulated unstressed material is underlined in Table 1. The accented syllable in all sentences is the CVC syllable /man/. For ease of reference, the PN sets are referred to as A and B and the N sets as Ȧ and Ḃ.

Table 1. *Speech material for PN and N syllables.*

	Preceding syllables	Following syllables
PN	(A) *(0) **Man** is an omnivore.* *(1) The **man** in the lobby is waiting for you.* *(2) There's a **man** in the lobby for you.* *(3) There was a **man** in the lobby for you.*	(B) *There's a **man** waiting for you. (0)* *There's a **man** in Navan. (1)* *There's a **man** in the lobby for you. (2)* *There's a **man** on the veranda for you. (3)*
N	(Ȧ) *(0) Mel has a new **man**dolin.* *(1) Mel plays the **man**dolin.* *(2) Mel is playing the **man**dolin.* *(3) Mel is playing on the **man**dolin.*	(Ḃ) *I'm meeting the **man**. (0)* *He's handsome and **man**ly. (1)* *It's an example of **man**liness. (2)* *We need some more **man**liness here.(3)*

Each test utterance was prompted separately on a computer screen. The utterances were randomized in five lists. Utterances unsuitable for analysis were excluded. Speakers were recorded in a quiet room using a digital recorder, ZOOM Handy Recorder H4. Data was analysed using the PRAAT software [17].

2.2. Measurements

Peak alignment (H) is expressed as a proportion and measured as the duration of the interval from the onset of the accented vowel /a/ to the peak location, relative to the duration of the vowel. For the PN accented syllables the alignment of the preceding low (l) was also measured. The alignment of (l) was measured as the duration of the interval from the onset of the accented syllable to the low location, relative to the duration of the onset-consonant, /m/.

The points of measurement are shown in Figure 2 where (H) is the location of f0 peak and (l) is the location of f0 low preceding the peak. The durations of the unstressed syllables and the segments of the accented syllable were measured and averaged across the seven speakers in (ms). It should be noted that alignment of the preceding low was measured only in PN accents as its location was not as clear in N accents. This was due to the declination effect that was apparent towards the end of the utterance. All measurements were hand-checked and corrected prior to analysis.

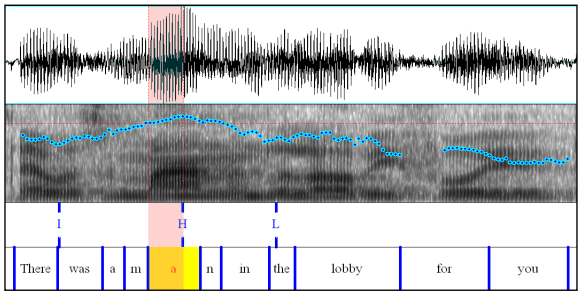

Figure 2: Peak alignment measurements taken for the peak on 'man' in 'There was a man in the lobby for you', The yellow bar is the accented vowel /a/. The pink bar is the interval from the onset of the accented vowel to the peak location (H). The point (l) is the low preceding the peak.

3. Results

Figure 3 shows peak alignment in the two different rhythmic conditions for both PN and N accents. Panels A and B show the influence of the number of unstressed syllables 'preceding' and 'following' PN /man/, respectively. Panels Ȧ and Ḃ show the N peak alignment in the 'preceding' and 'following' rhythmic conditions, respectively. The boxes in the figures represent the actual averaged syllable and segment durations in (ms). The percentages are the proportional peak alignment represented in upward arrows (↑). All unstressed syllables are shown in grey boxes. The manipulated unstressed syllables are provided in text. The fixed (not manipulated) unstressed syllables are marked as: s = following the accented syllable, sp = preceding the accented syllable. The beginning of the accented syllable is set to zero (ms) for ease of comparison across different conditions.

The general linear model is used for statistical analyses. The two fixed factors are 'the number of preceding/following unstressed syllables' and 'speaker'. The values used in the statistical analyses are the proportional values of (H).

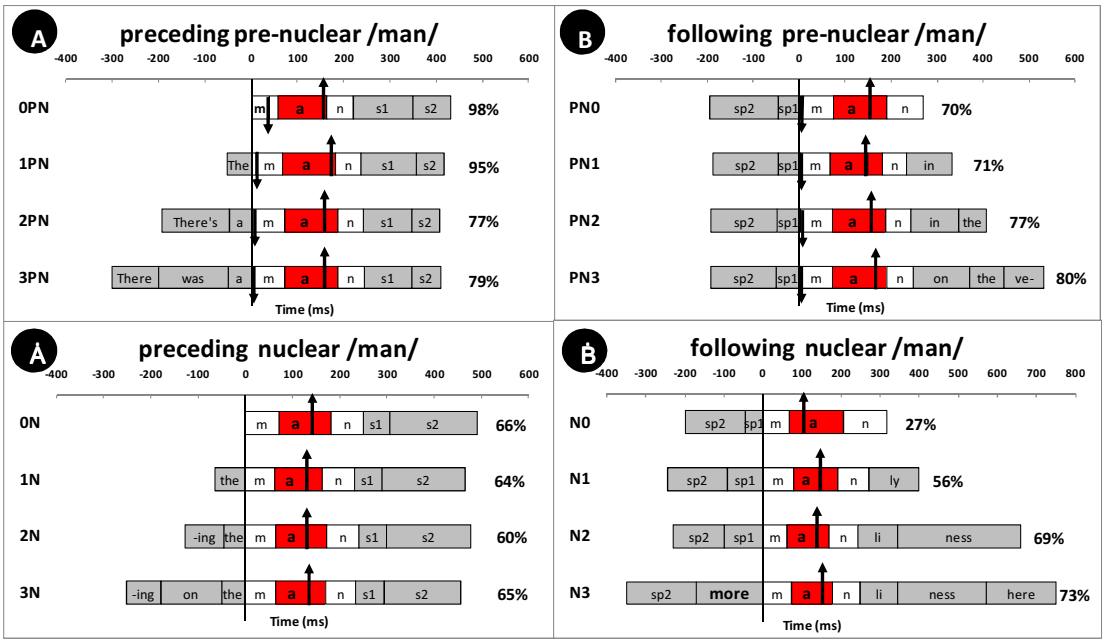

Figure 3: Averaged proportional alignment (l = ↓ , H = ↑) given in percentages for the accented syllable /man/ across the seven speakers. Unstressed syllables that are manipulated are provided in text. Fixed unstressed syllables are: s = following syllables, sp = preceding syllables.

3.1. Pre-nuclear peak alignment

In Panel A, it is observed that the (H's) in both 0PN and 1PN are different from those in 2PN and 3PN. Peak alignment values seem to reflect this difference. In B, peak alignment is constant across all conditions. This is observed when the accented PN syllable is preceded by a constant number of unstressed syllables in all conditions.The general trend of increasing (H) precentages in B is due to the PN intrinsic alignment of a single speaker whose alignment is later than all other speakers. When this speaker's values are excluded the values become: PN0 (68%), PN1 (73%), PN2 (72%), and PN3 (76%).

The general observations are confirmed by statistical tests. In A, the differences of (H) values between different conditions are significant: $F(3,106) = 7.527$, $p = 0.0001$. Bonferroni tests show that both 0PN and 1PN are different from both 2PN and 3PN. This is reflected in the averaged (H) values as shown in A. In B, the difference in (H) values across the conditions is not significant: $F(3,99) = 2.009$, $p = 0.118$.

3.2. Nuclear peak alignment

In Ȧ, it is noticed that (H) is aligned within the accented vowel and is not different across the different conditions. In Ḃ, (H) is also aligned within the accented syllable but in the N0 and N1 conditions, it is earlier than in the N2 and N3 conditions. It should be noted that in the test sentence in N3 (*We **need** some **more** **man**liness here*) the word *more* was unexpectedly accented. This changes the rhythmic context into a preceding tonal crowding instead of two preceding unstressed syllables as in all other conditions in Ḃ. This renders this sentence uncomparable with the other conditions in Ḃ. However, this sentence is not excluded as it appears that the preceding context in N accents does not significantly influence H* peak alignment as will be shown statistically.

In Ȧ, the differences of (H) values between different conditions are not significant: $F(3,82) = 0.6337$, $p = 0.5954$. In Ḃ, the differences in (H) values across the different conditions are significant: $F(3,87) = 38.674$, $p \leq 0.0001$. Bonferroni tests in Ḃ show that all conditions are significantly different from one another except for N2 and N3.

3.3. Pre-nuclear and nuclear peaks compared

The general observation is that the N peaks are earlier than the PN ones. The (H) values across PN and N accented syllables were compared within the individual rhythmic contexts i.e. for A vs. Ȧ and B vs. Ḃ, e.g. 1PN vs. 1N, PN2 vs. N2. Statistical results show that in all individual contexts the (H) values in N accents are significantly earlier than in PN peaks. The only two exceptions were in B-Ḃ. Peak alignment in N2 was earlier than that in PN2 but not significantly: $F(1,50) = 4.135$, $p = 0.473$. Also, peak alignment differences between PN3 and N3 were not significant: $F(1,36) = 2.17$, $p = 0.149$. It is not clear why the peaks are not significantly earlier in N3 than their corresponding PN3 peaks. It is unlikely that the accented syllable *more* is the reason. Preceding tonal clash/boundary did not show influence on nuclear peaks in the 0N condition in Ȧ where the peak alignment was within the vowel and similar to all other conditions in Ȧ.

4. Discussion

The statistical findings largely support the trends shown. Although, the peak remains within the accented syllable in all conditions in DroghE, the fine-detailed differences across different conditions are very systematic and indicative.

To answer the first question posited earlier, the findings in A show that it is the size of the anacrusis that is crucial for peak alignment of H* in PN accented syllables and not the

following unstressed material. When the size of the anacrusis is manipulated and the unstressed material following the PN syllable is kept constant, peak alignment shifts occur. The findings in B reinforce this conclusion. In B, no significant differences in alignment occur, even in the condition of a tonal crowding, PN0.

In A, peak alignment becomes stable when there is more than one unstressed syllable preceding the PN accented syllable. If another unstressed syllable is added, peak alignment remains at the same location within the segmental string. Therefore, in terms of the '$2^{+}A2^{+}$' formula, the base formula for obtaining intrinsic alignment in H* for PN accents in the accented syllable /man/ in DroghE is $2^{+}PN$.

The findings in Ḃ suggest that the size of the 'following' unstressed material is crucial for N peak alignment of H*, while that of the 'preceding' unstressed material is not. When the size of the following unstressed material is manipulated and the unstressed material preceding the N syllable is kept constant, peak alignment shifts occur. The findings in Ȧ reinforce this conclusion. In Ȧ, no significant differences occur even in the condition of a preceding tonal crowding, 0N.

Moreover, in Ḃ peak alignment in N0 is different from all other conditions suggesting that time pressure from the 'following' boundary tone necessitates that the nuclear peak be realised earlier. The (H) values in N3 are not significantly different form those in N2. However, the N1 condition is different from both N2 and N3. This shows that it takes more than two 'following' unstressed syllables for the nuclear peak of H* to be anchored at a specific location within the segmental string, in DroghE. Therefore, in terms of the '$2^{+}A2^{+}$' formula, the base formula for obtaining intrinsic peak alignment of H* in the nuclear accented syllable /man/ in DroghE is $N2^{+}$. If more than one unstressed syllable is added after the accented syllable, peak alignment manifests. If a third unstressed syllable is added, the peak remains at its location. From observations of other material from DroghE, the location of the H* peak stays within the accentual vowel even if more than three unstressed syllables followed the nuclear syllable.

The answer of question two appears to be that in DroghE the peak in H* for PN and N accents in CVC syllables is intrinsically aligned at a timepoint typically 70% ~80% and 60% ~75%, respectively, into the vowel of the accented syllable.

After establishing the intrinsic alignment of H* and its optimal context for PN and N accents in CVC syllables, question three can be answered and extrinsic alignment shifts can be evaluated. The observed peak shifts within the accented vowel in 0PN, 1PN, N0, and N1 show that peak shifts do not cross the accented syllable boundaries in DroghE. This small, but important, amount of variation leads us to characterise DroghE as dialect that shows relatively little extrinsic variation as a result of varying the rhythmic context. In this respect DroghE is similar to Dublin English [17] and Coir Fharraige Irish [5] which do not show extrinsic alignment shifts outside the boundaries of the accented syllable when the rhythmic context is varied. Conversely, DroghE is different from Belfast English [8] and Inis Oírr Irish [5], both of which show peak 'delay' outside the accented syllable boundaries when the rhythmic structure is varied.

5. Conclusions

Two distinct features of alignment were introduced and illustrated: intrinsic and extrinsic. These two types of tonal alignment features are seen as needed to better understand the working of alignment and intonation across and within dialects and languages.

Intrinsic alignments of PN and N peaks for H* accents in DroghE were investigated by testing the '$2^{+}A2^{+}$' formula. For H* accents in DroghE, the formula which shows PN intrinsic alignment is $2^{+}PN$, while that which shows N intrinsic alignment is $N2^{+}$. This entails that the size of 'preceding' unstressed material influences PN peak alignment while the size of 'following' unstressed material influences N peak alignment.

The intrinsic alignment of the H* accent in N peaks is earlier than that in PN peaks. This well established phenomenon in alignment studies is reported to be likely due to tonal repulsion from the following boundary tone. The present study confirms this reasoning.

6. References

[1] Bruce, G., "Alignment and composition of tonal accents: comments on Silverman and Pierrehumbert's paper," in *Papers in Laboratory Phonology I: between the grammar and physics of speech*, J. Kingston and M. Beckman, Eds., ed Cambridge: Cambridge University Press, 1990, pp. 107-114.

[2] Arvaniti, A., Ladd, R., and Mennen, I., "Stability of tonal alignment: the case of Greek prenuclear accents," *Journal of Phonetics,* pp. 3-25, 1998.

[3] Atterer, M. and Ladd, R., "On the phonetics and phonology of "segmental anchoring" of F0: evidence from German," *Journal of Phonetics,* vol. 32, pp. 177-197, 2004.

[4] Bruce, G., "Swedish word accents in sentence perspective," Ph.D dissertation, Lund, Gleerup, 1977.

[5] Dalton, M. and Ní Chasaide, A., "Melodic alignment and micro-dialect variation in Connaught Irish," in *Tones and Tunes: Studies in Word and Sentence Prosody*. vol. 2, C. Gussenhoven and T. Riad, Eds., ed Berlin: Mouton de Gruyter, 2007, pp. 293-315.

[6] Knight, R.-A. and Nolan, F., "the effect of pitch span on intonational plateaux," *JIPA,* vol. 36, pp. 21-38, 2006.

[7] Ladd, R., Faulkner, D., Faulkner, H., and Schepman, A., "Constant "segmental anchoring" of F0 movements under changes in speech rate," *Journal of the Acoustical Society of America,* vol. 106, pp. 1543-1554, 1999.

[8] Nolan, F. and Farrar, K., "Timing of f0 peaks and peak lag," in *ICPhS*, San Francisco, 1999, pp. 961-964.

[9] Silverman, K. and Pierrehumbert, J., "The timing of prenuclear high accents in English," in *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*, J. Kingston and M. Beckman, Eds., ed Cambridge: Cambridge University Press, 1990, pp. 72-106.

[10] Dalton, M., "The phonetics and phonology of the intonation of Irish dialects," Ph.D dissertation, Centre for Language and Communication Studies, Trinity College Dublin, Dublin, 2008.

[11] Ladd, D. R. and Schepman, A., ""Sagging transitions" between high pitch accents in English: experimental evidence," *Journal of Phonetics,* vol. 31, pp. 81-112, 2003.

[12] Dilley, L., Ladd, R., and Schepman, A., "Alignment of L and H in bitonal pitch accents: testing two hypotheses," *Journal of Phonetics,* vol. 33, pp. 115-119, 2005.

[13] Ladd, D. R., Schepman, A., White, L., Quarmby, L. M., and Stackhouse, R., "Structural and dialectal effects on pitch peak alignment in two varieties of British English," *Journal of Phonetics,* vol. 37, pp. 145-161, 2009.

[14] Dalton, M. and Ní Chasaide, A., "Tonal alignment in Irish dialects," *Language and Speech: Special Issue on Intonation in Language Varieties,* vol. 48, pp. 441-464, 2006.

[15] Kalaldeh, R., "Segmental and intonational analysis of Drogheda English," Ph.D dissertation, Centre for Language and Communication Studies, Trinity College Dublin, Dublin, 2011.

[16] Grabe, E., Post, B., and Nolan, F., "Modelling intonational variation in English: the IViE system," in *Proceedings of Prosody 2000*, Poznan, Adam Mickiewicz University, Faculty of Modern Languages and Literature, 2001, pp. 51-57.

[17] Kalaldeh, R., Dorn, A., and Ní Chasaide, A., "Tonal alignment in three varieties of Hiberno-English," in *Interspeech*, Brighton, 2009, pp. 2443-2446.