TIMING PATTERNS IN WELSH

Dafydd Gibbon & Briony Williams

Universität Bielefeld, Germany
gibbon@uni-bielefeld.de

University of Wales, Bangor, UK
b.williams@bangor.ac.uk

ABSTRACT

Studies of duration in Welsh have concentrated on specific environments, such as word-final, pre-final and pre-pre-final positions, in relation to the positions of linguistic stress and phonetic pitch accent.

The present approach is broader, in that it looks at the temporal patterning of syllable and foot sequences. First, general global and local measures of isochrony and irregularity are applied (standard deviation; pairwise variability). Second, since isochrony is a necessary but not a sufficient correlate of rhythm, a new algorithm is introduced for characterising grouping properties. Third, syllable sequences are segmented into feet using alternative duration relations and examined for their relation to grammatical groups.

It appears that Welsh duration patterns are neither syllable-timed nor foot-timed in the accepted sense, but have what may be called a "rallentando timing", in which sequences of increasing length mark grammatically relevant prosodic units.

Keywords: timing, duration, rhythm, Welsh

1. INTRODUCTION

The motivation for the present study lies in three areas:

1. further investigation of a little-studied phonetic domain, timing in Celtic languages, specifically in Welsh;
2. provision of systematic information on Welsh duration patterns for rule-based diphone synthesis;
3. trial of new concepts and algorithms for the study of timing and rhythm.

Intuitively, the rhythm of Welsh is very different from that of English; indeed speculation is rife about the typological similarity of Welsh prosody and the prosody of the Eastern Indo-European languages of India. Previous database-oriented work has either not considered prosody, as in the SpeechDat work by Jones & al. [4] or has restricted attention to specific environments such as the final, pre-final and pre-prefinal word-internal positions in relation to linguistic stress and phonetic pitch accent location, cf. Williams [8]. The same applies to experimental work by Williams [7].

The present study examines temporal patterning in syllable and foot sequences, with a long-term view to placing Welsh within a prosodic typological space, in relation to parameters such as rhythm, isochrony, and syllable or foot timing. Rather than using one measure as in a number of recent studies, a number of approaches are applied:

1. Evenness measures: phonetic studies have traditionally used evenness (conversely: irregularity) measures for determining rhythmic timing patterns. Measures such as those of Scott & al. [6] and Roach [5] are either global and variance-like, or local distance measures, e.g. the Pairwise Variability Index, PVI, of Nolan; cf. Grabe & Low [3]. Evenness is a necessary, but not a sufficient condition for rhythmic timing, so further measures are needed.
2. Another necessary but likewise not suffficient condition for rhythm is alternation. In addition to the PVI, a simple algorithm for chunking syllables into possibly alternating quantitative feet is therefore introduced.
3. The syllable chunks are further examined for linguistic functionality in relation to grammatical grouping.

2. WELSH PROSODY

The regular word prosodic system of Welsh is taken to have the following characteristics by Williams [9]. In polysyllabic content words, the default situation is as follows:

1. the penultimate syllable is associated with the linguistic stress;
2. the final syllable is associated with
    1. a high pitch peak,
    2. possibly a phonologically long vowel.

The "delayed" pitch peak is possibly a residue of an earlier Welsh pitch accent. A similar pattern is

familiar from Welsh-accented English, in which a delayed pitch peak occurs on the syllable following the lexically stressed syllable. In monosyllabic content words in Welsh, all features, e.g. pitch peak and length, occur on the single syllable.

Assuming that stressed syllables, if not immediately associated with a pitch peak, are indicated in some way by timing, three hypotheses are formed on the basis of this outline:

1. Welsh rhythm is not syllable-timed and may be foot-timed (or it may be more irregular than this or have some other foot-like pattern).
2. Welsh timing patterns tend to be iambic, or, in more general terms, right-heavy, i.e.
   |˘ - ˘ - ˘ -|, |˘ ˘ - ˘ ˘ - ˘ ˘ - |, ... etc.
   in traditional micron/macron notation.
3. Grammatical units are associated with clear timing patterns.

3. DATA

The speech data are taken from readings of lengthy newspaper and magazine extracts collected and annotated during the years 1996-2000; cf. Williams [8] and [10]. The database was constructed for basic phonetic research and is annotated with a hierarchy of phonetic and linguistic levels. For the present preliminary study, a selection was made: 1 speaker, female, age 19, non-smoker, no known voice or hearing problems; 173 utterances (actually 175, 2 unusable for technical reasons; syllable level annotations. This female speaker (recorded in 1998) was born in 1978, and brought up in Capel Curig, North Wales. Her mother spoke Welsh with a Cardiganshire accent, her father was an English speaker; she spoke Welsh and English at home. Her school education was in Welsh; she spoke English while at an English university, where she was active in the Welsh society. A data sample:
Ond pam fod cymaint o wleidyddion mor hoff o bregethu gwerthoedd teuluol heddiw?
"But why are so many politicians so fond of preaching family values today?"

4. METHODS

4.1. Even syllable timing: syllable durations

Table 1 summarises the main timing properties for the 3034 syllables in the corpus selection; the measurements are in seconds. In relation to the outliers, mean and median are fairly close, indicating a rough balance of durations. The standard deviation (SD) of 0.08 indicates considerable variability, so that a tentative conclusion can be drawn that variability of syllable duration in Welsh is too great for syllable timing.

Table 1: Syllable duration properties.

N:	3034
Min duration:	0.0091
Max duration:	0.5207
Mean duration:	0.1841
Median duration:	0.1748
SD of durations:	0.0872

4.2. Even syllable timing: standard deviation

Many variance-like methods have been used to examine syllable timing. The simplest is perhaps just variance, or standard deviation; more sophisticated is the PVI model. The approaches of Scott & al. [6] and Roach [5]; can be approximated using standard deviation; results are shown in Table 2 (units: seconds). Overall deviation of deviations is small, despite the low outlier.

Table 2: Variation in standard deviations.

N:	173
Min SD:	0.0009
Max SD:	0.0932
Mean SD:	0.0266
Median SD:	0.0225
SD of SDs:	0.0207

4.3. Even syllable timing: normalised PVI

A second method for investigating Welsh syllable timing was used: the popular normalised PVI model of local evenness of durations with rough normalisation for speech rate variation. The nPVI is a distance measure: the average normalised distances between 2 adjacent units in a sequence (in this case syllables), multiplied by 100 for readability. Normalised distance is the difference between durations of 2 adjacent units divided by their average duration. The nPVI values range from 0, i.e. totally even, towards an asymptote of 200, i.e. increasingly irregular; cf. Gibbon [1]. For the 173 utterances in the corpus selection, the nPVI results are shown in Table 3.

The nPVI range, 33-105=72, is rather large, and presumably due to rhetorical variation in the

reading performance of the textual material. However, with a standard deviation of 15, and very close mean and median, overall nPVI variation among the utterances is not excessive.

Table 3: Normalised Pairwise Variability Index (nPVI) properties.

N:	173
Min nPVI:	33
Max nPVI:	105
Mean nPVI:	61
Median nPVI:	60
SD of nPVIs:	15

The mean nPVI of 61 is higher than that found by Grabe & Low [3] for Welsh, with a vocalic interval nPVI of 48.2, n=152, and a non-normalised nPVI of 54.7, n=150 for non-vocalic intervals. Grabe & Low show a vocalic nPVI range of about 30-70, and an intervocalic nPVI range of about 40-80 for a wide range of language types. Welsh is located about mid-way, nearer to the French type than to the English or German types.

Grabe & Low do not show syllable nPVI or deviation values, but, speculating that syllable nPVI will be close to the mean of vocalic and intervocalic nPVI, and noting that the Grabe & Low values are within 1 SD of present values, the difference is not great, and perhaps due to different data types (with factors such as age, gender, dialect, genre).

4.4. Iambic and trochaic chunking

A simple chunking algorithm was used, which groups items of increasing length together by adding boundary markers. This simple algorithm promises more structural information than the evenness measures, and more direct interpretability than the complex tree construction algorithms introduced by Gibbon [1]. Initial inspection showed that this iambic chunking procedure shows a fairly close relationship to grammatical grouping, and this is the direction which was pursued further.

Tentative examination of the results of trochaic chunking (chunks of decreasing length) showed no relationship with grammatical structures. Consequently this variant was dropped; it is conceivable that it may relate to other prosodic facts, however. Iambic duration chunking is illustrated for the following utterance:
Dydy'r wraig ddim yn hapus ymhlith y bobl hon
"The woman is not happy among these people"
/dədɨr wraɪg ðɪm ən hapɨs əmɬɪθ ə boːbɒl hɒn/
Durations and labels are in square brackets and separated by a colon, followed by a chunk-final boundary marker:

```
[0.047:d@][0.218:dyr][0.378:rwaig]#
[0.2:dhim]#
[0.1:@n][0.1:hha][0.3:pYs]#
[0.1:@m][0.2:lhith]#
[0.04:@][.:bo][0.2:bol][0.3:hhon]#
```

The following results emerge. The 3034 syllables grouped into 1476 feet. This average does not mean, however, that duration patterns are largely binary: 1338 syllables are grouped in pairs, less than half, and there are 858 in 286 triples, 340 in 85 quadruples, 65 in 13 quintuples, 12 in 2 sextuples, and 421 feet are monosyllabic. Most of the end-heavy chunks are thus either binary (iambic) or ternary (anapaestic), but there are also longer and shorter chunks.

4.5. Grammatical grouping

In order to determine possible matches between quantitative feet and grammatical units, the chunks were manually tagged according to whether they ended in lexically a stressable word:

- Y: Foot unit ends in stressable word.
- N: Foot does not end in a stressable word (e.g. on non-final syllable of polysyllabic content word.
- U: Foot ends in stressable word but begins in non-initial syllable of polysyllabic content word (i.e. immediately after "N" foot).

The proportions of Y, N and U judgments, and alternative assignments of U judgments to Y or N, are shown in Table 4.

Table 4: Foot-grammar correspondences.

Tag	*Number*	*Percent*
Y:	729	49%
N:	444	30%
U:	303	21%
Total feet:	1476	100%
Y+U:	1032	70%
N+U:	747	51%

The Y tags show exact matches between duration feet and grammatical groups. Typical

matching grammatical groups are the following, where "Verbconj" means conjugated verb, eg. *gafodd*, while "Verbnoun" is the root form, e.g. *cael*:

(Det) N Adj
yn/wedi Verbnoun
yn Adj
PossPron N
(yn/wedi) Verbnoun Det N
Verbconj Det N

An exact match, MY, is found in about half the cases, 49%. This is a chance result if the U tags count as incorrect, with N+U as MNU=51%. However, if the U tags, which indicate partial matches, also count as correct, then Y+U, with MYU=70%, indicates: *A sequence of increasing syllable durations tends to end with a stressable word.* The converse is not true, of course: a stressable word does not necessarily end a sequence of increasing syllable durations. A middle road can be taken by weighting full and partial matches, e.g. with a model such as

$$MW=100*(|Y|+|U|/2)/n =60\%$$

This is not a spectacular match, but indicates a tendency which may be combinable with non-durational prosodic features for use in prosodic parsing.

5. CONCLUSION AND FURTHER WORK

Duration patterns in a single-speaker read corpus of Welsh were examined with different methods. The results show that Welsh syllable durations are moderately variable, but do not reach the variability of so-called foot-timed (stress-timed) languages. Within the limits of the data set, and referring to the three hypotheses detailed at the outset, the following has been tentatively shown. First, Welsh does not show syllable-timing and may be foot-timed in the sense of the quantitative feet defined here. Second, Welsh timing patterns tend to be right-heavy, but not solely binary (strictly iambic): they include feet of other lengths. Third, timing patterns tend to be associated with grammatical units, though this relationship may not necessarily be reversible.

The timing structure described in this contribution differs both from syllable-timing and foot-timing in the conventional isochrony-based sense. Syllable chunks of increasing length seem to be the relevant items, a low level "decelerando" phenomenon which is reminiscent of final lengthening phenomena. Further work is in progress on how generalisable this pattern is in relation to other data types, to Welsh syntax, and to other aspects of Welsh prosody such as the delayed pitch peak.

6. Acknowledgements

The speech recordings formed part of a project which was funded by the UK's Engineering and Physical Sciences Research Council, under research grant no. GR/K84356. The authors thank the two anonymous ICPhS reviewers for their very helpful input, most of which we have been able to accommodate.

7. REFERENCES

[1] Gibbon, Dafydd. 2006. Time types and time trees: prosodic mining and alignmment of temporally annotated data. In: Sudhoff, Stefan, Denisa Lenertová, Roland Meyer, Sandra Pappert, Petra Augurzky, Ina Mleinek, Nicole Richter & Johannes Schließer (eds.), *Methods in Empirical Prosody Research*. Berlin: Walter de Gruyter, 281-309.

[2] Gibbon, Dafydd & Flaviane Romani Fernandes. 2005. Annotation-mining for rhythm model comparison in Brazilian Portuguese. *Proc. Interspeech/Eurospeech* 2006, 329-3292.

[3] Grabe, Esther & Ee Ling Low. 2002. Durational Variability in Speech and the Rhythm Class Hypothesis. In Gussenhoven, C. and Warner (eds.), N. *Studies in Laboratory Phonology* VII, CUP, 515-546.

[4] Jones, Rhys James, John Mason, Robert Owen Jones, Louise Helliker & Mark Pawlewski. 1998. SpeechDat Cymru: a large-scale Welsh telephony database. *Proc. LREC Workshop on Language Resources for European Minority Languages*, Granada, Spain, 1998.

[5] Roach, Peter. 1982. On the distinction between 'stress-timed' and 'syllable-timed' languages. In: Crystal, David (ed.), Linguistic Controversies: *Essays in Linguistic Theory and Practice*, London: Edward Arnold, 73–79.

[6] Scott, Donia R., Stephen D. Isard, and Bénédicte de Boysson-Bardies. 1986. On the measurement of rhythmic irregularity: a reply to Benguerel. *Journal of Phonetics* 14, 327–330.

[7] Williams, Briony. 1985. Pitch and duration in Welsh stress perception: the implications for intonation. *Journal of Phonetics*, 13, 381-406.

[8] Williams, Briony. 1999a. A Welsh speech database: preliminary results. *Proceedings of Eurospeech 99*, Budapest, Hungary.

[9] Williams, Briony. 1999b. The phonetic manifestation of stress in Welsh. In: van der Hulst, Harry (ed.), *Word Prosodic Systems in the Languages of Europe*. Berlin & New York: Mouton de Gruyter.

[10] Williams, Briony. 2000. A practical method for deriving phonetic and linguistic data for minority languages. *Literary and Linguistic Computing*, 15, 445-463.